Graduate Thesis:

Incorporating Function into ββα-Motif Peptide Scaffolds

Michael Aaron Shogren-Knaak

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Division of Chemistry and Chemical Engineering

California Institute of Technology
Pasadena, California

30 May, 2000

In memory of my mother

Acknowledgements

First and foremost I would like to thank my advisor, Barbara Imperiali. Her enthusiasm and demand for excellence have driven me to become a better scientist. Her expertise and patience have allowed me to become one.

Thanks to Kevin McDonnell, my collaborator on the aldol condensation project. His scientific acumen and creativity never cease to amaze me, and he's made this past year and a half more fun than I would have ever expected.

Thanks to Jennifer Ottesen. Her expertise and help running and interpreting two-dimensional NMR has been invaluable.

Thanks to Ranabir Sinha Roy for all the time he spent helping me get started on the functional peptide project.

Thanks to all of the past and present members of the Imperiali group. When I was looking for a group to join, the Imperiali group drew me in because its scientific talent seemed unparalleled. Five years later, I'm struck not only by how correct my initial impressions were, but also by how lucky I was to find a group that is unparalleled on a personal level.

Thanks to Gabriel Brandt. I'll cherish our hours of discussion, scientific and otherwise, as much as anything I've accomplished at Caltech. I wish him the best of luck and will miss him.

Thanks to my father for his unflagging enthusiasm for chemistry. Our discussions and his encouragement have meant a lot.

Thanks to my wife, Traci Shogren-Knaak. Her unconditional support and help throughout graduate school have made all the difference in my making it through. I can't imagine being any luckier.

Thanks to Barbara Imperiali, Traci Shogren-Knaak, Kevin McDonnell, Gabriel Brandt, Sarah O'Connor, and Adam Mezo for all of their help with my thesis and research proposals.

Finally, thanks to the Howard Hughes Medical Institute for supporting my research with a Howard Hughes Medical Institute Predoctoral Fellowships in the Biological Sciences.

Abstract

The generation of functional biomolecules constitutes an important and challenging goal in bioorganic chemistry. Efforts to incorporate chemical functionality into a ββα-motif (BBA) peptide scaffold are presented. In the course of this work we have utilized this system to help determine to what extent folded peptide scaffolds can support chemical function, which strategies are best suited to the discovery of functional peptides, and what can be learned about biological catalysis. To address these issues several different strategies involving the BBA peptide scaffold have been explored.

Initial efforts focused on using the BBA peptide scaffold to modulate the reactivity of a non-natural pyridoxamine-associated amino acid (Pam). The Pam residue was incorporated at three different positions in the BBA peptide, and different potential general acid and general base residues were included. These peptides, BP1-BP6, showed improved transamination rates and stereoselectivity relative to a model pyridoxamine compound. Additionally, one peptide, BP5, maintained many of the secondary and super-secondary structural features of the BBA peptide scaffold. A second generation of Pam-containing peptides, peptides CBP01-CBP18, was constructed based on BP5 to investigate the effects of systematic amino acid changes in the vicinity of the pyridoxamine functionality. These peptides also exhibited enhanced rates of transamination and stereoselectivity and showed clear trends in stereoselectivity as a function of the amino acid substitutions. Finally, peptide systems utilizing other mechanisms of influencing pyridoxamine-mediated transamination were investigated.

Subsequent work, undertaken in collaboration with Kevin McDonnell, sought to address some of the limitations of the Pam amino acid approach. This effort focused on developing high-throughput techniques for finding BBA peptides capable of mediating aldol condensation. A library of over 100,000 different BBA peptides, the CPLB peptides, was designed to incorporate one of several possible basic residues within a core of many potential residues. These peptides were synthesized and assayed for their ability to sequester a fluorescent

probe, DMED, which resembles the reaction intermediates in aldol condensation. The "winners" of this screen exhibited a high degree of sequence similarity, exhibiting β-diaminoproprionic acid (Dap) exclusively as the basic residue, and favoring aromatic, hydrophobic amino acids as the neighboring residues. Structural and functional characterization of one of these "winners," peptide CPLB-A2, was performed. Despite the inclusion of the Dap group within the hydrophobic core, this peptide appeared to be a monomeric species with a high degree of the secondary and super-secondary structure expected of a BBA peptide. This peptide demonstrated an enhanced ability to sequester DMED relative to Dap-containing model peptides and had the capacity to generate an enamineone reaction intermediate.

In the course of these efforts, techniques were created to aid in the purification and characterization of *N*-terminally capped peptides. An affinity purification method involving a reversible biotin-based capping group was developed to aid in the isolation of *N*-terminal glycolyl-capped peptides. This method proved practical for the efficient parallel purification of peptides. A capping method involving an α-chloroacetyl group was developed to generate *N*-terminally capped peptides that were not only stable to standard assay conditions but also capable of being transformed to a form compatible with Edman sequencing. This method proved to be effective for the identification of unknown *N*-terminally capped peptides.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Schemes

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

Standard three-letter and one-letter codes are used for the naturally occurring amino acids.

2D	:	2-Dimensional
Abs	:	Absorbance
Ac	:	Acetyl
amu	:	Atomic Mass Unit
BBA	:	ββα-motif
Boc	:	*tert*-Butoxycarbonyl
Bzl	:	Benzyl
C-	:	Carboxy-
CAPS	:	3-Cyclohexylamino-1-propanesulfonic Acid
CAPSO	:	3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic Acid
Cbz	:	Carbobenzyloxy
CD	:	Circular Dichroism
Dap	:	β-Diaminoproprionic Acid
Dab	:	γ-Diaminobutyric Acid
DCM	:	Dichloromethane
DIEA	:	Diisopropylethylamine
DMED	:	Dansyl Dione Probe
DMF	:	*N, N*-dimethyformamide
DMSO	:	Dimethylsulfoxide
dPam	:	5′-Deoxypyridoxamine
EDTA	:	Ethylenediaminetetraacetic Acid
ee	:	Enantiomeric Excess
ESI-MS	:	Electrospray Ionization Mass Spectrometry
Fen	:	2-(1,10-Phenanthrol-2-yl)-L-alanine
Fmoc	:	9-Fluorenylmethoxycarbonyl
FmocOSu	:	9-Fluorenylmethoxycarbonylsuccinimidyl Carbonate
HATU	:	*O*-(7-azabenzotriazol-1-yl)-*N,N,N′,N*'-tetramethyluronium

		hexafluorophosphate
HBTU	:	*O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
Hcy	:	Homocysteine Amino Acid
HEPES	:	*N*-(2-Hydroxyethyl)piperazine-*N*'-(2-ethanesulfonic acid)
HEPPS	:	*N*-(2-Hydroxyethyl)piperazine-*N*'-(3-propanesulfonic acid)
HOAT	:	1-Hydroxy-7-azabenzotriazole
HOBT	:	1-Hydroxybenzotriazole
HPLC	:	High Pressure Liquid Chromatography
LDA	:	Lithium Diisopropylamine
N-	:	Amino-
Me	:	Methyl
MW	:	Molecular Weight
NMR	:	Nuclear Magnetic Resonance
NOE	:	Nuclear Overhauser Effect
NOESY	:	Nuclear Overhauser Effect spectroscopy
OPA	:	Orthophtaldialdehyde
OPfp	:	Pentafluorophenyl
Orn	:	Ornithine
Pal	:	Pyridoxal-Amino Acid
Pam	:	Pyridoxamine-Amino Acid
PEG	:	Polyethyleneglycol
PEGA	:	Polyethylene-glycol-dimethylacrylamide
Pip	:	Piperidine
Pos	:	Position
ppm	:	Parts per million
PS	:	Polystyrene
res	:	Residues
ROESY	:	Rotating-Frame Nuclear Overhauser Effect Spectroscopy
SPPS	:	Solid Phase Peptide Synthesis
TEA	:	Triethylamine

tert	:	Tertiary
TFA	:	Trifluoroacetic Acid
THF	:	Tetrahydrofuran
TLC	:	Thin Layer Chromatography
TOCSY	:	Total Correlation Spectroscopy
TOF	:	Time-of-flight
Tris	:	Tris-(2,6-dimethyphenyl)phosphate
UV	:	Ultraviolet
Vis	:	Visible
xs	:	Excess

Chapter 1: Introduction

"What I cannot create I cannot understand." *Richard Feynman, Caltech 1988.*

Enzymes are protein catalysts that nature has engineered to facilitate chemical reactions under mild conditions, and they perform this function with unparalleled specificity and rate acceleration.[1] The source of these remarkable features has long interested chemists, and continuing efforts have illuminated many of the structural and functional elements that contribute to enzyme activity. However, despite this tremendous progress, our current understanding of enzymes has proven largely insufficient to allow us to design biopolymer catalysts to facilitate any desired chemical reaction.

The ability to design biopolymer systems to catalyze particular reactions would be useful in a number of different contexts. From an industrial perspective, the potential to facilitate a reaction with catalytic quantities of biopolymer, as well as the possibility of using benign reaction conditions, make engineered biopolymers attractive for chemical production. From a biomedical perspective, the potentially high activity and selectivity of such systems might make them useful for modulating and monitoring cellular chemistry. Finally, from a scientific perspective, the ability to engineer such catalysts might provide new insights into the components and design of natural enzymes.

Historically, efforts to emulate enzymatic catalysis have been pursued from opposite directions. At one extreme, relatively small organic molecule systems have been used to approximate enzymatic activity.[2] For these systems a carbon skeleton such as a cyclodextrine[3] or cyclophane[4] form a structure scaffold. This scaffold is then elaborated with various functional groups. The orientation

of the groups around the scaffold provides the chemical functionality necessary to facilitate reactions. At the other end of the spectrum, whole proteins have been modified to create new functional activity.[5] In such approaches, the protein acts as the structural scaffold. This scaffold is then modified with new functional groups by techniques such as site-directed mutagenesis[6] or post-synthetic cysteine alkylation,[7, 8] generating a novel protein active site.

Recently, advances in generating small, folded protein scaffolds have been made with available structures including α-helix coiled-coils,[9] four helix bundles,[10] triple-stranded β-sheets,[11] and ββα-motif peptides.[12, 13] These systems now provide the means to explore an intermediate strategy, where the folded peptide provides the structural scaffold. By incorporating different amino acids at various positions along the peptide, functional groups can be introduced to mediate chemical activity.

Moreover, small folded peptide scaffolds offer many attractive features that combine advantages of both the small molecule and protein approaches. These systems can be assembled relatively easily using techniques of solid phase peptide synthesis (SPPS).[14] Because this technique uses modular synthetic monomers, a wide range of target molecules containing both natural and non-natural amino acids can be created, and many related compounds can be explored using combinatorial techniques.[15] Additionally, because these scaffolds are generally relatively small, often less than 50 amino acids in length, they are potentially simple enough to understand thoroughly. Techniques to structurally characterize these molecules include a wide range of standard chemical and biophysical methods.[16-18]

To capitalize on the attractive features peptide scaffolds provide for the design of catalytic systems, the Imperiali lab has developed a series of ββα-motif peptides (BBA peptides) (Figure 1.1).[12, 19, 20]

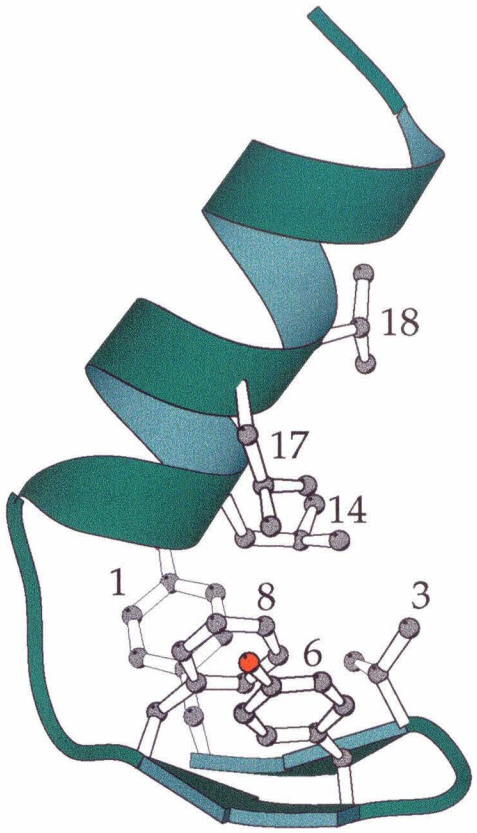


Figure 1.1. Structure of BBA peptide BBA4. Amino acids involved in hydrophobic core interactions are depicted and labeled.

The BBA peptides are 23-amino acid peptides based on the zinc-finger domain[21, 22] of zinc-finger proteins. However, unlike these domains, the BBA peptides do not require a zinc ion to adopt a ββα super-secondary structure. Instead, structure is maintained through stabilization of the two secondary structure

domains individually and by hydrophobic interactions between these domains. The β-hairpin region, amino acids 1-8, is stabilized by inclusion of a heterochiral type II′ turn element[19] and by inclusion of residues with β-sheet propensities.[23] The α-helix region, amino acids 12-22, is stabilized by residues with α-helical propensities[23] and by an α-helix capping sequence.[24] The packing of these two structural elements is promoted by hydrophobic interactions within the interior of the motif and hydrophilic residues on the exterior. In the BBA peptide BBA4,[20] both aliphatic and aromatic hydrophobic residues, Tyr1, Val3, Tyr6, Phe8, Leu14, Leu17, and Leu18, comprise the hydrophobic core.

Using the BBA peptides we have sought to incorporate function into a folded peptide scaffold and have tried to address several questions, including: What strategies are best suited for designing, synthesizing, and characterizing functional peptides? To what extent can small folded peptides support function? And what can be learned about enzyme function and design from modifying small folded peptides? The results of these efforts are detailed in Chapters 2-4.

Chapter 2 describes our efforts to incorporate transamination functionality into the BBA peptide scaffold using a pyridoxamine coenzyme-amino acid chimera strategy. In this strategy, the functionality of the pyridoxamine coenzyme is incorporated as the side chain of a non-natural amino acid (Pam), and the inherent reactivity of this group is modulated by the peptide scaffold. Attempts to modulate the activity of the pyridoxamine functionality included varying the site of Pam incorporation, introducing general acid and general base groups, and introducing sterically bulky groups. These efforts produced peptides

that showed increased rate of alanine transamination and increased stereoselectivity relative to pyridoxamine models.

Chapter 3 describes efforts performed in collaboration with Kevin McDonnell to incorporate the ability to mediate aldol condensation into the BBA peptide scaffold by introducing basic amino acids into the hydrophobic core of the BBA peptide. To find functionally interesting peptides, high-throughput methods were developed and employed. In the course of these studies a library of over 100,000 different peptides was synthesized and screened for members that stabilize analogs of aldol condensation intermediates. These peptides demonstrated interesting sequence trends. Further characterization of a promising library member suggested that this peptide had interesting structural and functional features.

Chapter 4 describes two techniques involving *N*-terminal capping groups that were developed to expedite generation and characterization of functional peptides. The first technique is a method for the rapid and parallel affinity purification of *N*-terminally capped full-length peptides using a reversible biotin capping group. The second technique is a method for capping peptides with a group that is stable toward functional assay conditions, but can be converted to a form compatible with Edman sequencing for peptide identification.

References

(1) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W H Freeman & Co.: New York, 1999; pp 631.

(2) Kirby, A. J. "Enzyme mechanisms, models, and mimics," *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 707-724.

(3) Breslow, R.; Dong, S. D. "Biomimetic reactions catalyzed by cyclodextrins and their derivatives," *Chem. Rev.*, **1998**, *98*, 1997-2011.

(4) Konig, B. "Carbon rich cyclophanes with unusual properties: an update." In *Carbon Rich Compounds* From *Topics in Current Chemistry Volume 196*; Springer-Verlag, Berlin, 1998; pp 91-136.

(5) Schmidtchen, F. P., Ed. *Implementation and Redesign of Catalytic Function in Biopolymers* From *Topics in Current Chemistry Volume 202*; Springer-Verlag, Berlin, 1999; pp 140.

(6) Harris, J. L.; Craik, C. S. "Engineering enzyme specificity," *Curr. Opin. Chem. Biol.*, **1998**, *2*, 127-132.

(7) Kaiser, E. T. "Synthetic approaches to biologically-active peptides and proteins including enzymes," *Acc. Chem. Res.*, **1989**, *22*, 47-54.

(8) Kuang, H.; Distefano, M. D. "Catalytic enantioselective reductive amination in a host-guest system based on a protein cavity," *J. Am. Chem. Soc.*, **1998**, *120*, 1072-1073.

(9) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. "High-resolution protein design with backbone freedom," *Science*, **1998**, *282*, 1462-1467.

(10) DeGrado, W. F.; Regan, L. "Characterization of a helical protein designed from 1st principles," *Science*, **1988**, *241*, 976-978.

(11) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. "Design of a 20-amino acid, three-stranded beta-sheet protein," *Science*, **1998**, *281*, 253-256.

(12) Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Design of a monomeric 23-residue polypeptide with defined tertiary structure," *Science*, **1996**, *271*, 342-345.

(13) Dahiyat, B. I.; Mayo, S. L. "De novo protein design: Fully automated sequence selection," *Science*, **1997**, *278*, 82-87.

(14) Fields, G. B.; Noble, R. L. "Solid-phase peptide-synthesis utilizing 9-fluorenylmethoxycarbony amino-acids," *Int. J. Peptide Protein Res.*, **1990**, *35*, 161-214.

(15) Lam, K. S.; Lebl, M.; Krchnak, V. "The "one-bead-one-compound" combinatorial library method," *Chem. Rev.*, **1997**, *97*, 411-448.

(16) Kamp, R. M.; Choli-Papadopoulou, T.; Wittman-Liebold, B. *Protein Structure Analysis : Preparation, Characterization, and Microsequencing*; Springer-Verlag: New York, 1997.

(17) Fasman, G. D. *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press: New York, 1996; pp 738 .

(18) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986; pp 292 .

(19) Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Economy in protein design: Evolution of a metal-independent ββα-motif based on the zinc finger domain," *Journal of the American Chemical Society*, **1996**, *118*, 3073-3081.

(20) Struthers, M.; Ottesen, J. J.; Imperiali, B. "Design and NMR analysis of compact, independently folded ββα-Motifs," *Folding and Design*, **1998**, *3*, 95-103.

(21) Berg, J. M. "Zinc finger domains - Hypotheses and current knowledge," *Annu. Rev. Biophys. Biophys. Chem.*, **1990**, *19*, 405-421.

(22) Michael, S. F.; Kilfoil, V. J.; Schmidt, M. H.; Amann, B. T.; Berg, J. M. "Metal-binding and folding properties of a minimalist Cys2His2 zinc finger peptide," *Proc. Natl. Acad. Sci. U. S. A.*, **1992**, *89*, 4796-4800.

(23) Fasman, G. D. *The Development of the Prediciton of Protein Structure*; Plenum Press: New York, 1989; pp 193-316 .

(24) Scholtz, J. M.; Baldwin, R. L. "The mechanism of α-helix formation of peptides," *Annu. Rev. Biophys. Biomol. Struct.*, **1992**, *21*, 95-118.

Chapter 2: Modulating Pyridoxamine-Mediated Transamination Through the ββα-Motif Peptide Scaffold

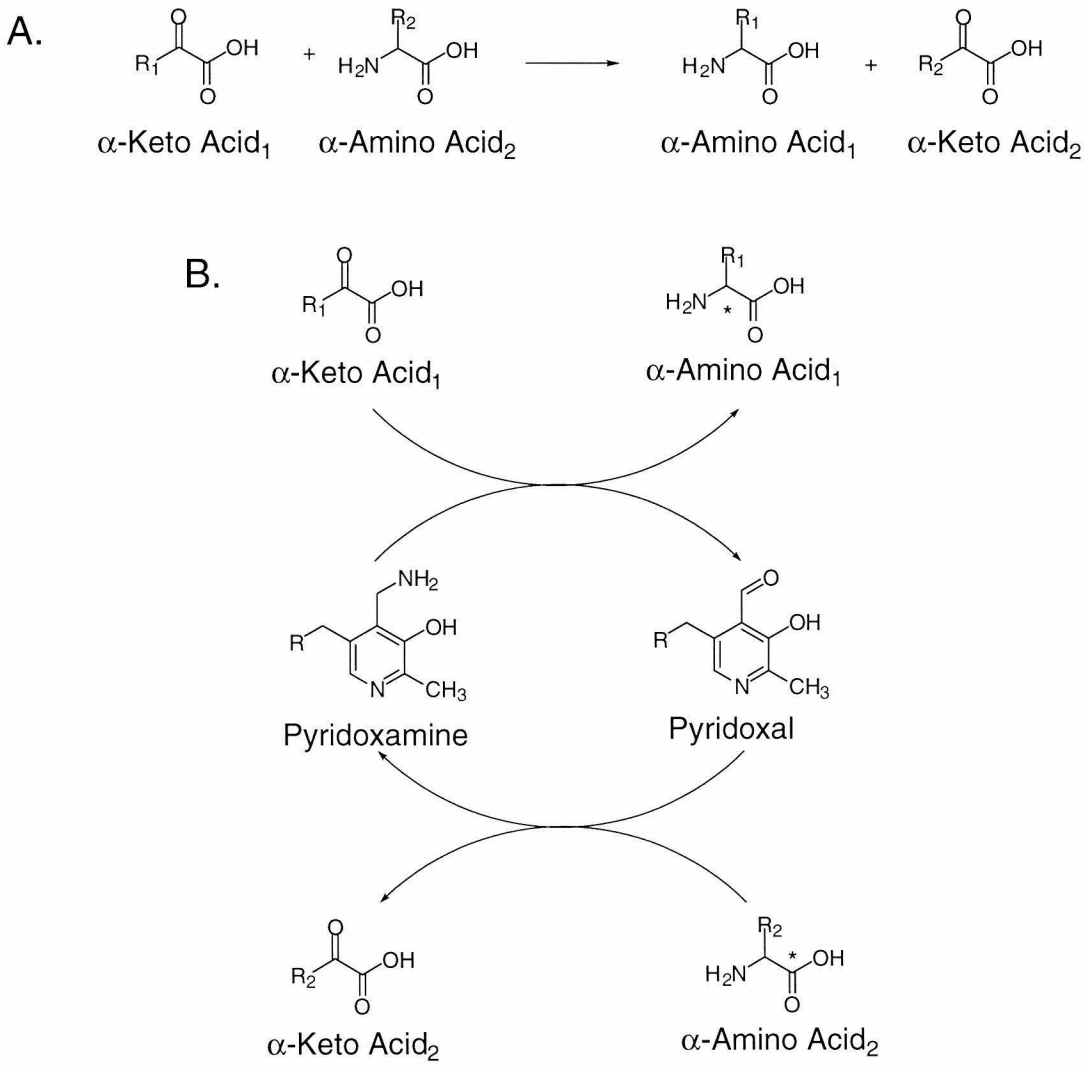
Introduction

The initial strategy for incorporating function into the BBA scaffold focused on using these peptides to modulate the reactivity of an inherently chemically active group. Many enzymes create chemical activity by directly utilizing the twenty standard amino acids. However, accomplishing such a feat is not trivial. In the case of an enzyme like a serine protease, the precise alignment of the catalytic triad of serine, histidine, and aspartic acid is necessary to improve the relatively poor nucleophilicity of the serine hydroxyl.[1] From the perspective of *de novo* peptide design, controlling the geometry of amino acid side chains with enough accuracy to generate activity is challenging. A simpler and more attractive strategy is to directly modulate baseline chemical activity using the peptide scaffold.

Coenzymes offer an attractive subject target for augmenting and directing inherent chemical reactivity. Coenzymes are small, densely functionalized organic molecules with intrinsic chemical reactivity.[2] These molecules are often found associated with enzymes, where the enzyme utilizes them to supplement the chemical reactivity offered by the side chains of the standard amino acids. In these enzymes the amino acids in the active site play a role in modulating the chemical reactivity of the coenzyme, and it was believed that the side chains of the BBA-peptide scaffold might also be able to perform a comparable role.

To investigate the ability of the BBA scaffold to influence coenzyme activity, the pyridoxamine coenzyme was chosen. Pyridoxamine, and the related compound pyridoxal (Scheme 2.1B), are coenzymes involved in a number of enzyme mediated reactions of amino acids.[2] Of these, one of the most chemically

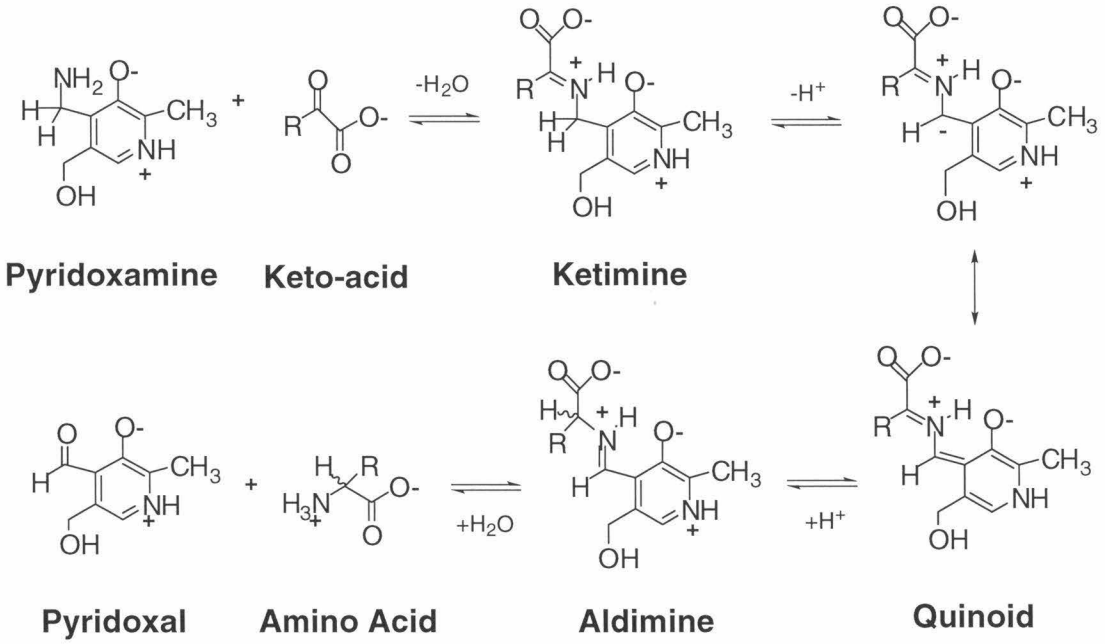
interesting and most thoroughly studied is the pyridoxamine-mediated process of transamination. Transamination of α-keto acids to α-amino acids involves exchange of the amine functionality of an amino acid with the carbonyl of a keto acid to generate a new keto acid and amino acid (Scheme 2.1A).



Scheme 2.1. A) Net transamination reaction. B) Coenzyme mediated transamination reaction.

The uncatalyzed reaction requires harsh conditions and results in poor yields.[3] However, the coenzymes pyridoxamine-phosphate and pyridoxal-phosphate facilitate this process by acting as catalytic intermediaries in the reaction and achieve the same net reaction in a manner that is more accessible kinetically (Scheme 2.1B).[4]

Mechanistically, the transamination half-reaction occurs in several steps, proceeding through ketimine, quinoid, and aldimine intermediates before generating products (Scheme 2.2).



Scheme 2.2. Mechanism of the pyridoxamine-mediated half-transamination reaction.

Each of these mechanistic steps is made feasible through the assistance of the coenzyme. Notably, the deprotonation of the ketimine and aldimine

intermediates, in the forward and backward half-reactions respectively, is facilitated by the ability of the pyridoxamine ring system to act as an "electron sink," stabilizing the deprotonated quinoid intermediate by resonance delocalization.

The ability of the pyridoxamine coenzyme to facilitate this reaction is further enhanced by the active site residues of the transaminase enzyme.[5, 6] A number of different amino acids provide binding energy for the coenzyme and the keto-acid substrate, facilitating imine formation. In the later mechanistic steps of the reaction, a lysine group acts as both a general base and acid, promoting ketimine deprotonation and stereoselectively controlling quinoid reprotonation. Additionally, protonation of the ring nitrogen of the coenzyme is controlled by an aspartic acid residue, stabilizing formation of the quinoid intermediate.

Because of pyridoxamine's catalytic potential, others have incorporated the coenzyme into non-natural scaffolds in attempts to influence its reactivity. In general, efforts to incorporate the pyridoxamine coenzyme into new scaffolds have focused on utilizing either small molecule organic systems,[7-9] including cyclophanes[10] and cyclodextrines,[9] or protein scaffolds.[11] Most of these systems influence pyridoxamine activity by promoting substrate binding, usually through the introduction of a hydrophobic domain.

The BBA peptide offered the opportunity to investigate the structural and functional effects of incorporating the pyridoxamine coenzyme into a small peptide scaffold. Because of its modular nature, the scaffold also provided a means to investigate strategies of influencing transamination activity beyond

enhancing substrate binding. Using this scaffold we sought to probe the extent to which we could increase the rate and enantioselectivity of transamination.[12]

The reactive functionality of the pyridoxamine coenzyme was incorporated into a range of BBA peptide scaffolds as the side chain of a non-natural amino acid, Pam. This amino acid was incorporated at one of three different positions in the BBA peptide scaffold, in the β-sheet region (BP5-BP6), the α-helix region (BP1), and the hydrophobic core (BP2-BP4), both in the presence and absence of an auxiliary general acid/base residue. Functional characterization of these peptides indicated that increases in the rate and stereoselectivity of alanine transamination were possible. Structural characterization of these peptides suggested that peptide BP5 demonstrated hallmarks of the BBA scaffold secondary and super-secondary structure. Peptide BP5 became the basis of a larger family of BBA peptides, CBP01-CBP18. These peptides were synthesized to investigate the effects of varying residues in the vicinity of the coenzyme functionality and of including general acid/base residues. Among these peptides, several showed interesting functional characteristics. Peptide CPB13 exhibited the greatest increase in transamination rate, and transamination mediated by peptide CBP10 resulted in the greatest levels of optical induction. Furthermore, all CBP peptides demonstrated interesting trends in enantioselectivity as a function of the core residues. In addition to the BP and CBP peptides, other peptides were investigated in an attempt to influence the pyridoxamine reactivity by other mechanisms.

Results and Discussion

2.1. Coenzyme-amino acid chimera strategy and synthesis.

To investigate the ability of the BBA peptide scaffold to influence the reactivity of the pyridoxamine functionality, it was necessary to incorporate this group into a peptide. Toward this end, we employed a coenzyme-amino acid chimera strategy, where the pyridoxamine group is the side chain of an amino acid (Figure 2.1A). This covalent attachment provides stable anchoring of the coenzyme to the peptide. Moreover, by synthesizing this amino acid, Pam, in a form compatible with Fmoc-based solid phase peptide synthesis, it is possible to easily incorporate the coenzyme into a peptide in a highly controlled fashion (Figure 2.1B).

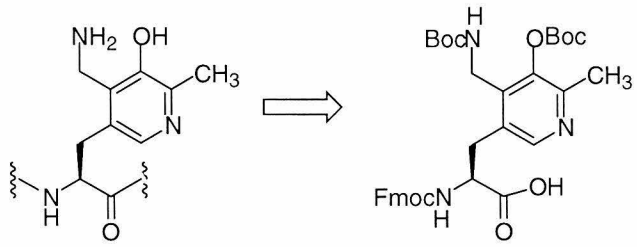
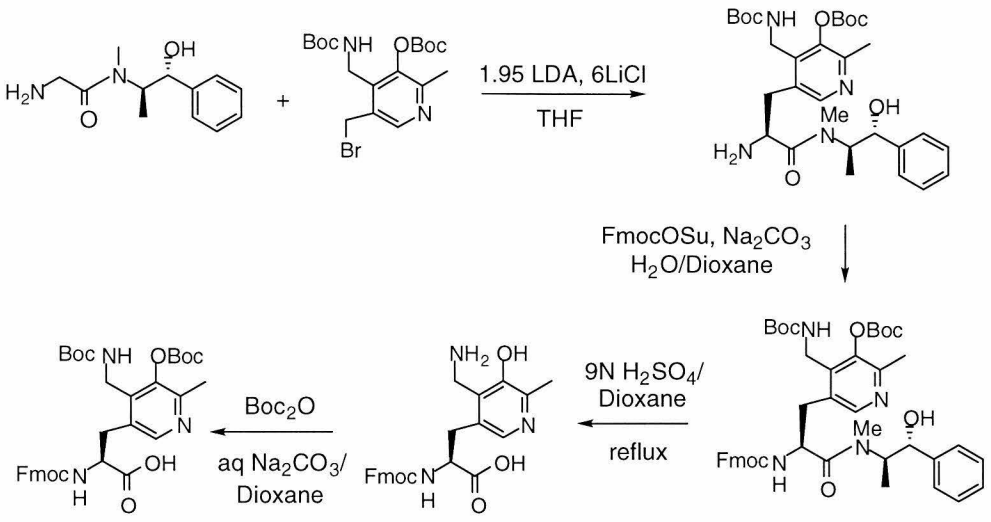


Figure 2.1. The pyridoxamine amino acid chimera (Pam) A) incorporated into a peptide and B) as its Fmoc-protected precursor.

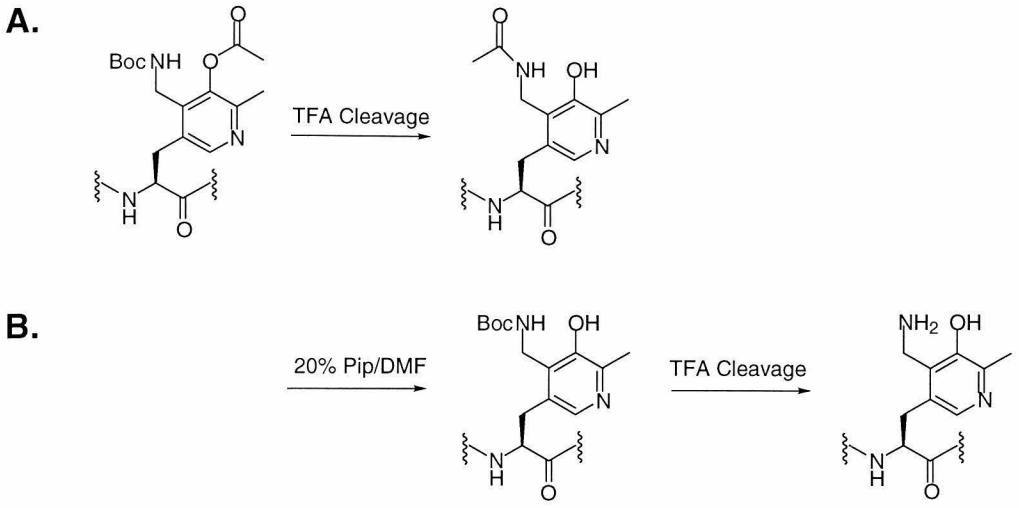
Synthesis of this amino acid was originally accomplished by Ranabir Sinha Roy[13] and was further optimized in the course of these studies (Scheme 2.3).



Scheme 2.3. Synthesis of the Fmoc-protected Pam amino acid.

In the synthesis, the Pam carbon skeleton was generated through asymmetric alkylation of a pseudoephedrine-glycinamide glycine-equivalent using a protected pyridoxamine halide. Removal of the chiral auxiliary and reshuffling of the side chain and main chain protecting groups resulted in the Pam amino acid, competent for Fmoc-based peptide synthesis.

The Pam amino acid could be incorporated into peptides by standard Fmoc-based solid phase peptide synthesis but required an additional post-synthesis transformation. It had been previously observed that the Boc protecting group of the phenolic oxygen of the Pam residue was labile under the Fmoc-SPPS conditions, subsequently resulting in acetylation of this group.[13] It was necessary to remove the acetyl group before side chain deprotection/resin cleavage to prevent irreversible acetylation of the 4′ pyridoxamine amine by acyl-shift (Scheme 2.4A).



Scheme 2.4. Effects of side chain deprotection/resin cleave on Pam A) in the absence of and B) in the presence of piperidine treatment.

Originally, this deacetylation was accomplished by treatment of the peptide with aqueous sodium carbonate immediately prior to acidic side chain deprotection/resin cleavage. However, while this treatment was compatible with most amino acids, it was discovered that it could lead to racemization of some of the amino acid residues in a peptide. Treatment of the peptide with piperidine/DMF instead of sodium carbonate was sufficient to remove the undesirable acetyl group while being gentle enough to avoid peptide racemization (Scheme 2.4B).

2.2. BP1-BP6 peptide family.[14]

Initial experiments on the effect of incorporating the Pam amino acid into the BBA peptide were performed on peptides BP1-BP6 (Figure 2.2). Previously, work in the Imperiali group had focused on incorporating the Pam and Pal (the

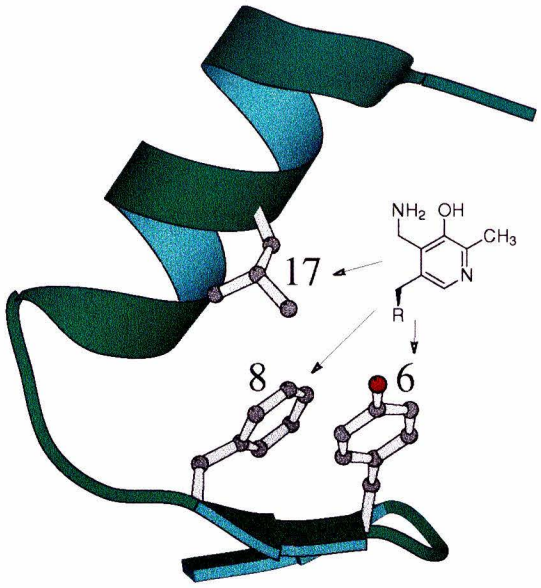
pyridoxal analogue of Pam) amino acids into small hexapeptide or semi-synthetic ribonuclease-S systems.[15, 16] In these studies, inclusion of the coenzyme and basic residues resulted in some enhancement in the rate of amino acid transamination and some enhancement in stereoselectivity. To determine whether the intrinsic reactivity of the pyridoxamine functionality could be further modulated, the Pam group was incorporated into three different sites in the BBA peptide scaffold, in the presence and absence of additional basic and acid residues. The structural and functional effects of the changes were then determined.

Peptides BP1 through BP6 were designed based on BBA peptides BBA1 and BBA4 and incorporated the Pam amino acid into either the α-helix, β-sheet, or hydrophobic core of the peptide scaffold (Figure 2.2). BBA1 peptide was used as the starting point for peptide BP1, while later peptides were able to use peptide BBA4, a BBA peptide that did not require the unnatural phenanthrolene amino acid to fold properly.[17]

In peptide BP1, Pam was incorporated at position 17 in order to access the inner face of the BBA helix and to be in close proximity to the basic His21 residue and the acidic Glu13 residue. Lys16 was replaced with an asparagine to reduce the number of basic residues present.

In peptides BP2-BP4, the Pam residue was incorporated into the BBA hydrophobic core at position 8. In BP4, the basic amino acid ornithine was incorporated above residue 8 at position 13. In peptide BP3, the acidic residue Asp replaced Val3 across the sheet from the Pam residue.

In peptides BP5-BP6, the Pam amino acid was incorporated into the sheet region of the BBA scaffold at position 6. In peptide BP5, the basic amino acid diaminobutyric acid (Dab) was used to replace Ser5.



	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
BBA1	Ac-Tyr-Thr- Val- dPro- Ser-Fen-Thr- Phe- Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BBA4	Ac-Tyr-Dap-Val- dPro- Ser-Tyr- Asp- Phe- Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BP1	Ac-Tyr-Thr- Val- dPro- Ser-Fen-Thr- Phe- Ser-Arg-Ser-Asp-Glu-Leu-Ala-Asn- Pam-Leu-Arg-Leu-His-Ala-Gly-NH_2
BP2	Ac-Tyr-Dap-Val- dPro- Ser -Tyr- Asp-Pam-Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BP3	Ac-Tyr-Dap-Asp-dPro- Ser -Tyr- Asp-Pam-Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BP4	Ac-Tyr-Dap-Val- dPro- Ser -Tyr- Asp-Pam-Ser-Arg-Ser-Asp-Orn-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BP5	Ac-Tyr-Dap-Val- dPro-Dab -Pam-Asp-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2
BP6	Ac-Tyr-Dap-Val- dPro- Ser -Pam-Asp-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Ala-Lys- Leu- Leu-Arg-Leu-His-Ala-Gly-NH_2

Figure 2.2. Structure of peptide BBA4 with the sites of Pam incorporation highlighted. Sequences of peptides BBA1, BBA4, and BP1-BP6, with the sites of sequence changes highlighted red. Dap: β-diaminoproprionic acid, Dab: γ-diaminobutyric acid, Orn: ornithine.

By incorporating Pam at these positions, we hoped to probe whether the peptide microenvironment influenced the transaminase activity and also to observe the structural consequences of these substitutions. By incorporating

basic amino acids in the vicinity of the Pam amino acid, we sought to determine to what extent the side chains of these groups could facilitate the general base and general acid catalysis observed in many mechanistic steps of enzymatic transamination, including ketimine deprotonation and quinoid reprotonation. Acidic amino acids were incorporated to determine the extent to which a carboxylic acid group might facilitate the transamination mechanism, especially through stabilization of ketimine ring protonation.

Peptides BP1 – BP6 were generated by standard Fmoc-based solid phase synthesis techniques. Peptides were purified by preparative reversed-phase high-pressure liquid chromatography. The identities of the peptides were confirmed by electrospray mass spectrometry. Since these peptides were prepared prior to the realization that the post-synthetic treatment of the peptides with aqueous sodium carbonate resulted in racemization, two isomeric products with the correct molecular weight were resolved and isolated in each synthesis. Both products were characterized further; these peptides tended to show similar properties. However, because racemization is not expected to exceed 50%, we assumed that the major product was the desired peptide, and it is the results with this product that are reported.

To investigate the functional effects the BBA peptide scaffold and residues provided, the rate and enantioselectivity of transamination was assayed. Originally, transamination was monitored by derivatizing the crude reaction mixture with amine-reactive Marfey's reagent.[18] This method did not afford easy resolution of the enantiomeric amino acid products. Orthophthaldialdehyde (OPA) derivatization proved superior; in this case resolvable diastereomeric

isoindol products were generated (Figure 2.3).[19, 20] Additionally, these products were fluorescent, providing a high degree of detection sensitivity.

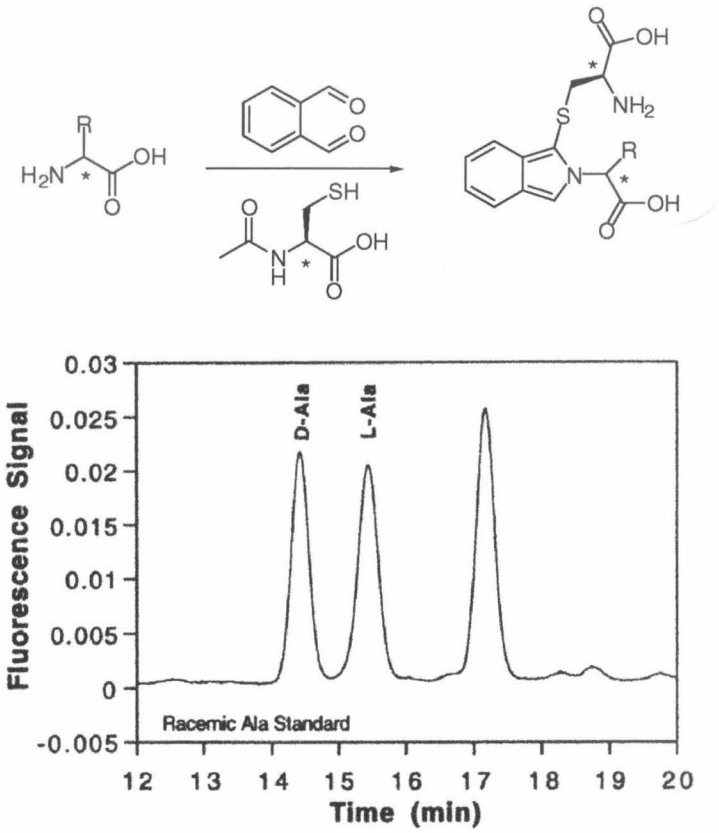


Figure 2.3. Orthophthaldialdehyde derivatization of amino acids, with an HPLC trace of the derivatization products of D- and L-alanine.

Typical transamination assays were run with 100 μM BP peptide, 10 mM pyruvic acid, and 10 mM D-phenylalanine at pH 8.0 at room temperature either in the presence of 100 μM Cu(II) ion or 100 μM EDTA. Over the course of 24 hours samples were collected, derivatized with OPA, and analyzed with a high pressure liquid chromatography (HPLC) instrument equipped with a fluorescence detector. The enantiomeric excess was calculated from the

difference in the amount of the alanine isomers produced. The rate of alanine production for the single turnover half-reaction was determined by fitting the total alanine production as a function of time to a first-order exponential build-up curve.

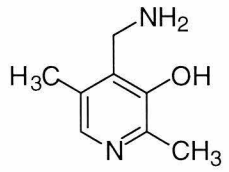
Assays were performed at pH 8.0 because this pH has been shown to be the optimal pH for pyridoxamine mediated transamination, and because a larger fraction of the basic residues would be deprotonated at this pH. Reactions were performed in the presence of excess amino acid, because in its absence rapid loss of optical induction was observed. The keto-acid and amino acids used were selected for consistency and for the ease of resolving the derivatized alanine product from the large excess of starting materials. Although other amino acids and keto-acids were studied, the products of these reactant pairs could not be resolved from the starting materials to a satisfactory degree. Assays were performed in both the presence and absence of copper(II) ion to assess the influence of metal ions on transamination. Metal ions are known to coordinate to pyridoxamine and the transamination intermediates, facilitating the reaction.[21]

Results of the transamination assays with the BP peptides and dPam, a 5′-deoxy-pyridoxamine model compound, are shown below (Table 2.1). From these results, it is clear that different peptides excelled under different conditions. The increase in the rate of alanine production relative to the dPam model under metal-free conditions ranged from 2- to 17-fold, with the greatest increase observed for peptide BP1. In the presence of copper(II) ion, transamination rate increases ranged from 7- to 19-fold, the greatest being seen with peptide BP4. For optical induction, in the absence of copper(II), peptide BP1 demonstrated a 35% enantiomeric excess of D-alanine relative to L-alanine. In the presence of

copper(II) ion, peptide BP1 also gave the greatest percent enantiomeric excess of alanine production. However, this level of optical induction continually decreased after reaching an initial high, while peptide BP5 maintained a high (30%) level of percent enantiomeric excess of D-alanine.

Peptide	Conditions	Initial Rate (μM/min)	%ee (xs D-amino acid)
dPam	Pyv/dPhe/EDTA	.0054	-7
	Pyv/dPhe/Cu^{+2}	.63	-10
BP1	Pyv/dPhe/EDTA	**.095**	**+35**
	Pyv/dPhe/Cu^{+2}	4.51*	+47*
BP2	Pyv/dPhe/EDTA	.020	+5
	Pyv/dPhe/Cu^{+2}	10.2	-11
BP3	Pyv/dPhe/EDTA	.016	+1
	Pyv/dPhe/Cu^{+2}	8.97	-13
BP4	Pyv/dPhe/EDTA	.043	+7
	Pyv/dPhe/Cu^{+2}	**12.06**	-8
BP5	Pyv/dPhe/EDTA	.019	+15
	Pyv/dPhe/Cu^{+2}	10.24	**+30**
BP6	Pyv/dPhe/EDTA	.012**	+12**
	Pyv/dPhe/Cu^{+2}	7.91**	+21**

* Values are based on initial behavior. After reaching a maximum at 25% of a turnover, the amount of product decreased.
** Values based on a mixture of 2 peptides-the full length BP6 and a truncation product lacking residues 1-3.



dPam

Table 2.1. Initial rates and enantiomeric excess of alanine production for dPam and the BP peptides. Largest values are highlighted. Structure of dPam shown at the bottom.

The source of these increases was not clear. In general, both the level of optical induction and rate enhancement changed depending on the site of Pam incorporation. In groups of peptides where a basic residue was either present or

absent (BP4 vs. BP2 and BP5 vs. BP6), it also appeared that the basic residue resulted in some increase in transamination rate. Inclusion of an acidic residue did not appear to exert a significant functional effect on the peptide pair where this difference was explicitly tested (BP3 vs. BP2). Overall, the magnitudes of the changes were modest and are difficult to rationalize without further study.

To complement the studies of the functional effects of Pam incorporation, the structural ramifications of including the Pam residue at different positions of the BBA scaffold were studied by circular dichroism (CD) and two-dimensional nuclear magnetic resonance (NMR) spectroscopy. Overall, peptide BP5, where the Pam residue was included in the sheet, showed signs of the greatest maintenance of the original BBA structure.

The secondary structural features of the BP family of peptides were probed by CD spectroscopy. Designed BBA peptides shown to adopt a ββα-secondary structure by 2D NMR share several common CD spectral features at pH 4.0 and 8.0.[22] While the magnitude of the mean residual ellipticity varies, the spectrum for each peptide shows an absolute minimum at 205-208 nm and a less intense minimum at 220-222 nm. The CD spectrum of BP5 exhibits similar characteristics with minima at 206 and 221 nm (Figure 2.4). At pH 4.0, the magnitude of the mean residual ellipticity at 206 nm was identical to that of parent peptide BBA4, while the magnitude of the mean residual ellipticity at 221 nm was about 16% greater. At pH 8.0, the pH at which transamination assays were performed, the CD spectra of BP1 and BBA4 were nearly identical, suggesting similar elements of secondary structure between peptide BP5 and BBA4 at both pH 4.0 and 8.0. In contrast, peptide BP1, where the Pam residue

was incorporated into the helix domain, showed a significant change in its CD spectrum upon a change in pH.

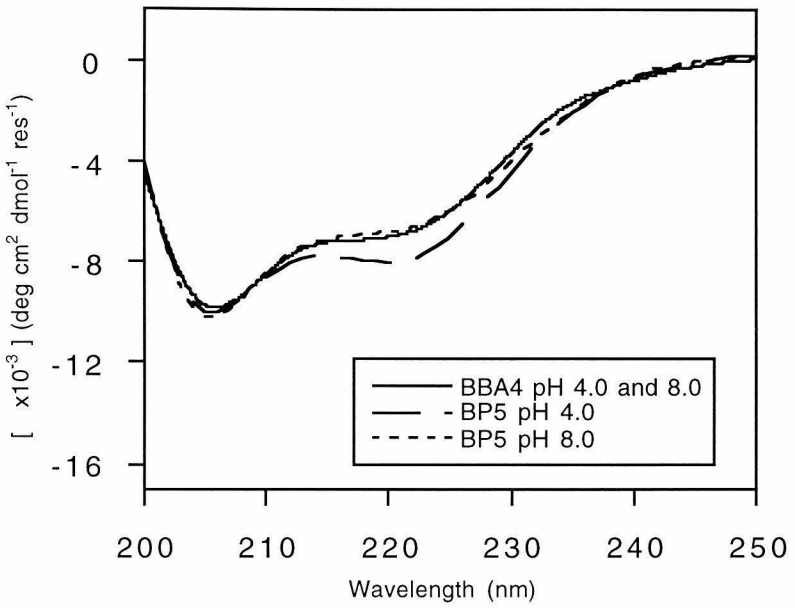


Figure 2.4. CD spectra of BBA4 and BP5 at pH 4.0 and 8.0.

To complement the CD analysis, two-dimensional ^{1}H NMR spectroscopy was used to probe the super-secondary structure of peptide BP5. The solution structure of peptide BBA4 has been solved by 2D NMR at pH 4.0.[17] A key feature of this structure is a set of packing interactions that are denoted by nuclear Overhauser effect (NOE) crosspeaks between hydrophobic aromatic residues located exclusively on the sheet region and hydrophobic aliphatic residues found only on the helix. These unique NOE crosspeaks offer a strong diagnostic tool for the formation of super-secondary structure, being present only when the β-sheet region is packed against the α-helix domain. While full structural

characterization of BBA4 has not been performed at pH 8.0, the pH at which transamination assays were performed, at this pH the peptide did exhibit aliphatic to aromatic NOE crosspeaks expected in a well-defined hydrophobic core. At pH 8.0, the spectra of peptide BP5 also showed aliphatic-to-aromatic side chain NOE contacts at relatively strong intensities (Figure 2.5).

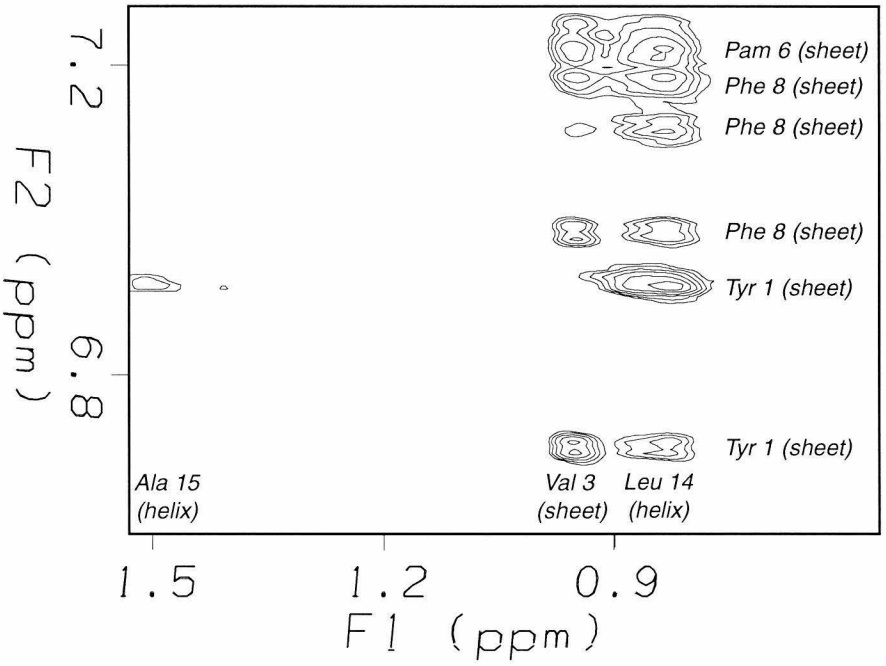


Figure 2.5. NOE side chain contacts for peptide BP5 in the aliphatic to aromatic region of the NMR spectrum at pH 8.0. Tentative NOE assignments based on BBA4. Chemical shifts are in ppm.

The presence of these contacts suggests that the hydrophobic core also exists in peptide BP5. In contrast, studies performed on peptide BP2 under the same conditions showed no aliphatic to aromatic NOE crosspeaks, suggesting that incorporation of the Pam amino acid in the hydrophobic core was detrimental to

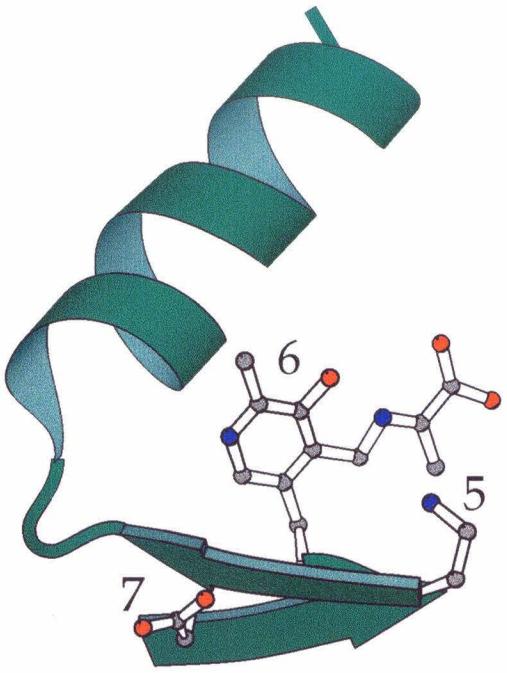
formation of stable super-secondary structure. Taken together, the structural analyses of peptide BP5 by CD and NMR spectroscopy suggest the overall ββα structure of peptide BBA4 is maintained under the conditions employed to assay transamination with the incorporation of both Pam and a basic amino acid.

Investigation of the BP family of peptides showed that the Pam amino acid could successfully be incorporated into the BBA peptide scaffold. These peptides demonstrated increases in the rate and stereoselectivity of transamination relative to a pyridoxamine model compound and showed features of BBA-peptide structure. Because peptide BP5 exhibited the hallmarks of the BBA-peptide scaffold in addition to possessing some interesting functional characteristics, this peptide was attractive as a platform for further development of functional Pam-containing BBA peptides. However, to improve the functional characteristics of this peptide, it seemed that exhaustive and targeted perturbations of the coenzyme environment were necessary.

2.3. CBP01-CBP18 peptide family.

Using BP5 as a template, a family of peptides, CBP01-CBP18, was designed to explore a range of coenzyme environments (Figure 2.6). The replacement of serine 5 with Dab in BP5 indicated that similar substitutions might be tolerated at this position with a minimal loss of structure. Six amino acids – histidine, lysine, ornithine, cysteine, homocysteine (Hcy), and Dab, were therefore placed at position five to investigate the ability of various basic residues to influence the course of transamination as either general acids or general bases. These amino acids provide basic side chains with a range of pK_a

values close to the pH of the transamination assays, along with a range of side chain length.



	5	6	7
BBA4	Ser	Tyr	Asp
BP1	Dab	Pam	Asp
CBP01	His	Pam	Asp
CBP02	Lys	Pam	Asp
CBP03	Orn	Pam	Asp
CBP04	Dab	Pam	Asp
CBP05	Cys	Pam	Asp
CBP06	Hcy	Pam	Asp
CBP07	His	Pam	Gly
CBP08	Lys	Pam	Gly
CBP09	Orn	Pam	Gly
CBP10	Dab	Pam	Gly
CBP11	Cys	Pam	Gly
CBP12	Hcy	Pam	Gly
CBP13	His	Pam	Dab
CBP14	Lys	Pam	Dab
CBP15	Orn	Pam	Dab
CBP16	Dab	Pam	Dab
CBP17	Cys	Pam	Dab
CBP18	Hcy	Pam	Dab

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
BP1	Ac	Tyr	Dap	Val	dPro	Dab	Pam	Asp	Phe	Ser	Arg	Ser	Asp	Glu	Leu	Ala	Lys	Leu	Leu	Arg	Leu	His	Ala	Gly	NH_2

Figure 2.6. Proposed structure of BP5/CBP04 based on the structure of peptide BBA4 with Pam depicted as the ketimine intermediate. Residues 5, 6, and 7 are shown to indicate substitution sites for peptides BP5 and CBP01-CBP18. Included are the full sequences of BP5 and the substitutions for peptides CBP01-CBP18. Hcy: homocysteine.

The placement of aspartic acid, glycine, and Dab, at position seven was also evaluated. This position flanks the C-terminal side of Pam on the exterior face of the β-sheet and the electrostatic, steric, and conformational differences of

these residues were expected to perturb the relative geometry of the coenzyme and the basic residue at position 5. However, these substitutions were predicted to not interfere with the super-secondary structure of the motif, because non-hydrophobic residues had previously been shown to be tolerated at this site.[17] Together the peptides with substitutions at positions five and seven comprised the 18 peptides CBP01-CBP18.

The advantages of the small peptide scaffold approach allowed the above designs to be realized rapidly. The common carboxy-terminal residues 8-23 were assembled on resin by solid phase peptide synthesis using an automated peptide synthesizer. The resin was then split into 18 smaller aliquots and the remaining residues were batch-coupled in parallel. Truncation peptides were observed following Pam coupling. To avoid time consuming HPLC purification, peptides CBP01-CBP18 were purified in parallel using a reversible biotin affinity tag purification method developed for this purpose which is described more fully in Section 4.1. This technique rapidly afforded N-terminal glycolyl-capped full-length peptides in 80% or greater purity and in quantities sufficient for a number of transamination assays.

For peptides CBP01-CBP18, peptide mediated transamination of an α-keto acid to the corresponding α-amino acid was probed as described for peptide BP5, except that initial amino acid production (less than 10% of turnover) and optical induction information were determined from a single time-point, run in duplicate (Table 2.2). Peptides CBP01-CBP18 demonstrated a range of functional behaviors in the transamination of pyruvic acid to alanine in the absence of metal ion.

			Pyv, D-Phe, EDTA		Pyv, D-Phe, Cu(II)	
Peptide	Pos. 5	Pos. 7	Ala /dPam	% ee L-Ala	Ala /dPam	% ee D-Ala
CBP01	His	Asp	3.4	10	24.4	8
CBP02	Lys	Asp	3.5	14	23.9	15
CBP03	Orn	Asp	4.4	10	25.4	19
CBP04	Dab	Asp	3.5	8	21.3	20
CBP05	Cys	Asp	3.0	4	23.3	10
CBP06	Hcy	Asp	3.0	10	24.6	7
CBP07	His	Gly	3.7	11	23.9	22
CBP08	Lys	Gly	3.8	18	23.4	27
CBP09	Orn	Gly	4.1	12	25.8	28
CBP10	Dab	Gly	4.2	2	26.2	**37**
CBP11	Cys	Gly	3.9	7	27.3	31
CBP12	Hcy	Gly	3.0	16	21.7	21
CBP13	His	Dab	**5.6**	24	**31.7**	3
CBP14	Lys	Dab	4.7	**27**	24.4	5
CBP15	Orn	Dab	3.8	23	20.1	7
CBP16	Dab	Dab	4.8	16	26.4	14
CBP17	Cys	Dab	4.8	8	25.1	13
CBP18	Hcy	Dab	4.1	17	22.1	6

Table 2.2. Sequence substitutions at positions five and seven for peptides CBP01-CBP18 relative to peptide BP5. Comparison of initial product formation and enantiomeric induction with pyruvate.

Relative to the model compound dPam assayed under identical conditions, all peptides showed at least a three-fold increase in the amount of alanine produced in 20 hours, with peptide CBP13 exhibiting a 5.6-fold increase. For all peptides a slight optical induction favoring L-alanine was observed, while no significant optical induction was observed for dPam under the same assay conditions. The level of optical induction varied from a slight excess of L-alanine to 27% ee for peptide CBP14.

In the presence of metal ions different functional properties were observed. The presence of copper(II) ion resulted in augmented production of

alanine for peptides CBP01-CBP18 relative to the dPam model compound in the presence of copper(II) ion. All peptides showed at least a 20-fold increase in product yield relative to dPam, with peptide CBP13 demonstrating a 31.7-fold increase in alanine production. All peptides also demonstrated some optical induction favoring D-alanine, with peptide CBP10 showing a 37% ee for this product.

While no clear trend was discernable in the rate data, interesting trends in stereoselectivity were seen for alanine production with peptides CBP01-CBP18. In the absence of metal ions, changing the residue at either position five or seven produced independent trends (Figure 2.7A). At position five, lysine gave the greatest optical induction, and in general longer side chain lengths resulted in greater optical induction. At position seven, aspartic acid and glycine afforded similar degrees of optical induction, while inclusion of Dab produced enhanced optical induction.

Additionally, the introduction of copper(II) ions resulted in dramatic and different stereochemical trends (Figure 2.7B). In contrast to the metal ion-free conditions, D-alanine production was favored in the presence of a stoichiometric amount of divalent copper ion. The degree of optical induction ranged from less than 5% to up to 37% enantiomeric excess of D-alanine for peptide CBP10. Furthermore, for the basic residues at position five, the greatest optical induction was produced by Dab, and in general shorter side chain lengths favored larger optical induction.

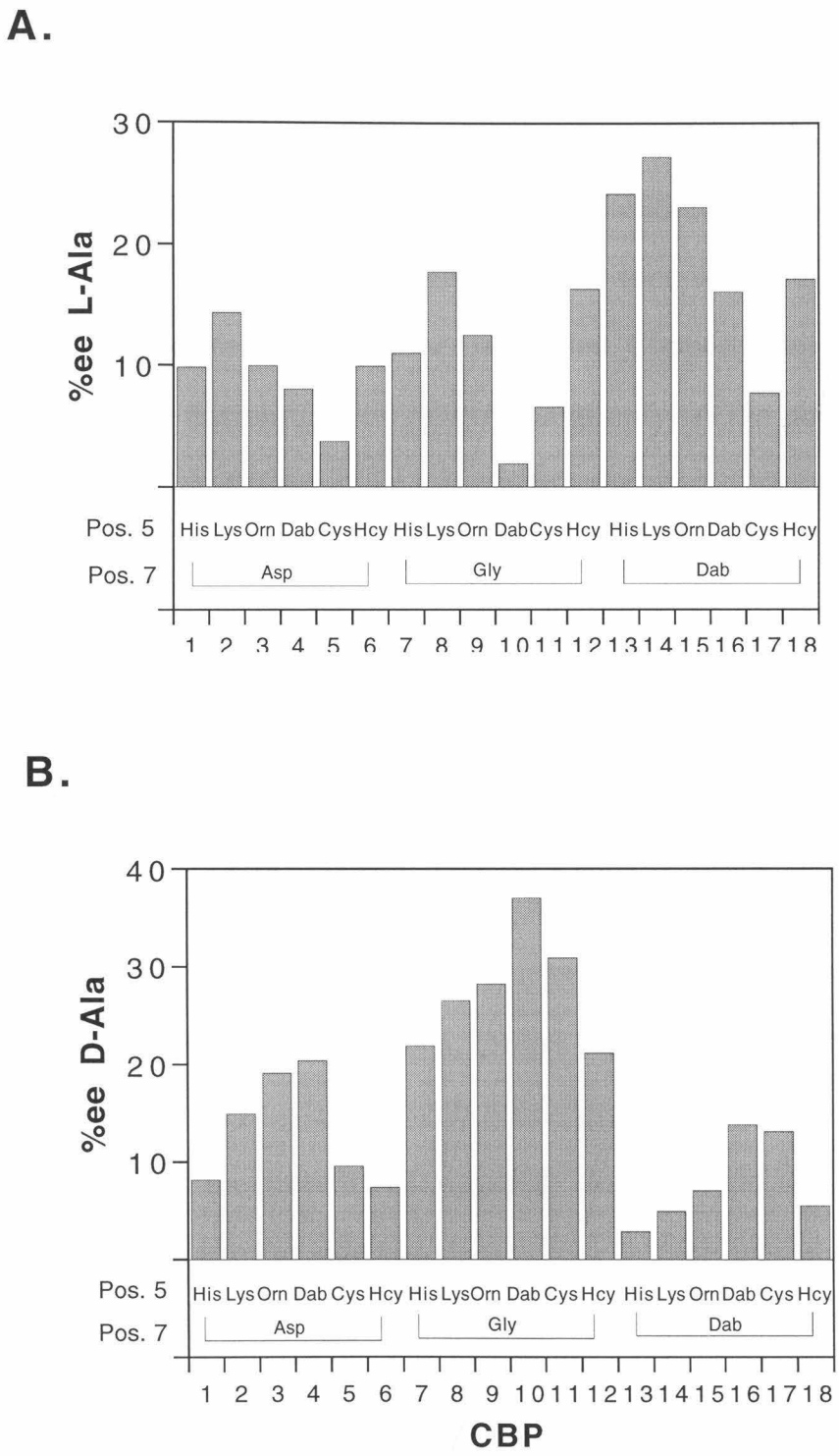


Figure 2.7. Observed optical induction in the production of alanine A) in the absence of metal ion and B) in the presence of copper(II) ion for peptides CBP01-CBP18. Peptide substitutions at position five and seven are indicated.

Peptides with glycine at position seven-resulted in the highest stereoselectivity, while those with Dab resulted in the lowest stereoselectivity.

From these studies it appears that the BBA peptide motif can influence the stereochemical outcome of the transamination reaction. In the absence of metal ions, L-alanine is the favored transamination product, while in the presence of copper(II), the opposite (D-) enantiomer is favored. Copper(II) ion is known to be coordinated by the transamination intermediates and has the potential to accept other ligands.[21] The difference in influencing the stereochemical outcome of transamination suggests that the BBA peptide motif has the capacity to interact differently with the metalated and unmetalated form of the coenzyme functionality. Furthermore, this optical induction is modulated by substitutions in the peptide sequence. Both in the presence and absence of divalent copper sequence changes at position seven resulted in different levels of optical induction independent of the nature of the basic residue at position five. Moreover, the identity of the basic residue at position five resulted in discernible trends. In the absence of metal ions, longer side chain residues resulted in increased production of L-alanine, with lysine producing the greatest degree of optical induction. In the presence of copper(II) ion, this trend was reversed, with basic residues with shorter side chains resulting in increased production of D-alanine, with the Dap residue producing the greatest increase. Modeling of the CBP peptides shows that the side chains of the basic residue at position five can interact with the pyridoxamine functionality. Although the complete structure has not been determined for each peptide, the observed trends do indicate that changes in the coenzyme environment can result in measurable functional differences between peptides.

The trends in optical induction observed with peptides CBP01-CBP18 suggest that the pyridoxamine functionality might exist within a defined peptide environment. To support this proposal, the structural features of peptide CBP14, the CBP peptide with the greatest optical induction in the absence of copper(II) ion, were investigated by CD and 2D NMR spectroscopy in a manner similar to that used for peptide BP5.

CD spectral analysis of CBP14 revealed spectral properties similar to peptide BP5 and previously studied BBA peptides (Figure 2.8).

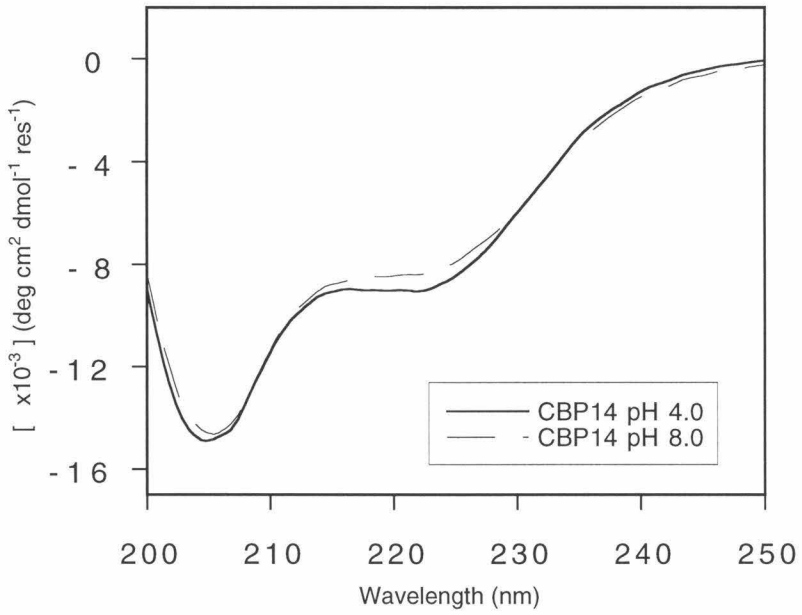


Figure 2.8. CD spectra of CBP14 at pH 4.0 and 8.0.

At both pH 4.0 and 8.0, peptide CBP14 exhibited identical absolute minima of mean residual ellipticity at 205 nm. Additionally, at both pH 4.0 and 8.0 a minimum was observed at 222 nm, with a 10% greater absolute mean residual

ellipticity at pH 4.0 relative to pH 8.0. While the absolute mean residual ellipticities at 205 and 222 nm were greater than those observed for either BP5 or BBA4, other BBA peptides shown to have ββα-structure have varied in the magnitude of the observed mean residual ellipticities.[17] Thus, the CD spectral data suggest that peptide CBP14 shares secondary structural features with peptides BP5 and BBA4.

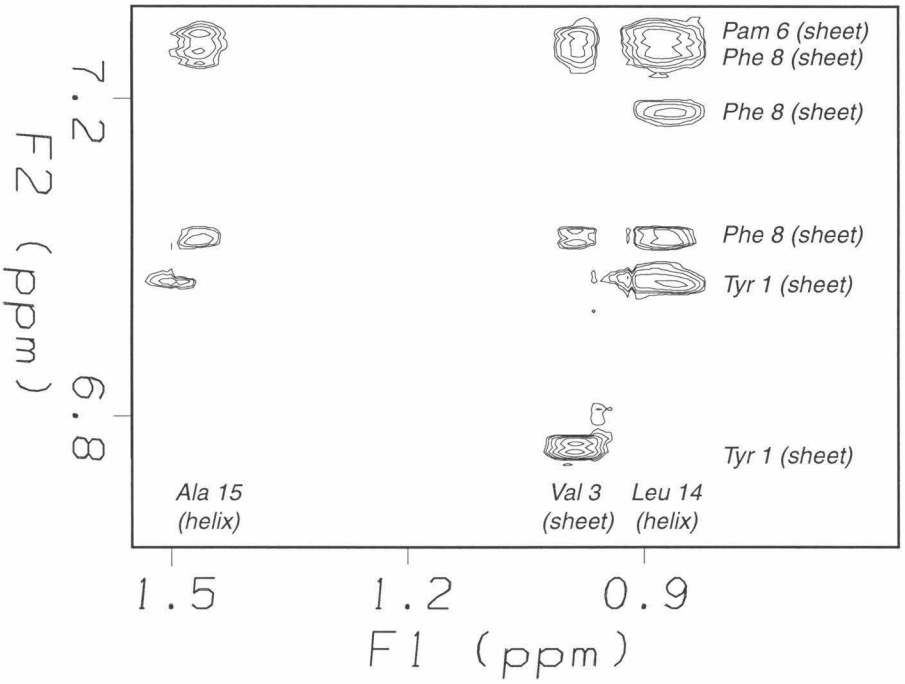


Figure 2.9. NOE side chain contacts for peptide CBP14 in the aliphatic to aromatic region of the NMR spectrum at pH 8.0. Tentative NOE assignments based on BBA4. Chemical shifts are in ppm.

The 2D NMR spectrum of CBP14 at pH 8.0 also suggests that peptide CBP14 adopts the super-secondary structure of the ββα motif peptides (Figure

2.9). Like BBA4 and BP5, the aliphatic to aromatic portion of the nuclear Overhauser effect spectroscopy (NOESY) spectrum of peptide CBP14 shows NOE crosspeaks between aliphatic and aromatic side chains. Because these crosspeaks are only expected for a species in which the sheet and helix domain are packed, the presence of these crosspeaks suggests that CBP14 adopts the ββα-motif super-secondary structure. Together these data suggest that peptide CBP14, and potentially the remaining CBP peptides, conserve the overall BBA peptide structure.

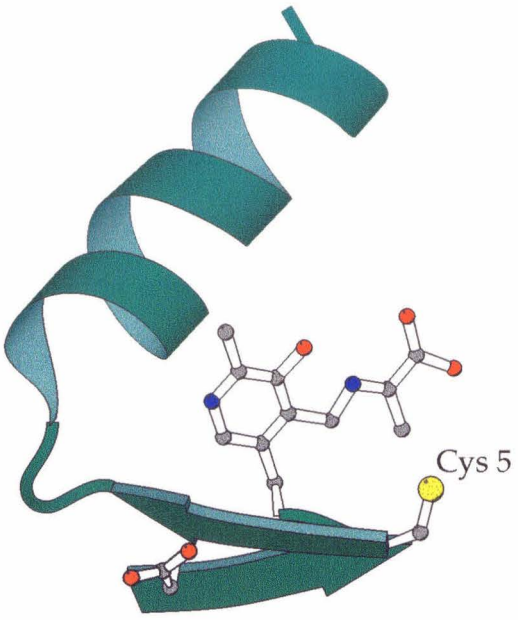
Altogether, studies of the CBP family of BBA peptides resulted in Pam containing peptides with interesting functional and structural properties. Peptide CBP13 demonstrated a 5.6-fold increase in alanine production relative to the dPam model compound in the absence of copper (II) ion, and a 31.7-fold increase in its presence. The copper-dependent increase in transamination rate is the largest we have observed for any Pam containing peptide or protein. Additionally, modest levels of optical induction in transamination were observed with peptides CBP10 (37% ee for D-Ala) and CBP14 (27% ee for L-Ala) in the presence and absence of copper ion, respectively. Moreover, the observed levels of optical induction followed discernable trends based on the substitutions made at positions five and seven. Structural studies of peptide CBP14 indicated that this peptide maintained much of the BBA peptide scaffold structure and suggested that the other CBP peptides also maintained this structure.

2.4. Other Pam containing peptides.

Several other strategies were explored to modulate transamination activity in Pam containing BBA and hairpin peptides. In an attempt to influence

the stereoselectivity of amino acid production a collection of post-synthetically modified BBA peptides was constructed. In these peptides a bulky group was placed near the pyridoxamine functionality to potentially occlude one face of the quinoid transamination intermediate, favoring reprotonation of this intermediate from the other face. These peptides were constructed based on peptide CBP17, where positions five, six and seven were cysteine, Pam, and Dap residues respectively (Figure 2.6). The cysteine group was alkylated by various organic halides using standard alkylation techniques.[23] These alkylating groups were all aromatic residues with either one or two rings and included the benzyl-, 1-naphthyl-, 2-naphthyl, 2-nitrobenzyl-, 3-nitrobenzyl-, and 4-nitrobenzyl-groups. Alkylation proceeded efficiently and specifically as determined by analytical HPLC and electrospray mass spectral analysis. Stereospecificity of alanine production was assayed in a manner similar to the CBP peptides (Section 2.3) and compared to the parent CBP17 peptide (Table 2.3). The optical induction observed for these peptides was minor and not significantly different from the parent CBP17 peptide. Although other alkylating groups and positions for the cysteine group were possible, synthesis and characterization of these peptides were not pursued.

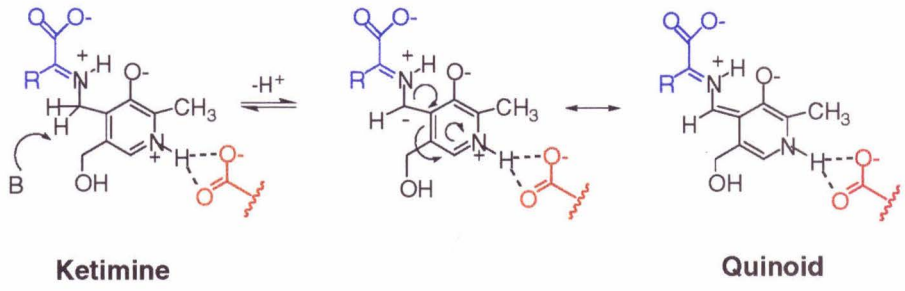
In an attempt to modulate the rate of transamination, hairpin peptides were designed and characterized for their ability to stabilize the protonation of the ketimine and aldimine intermediates. As described previously (Chapter 2-Introduction), the rate of transamination is in part dictated by the stability of the quinoid intermediate.



Alkyl Group	% ee D-Ala
-	-8%
Benzyl	3
1-Naphthyl	2
2-Naphthyl	7
2-Nitrobenzyl	5
3-Nitrobenzyl	8
4-Nitrobenzyl	11

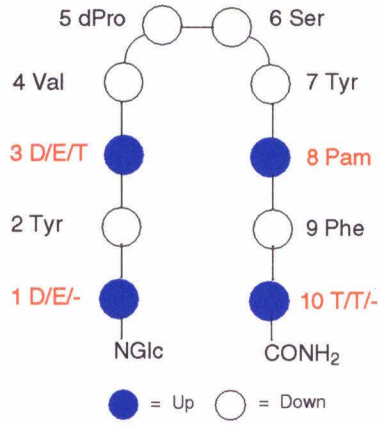
Table 2.3. Proposed structure of peptide CBP17 with cysteine 5 highlighted. Groups used to alkylated position 5 of peptide CBP17 and the associated enantiomeric excess for D-alanine observed for these peptides.

Natural transaminase enzymes facilitate formation of this intermediate in a number of ways, including stabilizing protonation of the coenzyme ring-nitrogen (Scheme 2.5).[5]



Scheme 2.5. Quinoid intermediate formation as facilitated by stabilization of ketimine protonation.

While this strategy had been explicitly explored to some degree with the BP family of peptides, we thought a more systematic investigation might be illuminating. A family of nine hairpin peptides, CHP1-CHP9, was designed to incorporate the Pam residue on the amino-terminal strand of the hairpin. Zero to two amino acids with carboxylic acid side chains were incorporated across the sheet from Pam (Table 2.4).



	1	2	3	4	5	6	7	8	9	10		pKa
CHP1	Glc-	D	- Y -	D	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.52
CHP2	Glc-	D	- Y -	E	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.60
CHP3	Glc-	D	- Y -	T	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.48
CHP4	Glc-	E	- Y -	D	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.57
CHP5	Glc-	E	- Y -	E	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.55
CHP6	Glc-	E	- Y -	T	- V -	dP	- S -	Y	-	Pam	- F -	T	-NH_2	8.47
CHP7	Glc-		- Y -	D	- V -	dP	- S -	Y	-	Pam	- F -		-NH_2	8.50
CHP8	Glc-		- Y -	E	- V -	dP	- S -	Y	-	Pam	- F -		-NH_2	8.49
CHP9	Glc-		- Y -	T	- V -	dP	- S -	Y	-	Pam	- F -		-NH_2	8.15
dPam														8.84

Table 2.4. Proposed structure and sequences of peptides CHP1-CHP9. Associated pyridoxamine ring nitrogen pK_a is shown to the right.

These peptides were based on the sheet region of the BBA peptide scaffold, and structural studies performed by Jennifer Ottesen on similar peptides suggested that such peptides would adopt a hairpin structure (unpublished result). The CHP peptides were synthesized and purified in parallel in a manner similar to that used for BBA peptides CBP01-CBP18.

Studies were performed to determine the functional effect of including the carboxylic acid groups across the sheet from the Pam amino acid. For the CHP peptides, the pK_a of the ring nitrogen of Pam was determined spectroscopically.

When protonated the Pam functionality exhibits a different maximal UV-Vis absorption relative to the unprotonated form.[4] By following the absorption spectrum of the CHP peptides as a function of pH, it was possible to determine the pK_a of the ring nitrogen (Table 2.4). From these studies, it appeared that inclusion of even one acidic amino acid across from the pyridoxamine functionality was sufficient to raise the pK_a of the ring nitrogen. In the presence of two acidic residues the pK_a was 8.60, almost a 0.50 increase. However, when studies were performed to determine the effect of this difference on the rate of alanine production, the effect of pH and pK_a on transamination rate was too complicated to draw any clear conclusions.

Conclusion

Peptides from the BP and CBP families of peptides constitute the first use of the BBA peptide scaffold to generate functional peptides and demonstrate the potential utility of the scaffold. From these studies it appears that the Pam amino acid can be incorporated into the BBA peptide scaffold with maintenance of overall structure. This incorporation of the Pam amino acid results in improved amino acid production and stereoselectivity, with magnitudes at least comparable to other peptides and small molecule systems, and demonstrates that elements of peptide structure can provide an environment to influence various aspects of the transamination process.

Despite these successes, it was felt that the BBA peptide scaffold could do more to augment chemical function. Work with the Pam amino acid highlighted

some of the difficulties in incorporating function into a folded peptide scaffold. Specifically, this work raised the questions as to how a peptide should facilitate improved function and what specific amino acids should be used to carry this out. Furthermore, could this process be performed more rapidly to improve the likelihood of finding interesting peptides? An effort to address these questions motivated our further attempts to incorporate function into the BBA peptide scaffold.

Experimental

Standard N^{α}-9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids and amino acid reagents were purchased from Millipore or PerSeptive Biosystems. The Pam residue was synthesized as previously described.[13] PAL-PEG-PS resin was purchased from PerSeptive Biosystems. The Vac-man vacuum system was purchased from Promega, and the disposable polypropylene columns and valves were obtained from BioRad.

Modeling and synthesis of Pam-containing peptides.

Peptide design and modeling were performed using Insight II software (Biosym Technologies, San Diego, CA, USA).

Peptides BP1-BP6 and the parent batch of resin for peptides CBP01-CBP18 (residues 8-23) were all made by solid phase peptide synthesis performed on a Milligen 9050 automated peptide synthesizer. Typically, peptides were synthesized on a 200 μmole scale from PAL-PEG-PS resin (1.0 g, 0.2 mmole/g).

Amino acids were double coupled in four-fold excess, except for Pam, which was double coupled at 2.5-fold and 1.5-fold excess. Peptide coupling was performed with HOBT/HBTU chemistry, with the exception of Pam, which was coupled by HOAT/HATU chemistry. The Fmoc-amines were deprotected using a 20% piperidine/DMF wash. Between each new residue coupling, free amines were capped with a 0.3 M acetic anhydride/HOBT solution in 9:1 DMF/DCM.

Residues 1-7 of peptides CBP01-CBP18, and all residues of peptides CHP1-CHP9, were synthesized in parallel using batch solid phase peptide synthesis techniques. The synthesis of the CBP peptides is representative. In this synthesis, the parent batch of resin was divided into 18x50 mg (6.4 μmole theoretical) portions. All residues were double coupled with 4 eq. amino acid (0.256 M), 4 eq. HOBT (0.256 M), 4 eq. HBTU (0.256 M) and 9 eq. diisopropylethylamine (0.576 M) in 89 μl DMF (Pam was coupled with a 2.5/1.5-fold excess with HOAT/HATU chemistry) for 2 hours. Coupling of the *N*-terminal reversible biotin affinity tag was accomplished under conditions described in Chapter 4. The extent of residue coupling was examined by Kaiser test on one of the 18 peptides after each coupling round. Acetylation steps were performed for 30 min. using the acetylation solution described earlier. Deblocking steps were performed for 30 min. with a 20% piperidine/DMF solution. The resin was washed with 20 ml of DMF between reaction steps and was occasionally agitated manually during the coupling step.

All Pam containing peptides were deacetylated prior to acid cleavage to remove an acetyl group from the phenolic oxygen of the Pam residue. Applying the previously described deacetylation conditions to peptide BP5 resulted in

isomeric products.[13] Subjecting the peptide resin to 20% piperidine in DMF for 12 hours resulted in full deacetylation without any isomerization. Resin was subsequently washed thoroughly with DMF and DCM and lyophilized to dryness. Peptide was cleaved from the resin with an excess of reagent K[24] (trifluoroacetic acid/phenol/H_2O/thioanisole/ethanedithiol, 82.5:5.0:5.0:5.0: 2.5) for two hours to afford the deprotected peptide. Excess TFA was removed with nitrogen gas and was triturated with 2:1 ether/hexane. The resulting white pellet was resuspended in water and further purified by reversed-phase HPLC. Peptides CBP01-CBP18 were rapidly purified in parallel using the avidin affinity column method described in Chapter 4. Post-synthetic alkylation of the CBP17 peptides by the various organic halides was accomplished using standard alkylation conditions on the purified peptides.[23] The identity of all peptides was confirmed by electrospray mass spectrometry.

Quantification of Pam-containing peptides.

The concentration of purified peptide stocks were quantified by UV-Vis spectroscopy under denaturing conditions using a Beckman DU 7500 Spectrophotometer in a 1.0 cm path length quartz cuvette. Typically, peptide stock was mixed with an equal volume of a 4.8 M guanidine hydrochloride solution buffered at pH 7.5 with 0.2M HEPPS. The concentration of the original stock was calculated from the absorbance at 322 nm with an ε_{322} of 5610 $M^{-1}cm^{-1}$, where ε was initially determined from a Pam-containing peptide stock with the concentration determined by quantitative amino acid analysis.

Circular dichroism studies.

CD experiments were performed on a Jasco J-600 spectropolarimeter or an Aviv 62DS Circular Dichroism Spectrometer. CD samples were prepared by dilution of peptide stocks to concentrations between 50 and 200 μM in degassed water for experiments performed at pH 4.0 or in 0.01M pH 8.0 Tris buffer for experiments performed at pH 8.0. The pH of the sample was adjusted to either 4.0 or 8.0 by addition of 0.01N NaOH. CD measurements were performed at room temperature in a 0.1 cm path length quartz cell with the optical chamber continually flushed with dry N_2 gas. Scans were collected the range from 195 nm to 250 nm, with a band width of 1.5-2.0 nm, a sensitivity of 50 mdeg, a scan speed of 50 nm/min, a time constant of 0.5 sec, and a scan resolution of 0.2 nm. Four accumulations per sample were performed. CD spectra were analyzed in KaleidaGraph software version 2.11 and are reported in (deg x $cm^{2)}$/(dmol x amide bonds).

Nuclear magnetic resonance studies.

2D NMR experiments were performed on a 500 MHz Brucker Instrument AMX500 spectrometer or a Varian 600 MHz Unity Plus instrument. NMR samples were prepared in either a 90:10 H_2O/D_2O solution or a 99.98% D_2O solution with dioxane or dimethyl sulfoxide as an internal standard. Peptide concentration ranged from 2-10 mM. The pH was adjusted to 8.0 with small additions of 0.1N NaOH or 0.1N NaOD. NOEs were detected using spin-locked rotating-frame nuclear Overhauser effect spectroscopy (ROESY) with a 400 ms

mixing time,[25] or NOESY with a 250 ms mixing time.[26] Water suppression was obtained using presaturation during the relaxation delay. Spectra were processed and analyzed in Felix 95 (Biosym Technologies, San Diego, CA, USA) and NOESY spectra were baseline corrected.

Pam amino acid pK_a determination.

Solutions of buffer at 0.2M concentration were prepared for every 0.5 pH unit between pH 6.5 and 11.5 (pH 6.5-7.5: HEPES; pH 8.0-9.5: CAPSO; pH 10.0-11.5: CAPS). Peptide samples at each pH were prepared by diluting 100 μL of CHP peptide stock into 150 μL of buffered solution and then were transferred to a 1.0 cm path length quartz cuvette. Using a Beckman DV 7500 spectrophotometer, the UV-Vis absorption spectrum between 250 and 450 nm for each peptide sample at each pH was collected with the appropriate blank spectrum subtracted. The absorption at 326 nm was plotted as a function of pH and fit to an expression of percent protonation as a function of pH using KaleidaGraph software version 2.11. From this fit, the pK_a of the ring nitrogen of the Pam amino acid in each peptide was determined.

Transamination assays.

For the BP peptides, a typical assay was performed as follows: 2 μl of a 10 mM EDTA solution (100 μM), a 10 μl of a 2 mM BP5 stock (100 μM), and 148 μl of a pH 8.0 0.1 M HEPPS buffer with 0.02% NaN_3 and an ionic strength of 0.29 were added to a small eppendorf. The solution was equilibrated at room temperature for 10 minutes, and the reaction was initiated by the addition of 40 μl of a fresh

50 mM sodium pyruvic acid/50 mM D-phenylalanine solution (10 mM final concentration of each). Assays proceeded at room temperature and over 24 hours a number of 10 µl time-points were collected. For the CBP01-CBP18, CHP1-CHP9, and alkylated CBP17 peptides, the total volume, temperature, and order of addition were the same. However, peptide and EDTA or $CuCl_2$ were added to a final concentration of 10 µM. Reactions were initiated with 40 µl of a 50 mM sodium pyruvic acid/50 mM D-phenylalanine solution (10 mM of each) and were stopped before reaching 10% of a single turnover. Reactions with copper(II) ion were quenched after one hour with a 5-fold excess of EDTA. Metal free assays were stopped after 20 hrs. All samples not derivatized immediately were flash frozen and stored at -80° C. Transamination assays for the dPam model compound were performed both in the presence and absence of copper(II) ion under the conditions described for the other Pam-containing peptides.

The amino acids generated in the assay were derivatized with saturated sodium borate buffer and freshly prepared N-acetyl cysteine (5 mg/ml)/OPA Incomplete Solution from Sigma according to literature procedure.[19, 20] To derivatize the samples generated in the presence of peptides BP1-BP6, 20 µl of the saturated sodium borate buffer was added followed by 40 µl of the OPA solution. The reaction mixture was vortexed and then allowed to sit for exactly five minutes at room temperature before one-half of the reaction mixture was injected onto the HPLC (36 µl). For samples from all other peptides, 100 µl of sodium borate buffer and 500 µl of the OPA solution was added, mixed, and allowed to react for five minutes at room temperature before the full volume of the sample was injected.

Resolution of derivatized products was performed using a Beckman HPLC equipped with either a 5 micron 4.6 x 150 mm Beckman or 5 micron 4.6 x 250 mm Microsorb C_{18} Reverse Phase Column. Methanol and a solution of 80 mM sodium citrate, 20 mM sodium phosphate, with 0.02% sodium azide at pH 6.8 were employed as co-solvents. Gradients were optimized for each reaction condition tested. The fluorescent signals were detected on a Waters 470 scanning fluorescence detector with an excitation wavelength of 344 nm and an emission wavelength of 443 nm.

Quantities of the amino acid enantiomers produced were calculated by integration of peaks from calibration curves of D-, L- amino acid standards generated over a range of concentrations. For the BP1-BP6 peptides, the initial reaction rate was determined by fitting the total amino acid product for each time-point to a pseudo first-order exponential build-up curve. From this fit the initial rate was extracted. For all other peptides, the extent of reaction was calculated from the total quantity of amino acid produced. The enantiomeric excess was calculated from the difference in quantity of amino acid enantiomers produced relative to the total quantity of amino acid produced. All samples were performed in duplicate.

References

(1) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman and Company: New York, 1985; pp 475 .

(2) Stryer, L. *Biochemistry*; W. H. Freeman and Company: New York, 1995; pp 1063.

(3) Smith, P. A. S. *The Chemistry of Open-Chain Organic Nitrogen Compounds*; W. A. Benjamin: New York, 1965.

(4) Akhtar, M.; Emery, V. C.; Robinson, J. A. Pyridoxal Phosphate Dependant Enzymatic Reactions: Mechanism and Stereochemistry. In *The Chemistry of Enzyme Action*; Page, M. I.; Elsevier Science Publishers B. V., Amsterdam, 1984.

(5) Hayashi, H. "Pyridoxal enzymes: Mechanistic diversity and uniformity," *J. Biochem.*, **1995**, *118*, 463-473.

(6) Johns, R. A. "Pyridoxal dependent enzymes," *Bioch. Bioph. Acta.*, **1995**, *1248*, 81-96.

(7) Ando, M.; Tachibana, Y.; Kuzuhara, H. "Chemistry of chiral vitamin-B6 analogs. 4. Synthesis of chiral pyridoxal and pyridoxamine analogs having a branched ansa chain between 2'-positions and 5'-positions," *Bull. Chem. Soc. Jpn.*, **1982**, *55*, 829-832.

(8) Breslow, R.; Czarnik, A. W.; Lauer, M.; Leppkes, R.; Winkler, J.; Zimmerman, S. "Mimics of transaminase enzymes," *J. Am. Chem. Soc.*, **1986**, *108*, 1969-1979.

(9) Kikuchi, J.; Zhang, Z. Y.; Murakami, Y. "Enantioselective catalysis by a supermolecular bilayer-membrane as an artificial aminotransferase-Stereochemical roles of an L-lysine residue and L-phenylalanine at the reaction site," *J. Am. Chem. Soc.*, **1995**, *117*, 5383-5384.

(10) Breslow, R.; Canary, J. W.; Varney, M.; Waddell, S. T.; Yang, D. "Artificial transaminases linking pyridoxamine to binding cavities-Controlling the geometry," *J. Am. Chem. Soc.*, **1990**, *112*, 5212-5219.

(11) Kuang, H.; Distefano, M. D. "Catalytic enantioselective reductive amination in a host-guest system based on a protein cavity," *J. Am. Chem. Soc.*, **1998**, *120*, 1072-1073.

(12) Shogren-Knaak, M. A.; Imperiali, B. "Modulating pyridoxamine-mediated transamination through a ββα-motif peptide scaffold," *Biorganic Med. Chem.*, **1999**, *7*, 1993-2002.

(13) Sinha Roy, R.; Imperiali, B. "Stereoselective synthesis of a pyridoxamine coenzyme-amino acid chimera-assembly of a polypeptide incorporating the pyridoxamine moiety," *Tetrahedron Lett.*, **1996**, *37*, 2129-2132.

(14) Sinha Roy, R.; Imperiali, B. "Pyridoxamine amino-acid chimeras in semisynthetic aminotransferase mimics," *Protein Eng.*, **1997**, *10*, 691-698.

(15) Sinha Roy, R.; Imperiali, B. "Coenzyme-amino acid chimeras: New residues for assembly of functional proteins," *J. Am. Chem. Soc.*, **1994**, *116*, 12083-12084.

(16) Sinha Roy, R.; Imperiali, B. "Stereoselective synthesis and peptide incorporation of a pyridoxal coenzyme-amino acid chimera," *J. Org. Chem.*, **1995**, *60*, 1891-1894.

(17) Struthers, M.; Ottesen, J. J.; Imperiali, B. "Design and NMR analysis of compact, independently folded ββα-Motifs," *Fold. Des.*, **1998**, *3*, 95-103.

(18) Marfey, P. "Determination of D-amino acids. II. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene," *Carlsberg Res. Comm.*, **1984**, *49*, 591-596.

(19) Buck, R. H.; Krummen, K. "High-performance liquid-chromatographic determination of enantiomeric amino-acids and amino-alcohols after derivatization with ortho-phthaldialdehyde and various chiral mercaptans-Application to peptide hydrolysates," *J. Chromatogr.*, **1987**, *387*, 255-265.

(20) Nimura, N.; Kinoshita, T. "Ortho-phthaldialdehyde-N-acetyl-L-cysteine as a chiral derivatization reagent for liquid chromatographic optical resolution of amino acid enantiomers and its application to conventional amino-acid-analysis," *J. Chromatogr.*, **1986**, *352*, 169-177.

(21) Martell, A. E. "Vitamin B6 catalyzed reactions of α-amino and α-keto acids: Model systems," *Acc. Chem. Res.*, **1989**, *22*, 115-124.

(22) Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Economy in protein design: Evolution of a metal-independent ββα-motif based on the zinc finger domain," *J. Am. Chem. Soc.*, **1996**, *118*, 3073-3081.

(23) Novabiochem *1999 Novabiochem Catalog and Peptide Synthesis Handbook*; Novabiochem: San Diego, 1999.

(24) King, D. S.; Fields, C. G.; Fields, G. B. "A cleavage method which minimizes side reactions following Fmoc solid-phase peptide-synthesis," *Int. J. Pept. Protein. Res.*, **1990**, *36*, 255-266.

(25) Bax, A.; Davis, D. G. J. "Practical aspects of two-dimensional transverse NOE spectroscopy," *J. Magn. Reson.*, **1985**, *63*, 207-213.

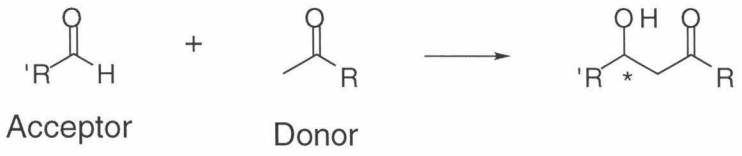
(26) States, D. J.; Haberkorn, R. A.; Ruben, D. J. "A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants," *J. Magn. Res.*, **1982**, *63*, 207-213.

Chapter 3: Incorporating Aldol Condensation Activity into ββα Motif Peptides

Introduction

To address some of the limitations demonstrated by the coenzyme amino acid chimera strategy, it was necessary to explore a new system and strategy for incorporating function into BBA scaffold peptides. In choosing a system, it was important to focus on a reaction that was chemically interesting and had a precedent for a simple mechanism of generating catalysis. However, even if such a simple mechanism were available, choosing the appropriate amino acids to facilitate this mechanism might still be difficult. Thus, in choosing a strategy, it was important to focus on peptides that could be easily generated and assayed, allowing for rapid characterization of potential peptide candidates. All of these criteria could be met in peptides designed to mediate the aldol condensation reaction.

The aldol condensation reaction has long interested chemists. In this reaction, a donor and acceptor carbonyl containing substrate are condensed to form a β-hydroxy-carbonyl product (Scheme 3.1)

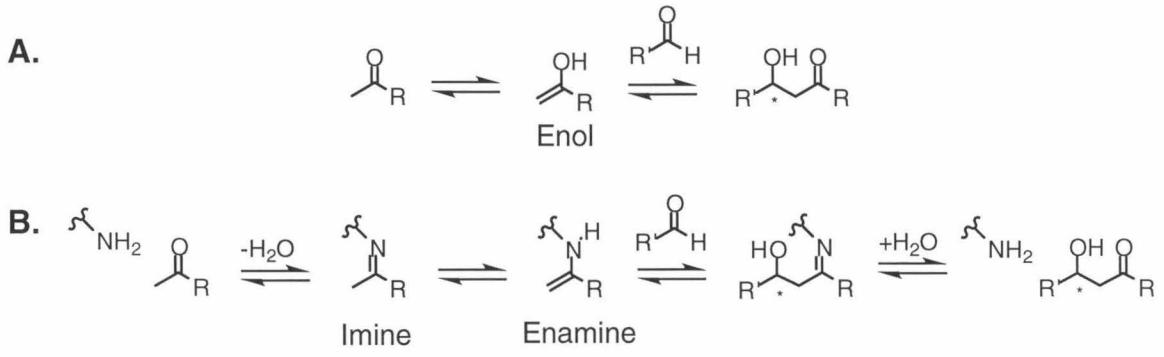


Scheme 3.1. Aldol condensation reaction.

This reaction is a robust means of forming carbon-carbon bonds and has the potential to introduce a chiral center at the β-alcohol moiety. From a synthetic organic perspective, this reaction has proven very useful,[1] and a great deal of

attention has been devoted to developing catalytic and asymmetric catalytic reagents to facilitate it.[2] From a biological perspective, aldol condensation chemistry is one of the primary means of forming carbon-carbon bonds anabolically.[3]

Aldolase enzymes are extremely efficient and effective in catalyzing this condensation and do so through two different mechanisms.[4] The type I aldolase enzymes facilitate this reaction through the use of a primary amine (Scheme 3.2B), while type II aldolases utilize a bound zinc ion.



Scheme 3.2. Mechanism of A) uncatalyzed and B) amine-catalyzed aldol condensation.

In this mechanism, the initial step is formation of an imine group through nucleophilic attack of a side-chain amine on the donor substrate. Relative to the uncatalyzed reaction (Scheme 3.2A), this Schiff base is activated toward deprotonation of the carbon α- to the former carbonyl, and removal of this proton results in the generation of the enamine intermediate. This intermediate can act as a nucleophile toward the acceptor substrate to generate the β-alcohol. Hydrolysis of the resulting imine liberates the β-hydroxy-carbonyl product and

regenerates the primary amine group. This amine can then react with a new substrate donor, to reinitiate the catalytic cycle.

The active site composition of many type I aldolase enzymes appears to be relatively simple. From the crystal structure of the *Esherichia coli* aldolase, *N*-acetylneuraminate lyase,[5] it can be seen that the catalytic primary amine is located on the side chain of the lysine 165 (Figure 3.1).

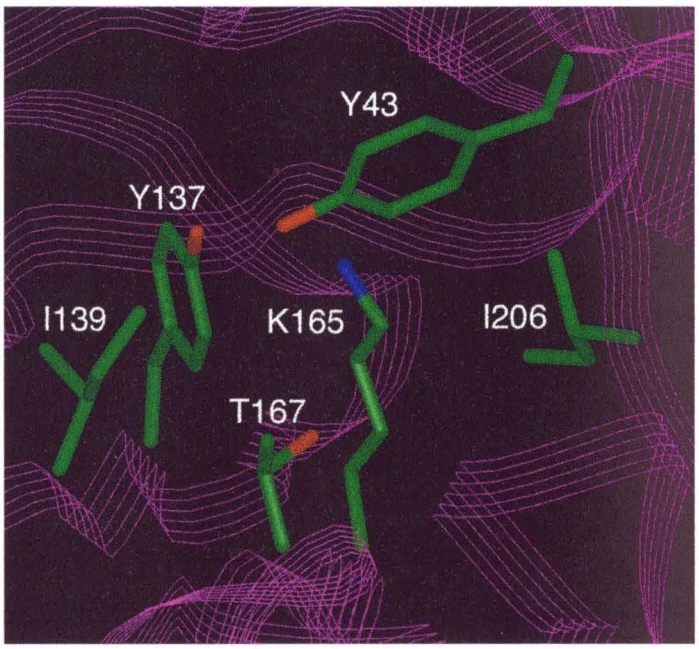


Figure 3.1. Structure of the *N*-acetylneuraminate aldolase active site.

Surrounding the amine are largely hydrophobic residues, including isoleucine, tyrosine, and threonine. It appears that the combination of this amine group and the hydrophobic core provide many of the factors necessary to facilitate aldol condensation by the amine-catalyzed mechanism. The primary amine of lysine allows for the formation of the imine and enamine reaction intermediates.

Moreover, because a largely hydrophobic core surrounds this amine, protonation of the amine is disfavored. In its unprotonated form, this amine is a more effective nucleophile for attack on the donor substrate. This hydrophobic core also facilitates the condensation by entropic effects. Aldol condensation is a bimolecular process, and by sequestering hydrophobic substrates, pseudo-unimolecular reaction kinetics are exhibited.

Due to the relatively simple mechanism type I aldolase enzymes employ to facilitate aldol condensation, approximating this strategy in other systems is attractive and has been demonstrated. Reymond and coworkers have shown that a primary amine-containing molecule can be incorporated (Figure 3.2), into an antibody.[6]

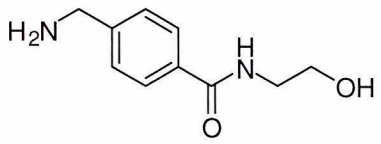


Figure 3.2. Primary amine incorporated into an antibody by Reymond and coworkers.

This antibody complex accelerates a representative aldol condensation, increasing k_{cat} relative to k_{uncat} by a factor of 10-100 and sequestering substrate to provide a millimolar K_m. This idea was explored further by Lerner and coworkers, finding that the lysine group of an antibody directly provides the catalytic amine.[7, 8] In this system, substrate binding and turnover is improved. Like the type I aldolases, the basis for this activity is a primary amine in a

hydrophobic core, as seen in a crystal structure of aldolase antibody 33F12 (Figure 3.3).[9]

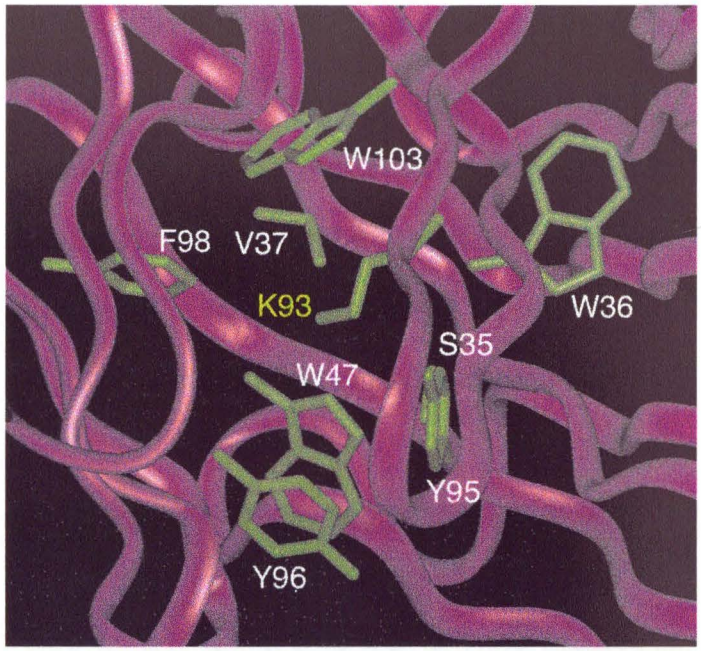


Figure 3.3. Structure of the aldolase antibody 33F12 active site.

In this structure, lysine 93 acts as the catalytic amine and is surrounded by hydrophobic amino acids including tryptophan, phenylalanine, tyrosine, and valine. By incorporating the amine into this core, the pK_a of the lysine is reduced from 10.8 to 6.5, making it a more effective nucleophile at the reaction pH.

The relatively simple active sites of the natural type I aldolase enzymes and non-natural aldolase antibodies suggested that these features could also be incorporated into the hydrophobic core of the BBA scaffold. However, choosing which specific amino acids to use to create such a core was not trivial. Work with the CBP family of peptides had revealed the structural and functional difficulties

inherent in changing the sequence of the BBA peptide scaffold. Changes in the primary sequence of the BBA scaffold, especially in the hydrophobic core of the peptide, could result in loss of secondary and tertiary structure (Section 2.2). Even those peptides that did maintain structural aspects of the desired ββα-motif did not necessarily access a well-defined constellation of functional groups that are required to mediate coenzyme reactivity (Section 2.2-2.3). Thus, a strategy in which many different peptides could be generated and screened was sought.

In the highly parallel screening strategy we envisioned (Figure 3.4), a large library of BBA peptides would be generated synthetically.

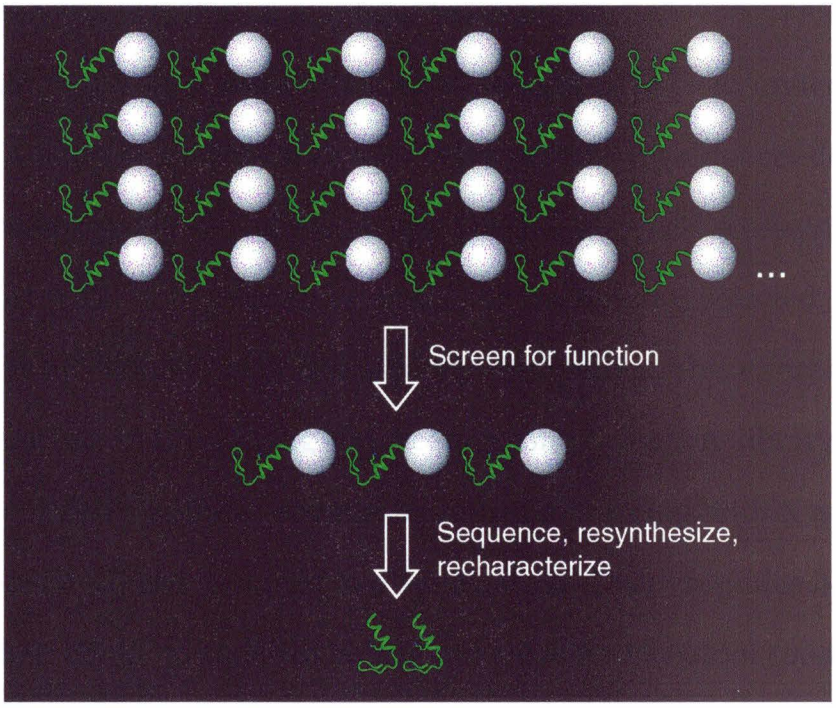


Figure 3.4. High throughput functional screening strategy.

This library would be associated with the synthetic support and screened for functional features. From the subset of selected peptides, further characterization and screening would be performed to find the best peptides. By focusing on peptides that use a basic amino acid in a hydrophobic core to mediate aldol condensation, such a highly parallel screening strategy seemed feasible.

Because the desired active site of the peptides was relatively simple, consisting of an amino acid with a primary amine side chain and hydrophobic amino acids, synthesis of a large group of peptides should be routine. Additionally, because a number of different techniques for characterizing aldol condensation intermediates and products were available, it seemed feasible to extend these techniques to high throughput screening.

In collaboration with Kevin McDonnell, efforts have been undertaken to impart the BBA peptide scaffold with the ability to facilitate aldol condensation. This effort is comprised of two related parts: 1) generating and screening a large BBA library for functional properties and 2) characterizing selected peptides structurally and functionally to understand their properties. This effort is not yet complete. However, the following has been accomplished. A library of over 100,000 different BBA peptides containing a primary amine within a variable amino acid core have been designed and synthesized. These peptides have been screened for their ability to stabilize imine- and enamine-like aldol condensation intermediates, producing a subset of about 100 potential peptides. Approximately 20 of these peptides have been sequenced to reveal interesting sequence trends. Methods have been developed to perform further functional screening of these peptides. A promising peptide discovered early in primary screening has been more fully characterized structurally and functionally. These

studies have suggested that this peptide can incorporate a primary amine into a hydrophobic core without disturbing the secondary or super-secondary structure. In terms of function, this peptide possesses a fairly acidic amine group and can stabilize enamine formation relative to a model peptide to a high degree, facilitating enamine formation.

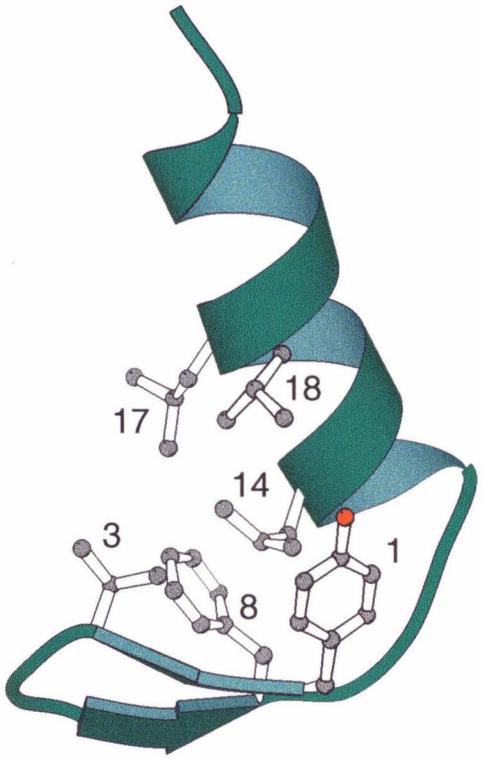
Results and Discussion

3.1. CPLB peptide library design.

The CPLB family of peptides was designed to explore the ability of a diverse set of BBA peptides to mediate the aldol condensation, where reactivity is achieved by incorporating an amino acid with a primary amine side chain into a hydrophobic core. As described above, natural and antibody aldolase enzymes facilitate aldol condensation through a two-fold strategy. At the core of these enzymes a primary amine, such as the side chain of lysine, acts as a nucleophile. By reacting with the electrophilic carbonyl substrate, the resulting imine is both sequestered covalently, and activated chemically (Scheme 3.2). Surrounding this amine is a core of largely hydrophobic amino acids. These residues serve to help sequester both substrates, paying part of the entropic cost of the bimolecular reaction, and increasing the nucleophilicity of the δ-amine of lysine by encouraging its deprotonation.

The BBA peptide scaffold appeared potentially well suited to support both a basic amino acid and a hydrophobic core. At the interface of the β-sheet

and α-helix, the BBA peptides possess a hydrophobic core of six key residues: Tyr1, Val3, Phe8, Leu14, Leu17, and Leu18 (Figure 3.5A).[10-12] At the center of this core is leucine 14. By substituting this leucine with an amino acid with a primary amine side chain, this residue would be well positioned to approximate the catalytic core of the aldolase proteins, both sequestering and activating the aldol condensation substrates.



Pos	BBA5	CPLB
0	Ac	G
1	**Y**	**Y, L, I, F, A, W, V, H, N**
2	R	R
3	**V**	**V, I, T**
4	dP	dP
5	S	S
6	Y	Y
7	D	D
8	**F**	**Y, L, I, F, A, W, V, H, N, R**
9	S	S
10	R	R
11	S	S
12	D	D
13	E	E
14	**L**	**Dap, Dab, Orn, K**
15	A	A
16	K	K
17	**L**	**Y, L, I, F, A, W, V, H, N, R**
18	**L**	**Y, L, I, F, A, W, V, H, N, R**
19	R	R
20	Q	Q
21	H	H
22	A	A
23	G	G

Figure 3.5. A) Structure of the BBA5 peptide. B) Sequences of the BBA5 peptide and the CPLB family of peptides.

While this general strategy seemed promising, the choice of the specific basic amino acid and surrounding residues was not obvious. From a functional perspective, it was unclear which amino acids should be utilized. Enzyme active

sites often require the precise alignment of functional groups to maximize chemical activity[13] and it is not trivial to predict *de novo* which surrounding core residues would effectively tune the reactivity of the nucleophilic amine and sequester the condensation substrates. To further complicate this decision, because peptides were to be generated by solid phase peptide synthesis, non-natural amino acids needed to be considered. From a structural perspective the choice of amino acids was also important. Leucine 14, along with tyrosine 1 and phenylalanine 8, are the three core hydrophobic residues that are strongly conserved in natural BBA motifs from zinc-finger domains.[14] Our own experience with substituting Pam at position 8 suggests that these residues cannot be replaced easily without undesirable structural consequences (Section 2.1). The substitution of the hydrophobic leucine 14 residue with a polar and potentially charged basic residue might also be structurally unfavorable. Thus, a wide range of amino acid substitutions was explored to find BBA peptides that exhibit the desired structural and functional characteristics (Figure 3.5B).

Peptide BBA5 was chosen as the parent peptide for the CPLB library of peptides (Figure 3.5B). BBA5 is part of the BBA family of designed peptides, has been structurally well characterized, and demonstrates well-defined secondary and tertiary structure. Like other members of this family, it possess a heterochiral DPro-Ser type II′ turn. However, unlike the other structured BBA peptides, BBA1 and BBA4, the DPro is the only non-natural amino acid.

Four basic amino acids, lysine, ornithine (Orn), γ-diaminobutyric acid (Dab), and β-diaminoproprionic acid (Dap), were selected to replace leucine 14 (Figure 3.5B). These primary amine containing amino acids span a range of side

chain lengths, having between one (Dap) and four (Lys) methylene units between the amino acid α-carbon and side chain amine. They also encompass a range of basicities, with pK_a's of the side chain amine of the free amino acid ranging from 9.7 for Dap to 10.8 for Lys[15] (pK_a values for these groups in peptides were investigated and are discussed later in Section 3.7). To what extent achieving the desired structural and functional properties requires facile deprotonation versus the need to exclude the amine group from direct contact with the hydrophobic core was unclear and necessitated the investigation of the full range of amino acids.

Hydrophobic core residues 1, 3, 8, 17, and 18 of BBA5 were replaced with a range of hydrophobic, neutral polar, and positively charged amino acids (Figure 3.5B). These five residues are in close proximity to position 14 and form a microenvironment for the basic group. These core residues were selected to perturb the pK_a of residue 14, to generate a core capable of sequestering the substrates, and to maintain the packing and structure of the BBA peptide. At position 3 in BBA and in model peptides it has been shown that β-branched residues are required to achieve the type II' turn of the sheet region. To maintain this feature, only β-branched amino acids were incorporated into this site.

The combination of all six of these substitutions results in a library of peptides with more than 108,000 unique members.

3.2. CPLB peptide library synthesis.

A combination of conventional and split and pool solid-phase peptide synthesis were employed to generate the 108,000 different BBA peptide

sequences. Split and pool solid-phase synthesis is an ideal method for generating diversity in a controlled fashion (Figure 3.6).[16]

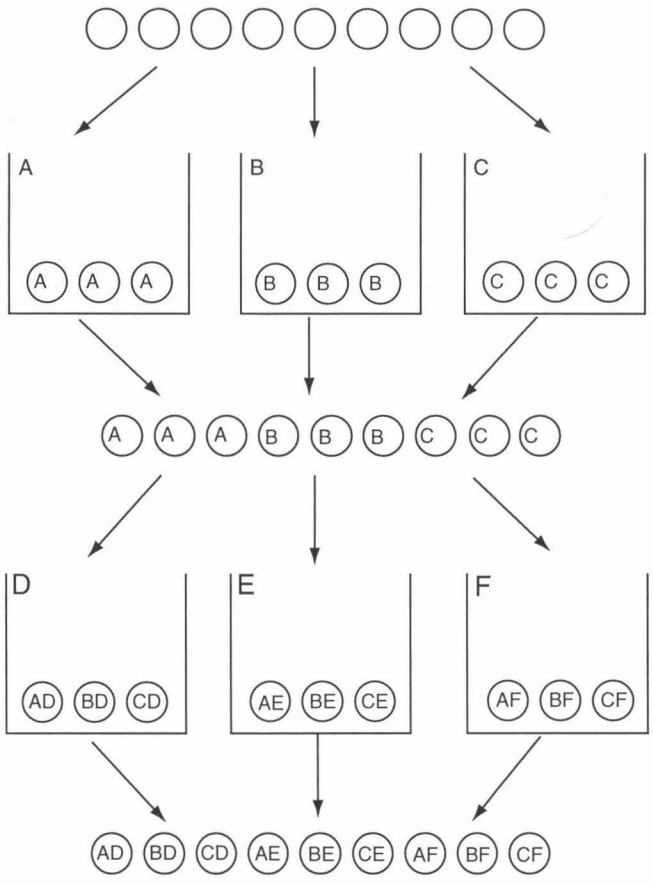


Figure 3.6. Split and pool synthetic method.

As applied to peptide synthesis, peptide beads are separated at the positions where sequences differ, and the desired amino acids are coupled to the peptides. All beads are then re-pooled and split again. Using this approach it is possible to cover a significant portion of sequence space in few chemical steps. Furthermore, because each peptide bead is exposed to only one amino acid at every coupling step, each bead carries a unique sequence.

Polyethylene-glycol-dimethylacrylamide (PEGA) resin was utilized as the synthetic support.[17] For initial functional screening, it was desirable to screen directly from the beads to make handling of the peptides more practical and rapid. Thus, despite the variability of bead size, PEGA resin was chosen as the synthetic support because it has been shown to have good water accessibility and to provide a solution-like environment for peptides.[18] Because initial screening (and subsequent sequence readout) was performed directly on the peptide-beads, peptides were coupled directly to the resin without a cleavable linker.

The use of an excess of resin beads relative to peptide sequences was employed to ensure good statistical representation of the possible sequences. In the split and pool method, at a position with sequence variation the pooled resin beads are split amongst reaction vessels containing one of the possible amino acids at that position. This splitting occurs randomly, so if beads of a given sequence do not happen to be put into all of the possible coupling vessels, some of the designed sequences are not represented in the final pool of peptides. To address this issue, it is possible to use a large excess of resin beads relative to the number of sequences. However, the problem then becomes that a large redundancy of each sequence is present, requiring additional screening. The coverage of possible sequences as a function of number of beads can be best described by a Poisson distribution (Figure 3.7),[19] where k is the number of times a given library sequence occurs, and m is the number of beads divided by the number of library sequences.

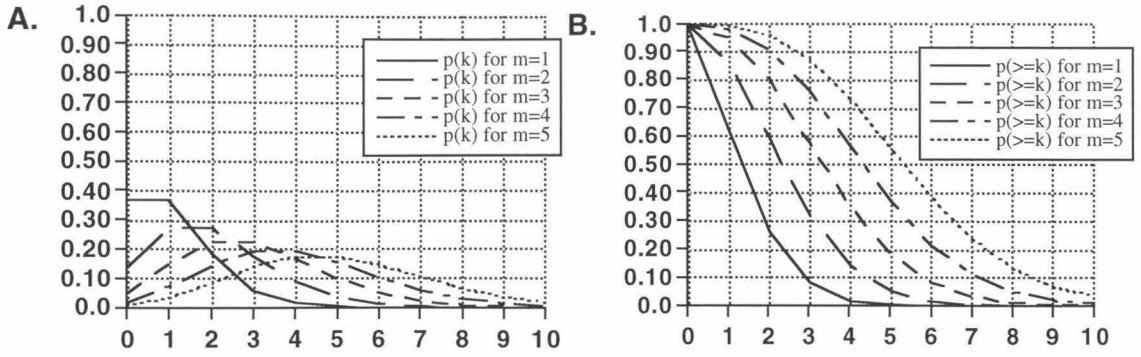


Figure 3.7. A) Poisson equation and distribution. B) Summed distribution.

This distribution suggests that for 98% coverage of all sequences, a four-fold excess of resin beads to sequences is necessary. In the synthesis of the CPLB peptide library, enough resin was used to generate four libraries each with approximately a four-fold excess of beads to sequence.

Couplings were performed with standard Fmoc-amino acids using HOBT/HBTU chemistry with extended coupling times. In the synthesis of the CPLB peptide library, long stretches of common sequence (for example residues 23-19) were generated using an automated peptide synthesizer. Shorter stretches of common sequence were synthesized by standard batch coupling protocols. At positions where the amino acid sequence varied, peptide resin was divided in equal portions reflecting the number of different amino acids at that position. The desired amino acid was then batch coupled in parallel. Extent of coupling was monitored spectroscopically for the semi-automated coupling and by a single batch Kaiser test[20] during batch coupling. No acetyl capping was performed between coupling steps. After the final amino acid coupling, the amino terminus of the peptide was capped with an α-chloroacetyl group. This

group was included to insure that a free amino *N*-terminus would not alter the structural properties of the scaffold and would not compete as a nucleophilic amine in the reaction assays. Section 4.2 contains a more complete discussion of the motivation for using the α-chloroacetyl capping group.

To assess the quality of the CPLB library, random library beads were analyzed by Edman degradation sequencing after conversion of the α-chloroacetyl group to glycine (Table 3.1). Sequencing was performed on the beads using an automated "sequenator" and standard protocols. The sequencing data indicated that 50 and 100 pmoles of peptide were present on each bead with little or no truncation products (The Dap-containing peptide analysis was slightly more complicated, but suggested the same result. See Section 4.2). All peptide beads gave expected sequences at positions that were common, and reasonable sequences for the positions that varied. These sequencing results suggested that the quality of the synthesized library was sufficient for further experiments.

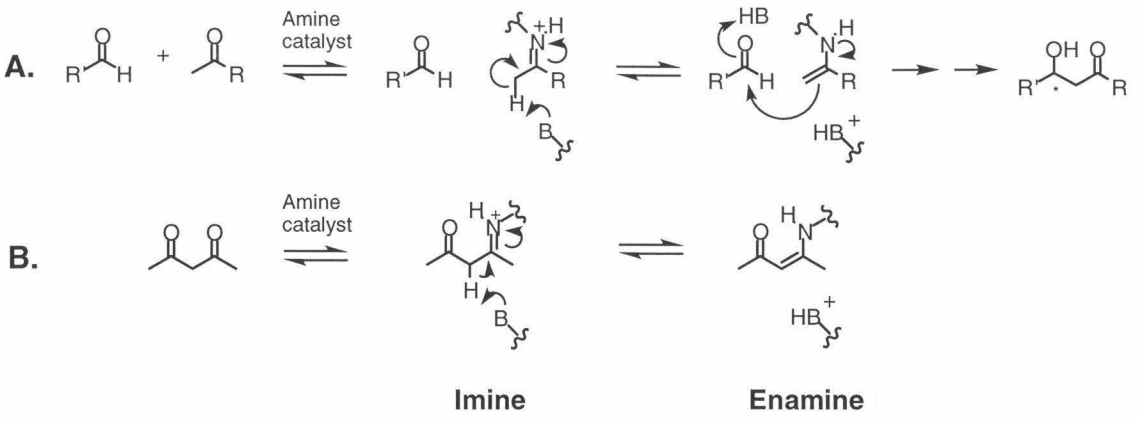
3.3. Primary screening of the CPLB peptide library.

Primary screening of the CPLB peptide library was performed to find peptides capable of stabilizing key intermediates in the aldol condensation process. This screening was accomplished by looking for association of a fluorescent dione probe with individual peptide beads.

Pos	BBA5	CPLB	CPLB-R1	CPLB-R2	CPLB-R3
0	Ac	G	G	G	G
1	**Y**	**Y, L, I, F, A, W, V, H, N**	**F**	**I**	**Y**
2	R	R	R	R	R
3	**V**	**V, I, T**	**T**	**T**	**I**
4	dP	dP	P	P	P
5	S	S	S	S	S
6	Y	Y	Y	Y	Y
7	D	D	D	D	D
8	**F**	**Y, L, I, F, A, W, V, H, N, R**	**R**	**Y**	**R**
9	S	S	S	S	S
10	R	R	R	R	R
11	S	S	S	S	S
12	D	D	D	D	D
13	E	E	E	E	E
14	**L**	**Dap, Dab, Orn, K**	**K**	**Dap**	**K**
15	A	A	A	A	A
16	K	K	K	K	K
17	**L**	**Y, L, I, F, A, W, V, H, N, R**	**W**	**F**	**W**
18	**L**	**Y, L, I, F, A, W, V, H, N, R**	**N**	**R**	**I**
19	R	R	R	R	R
20	Q	Q	Q	Q	Q
21	H	H	H	H	H
22	A	A	A	A	A
23	G	G	G	G	G

Table 3.1. Sequencing results of random CPLB library beads.

To find peptides capable of promoting the aldol condensation reaction it was necessary to devise a screen that would report on the ability of these peptides to form condensation pathway intermediates. To be of practical utility, this screen also needed to be fast, sensitive, and accurate. In the amine-mediated reaction pathway a number of chemical intermediates exist (Scheme 3.3A).



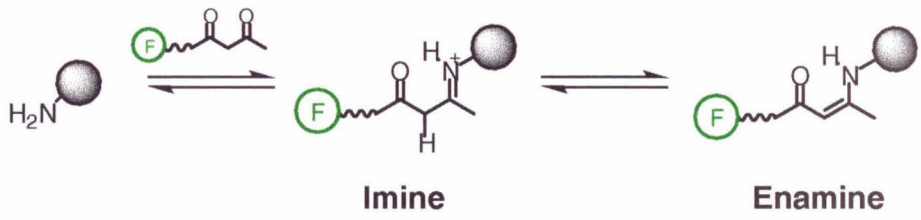
Scheme 3.3. Reaction intermediates in the aldol condensation mechanism for A) the condensation substrates and B) dione inhibitor.

For the initial screen, we chose to focus on the formation of imine and enamine intermediates, because of the ease with which these intermediates could be probed.

To probe formation of the imine and enamine intermediates, a dione probe was chosen. Molecules similar to beta diketones are known mechanism-based inhibitors of enzymes related to aldolase.[21] This probe is effective as an inhibitor because it both physically resembles the aldolase acceptor and donor substrates and because it can undergo a similar mechanism as the substrates (Scheme 3.3B). In the inhibition pathway, the dione reacts by first forming the imine intermediate. This species then tautomerizes to the stable enamineone. As this species, the dione is covalently trapped by the enzyme and is not carried forward on the aldol reaction pathway. This functional group could thus be used to search for peptides that could facilitate this covalent association, and thereby favorably catalyze the aldol condensation pathway. In fact, use of this functional group for this purpose had already been demonstrated. In their search for

antibodies capable of facilitating the aldol condensation, Lerner and coworkers utilized a dione conjugated to a hapten.[7] Immunization with this compound produced antibodies which were capable of covalently cross-reacting with the dione by enamineone formation. These antibodies were then shown to facilitate the aldol condensation with a range of acceptor and donor substrates. Thus screening the library for peptides capable of forming stable association with a dione probe seemed a promising means of also finding peptides capable of facilitating the aldol condensation reaction.

To make screening of dione association feasible it was necessary to detect this association. Fluorescence detection was attractive because of its sensitivity and robustness.[22] Attachment of a fluorophore to the dione would allow for direct detection of peptide-dione association. In this scheme, attachment of the fluorescent dione probe to the bead bound peptide localizes the fluorescent species to the bead (Scheme 3.4).



Scheme 3.4. Fluorescence detection strategy for dione association.

Thus, screening for dione-peptide association would only require finding the most fluorescent beads.

The choice of a fluorescent dione tag required a compound that was highly fluorescent, chemical unreactive, and small. Initial efforts focused on the attachment of an anthranilic acid fluorophore.[23] This compound was synthesized by Kevin McDonnell (Figure 3.8).

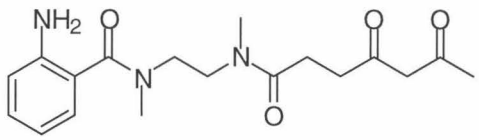
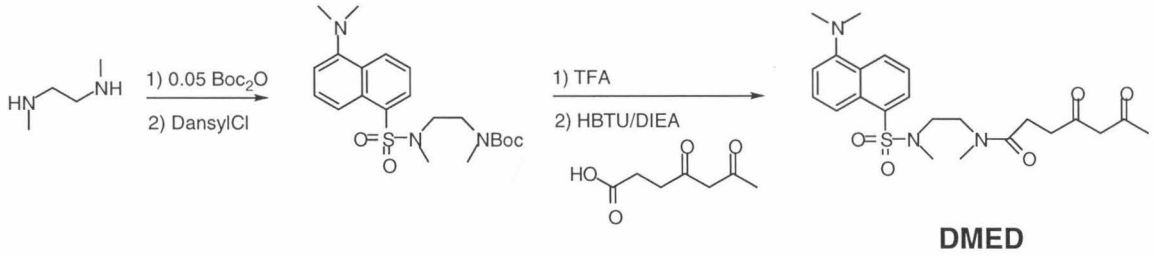


Figure 3.8. Anthranilic acid containing dione probe.

Despite the attractive features this fluorophore offered, it proved chemically self-reactive at the exocyclic amine. Attention was then turned to the attachment of a dansyl fluorophore to the dione. The dansyl group is relatively small and has reasonable photophysical properties, with an extinction coefficient in methanol of 4.2 x 10^{-3} $cm^{-1}M^{-1}$ at 335 nm and emission above 500 nm.[22]

Synthesis of the dansyl-associated dione (dansyl methyl ethylenediamine dione, DMED) was accomplished by Kevin McDonnell (Scheme 3.5).



Scheme 3.5. Synthetic scheme for the dansyl containing dione probe, DMED.

In this scheme, the diamine linker was Boc-protected at a single nitrogen and then reacted with the unprotected dioxoheptanoic acid compound using standard coupling chemistry conditions. Deprotection of the Boc group, followed by coupling to the activated dansyl-chloride species, resulted in the desired DMED product. The methylated diamine linker was used in place of a simple diamine linker because the unmethylated amide nitrogens reacted intramolecularly with the carbonyl groups.

DMED exhibited the desired characteristics to serve as a probe. The UV-Vis absorption and fluorescence spectrua resemble those of the parent dansyl fluorophore. Like the dansyl group DMED shows an absorption maximum at 335 nm with comparable extinction coefficients (Figure 3.9A).

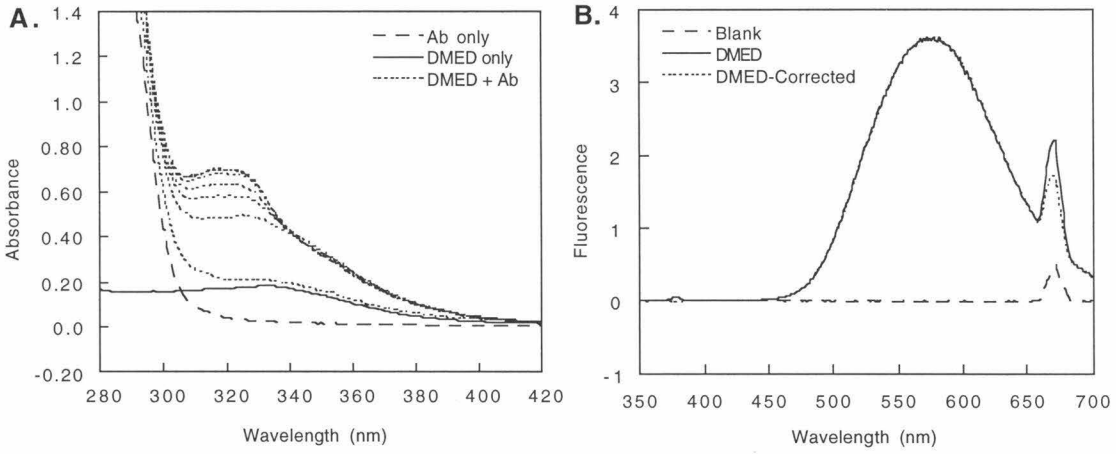


Figure 3.9. A) Absorption spectra of the DMED probe alone and in the presence of aldolase antibody 38C2 over the course of 10 minutes. B) Fluorescence spectrum of the DMED probe in water/5% methanol with excitation at 335nm.

Excitation at 335 nm results in strong emission with a maximum at 578 nm, similar to the dansyl group (Figure 3.9B). Furthermore, the fluorophore does not interfere with the ability of the dione to form enamine with reactive amine group. The extent of enamineone formation can be monitored spectroscopically by the UV absorption at 315nm (Figure 3.9A).[7] When 100 μM of DMED is added to 20 μM of commercially available aldolase antibody 38C2, a rapid increase in 315 nm absorption is observed. The magnitude of this absorption change is comparable to that observed in the reaction of 2,5-pentanedione and aldolase antibody, and suggests that the addition of the fluorophore tag does not preclude reaction between amines and the dione portion of DMED. Thus, the DMED probe exhibited the desired functional properties.

Fluorescent probe molecules associated with PEGA beads were visualized with an inverted fluorescent microscope attached to a color digital video camera and the excitation and emission signals were optimized. Excitation of the fluorophore was accomplished by a 100 W mercury lamp source. The excitation signal was filtered by an excitation cube with a bandpass wavelength of 330-385 nm. This filter was chosen for optimal dansyl excitation and minimal excitation below 300 nm. The intensity of this filtered excitation radiation caused rapid and irreversible photobleaching of the dansyl group. Thus, a neutral density filter was utilized to allow transmittance of only 12% of the original light intensity, preventing any significant photobleaching over several minutes of continuous exposure. To compensate for loss in excitation intensity, the camera gain was maximized. Emission was filtered by the excitation cube, with a cut off wavelength of approximately 400 nm. Emission signal was further filtered

digitally through the acquisition software. The blue signal channel was eliminated entirely to limit shorter wavelength background fluorescence. The dynamic range of the green channel was also modified. This modification introduced a minimum cut off intensity to facilitate more rapid screening and maximized the intensity difference between the most fluorescent beads. Quantitative assignments of bead intensity were made using image analysis software, where the average density of pixels contained within a bead was calculated.

These optical conditions gave good detection sensitivity for a wide range of bound dansyl molecules. A fluorescence intensity ladder was generated by derivatizing PEGA beads with zero to eight equivalents of activated dansyl-chloride molecules. Increasing intensity could be observed throughout this range, with no observable intensity for PEGA beads derivatized with zero equivalents of dansyl-chloride, and saturating intensities at eight equivalents (Figure 3.10). Furthermore, the sensitivity of the method was sufficient to detect beads derivatized with less than 0.1 equivalents of dansyl molecules, lower signal than we expected to investigate. Interestingly, variability in signal intensity was not clearly correlated with bead size, suggesting that it was not necessary to normalize bead signal intensity as a function of bead size. Fluorescent intensity did vary somewhat between beads for a given derivatization in a seemingly random fashion, suggesting a limitation in the ability to rank beads based on single bead intensities. Altogether, however, these results validated the approach of using fluorescence as a means of discriminating different amounts of bound dansyl molecules.

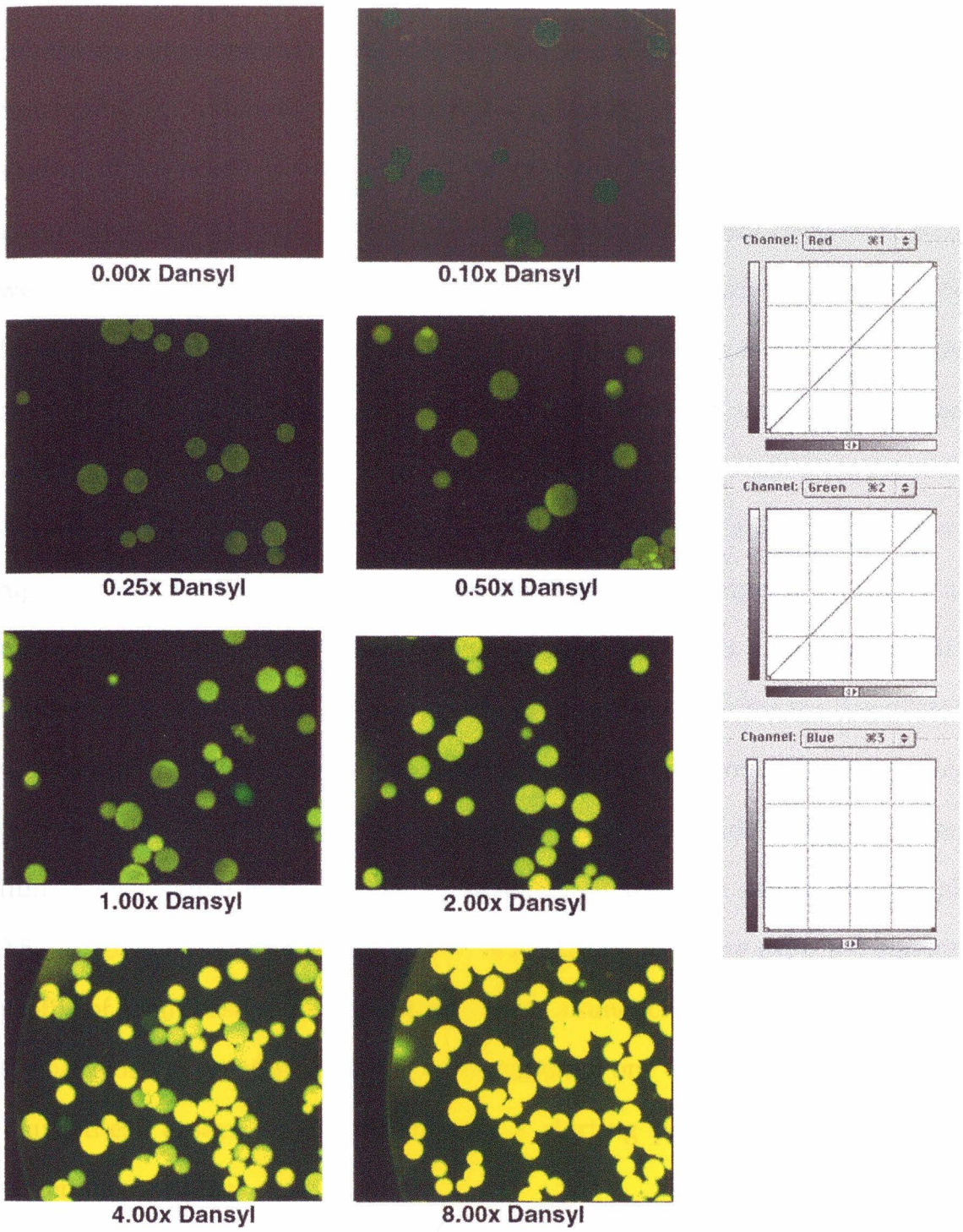


Figure 3.10. Fluorescence signal of dansyl-ladder beads in the presence of the 12% neutral density filter, maximum gain, and the blue color channel set to zero. Color channels are shown to the right.

Initial studies performed on a range of model compounds over a range of pHs conditions suggested that DMED association could be observed at high DMED concentrations, relatively long incubation periods, and at basic pH. In these studies, PEGA beads derivatized with a dansyl group (positive control), an acetyl group (negative control), and a lysine containing tripeptide (AcGKG-PEGA) were incubated for 24 hours either in the presence and absence of 250 μM DMED from pH 6.50 to pH 10.5 (Figure 3.11). From these data it is clear that in the absence of beads or DMED, relatively little fluorescence signal is observed. Acetylated PEGA beads in the presence of DMED provides some non-specific fluorescent signal. Dansylated PEGA beads in the presence of DMED provides a high level of fluorescent signal, and the lysine containing tripeptide exhibits an intermediate fluorescence intensity that appear to vary as a function of pH. Determining and plotting the average fluorescence intensity for the dansylated, acetylated, and peptide beads as a function of pH confirms this observation (Figure 3.11). Furthermore, the data show that tripeptide-bead associated fluorescence signal increase from pH 6.50 to 9.50 and then decreases from pH 9.50 to 10.50. This trend in association is consistent with the expected stability profile of Schiff base or enamine formation as a function of pH and suggests that the basis of this association is due to formation of one of these chemical species. Thus, these assays conditions provide a starting point for more stringent screening conditions.

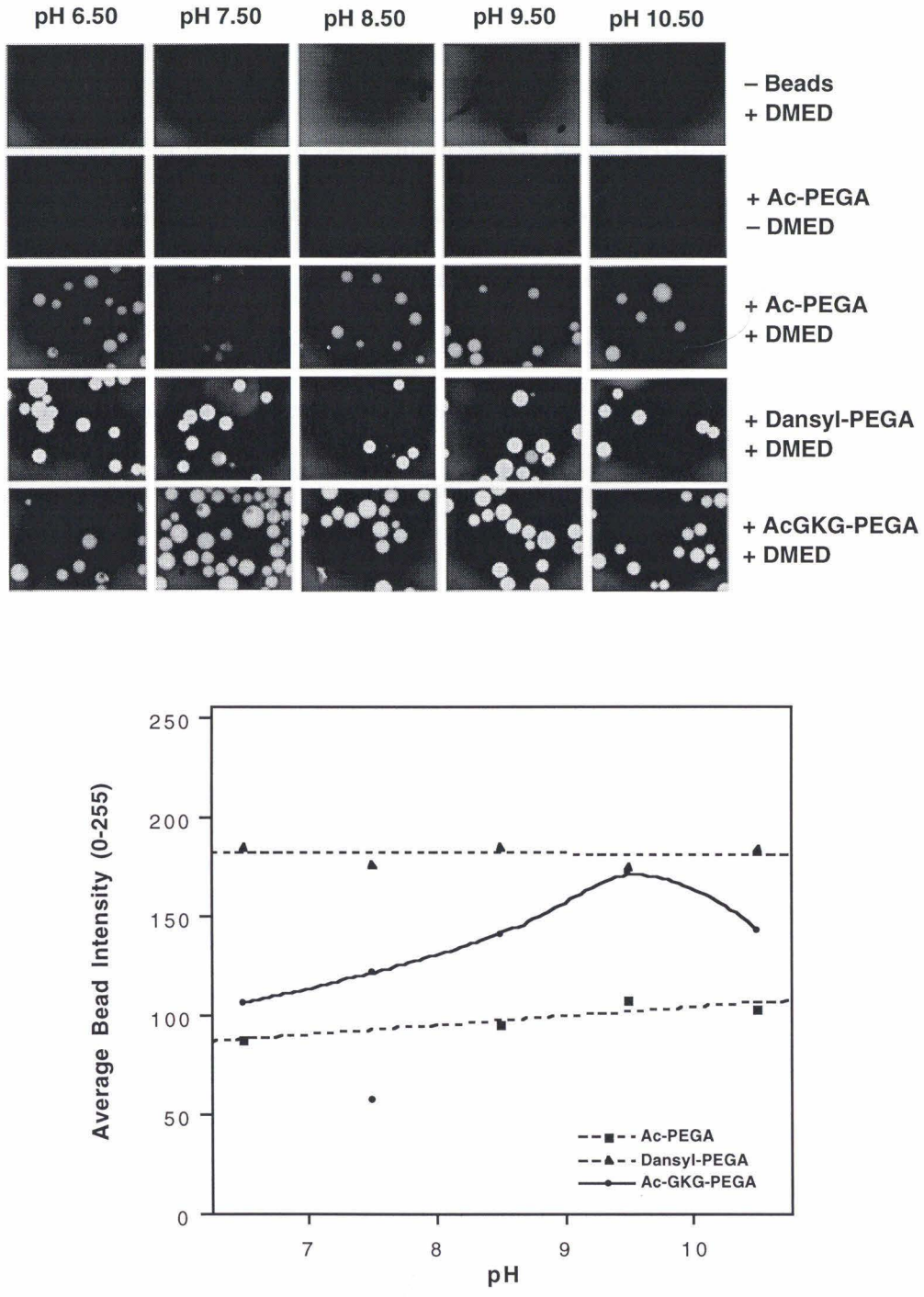


Figure 3.11. Results of the pH dependent association of DMED with model compounds and the plot of the associated intensities. 250 μM DMED, 0.05 M phosphate buffer, 24 hours.

Prior to screening the CPLB library beads for their ability to associate with the DMED probe, samples of the library were suspended in pH 7.50 buffer solution and prescreened. Screening of CPLB beads in the absence of DMED probe showed that the beads had some degree of background fluorescence over a range of intensities. A small portion of the beads (from 5-10 per approximately 4,000 beads) demonstrated sufficient inherent fluorescence to be visibly beyond the intensity cut off imposed by filtering of the green fluorescence channel. The number of inherently fluorescent beads exceeding the intensity threshold did not increase after a 24 hour incubation, suggesting that DMED independent drift in bead fluorescence was not an issue. These inherently fluorescent beads were removed immediately prior to DMED addition to eliminate them from consideration during the screening.

Screening for dione association was performed after the DMED probe was added to the CPLB bead suspension to a final concentration of 1 μM, and the mixture was allowed to equilibrate for two hours. These conditions seemed reasonable based on preliminary studies with model compounds.

Pilot studies performed on the CPLB peptide library suggested that more stringent screening conditions could be used successfully. Screening of 1% of the CPLB peptide library was performed under a much lower concentration of DMED probe, 1 μM instead of 250 μM, and at a milder pH of 7.50. Screening under these conditions after two hours generally revealed 1-2 very bright beads in the pool of about 4,000 beads, with another 5-10 bright beads (Figure 3.12).

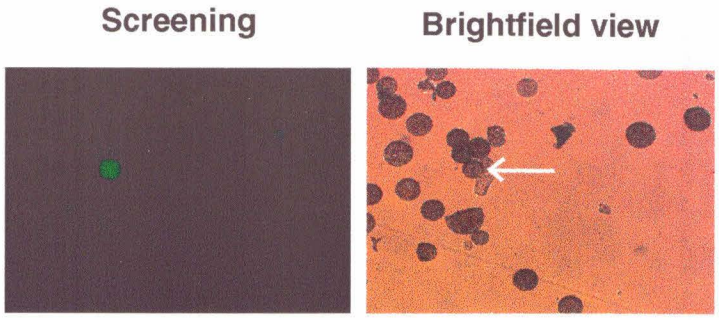


Figure 3.12. Image captured from an initial 1% screen under both screening and brightfield views. pH 7.50 0.1M phosphate, 1.0 μM DMED, 2 hours.

The properties of these beads suggested that the screening results reflected the expected association between peptides and the DMED probe. The magnitude of fluorescent increase for "winning" beads was larger than other beads in an absolute and relative sense. Because beads were screened in a large pool, it was not possible to follow the magnitude of fluorescence change for every individual bead. Thus, selected beads tended to be those that exhibited the largest absolute fluorescence signal on screening. To study the relative change in fluorescence magnitude, it was necessary to rescreen individual "winning" beads and compare these results to those of 48 randomly selected bead (Figure 3.13). From these results it was observed that with DMED, the randomly selected beads typically showed some increase in fluorescence. However, the magnitude of this change tended to be smaller than those observed for the "winning" beads. Thus, it appeared that beads selected as the absolute "winners" of the fluorescence screen reflected beads that also exhibited the greatest change in fluorescence.

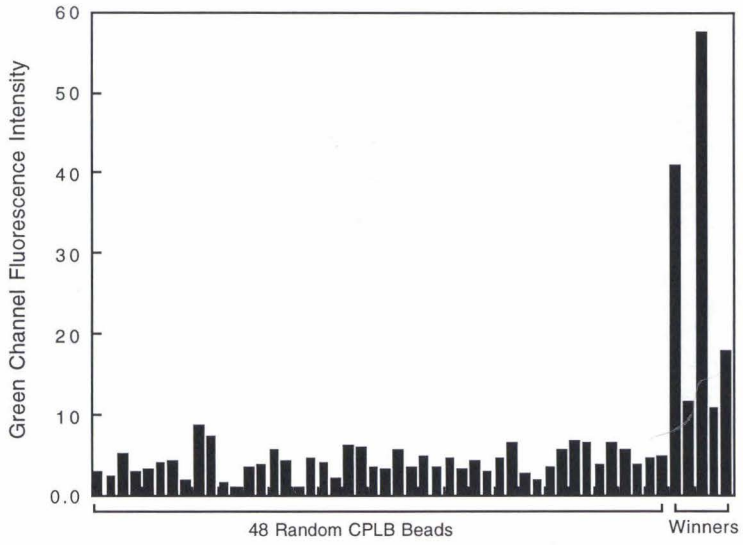


Figure 3.13. Plot of the fluorescence intensity differences between rescreening and rebuffering of peptide beads over the green color channel. The last five data points (on the right side) represent winning beads from initial screening. The remaining data points are from randomly selected CPLB library beads.

The screening results also appeared to reflect association between DMED probe and peptide, as the increase in fluorescence was reversible. Treatment of the "winning" CPLB beads with a 0.1% TFA solution and then "rebuffering" in a pH 7.50 solution resulted in significant release of the DMED probe (Figure 3.14). Under these conditions, loss of fluorescence was rapid (less than 5 minutes). However, this decrease reflected quenching of fluorescence by protonation of the dansyl exocyclic amine and complete release of the DMED probe was slower. Incubation of the CPLB peptide beads with the DMED probe for 18 hours resulted in even larger signals. The relative distribution of intensities did appear unchanged, however, and the brightest beads exhibited less reversible release of the fluorescent probe.

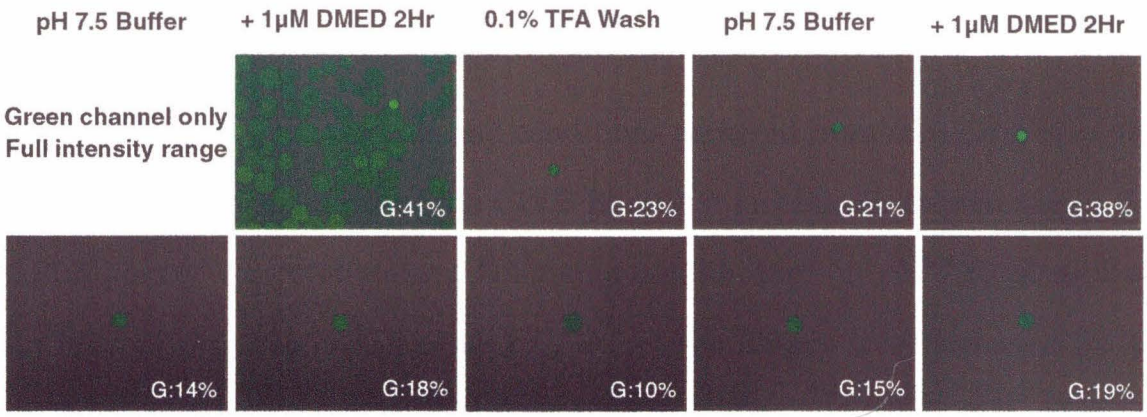


Figure 3.14. Fluorescence intensity of beads from a 1% screens under different solution conditions. Upper row shows a winning bead selected from a field of beads. Lower row shows a control bead (non-winner). Inset – G-%'s indicates relative green channel intensities.

Increased fluorescence in the presence of DMED appeared not only reversible, but also reproducible. Beads that exhibited the greatest increase in fluorescence upon DMED addition exhibited a comparable increase upon reintroduction of the fluorescent probe (Figure 3.14) Similarly, beads that were not selected as "winner" showed reproducible behavior. The 48 randomly selected CPLB library beads showed low levels of fluorescence increase upon screening and rescreening.

However, despite these properties, not all selected beads were utilized. A portion of the "winning" beads could not be sequenced. These beads tended to be smaller and brighter than the other "winners." They demonstrated different properties when subjected to the 0.1% TFA wash, not swelling like the other beads. This inability to swell was found to be an accurate predictor of beads that

could not be sequenced, and thus was used as an additional step to discriminate selected beads.

Combining the insights and conditions gathered from screening of model compounds and fractions of the CPLB library, a large-scale sequencing of the CPLB library was undertaken. Following equilibration in pH 7.50 phosphate buffer and prescreening, batches of 1-2% of the library were screened in the presence of 1 μM DMED probe after two hours. The 5-10 beads that exhibited fluorescence exceeding our screening cutoff were isolated and incubated in 0.1% TFA. Beads that did not swell were noted and later eliminated. After reequilibration in pH 7.50 buffer, the individual beads were rescreened in the presence of 1 μM DMED. From digital images of the individual beads, the change in fluorescence between rescreening and rebuffering was calculated and used to rank beads.

Altogether, one full 4-fold excess copy of the CPLB peptide was screened, containing over 400,000 beads. A total of 280 beads (approximately 0.07%) were selected based on cut-off data. Of these, 128 beads (approximately 0.03%) showed sufficient reswelling to be considered further (Figure 3.15). No single bead or collection of beads demonstrated a change of fluorescence significantly greater than other winners. Twenty beads (the top 15 of the first half of the screening and the top 5 of the second half of the screening) were selected for sequencing. These beads constitute a small fraction of the original CPLB peptide library (approximately 0.005%) and represent peptides that can associate with the DMED dione probe rapidly at a low concentration and mild pH.

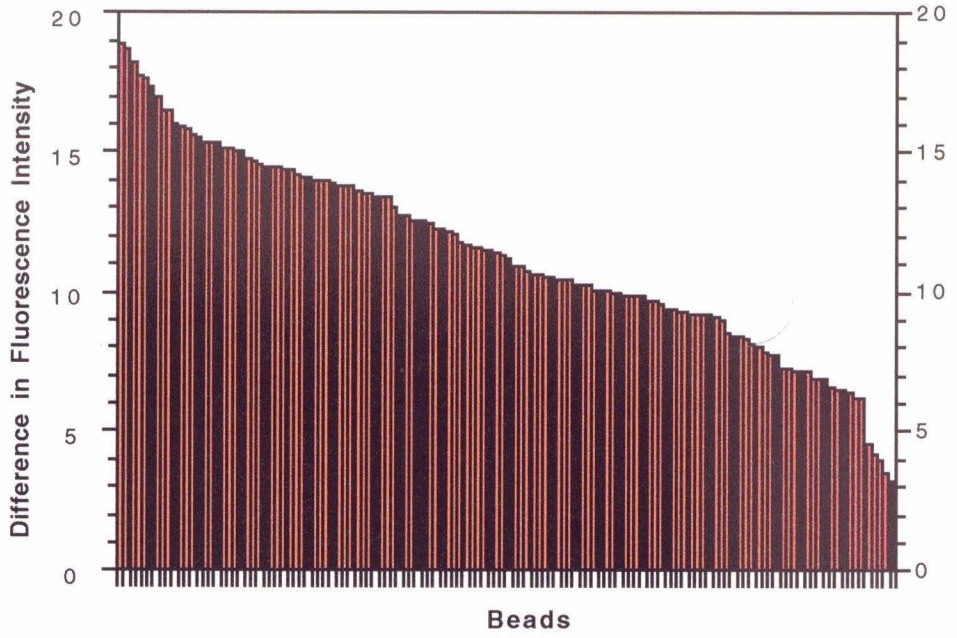


Figure 3.15. Ranked green channel fluorescence intensity for all beads selected from the CPLB peptide library.

3.4. Sequencing and resynthesis of the association screen "winners."

Sequence determination of the association screen winners was performed using automated peptide sequencing at an external facility.

The results are listed below (Table 3.2). Of the 20 beads submitted, 19 were sequenced and one was lost. The sequence of a winning bead from an earlier pilot screening, CPLB-A2, is also included. For the peptides sequenced, the first residue was always glycine, demonstrating the conversion of the chloroacetyl cap group to glycine. At the sequence positions with constant amino acids (all but positions 1, 3, 8, 14, 17, and 18), expected residues were observed.

Pos	BBA5	A2	C12	C11	C10	C8	C7	C6	C5	C4	C3	CPLB
0	Ac	G	G	G	G	G	G	G	G	G	G	ClAc
1	**Y**	**F**	**W**	**F**	**W**	**W**	**F**	**W**	**W**	**W**	**W**	**Y, L, I, F, A, W, V, H, N**
2	R	R	R	R	R	R	R	R	R	R	R	R
3	**V**	**V**	**T**	**I**	**I**	**I**	**V**	**I**	**I**	**V**	**I**	**V, I, T**
4	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP
5	S	S	S	S	S	S	S	S	S	S	S	S
6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
7	D	D	D	D	D	D	D	D	D	D	D	D
8	**F**	**F**	**F**	**W**	**F**	**F**	**W**	**L**	**W**	**I**	**F**	**Y, L, I, F, A, W, V, H, N, R**
9	S	S	S	S	S	S	S	S	S	S	S	S
10	R	R	R	R	R	R	R	R	R	R	R	R
11	S	S	S	S	S	S	S	S	S	S	S	S
12	D	D	D	D	D	D	D	D	D	D	D	D
13	E	E	E	E	E	E	E	E	E	E	E	E
14	**L**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap, Dab, Orn, K**
15	A	A	A	A	A	A	A	A	A	A	A	A
16	K	K	K	K	K	K	K	K	K	K	K	K
17	**L**	**W**	**W**	**W**	**F**	**L**	**W**	**W**	**W**	**W**	**F**	**Y, L, I, F, A, W, V, H, N, R**
18	**L**	**W**	**W**	**Y**	**W**	**W**	**L**	**L**	**I**	**W**	**F**	**Y, L, I, F, A, W, V, H, N, R**
19	R	R	R	R	R	R	R	R	R	R	R	R

Pos	BBA5	C2	C1	B12	B10	B9	B8	B7	B6	B5	B4	CPLB
0	Ac	G	G	G	G	G	G	G	G	G	G	ClAc
1	**Y**	**W**	**W**	**W**	**W**	**W**	**W**	**W**	**W**	**W**	**F**	**Y, L, I, F, A, W, V, H, N**
2	R	R	R	R	R	R	R	R	R	R	R	R
3	**V**	**V**	**I**	**V**	**V**	**V**	**V**	**V**	**I**	**I**	**V**	**V, I, T**
4	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP	dP
5	S	S	S	S	S	S	S	S	S	S	S	S
6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
7	D	D	D	D	D	D	D	D	D	D	D	D
8	**F**	**L**	**V/A**	**W**	**W**	**Y**	**I**	**F**	**W**	**W**	**W**	**Y, L, I, F, A, W, V, H, N, R**
9	S	S	S	S	S	S	S	S	S	S	S	S
10	R	R	R	R	R	R	R	R	R	R	R	R
11	S	S	S	S	S	S	S	S	S	S	S	S
12	D	D	D	D	D	D	D	D	D	D	D	D
13	E	E	E	E	E	E	E	E	E	E	E	E
14	**L**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap**	**Dap, Dab, Orn, K**
15	A	A	A	A	A	A	A	A	A	A	A	A
16	K	K	K	K	K	K	K	K	K	K	K	K
17	**L**	**W**	**W**	**W**	**F**	**W**	**W**	**F**	**W**	**F**	**F**	**Y, L, I, F, A, W, V, H, N, R**
18	**L**	**W**	**W**	**F**	**L**	**W**	**W**	**W**	**F**	**W**	**W**	**Y, L, I, F, A, W, V, H, N, R**
19	R	R	R	R	R	R	R	R	R	R	R	R

Table 3.2. Sequences of the peptides selected from the CPLB library from residues 1-19, with the variable positions highlighted. Included are the sequence of parent peptide BBA5 and the potential sequences of the CPLB library peptides.

At the positions with sequence variability (amino acids 1, 3, 8, 14, and 18), only variable residues expected at those positions were observed.

Several trends in average sequence identity at the variable positions were observed. At position 14, the center of the active site core, the basic amino acid Dap was found exclusively. The preponderance of this residue at this position did not appear to be due to some bias in the original library itself, as sequencing of random CPLB peptide library beads resulted in other basic amines at this position (Table 3.1). The Dap residue is the basic amine with the shortest side chain and most acidic amine. The extent to which these factors contributed to their exclusive inclusion in the selected peptides is investigated and discussed further in Section 3.7.

For the other positions, while no site contains one amino acid exclusively, hydrophobic residues tended to predominate. At position three, just prior to the type II' turn, isoleucine and valine are found almost exclusively, with threonine occurring only once. At positions 1, 8, 17 and 18, the amino acids that directly surround Dap 14, aromatic hydrophobic residues dominate. Tryptophan is the most common residues found at these positions, with at least one always present and an average of almost 2.5. Phenylalanine is the next most common residue, with an average of almost 1. Interestingly, despite the preponderance of aromatics at these positions, tyrosine is relatively rare, present only twice in a total of 76 potential sites. After the aromatic hydrophobic residues, the aliphatic hydrophobic residues are next most common with leucine appearing 6 times, isoleucine 3 times, and alanine occurring once. The polar and charged residues histidine, asparagine, and arginine do not appear at all. The source of this bias toward hydrophobic residues is not entirely clear. It is possible that tryptophan

may be common because of its ability to participate in fluorescence resonance energy transfer with the dansyl group, despite our efforts to filter out light close to the excitation energy around the absorption wavelength of tryptophan. Functional experiments tend to suggest that the general bias toward hydrophobic residues may be because of a functional role they play.

Not only were average trends in sequences apparent, trends in sequence identity appeared to be common. Peptides CPLB-B8 and CPLB-C4 had identical sequences. This is perhaps less than what might be expected from the statistics of the CPLB library. However, the fact that the absolute differences of fluorescence between the top peptides was relatively small, and the fact that some bead to bead variation was observed even in the model systems, is more than sufficient to account for this. Moreover, closely related peptides were very common in the sequenced peptides. 11 peptides (CPLB-B5, -B6, -B7, -B8, -B9, -B12, -C2, -C3, -C5, -C8, -C10) were only one amino acid different from another sequence. Thus, it appears that the primary screening provided a relatively stringent selection methodology for closely related peptide sequences.

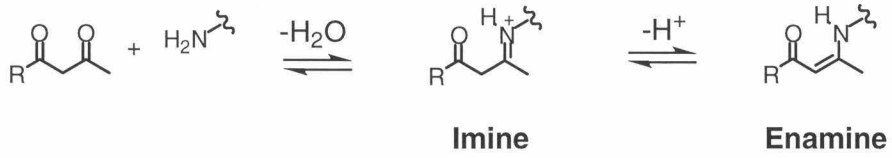
The full set of 19 peptides has been resynthesized in quantity on noncleavable PEGA resin for further solid-phase based studies. Additionally, these peptides have been synthesized on cleavable PAL-PEG-PS support. These peptides have been HPLC purified and are ready for solution phase characterization.

3.5. Secondary screening of CPLB peptides.

At this stage it was necessary to further characterize the selected peptides in order to find those best able to facilitate the aldol condensation process. Two

potential strategies are available - screening directly for enamine formation and screening for aldol condensation product.

Screening with the DMED probe is effective for reducing the size of the initial peptide library by rapidly finding peptides that form a strong association with the dione molecule. This association can reflect the stabilization of a number of species, including species that resemble the imine and enamine intermediates of aldol condensation (Scheme 3.6).



Scheme 3.6. Condensation products of the DMED probe and a peptide amine.

Performing screening for enamine formation was initially pursued because of the attractive features it offered. In their search for aldolase antibodies, Lerner and coworkers showed that antibodies capable of stabilizing enamine formation with the dione, also catalyzed the aldol condensation.[7] Thus, it appeared that screening for formation of this reaction intermediate would also be a good strategy for finding peptides capable of facilitating the aldol condensation. Furthermore, because the enamineone is a covalent intermediate, it was potentially amenable to bead-based screening, allowing for more rapid analysis.

Several different strategies to rapidly screen for enamine formation were investigated. As described in Section 3.3, the enamineone species is

chromophoric, with an absorption maximum at 315 nm. This absorption is easily monitored when this species is in solution. Extending this detection to bead-bound species seemed possible. However, equipment for direct visualization of UV-Vis absorption in a spatially resolved fashion was not available. Furthermore, the absorption wavelength and extinction coefficient for the enamineone group is not sufficient to allow differentiation of the peptide-bead and the dansyl background.

Another strategy for on-bead detection of the enamine was to chemically differentiate between imine and enamine. In such a scheme (Scheme 3.7), beads associated with the dione would be detected with the DMED probe.

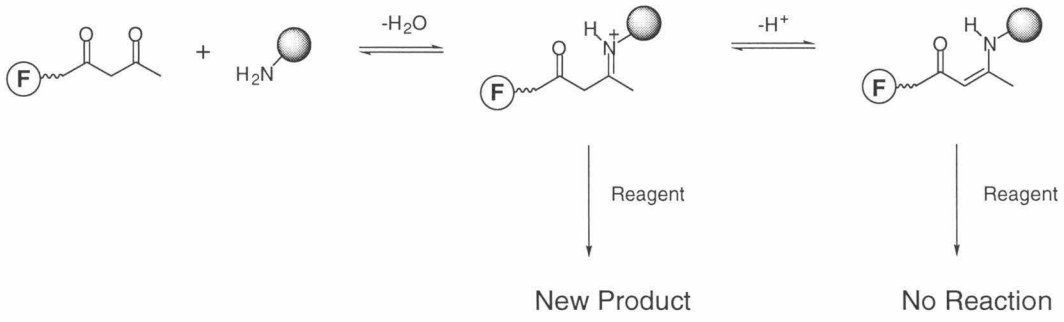


Figure 3.7. Potential scheme for descriminating imine from enamine species.

Differentiating the imine from enamine form would then be performed by subjecting the beads to new chemical conditions, such as a change in pH, where the difference in reactivity or stability of the two forms would result in some detectable change. To investigate the relative stability of the imine and enamine species at basic pH, solution phase and solid phase experiments were performed.

To study the stability of enamine in solution, the aldolase antibody 38C2 was used to generate the enamineone from 2,5-pentanedione at pH 7.50. Formation of this species proceeded rapidly as monitored by UV-Vis absorption spectroscopy. The pH of this solution was then rapidly raised to pH 10.0 or 11.0 by addition of a buffered high pH solution. This pH jump resulted in a shift in the absorption profile. However, the change could not be readily interpreted. Under the assumption that the complicated behavior was due to protein effects, enamine formation and stability was studied with simple tripeptides or pyrrolidine. However, with these amine-containing systems, no spectroscopic signal consistent with enamine absorption was observed over time for a range of different pH values and dione concentrations.

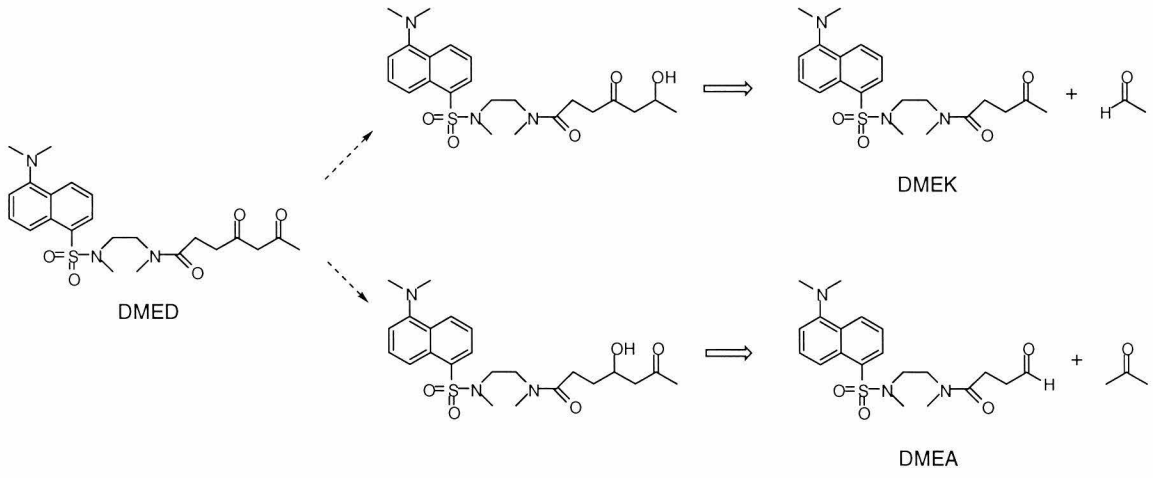
As an alternative to the solution phase experiments, attempts to localize and detect enamine species on a solid support were also performed. To find a system capable of generating localized enamine, attempts were made to derivatize ELISA plates[24] and PEGA beads with the 38C2 aldolase antibody. However, neither approach resulted in derivatization at a level high enough to observe association.

Thus, as of yet, a rapid means of screening for enamine formation has not been found. However, spectroscopic techniques for monitoring solution-phase enamine formation are still practical for medium sized sets of peptides once these peptides have been synthesized and purified. As such, this technique is currently being applied by Kevin McDonnell to screen the resynthesized "winners" of the DMED association screen.

As an alternative and supplement to screening for enamine formation, screening for product formation has been investigated. In general this approach

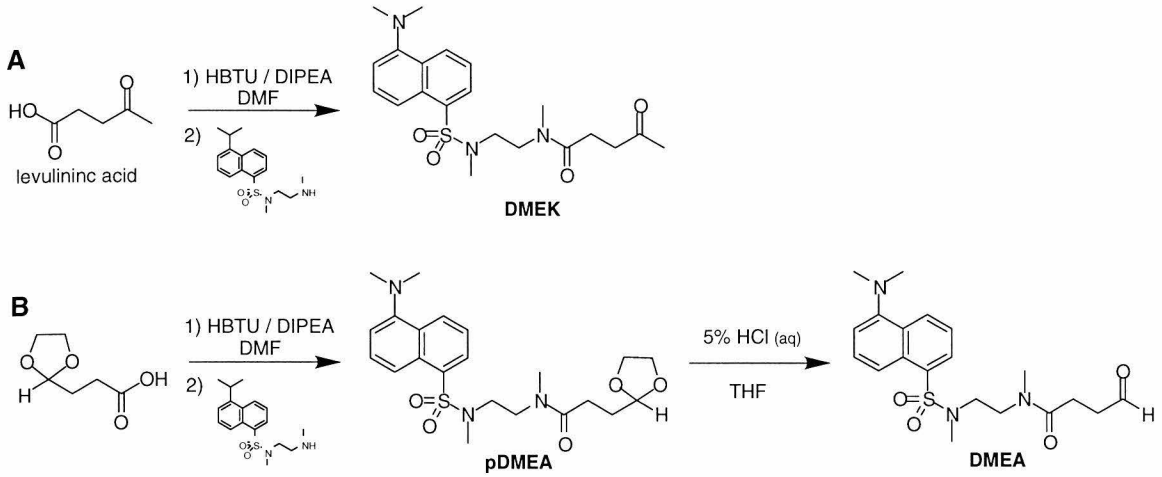
requires the isolation of individual peptides because the products are released, limiting its use in high-throughput screening. However, the advantage of such an approach is that it screens directly for peptides that mediate aldol condensation. An UV-Vis/HPLC based assay has been investigated for a range of acceptor/donor pairs.

One set of donor/acceptor pairs was based on the DMED probe. The DMED probe is designed to resemble reaction intermediates in the condensation process. Peptides that form favorable interactions with this molecule could presumably form similar interactions with molecules similar to DMED. From DMED, two potential donor/acceptor pairs can be generated, where the bulk of the DMED molecule is either attached to a ketone or aldehyde moiety (Scheme 3.8).



Scheme 3.8. Potential aldol condensation product resembling DMED and their synthetic precursors, DMEA and DMEK.

Synthesis of both the aldehyde (DMEA) and ketone (DMEK) forms depicted above have been accomplished by Kevin McDonnell (Scheme 3.9).



Scheme 3.9. Synthesis of A) DMEK and B) DMEK.

Levulinic acid is commercially available and the protected aldehyde can be generated from commercially available 3-bromopropionaldehyde acetal. The condensed dansyl-linker molecule is generated as described in the synthesis of DMED. Coupling of these fragments using standard HBTU activation chemistry affords DMEK and the protected DMEA. Acid deprotection of the acetal group results in the free DMEA molecule.

Characterization of the molecules indicated that they possessed absorption and fluorescence properties similar to the dansyl group. Moreover, both compounds could be detected to a high degree of sensitivity by UV-Vis, allowing as little as 100 pmoles of molecule to be detected reproducibly. This small quantity of compound is comparable to the amount of peptide expected on

a single PEGA bead, making it feasible to perform product assays on single CPLB library beads. However, this possibility was not pursued, as it was unclear to what extent peptide-beads could be recovered for sequencing and further assaying.

DMEK, DMEA, and commercially available aldehydes and ketones were tested to find appropriate substrate/donor pairs for assays. In addition to DMEK and DMEA, these aldehydes and ketones included acetone, cyclopentanone, 2-butanone, acetaldehyde, benzaldehyde, and *p*-nitrobenzaldehyde. The assay conditions utilized by Lerner and coworkers with the aldolase antibodies were 5% (v/v) ketone, 1 mM aldehyde, and 1 uM antibody,[8] where high substrate concentrations were necessary to observe rapid turnover. Some of the aldehydes and ketones of interest displayed limited solubility in aqueous pH 7.50 buffer, either eliminating them from further consideration or requiring their use at reduced concentrations. Analytical HPLC conditions were optimized to maximize the separation of various acceptor /donor pairs. Screening under these conditions has not been performed on all of the "winners" of the DMED-based screen. However, assaying of one of the initial winners, peptide CPLB-A2, and various model compounds has been accomplished and is described further in Section 3.7.

3.6. Structural characterization of the CPLB-A2 peptide.

In order to develop the techniques required to perform further screening of the association-screen "winners" and to understand the properties of these peptides, structural and functional characterization was performed on peptide CPLB-A2, a peptide selected during the early rounds of DMED-based screening.

In the experiments focusing on structural characterization, the goal was to determine to what extent the structure of the BBA scaffold had been maintained.

The primary structure of peptide CPLB-A2 is as follows (Figure 3.16):

Pos	BBA5	A2
0	Ac	G
1	**Y**	**F**
2	R	R
3	**V**	**V**
4	dP	dP
5	S	S
6	Y	Y
7	D	D
8	**F**	**F**
9	S	S
10	R	R
11	S	S
12	D	D
13	E	E
14	**L**	**Dap**
15	A	A
16	K	K
17	**L**	**W**
18	**L**	**W**
19	R	R
20	Q	Q
21	H	H
22	A	A
23	G	G

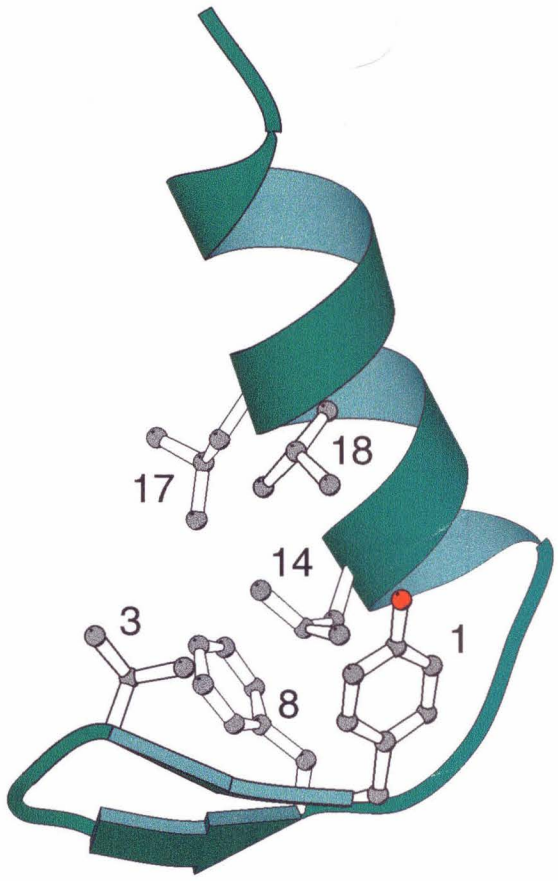


Figure 3.16. Primary sequences of peptides BBA5 and CPLB-A2 with differences highlighted. Solution structure of peptide BBA5.

This sequence is similar to the parent peptide BBA5.[12] In the sheet region at position one the tyrosine residue is replaced by a structurally analogous phenylalanine amino acid, while positions three and eight are conserved as

valine and phenylalanine, respectively. At position 14, the basic amino acid that replaces leucine is Dap, the shortest of the possible bases. In the helix, the leucine residues at position 17 and 18 are both replaced by tryptophan residues. Thus, if the structure of CPLB-A2 is similar to that of parent peptide BBA5, the Dap amine side chain is surrounded by four amino acids with aromatic side chains, with two tryptophans above, and two phenylalanines below.

To study the secondary structure of peptide CPLB-A2, both CD and NMR spectroscopy were employed. CD spectroscopy was performed from 195–400 nm at pH 7.50, the pH of the primary screen, at a concentration of 50 μM peptide (Figure 3.17).

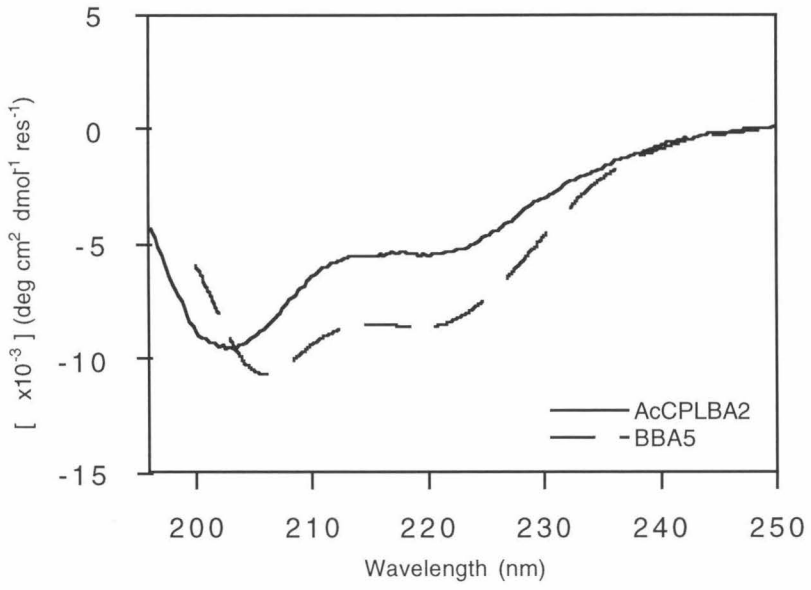


Figure 3.17. Circular diochroism spectra of peptides CPLB-A2 and BBA5.

From this spectrum, it is apparent that peptide CPLB-A2 share characteristics with parent peptide BBA5. Both possess two minima in mean residual ellipticity, with the absolute minimum occurring between 200 and 210 nm and the other minimum occurring around 220 nm. However, the spectra are not identical, as the position and magnitude of these ellipticities differ to some extent. CD spectral information at these wavelengths tends to reflect secondary structure content, but tend to be difficult to interpret with small peptides.[25] However, the similarity between the two spectra suggests that a fair amount of secondary structure was maintained in A2, with more certainty in the conservation of helical character.

To supplement the CD spectral analysis of peptide CPLB-A2, two-dimensional NMR spectroscopy and analysis was performed by Jennifer Ottesen. Secondary structure was probed at pH 4.50. This pH was utilized because amide proton exchange under more basic conditions is sufficiently rapid to prevent detection of many important NOE crosspeaks.[26] A full spin system assignment was performed at this pH from the TOCSY spectrum of a 5 mM solution of the peptide (Table 3.3).

Pos	Residue	HN	Hα	Hβ	H-Others
0	Ac	7.461 or 7.142	-	-	-
1	Phe	8.273	4.567	3.014,2.915	
2	Arg	8.316	4.444	1.710,1.634	γ:1.440, δ:3.066, ζ:7.186
3	Val	8.373	4.298	2.048	γ:0.936
4	dPro	-	4.448	2.218,1.936	γ:3.707
5	Ser	8.175	4.378	3.756	-
6	Tyr	8.293	4.423	2.819	γ:6.943, δ:6.732
7	Asp	8.241	4.607	2.591,2.448	-
8	Phe	8.320	4.533	3.148,2.933	
9	Ser	8.466	4.340	3.899,3.830	-
10	Arg	8.246	4.320	1.882,1.710	γ:1.564, δ:3.090, ζ:7.288
11	Ser	8.301	4.382	3.858	-
12	Asp	8.293	4.595	2.750	-
13	Glu	8.568	3.971	2.011,1.927	γ:2.301,2.139
14	Dap	8.644	4.607	3.431	-
15	Ala	8.100	4.266	1.365	-
16	Lys	8.311	4.002	1.704	γ:1.354,1.247, δ:1.607, ε:2.891, ε2:7.616
17	Trp	7.966	4.455	3.301,3.245	2:7.073, 4:7.322, 5:6.854, 6:7.077, 7:7.299, N:10.157
18	Trp	7.842	4.358	3.164	2:7.172, 4:7.422, 5:7.109, 6:7.203, 7:7.441, N:10.185
19	Arg	7.907	4.029	1.753,1.677	γ:1.451, δ:3.087, ζ:7.229
20	Gln	7.984	4.081	1.871	γ:2.301,2.195, N:7.502,6.887
21	His	8.181	4.493	3.058,2.650	8.245,6.994
22	Ala	8.279	4.225	1.281	-
23	Gly	8.399	3.838	-	-

Table 3.3. Chemical shift assignment of peptide CPLB-A2 at pH 4.50.

Spectral crosspeaks suggestive of sheet and helix formation were observed at pH 4.50 in the CPLB-A2 peptide NOESY spectrum (Figure 3.18).[26]

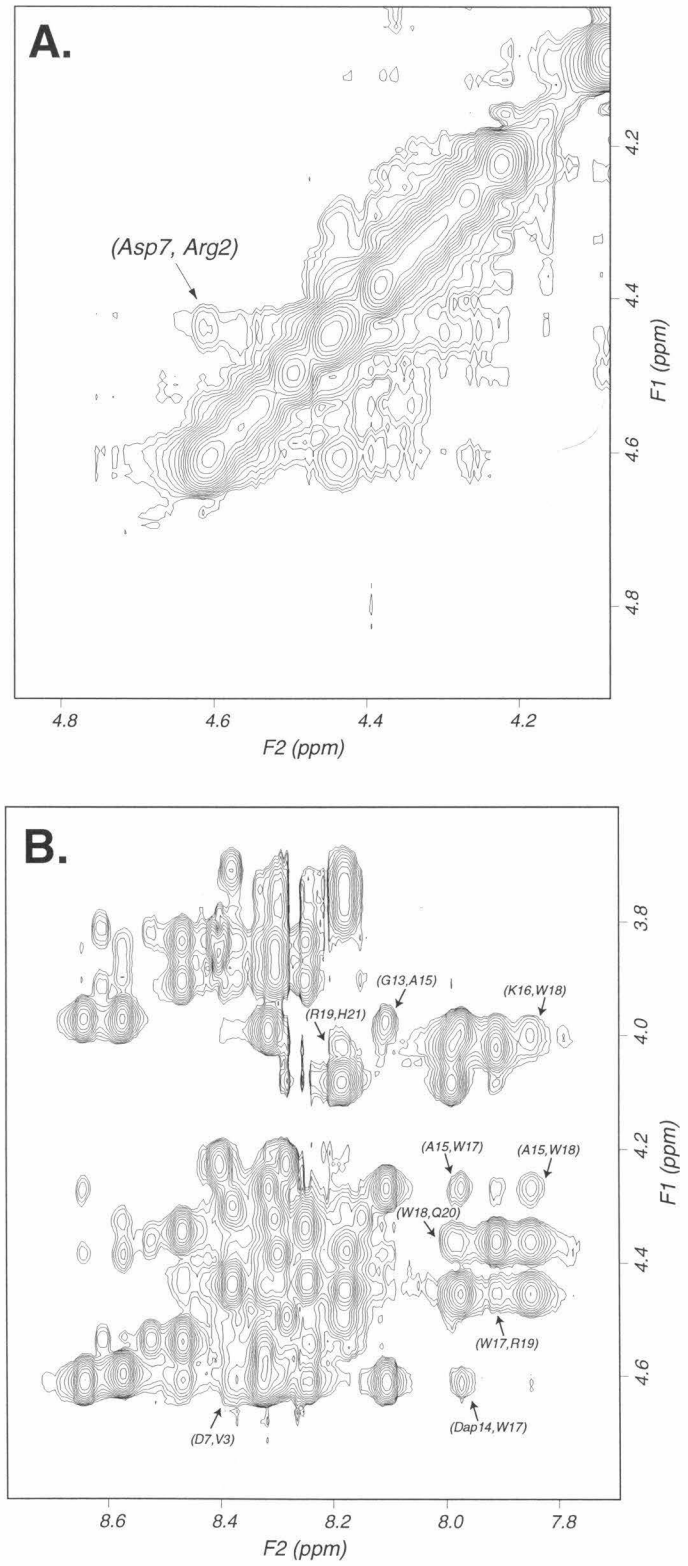


Figure 3.18. NOE crosspeaks of peptide CPLB-A2 at pH 4.50 in the A) α to α and B) α to amide (NH) region of the NOESY spectrum.

In the sheet region, cross-sheet NOE connectivities between the α-protons of Arg2 and Asp7 were observed (Figure 3.18A), as well as between the α-proton of Asp7 and the amide proton of Val3 (Figure 2.18B). Other expected cross-sheet NOE crosspeaks are difficult to discern, as they overlapped significantly. These data strongly suggested that the sheet region of the CPLB-A2 peptide was intact.

Throughout the helical region, residues 12-21, (i, i+2) alpha to amide (Figure 2.18B) and (i, i+2) amide to amide NOE connectivities were observed. Some (i, i+3) alpha to amide NOE connectivities were also observed, although less than might be expected from a well-formed helix. Taken together, the observed NOE crosspeaks suggest helix formation. The NMR results and the CD spectral analysis together therefore suggested the formation of the expected secondary structural element for the BBA peptide, CPLB-A2.

The degree to which peptide CPLB-A2 adopted the super-secondary structure of the BBA scaffold was also investigated by two-dimensional NMR. At pH 4.50 and pH 8.50, the NOESY spectra of peptide CPLB-A2 were well dispersed with tight peak widths. Additionally, few minor NOE spin systems were observed at these pH values (although multiple spin systems were seen at pH 7.50). These data suggested that the peptide exhibited some degree of higher order structure as a single species. At pH 4.50 the degree of super-secondary structure was less clear, as no sheet to helix cross-peaks were easily discernable. However, due to the high redundancy and chemical shift similarity of the aromatic groups involved in these contacts, it was not possible to rule out these interactions. At pH 8.50 sheet to helix contacts were more discernable. While

resolving aromatic to aromatic NOE contacts was still difficult, aromatic to aliphatic NOE contacts from the sheet to helix could be detected (Figure 3.19).

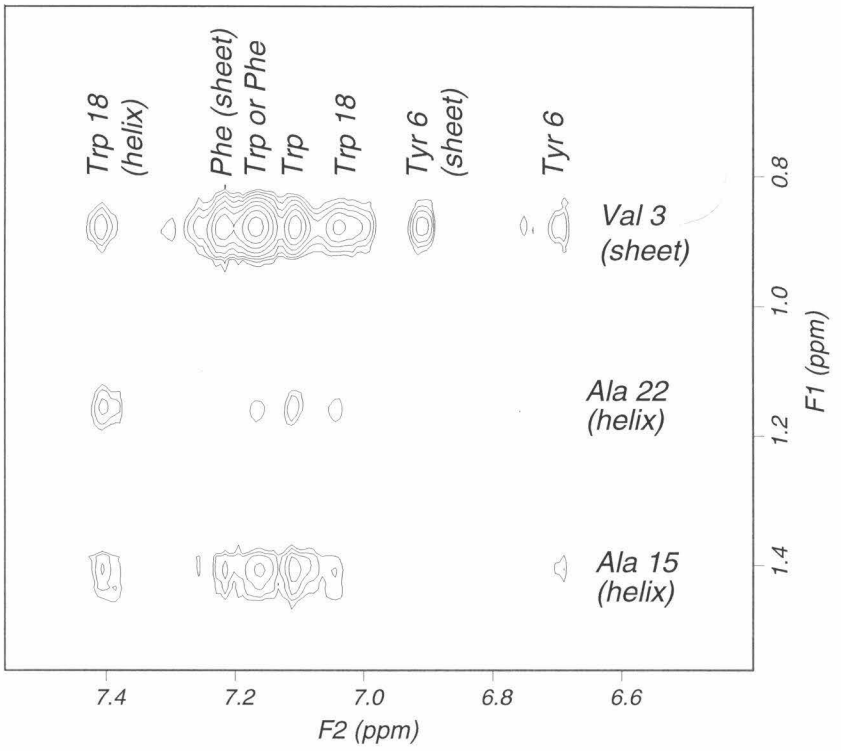


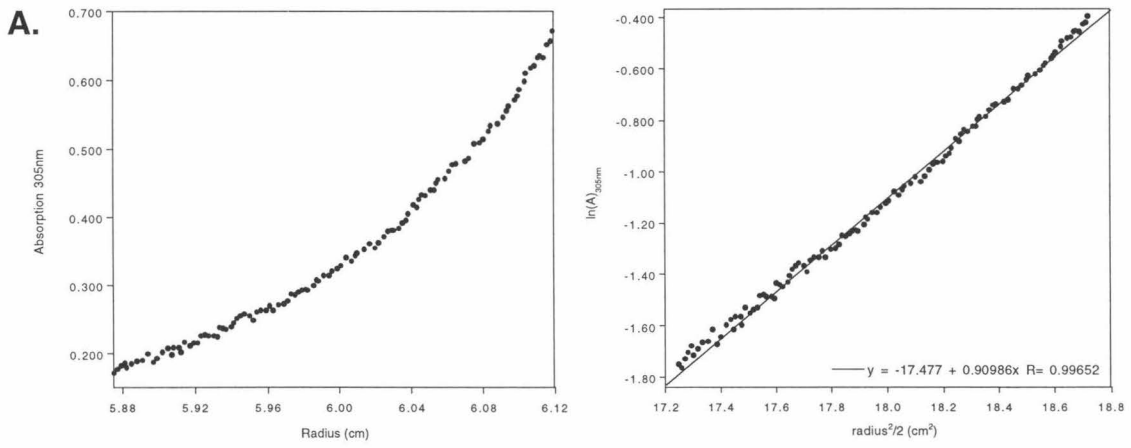
Figure 3.19. NOE crosspeaks in the aliphatic to aromatic region of the NOESY spectrum of peptide CPLB-A2 at pH 8.50.

Specifically, NOE contacts between Trp (helix) and Val (sheet), Phe (sheet) and Ala (helix), and Tyr (sheet) and Ala (helix) were observed (the Tyr-Ala contact is quite weak). Thus, at pH 850, it appears that some degree of super-secondary packing of the sheet against the helix occurs. A full solution structure of the peptide has not yet been attempted.

To probe for the presence of quaternary structure in the CPLB-A2 peptide, a range of complementary techniques were employed. Because the CPLB-A2

peptide possessed a large hydrophobic core, it was possible that specific or nonspecific intermolecular interaction between individual peptides might be occuring - especially if the peptide was not well-folded. Analysis of the two-dimensional NMR spectra of CPLB-A2 peptide at 5 mM suggested that this was not the case. At pH 4.50 and 8.50 the TOCSY and NOESY spectra exhibited well-resolved peak widths and few unexpected or minor spin systems, suggesting limited specific or non-specific intermolecular contacts.

More definitive evidence that peptide CPLB-A2 existed in a monomeric state came from analytical ultracentrifugation analysis. Using equilibrium sedimentation experiments, the peptide was analyzed at pH 4.50 (0.5 mM) and at pH 7.50 (50 μM and 0.5 mM) (Figure 3.20.).



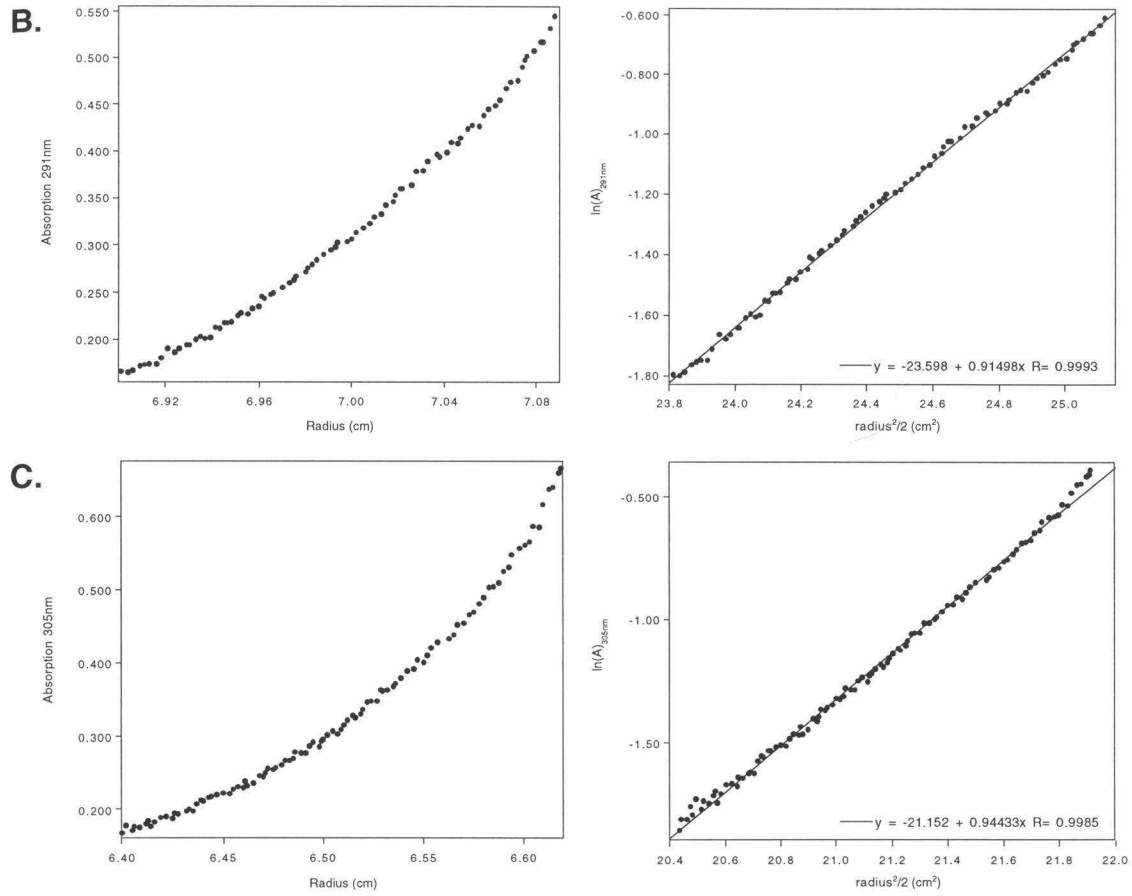


Figure 3.20. Equilibrium sedimentation analysis of peptide CPLB-A2 at A) 0.5 mM, pH 4.50, 305 nm, B) 50 μM, pH 7.50, 291 nm, and C) 0.5 mM, pH 7.50, 305 nm. Raw data is shown to the left and was collected after 42 hours at 48,000 rpm. Linear fit of the data is shown to the right.

At equilibrium, a peptide is distributed along the radius of an analytical cell as a function of a number of factors, including molecular weight, according to the following relationship[27] (Figure 3.21). By monitoring this distribution spectroscopically and then calculating the molecular weight of the analyzed species, the oligomeric state of the peptide can be determined. From these data the molecular weight of peptide CPLB-A2 was calculated (Table 3.4).

$$ln(A) = M\left(\frac{bw^2r^2}{2RT}\right)$$

Figure 3.21. Relationship to calculate molecular weight upon reaching sedimentation equilibrium. A: absorption, M: molecular weight, b: buoyancy factor, ω: angular velocity, r: radius, R: gas constant, and T: temperature.

Conditions	Calculated MW	Experimental MW	Oligomeric State	% of Expected
500 μM peptide, pH 4.50	2854	2713	.95	95
50 μM peptide, pH 7.50	2854	2759	.97	97
500 μM peptide, pH 7.50	2854	2847	1.0	100

Table 3.4. Calculated molecular weight, oligomeric state, and percent expected for peptide CPLB-A2.

Under all three conditions, the calculated molecular weight is in good agreement with the presence of the monomeric species, having 95-100% of the expected molecular weight. The only inconsistency in this result is found in the sample with 0.5mM peptide at pH 7.5. Some deviation from the monomeric curve fit is observed at the radius where the highest concentration of peptide is found. However, this species constitutes a small fraction of the overall population and is only seen at this concentration, a concentration higher than that ultimately used for functional assays.

Complementing the analytical ultracentrifugation results for lower concentrations were the results of CD analysis. CD spectral analysis was

performed on the peptide at pH 7.50 at concentrations from 2 to 50 μM (Figure 3.22).

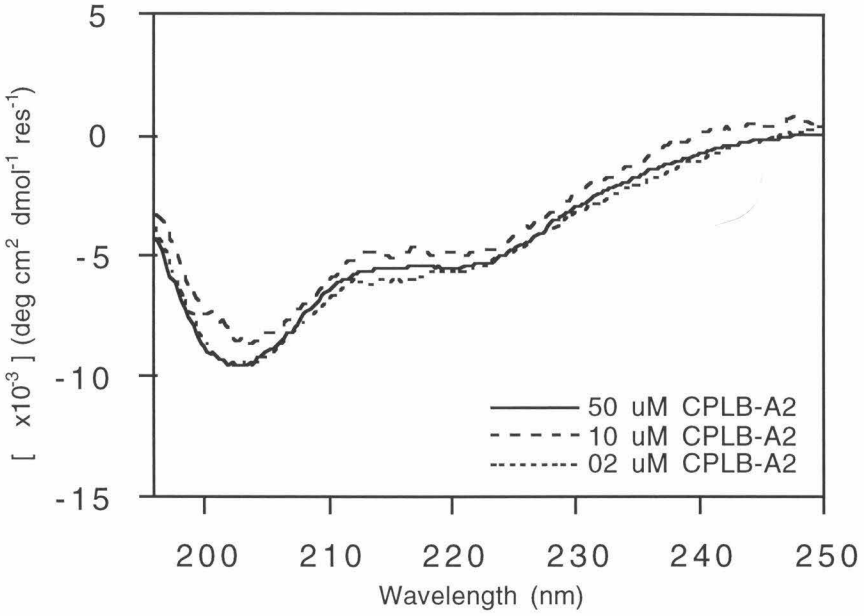


Figure 3.22. Concentration-dependent CD spectra of peptide CPLB-A2 at pH 7.50.

In these spectra, little change in line shape was observed. While this result only suggests that the secondary structure does not change at these concentrations, the ultracentrafugation experiments performed at and above these concentrations confirm that CPLB-A2 exists as a monomer throughout the lower concentrations. Thus, experiments support the existence of peptide CPLB-A2 as a monomeric species in solution over a fairly broad range of concentrations.

Structural analysis of the CPLB-A2 peptide therefore suggests that the peptide exists as a monomer with secondary and super-secondary structural features similar to those seen in the parent peptide BBA5. The fact that the

peptide can adopt such a structure in the presence of an amine group at position 14 was surprising. Functional studies have been performed to determine the protonation state of this amine and to characterize the functional properties of this folded BBA peptide.

3.7. Functional characterization of the CPLB-A2 peptide.

Functional characterization of peptide CPLB-A2 was performed to determine the pK_a of the core Dap amine, the association constant for the DMED molecule, the nature of the dione association, the degree and rapidity of enamine formation, and the ability of the peptide to mediate the aldol condensation process. These studies not only illuminated the functional properties of the peptide, but also provided insights into the reason for the observed screening trends and structural properties.

In amine-mediated aldol condensation, formation of the imine intermediate requires nucleophilic attack by the amine on the carbonyl center.[4] For the amine to function effectively as a nucleophile, it must be deprotonated. In the aldolase enzymes and antibodies, hydrophobic amino acids facilitate deprotonation of the amine.[9] The extent to which the pK_a of the Dap amine of CPLB-A2 had been perturbed and was important for structural and functional purposes and needed to be addressed.

To determine the proton affinity of amine side chain of Dap 14, NMR studies were undertaken in collaboration with Jennifer Ottesen. In these experiments, the two-dimensional TOCSY spectrum of purified CPLB-A2 peptide was collected at 7 °C over a range of pH values, where the pH was set

using either buffering or non-buffering conditions. From the original spin system assignments of the peptide at pH 4.50 it was possible to assign the chemical shift of the alpha and beta proton of the Dap group for the various pH values. Plotting this data against pH gave an experimental titration curve (Figure 3.23).

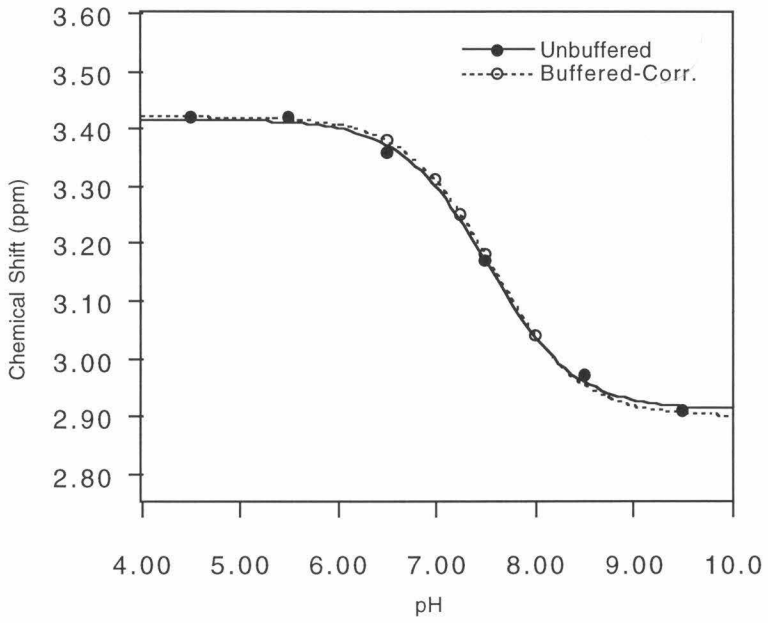


Figure 3.23. Chemical shift as a function of pH for the β-protons of CPLB-A2 peptide. Data fit to the relationship described below.

Fitting this data with the theoretical chemical shifts predicted by a rearrangement of the Henderson-Hasselbach equation[28] (Figure 3.24) provided the pK_a of the β-amine. For experiments performed under buffered conditions, it was necessary to correct for ionic strength effects using the Debye-Hückel relationship.[29]

$$\delta = \delta_{max}\left[\frac{e^{2.3(pKa - pH)}}{1 + e^{2.3(pKa - pH)}}\right] + \delta_{min}$$

Figure 3.24. Relationship between chemical shift and pH.

Experiments performed under buffered and non-buffered conditions were in good agreement, suggesting that the pK_a of the side chain amine of Dap was 7.5. The accuracy of this method was validated by demonstrating agreement between the determined and published pK_a values for the model peptide (Ac-G-Dap-G-Y-NH_2) tyrosine hydroxide.[15] The pK_a for the Dap group in CPLB was low for an amine and suggested a compelling reason for the observed screening and structural results. Screening was performed at pH 7.50. Under these conditions the Dap amine of peptide CPLB-A2 would be approximately 50% deprotonated, a sufficiently large fraction to facilitate imine formation. Similarly, with a significant population of unprotonated and protonated peptides occurring dynamically, the side chain amine might be sufficiently uncharged to allow stable formation of the BBA tertiary structure.

The extent to which the pK_a of the Dap amine was perturbed from the pK_a of a native Dap amine was investigated by determining the pK_a of a Dap-containing model peptide. It was known that the pK_a of the side chain amine of a free Dap amino acid was 9.6.[15] It was expected that in a peptide this pK_a would be reduced. To determine the extent of this difference a model peptide (Ac-G-Dap-G-Y-NH_2) was synthesized and studied in the same fashion as the CPLB-A2 peptide. From these studies, it was seen that the Dap side chain amine in the

peptide was itself quite low, being approximately 7.7. Thus, it appeared that incorporating the Dap amino acid into a peptide was sufficient to reduce its pK_a. This result also potentially explained why Dap was the only amino acid found at position 14 in the "winners" of the screened CPLB library. With such a low pK_a, these peptides had a substantial initial advantage over the other peptides, potentially making them the most reactive. However, it is important to note that incorporation of a Dap amino acid into a peptide alone was not sufficient to favor selection. All of the selected CPLB peptides also incorporated a similar hydrophobic core. Additionally, only the pK_a of the CPLB-A2 peptide has been characterized; the other selected CPLB peptides might exhibit different characteristics.

To quantify the nature of the association between the CPLB-A2 peptide and DMED probe, affinity assays were determined by developing a bead-based assay. In this system, individual CPLB beads were subjected to increasing concentrations of DMED probe, from 0.01 μM to 100 μM. Digital images of the fluorescence intensity of the beads were captured at timed intervals until little change in fluorescence intensity was observed. The fluorescence intensity of each bead for a given concentration of DMED as a function of time was plotted. Fitting each individual curve allowed the calculation of a time infinity equilibrium intensity (Figure 3.25A). These equilibrium intensities were then plotted as a function of concentration (Figure 3.25B). Fitting these data to the theoretical derived relationship between concentration and fraction bound, $I = I_{max}/[1 + (K_D/[DMED])]$,[28] allowed a rough estimate of the dissociation constant for the peptide to be calculated.

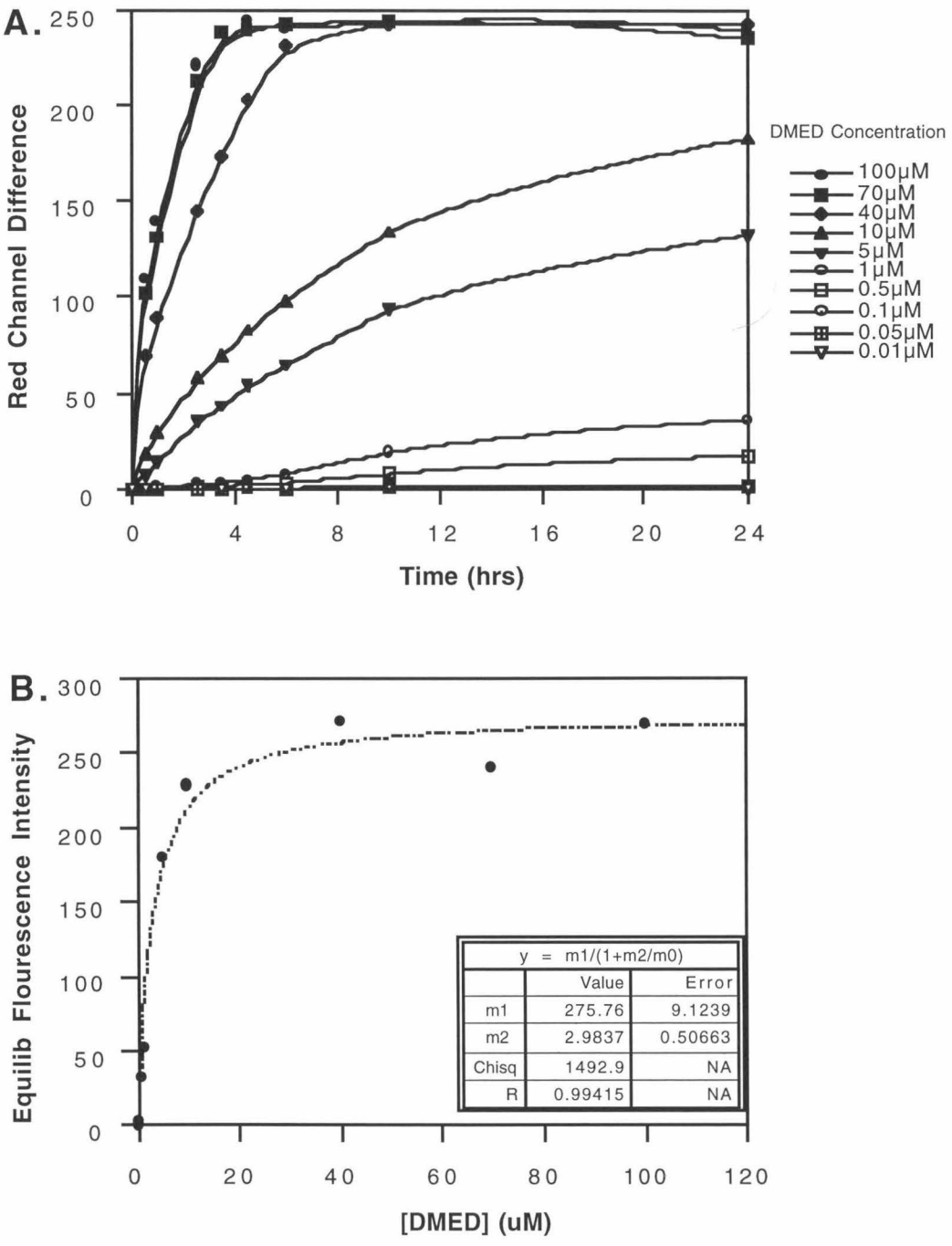


Figure 3.25. A) Average fluorescence intensity of peptide CPLB-A2 as a function of time at different concentrations. B) Equilibrium fluorescence intensity as a function of concentration.

For peptide CPLB-A2, the K_D was determined to be about 3 μM. This value was consistent with screening results, as no peptide appeared to exhibit binding saturation under the 1 μM concentration of DMED used for screening. More importantly, the relative accuracy of the method was demonstrated with known substrates (Adam Mezo and Kevin McDonnell-unpublished results). In these experiments, the dissociation constant for the coiled-coiled helical peptide dimer Fos-Fos was shown to be about 4 μM. This value is fairly close to the known dissociation constant of 6 μM.[30]

The degree to which peptide CPLB-A2 showed improved affinity toward the DMED probe was compared to a Dap-containing model peptide. A PEGA-bound, Dap-containing dipeptide (Dap-Dap-PEGA) was synthesized and subjected to the same experimental conditions applied to the CPLB-A2 peptide. The equilibrium dissociation constant could not be calculated using these techniques because concentrations of DMED in excess of the capabilities of the fluorescent camera were required. However, a lower limit K_D value of 100 μM for the Dap group and DMED was estimated. The source of the increased binding affinity for the CPLB-A2 peptide relative to the Dap-containing model peptide is probably due to the participation of the hydrophobic core residues in sequestering the hydrophobic DMED substrate. However, this hydrophobic interaction alone is not sufficient, as the hydrophobic dansyl group alone does not show significant association with the CPLB-A2 peptide under the screening conditions employed.

To determine more about the nature of the association between the CPLB-A2 peptide and the DMED probe, attempts were made to isolate the reaction

products. Peptide CPLB-A2 and DMED exhibit tight association. However, the exact nature of this association was unclear and could potentially be due to nonspecific binding or unexpected side-reactions. Initial attempts to reductively trap[5] the imine or enamine species were unsuccessful. However, electrospray mass spectral analysis of the crude mixture of 20 μM CPLB-A2 peptide and 100 μM DMED after 48 hours revealed the presence of a new reaction intermediate as well as the presence of CPLB-A2 alone (Figure 3.26).

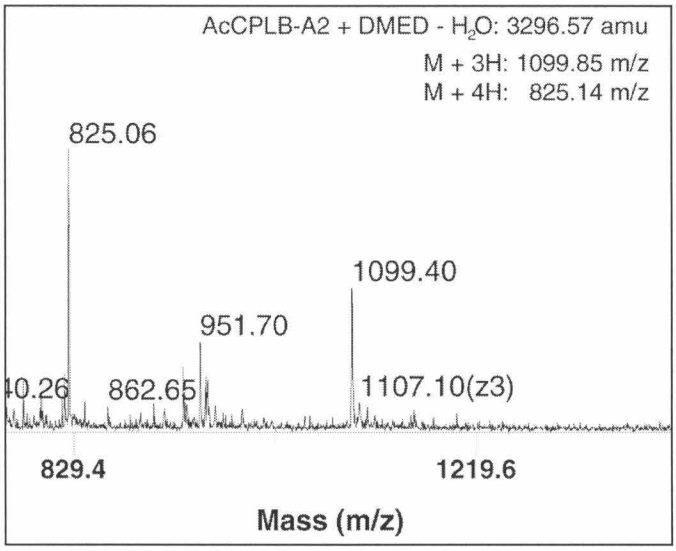


Figure 3.26. Electrospray mass spectrum trace of crude CPLB-A2 peptide and DMED probe at pH 7.5 after 48 hours.

Electrospray analysis conditions, however, are not identical to solution phase reaction conditions. Electrospray mass spectrometry is a gentle ionization technique and tends not to produce new covalently-bound species that are not present in solution phase,[31] suggesting that this intermediate is a significant reaction species. The mass of this species corresponded to the condensation

product of the peptide and the DMED probe with loss of a single water molecule. This intermediate is consistent with either the imine or enamine species, and suggests that the fluorescent signal observed in the initial screens and association studies was due to formation of one of these chemical species.

To probe the extent to which peptide CPLB-A2 and DMED form the enamineone intermediate, UV-Vis absorption studies were performed. In these studies 100 μM dione was added to 20 μM peptide buffered at pH 7.50. Diones investigated included 2,5-pentanedione and DMED. Full spectra from 250 to 400 nm were collected at regular time intervals. Over the course of several days, discernable growth in spectral signal was observed around 315 nm (Figure 3.27).

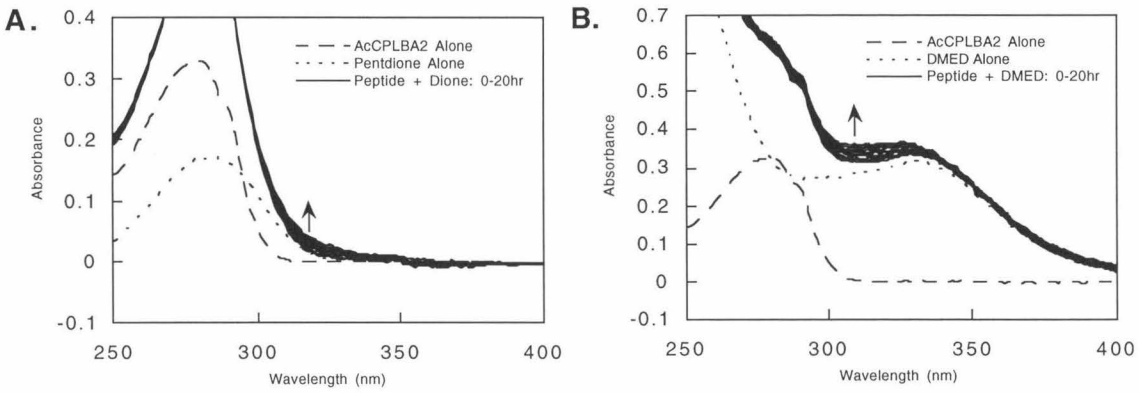
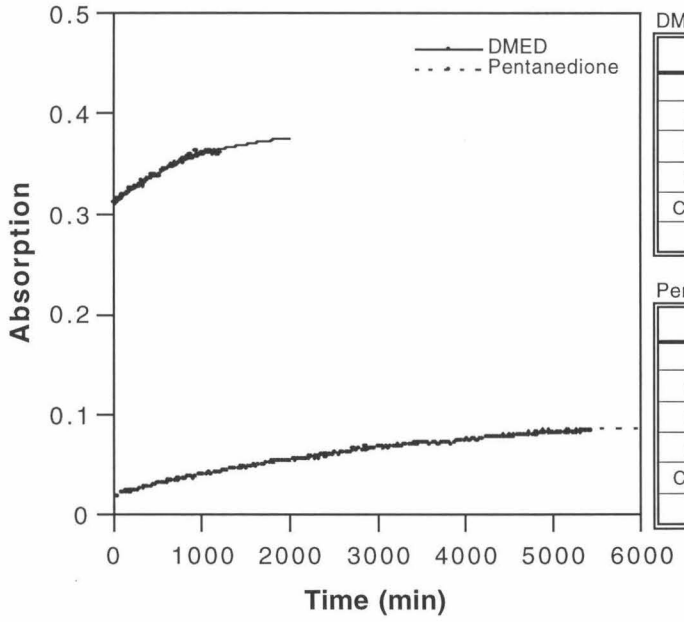


Figure 3.27. UV-Vis absorption spectra of peptide CPLB-A2 with A) 2,5-pentanedione and B) DMED probe at pH 7.5 for the first 20 hours.

Reaction with 2,5-pentanedione was monitored for four days, and little additional change in absorption was observed. Reaction with DMED was monitored for two days or less, as DMED did not remain soluble. From these

data, the change in absorption at 315 nm as a function of times was plotted and fitted to a simple first-order exponential increase curve (Figure 3.28).



DMED

y = m1 + m2*(1-exp(m3*m0))		
	Value	Error
m1	0.30895	0.00072399
m2	0.07462	0.0025623
m3	-0.0011487	8.6098e-05
Chisq	0.00024814	NA
R	0.99379	NA

Pentanedione

y = m1 + m2*(1-exp(m3*m0))		
	Value	Error
m1	0.020656	0.00026158
m2	0.082697	0.00079022
m3	-0.00027188	6.0042e-06
Chisq	0.00016086	NA
R	0.9986	NA

Figure 3.28. Change in 315 nm absorption as a function of time for peptide CPLB-A2 in the presence of either 2,5-pentanedione or DMED at pH 7.50.

These results indicate formation of the enamineone species for both the pentanedione and DMED molecules, further validating the potential of the CPLB-A2 peptide to mediate aldol condensation.

Although more studies need to be performed, these results also suggest several things about the nature of CPLB-A2 mediated enamine formation. Comparison of the binding of DMED and 2,5-pentanedione to the peptide suggest that the degree of enamineone formation is roughly the same for both substrates, but proceeds slightly more rapidly (four-fold) for the DMED substrate. This may potentially be due to the more hydrophobic nature of the

DMED substrate. The time required for peptide CPLB-A2 and DMED to reach equilibrium with respect to enamineone formation appears longer than that seen in the DMED-association screen by two orders of magnitude, suggesting that the association observed in the primary screens is due, to a significant degree, to pre-enamineone intermediates. Finally, in comparing the rate of enamineone formation between peptide CPLB-A2 and the aldolase antibody 38C2 under similar conditions, it is clear that the antibody enhances this formation at least three orders of magnitude more efficiently than the peptide (presumably a comparison of a k_{cat}-like factor). Nonetheless, this result does not preclude peptide CPLB-A2 from acting as an effective catalyst for aldol condensation relative to standard Dap-containing peptides. In particular, a significant portion of the catalytic enhancement that the aldolase antibody provides comes from its ability to sequester low concentrations of reaction substrates (a K_m effect),[8] and peptide CPLB-A2 or some other winner peptide may also perform this task effectively.

To assess the ability of the CPLB-A2 peptide to facilitate full aldol condensation, HPLC-based product assays were performed. In these assays donor and acceptor substrates described in Section 3.5 were utilized. In addition to testing the CPLB-A2 peptide, assays were run in the presence of a Dap-containing model peptide (Ac-Y-G-Dap-G-NH_2) or in the absence of peptide. Reactions were run at pH 7.50 with 1-30 μM peptide, 1 mM aldehyde, and 1 mM-5% (v/v) ketone. The progress of the reaction was monitored over the course of several days by analytical reversed-phase HPLC analysis of the crude reaction mixture using the optimized gradient conditions previously determined (Section

3.5). Despite the use of a number of different donor/acceptor pairs, no discernable increase in aldol condensation rate was observed in the presence of the CPLB-A2 peptide relative to the Dap-containing model peptide or to assays run without any peptide. Current efforts to follow aldolase activity are focused on using a higher throughput and a more sensitive fluorescence assay.

Conclusion

In our efforts to find BBA peptides capable of facilitating the aldol condensation, methods have been developed to generate, assay, and characterize new functional peptides. Applying these methods has allowed us to design and generate a library of over 100,000 different peptides (the CPLB peptides), screen this library for members that facilitate binding of a aldol condensation reaction analog, and further characterize these members. In the screening process a family of BBA peptides has been discovered with high sequence preference for aromatic hydrophobic residues and for the basic Dap amino acid. The best characterized of these peptides, peptide CPLB-A2, has demonstrated interesting structural and functional properties. This peptide appears to be a monomeric species with a high degree of ββα-motif secondary and super-secondary structure. Functionally, it possesses a primary amine base with a relatively low pK_a, shows low micromolar affinity for dione molecules, and can facilitate enamineone formation. Further studies are being performed to better characterize this peptide and the other selected CPLB peptides. Additionally, because many of the techniques and ideas developed in the course of this work appear to be of a

general utility, future efforts may be applied to peptide scaffolds beyond the BBA peptides and to reactions outside of aldol condensation. Continuing efforts in this area should provide new and interesting functional peptides and will further our understanding of designing functional peptides.

Experimental

All syntheses of small molecules were performed by Kevin McDonnell.

Synthesis of mono-Boc *N,N'*-dimethylethylenediamine (1).

15.4 mL of *N-N'*-dimethylethylenediamine (145 mmoles, 12.8g) were dissolved into 25 mL absolute ethanol with stirring and cooled on an ice/water bath. 1.6g di-*tert*-butyldicarbonate (7.34 mmoles) were dissolved into 5 mL absolute ethanol and dripped slowly (over 30 minutes) into the diamine solution. The reaction was stirred for 2 additional hours at room temp. The reaction was then concentrated *in vacuo* to a wet solid. This residue was dissolved in 50 mL of water and extracted with 7 x 30 mL ethyl acetate. The organic layers were combined, washed once with 30 mL brine, dried over Na_2SO_4(s), and filtered. The solvents were removed *in vacuo* to give a clear oil. ^{1}H NMR (CD_3OD): δ 2.08 (s) 9H, 3.00 (s) 3H, 3.30 (t) 2H, 3.49 (s) 3H, 3.97 (t) 2H ppm.

Coupling of dioxoheptanoic acid to Boc-*N,N'*-dimethylethylenediamine (2).

150 mg 4,6-dioxoheptanoic acid (0.95 mmoles) were dissolved in 3 mL DMF with 360.2 mg HBTU (0.95 mmoles). 1 equivalent DIPEA (165.5 µL) was added, and the reaction was stirred for 30 minutes. 221.3 mg mono-Boc N,N'-dimethylethylenediamine (**1**) (0.95 mmoles) were dissolved in 3 mL DMF with 6 equivalents DIPEA. This mixture was added to the heptanoic acid/HBTU mixture with stirring. After 2 hours, the reaction was concentrated *in vacuo* with gentle heating (40-50°C). The residue was dissolved in $CHCl_3$ and purified by silica gel column with 90:10 $CHCl_3$/MeOH. A second column with 96:4 $CHCl_3$/MeOH was required to obtain the desired product with good purity (30% yield). ^1H NMR ($CDCl_3$): δ 1.41 (s) 9H, 2.02 (s) 2H, 2.25 (s) 1H, 2.65 (m) 3H, 2.8 (m) 1H, 2.88 (m) 3H, 2.95 (m) 2H - DMF, 3.05 (2 s) 2H, 3.35 (m) 2H, 3.48 (m) 2H, 3.68 (s) 0.4H, 5.60 (s) 0.4H, 15.30 (broad s) 0.4H ppm. (Clean product appears to have several conformations or states - a result of tautomerization of beta-dione and also cis/trans amide exchange.)

Deprotection of Dione-NMEDA-Boc (3).

93.6 mg of **2** (0.28 mmoles) were dissolved in 10 mL DCM. 3 mL TFA were added with stirring, and the reaction was allowed to proceed for 30 minutes. Solvents were removed *in vacuo*, and the residue was dissolved in CH_3OH/toluene (50:50) and dried down 3 times to remove excess TFA, resulting in 128 mg of a clear oil (131% yield - "wet"). ^1H NMR (CD_3OD): δ 1.96 (s) 2H, 2.15 (s) 1H, 2.65 (2 s) 7H, 3.07 (2 s) 3H, 3.14 (m) 2H, 3.32 (s) 3H, 3.62 (m) 2H, 5.62 (s) 0.3H ppm.

Coupling of Dione-NMEDA to dansyl chloride (DMED) (4).

75 mg dansyl chloride (0.28 mmoles) were dissolved in 25 mL DMF. 49 µL DIPEA were added (0.28 mmoles) followed by 128 mg of **3** (0.28 mmoles) in 4 mL DMF. The pH of the reaction was brought to 8.5 by the addition of 3 more equivalents DIPEA (147 µL), and the reaction was stirred for 2 hours. 0.5 additional equivalents of dansyl chloride were added. DMF was then removed *in vacuo* with gentle heating, and the residue was dissolve in $CHCl_3$ and purified by silica gel column with 97:3 $CHCl_3$/MeOH. The strongest fractions were pooled, and the solvent was evaporated, yielding 126 mg of a light green oil, DMED (96% yield). Extensive drying provided 103.5 mg product (79% yield). ESMS: Expected 461.20. Found M+1=462.1, M+Na=484.2. Elemental Analysis: Expected C 59.85%, H 6.77%, N 9.10%. Found C 59.55%, H 6.22%, N 8.65%. ^{1}H NMR ($CDCl_3$): δ 1.98 (2 s) 2H, 2.20 (s) 1H, 2.40 (m) 2H, 2.60 (m) 2H, 2.80-2.95 (m,s) 12H, 3.32 (m) 2H, 3.51 (m) 2H, 3.63 (s) 0.5H, 5.56 (s) 0.5H, 7.14 (m) 1H, 7.50 (m) 2H, 8.0 (s) 0.5H (DMF?), 8.10 (m) 1H, 8.23 (m) 1H, 8.50 (m) 1H, 15.30 (broad s) 0.5H ppm.

Synthesis of Dansyl-NMEDA-Boc (DMEB) (5).

1.45g of Boc-*N*,*N*'-dimethylethylenediamine (**1**) (7.34 mmoles) were dissolved in 50 mL dry THF. 2.6 mL DIPEA were added with stirring followed by 4 g dansyl chloride (14.8 mmoles). The reaction was stirred for 1 hour. Solvents were removed *in vacuo*, and the residue was dissolved in DCM and extracted with NH_4Cl (aq), then brine. Organics were dried over Na_2SO_4 (s) and filtered The solvents were then removed *in vacuo*. The residue was dissolved in CH_3Cl and purified by silica gel column with 3:1 hexane/ethyl acetate.

Collection of early eluting fluorescent fractions and subsequent removal of solvent yielded 1.75g of a bright green oil (58% yield). ^{1}H NMR ($CDCl_3$): δ 1.42 (s) 9H, 2.81 (broad d) 3H, 2.90 (broad s) 12H, 3.38 (m) 4H, 7.2 (d) 1H, 7.54 (q) 2H, 8.14 (broad t) 1H, 8.35 (broad d) 1H, 8.57 (broad d) 1H ppm.

Deprotection of DMEB (6).

92 mg DMEB (**5**) (0.216 mmoles) were dissolved in 10 mL DCM. 3 mL TFA were added with stirring, and the reaction was allowed to proceed for 30 minutes. Solvents were removed *in vacuo*, and the residue was dissolved in CH_3OH/toluene (50:50) and dried down 3 times to remove excess TFA, resulting in a light green oil. ^{1}H NMR (CD_3OD): δ 2.73 (s) 3H, 2.93 (s) 3H, 2.98 (s) 6H, 3.20/3.40 (m) 4H, 7.45 (d) 1H, 7.65 (q) 2H, 8.12 (d) 1H, 8.55 (t) 2H ppm.

Coupling of levulinic acid to DME (DMEK) (7).

82 mg HBTU (0.216 mmoles), 25 mg levulinic acid (0.216 mmoles), and 38 μL DIPEA (0.216 mmoles) were dissolved in 3 mL DCM and stirred for 20 minutes. All of **6** (assumed quantitative, 0.216 mmoles) was dissolved in 2 mL DCM with 6 equivalents DIPEA (228 μL). This mixture was added to the levulinic/HBTU reaction with stirring and allowed to react for 1 hour. Reaction was then concentrated *in vacuo* and dissolved in $CHCl_3$ and purified by silica gel column with 100% $CHCl_3$, then with 9:1 $CHCl_3$/MeOH. The desired fractions were pooled and the solvents were removed, yielding 91.2 mg product, DMEK (quantitative yield). ESMS: Expected 419.2. Found M+1=420.1, M+Na^+=442.2. ^{1}H NMR ($CDCl_3$): δ 2.18 (s) 3H, 2.38 (t) 2H, 2.70 (t) 2H, 2.78 (s) 3H, 2.86 (broad s) 6H,

2.90 (broad s) 3H, 3.33 (t) 2H, 3.52 (t) 2H, 7.18 (m) 1H, 7.51 (m) 2H, 8.10 (m) 1H, 8.30 (m) 1H, 8.55 (m) 1H ppm.

Synthesis of (2-Dioxolanyl)proprionic acid (8).

2g 3-bromoproprionaldehyde acetal (11.0 mmoles, 1.3 mL) were dissolved in 5 mL dry THF and added dropwise to a suspension of 0.48g Mg turnings in 25 mL dry THF. A short pulse of heat from a heat gun was applied and after 1 minute the reaction was refluxing freely. Bromide was added continuously over 30 minutes. The reaction was allowed to proceed for 60 minutes more then cooled to -40°C. An excess of granular dry ice (CO_2) was added to the reaction with stirring. After 10 minutes the cooling bath was removed, and 10 mL saturated NH_4Cl (aq) were added followed by 10 mL water. The resulting mixture was filtered through glass wool. The mixture was extracted with 20 mL diethyl ether, and then the aqueous layer was acidified to pH 2-3 with 5% HCl (aq). The acidified aqueous layer was extracted 6 times with 20 mL diethyl ether. Ether layers were combined and dried over $MgSO_4$(s). The solvents were removed *in vacuo*, producing 856mg of a clear oil (53% yield). ^1H NMR ($CDCl_3$): δ 2.00 (m) 2H, 2.47 (t) 2H, 3.85 (m) 2H, 3.90 (m) 2H, 4.96 (t) 1H, 10.30 (broad s) 1H ppm.

Coupling of DME to 3-(2-dioxolanyl)proprionic acid (pDMEA) (9).

270 mg 3-(2-dioxolanyl)proprionic acid (1.85 mmoles) and 700 mg HBTU (1.85 mmoles) were dissolved in 30 mL DCM followed by 343 µL DIPEA (1.85 mmoles). This reaction was allowed to proceed for 30 minutes. Deprotected

DMEB (**6**) (815 mg, 1.85 mmoles) was dissolved in 10 mL DCM with 5 equivalents DIPEA (1.715 mL). This solution was added to the acid/HBTU reaction with stirring. Two more equivalents DIPEA were added to bring the final pH to >8. The reaction was allowed to proceed for 2 hours then was concentrated *in vacuo*. The residue was dissolved in $CHCl_3$ and purified by silica gel column with $CHCl_3$ followed by 95:5 $CHCl_3$/MeOH. Higher Rf fractions were collected and pooled, and the solvents were removed, producing 748 mg pDMEA (84% yield). ^1H NMR ($CDCl_3$): 1.96 (m) 2H, 2.25 (t) 2H, 2.80 (s) 3H, 2.87 (broad s) 6H, 2.92 (broad s) 3H, 3.35 (t) 2H, 3.55 (t) 2H, 3.82 (m) 2H, 3.95 (m) 2H, 4.90 (t) 1H, 7.17 (broad d) 1H, 7.53 (m) 2H, 8.11 (m) 1H, 8.30 (broad d) 1H, 8.55 (m) 1H ppm.

Deprotection of pDMEA (DMEA) (10).

89.4 mg pDMEA (**9**) (0.20 mmoles) were dissolved in 1 mL THF. 0.5 mL 5% HCl (aq) was added dropwise, and the reaction was stirred for 24 hours under N_2 atmosphere. The reaction was concentrated *in vacuo*, and the residue was dissolved in toluene and dried 3 times. The final residue was dissolved in 97:3 $CHCl_3$/MeOH and purified by silica gel column with 97:3 $CHCl_3$/MeOH. Product fractions were combined and concentrated *in vacuo* to produce 35.4 mg of DMEA (44%). ^1H NMR ($CDCl_3$): δ 2.46 (t) 2H, 2.75 (t) 2H, 2.90 (broad s) 6H, 2.92 (s) 3H, 2.94 (s) 3H, 3.37 (m) 2H, 3.56 (m) 2H, 7.18 (broad d) 1H, 7.54 (m) 2H, 8.13 (m) 1H, 8.27 (broad d) 1H, 8.55 (broad d) 1H, 9.82 (s) 1H ppm.

CPLB library and bead-based model compound synthesis.

Peptide design and modeling was performed using Insight II software (Biosym Technologies).

The CPLB peptide library was made by split and pool solid phase peptide synthesis. For 98% statistical coverage of the potential CPLB sequences a four-fold excess of beads to sequences was required. The number of PL-PEGA beads (Polymer Labs) for a given weight was estimated as follows: By measurement under a light microscope, the average diameter of the PEGA beads swollen in DMF was determined to be 0.200 mm/bead. Assuming the beads were packed in a simple cubic crystal (which would err favorably on the side of underestimating the number of beads for a given volume), it was calculated that approximately 125,000 beads were contained in 1 ml of beads. One gram of dry PEGA resin was found to swell to 10.8 ml, suggesting that approximately 1,350,000 beads were contained in 1 gram of dry PEGA support. However, because PEGA beads do not perform satisfactorily after drying (PL-PEGA resin is supplied pre-swollen in methanol), the approximate dry weight of the pre-swollen PEGA was estimated by drying down a small sample.

The CPLB peptide library was synthesized with 1.333 g dry equivalents of PL-PEGA resin substituted at 0.40 mmol/g, for a total of 0.5332 mmoles (each bead has been shown to contain from 50-100 pmoles peptide). This was enough resin to provide four copies of the library, where for each there was a four-fold excess of beads to sequences, to give approximately 98% sequence coverage. Amino acid coupling for long common sequences was performed on a Milligen

automated 9050 synthesizer (P.E. Biosystems), while positions that varied were coupled by parallel batch coupling.

For automated synthesis of common sequences, standard Fmoc-protected amino acids at 5-fold excess (0.256 M) were double coupled using HOBT/HBTU/DIPEA (5eq-0.256M:5eq-0.256M:11eq-0.576M) activation chemistry and two hour coupling times. Capping of uncoupled *N*-terminal amines was not performed following the second double coupling. Fmoc-amines were deprotected using a 20% piperidine/DMF wash.

For batch synthesis of variable positions, the resin was split into portions equal in number to the number of different amino acids present at that position, and each portion was placed in its own column. The Fmoc-amines of the peptides were deprotected for 30 minutes and then the resin was thoroughly washed with DMF. Each Fmoc-amino acid was coupled at a 6-fold excess (0.256 M) in its own individual column using HOBT/HBTU/DIPEA (6eq-0.256M:6eq-0.256M:6eq-0.576M) activation chemistry. Couplings were allowed to proceed for six hours with either nitrogen bubbling or inversion mixing before a Kaiser test was performed on a single column. If coupling was incomplete, a second identical coupling was performed and was reacted until coupling was complete. Following coupling, the peptide resin was thoroughly washed with DMF, pooled, mixed for 15 minutes, and then split if necessary. The final Fmoc-amine of the sequence was not deprotected and the library was stored at -80° C.

One full CPLB library (0.133 mmoles) was capped with the *N*-terminal α-chloroacetyl cap. The Fmoc-amine protecting group was removed with a one hour treatment of 20% PIP/DMF. After a DMF wash, capping was performed in 10 ml of a 9:1 DMF/DCM solution of monochloroacetic anhydride (0.414 g, 0.30

M) and HOBT (0.513 g, 0.27 M) for six hours. A Kaiser test at this point was negative. The resin was then washed with DMF and methanol and then dried on the lyophilizer.

Removal of the side chain protecting group was accomplished with a two hour treatment of the resin with an excess of reagent K (trifluoroacetic acid/ phenol/H_2O/thioanisole/ethanedithiol, 82.5.0:5.0:5.0:5.0:2.5). The resin was then thoroughly washed with TFA and methanol and then stored at -80° C.

Syntheses of peptide model compounds were accomplished using the same method described for the CPLB peptides. Synthesis of the dansyl-ladder beads was accomplished using similar chemistry. However, the dansyl sulfonamide linkage to the PEGA resin was achieved using activated dansyl sulfonyl chloride, and the ladder beads were not treated with an acidic cleavage reagent.

DMED association screening.

Screening for the association of the fluorophoric probe, DMED, with the CPLB peptide-bead library was accomplished using an IX50 Inverted System Microscope (Olympus America Inc.). A 100 W extra-high pressure mercury lamp was used as the excitation source, where the excitation frequency was filtered for transmission between 330 and 385 nm by a U-MWU wide-band excitation cube (Olympus America Inc.). Excitation intensity was reduced to 12% of the source intensity by a 32ND12 neutral density filter (Olympus America Inc.) in order to effectively prevent observed irreversible photobleaching of the DMED probe. Emission was filtered optically to wavelengths above 400 nm by the same wide

band excitation cube. Real-time digital capture of the fluorescence signal was accomplished using a DVC-1300 digital color camera (DVC Company) driven by a PC system (500 MHz Pentium III processor, 128 MB RAM, ATI-RagePro graphics card, Windows 98 operating system) equipped with a PIXCI-D image processing board configured for the camera (EPIX Inc.) and with XCAP-Lite image-capture and processing software (EPIX Inc.)

During library screening, the camera was typically calibrated for white balance and set to 80% of maximal gain for strong fluorescence signal. Real-time fluorescent signal was usually filtered digitally using the XCAP-Lite software to both improve signal relative to background and to make differentiation of the fluorescence intensities more amenable to rapid screening (although data was captured without any digital filtering). The typical procedure for digital filtering was to set the red and blue channels to zero across their full intensity range. The green channel was set to zero for the lower 20% of the intensity range and to one for the upper 40% of the range.

Screening of the CPLB library beads was performed in batches comprising 1-2% of the library, which theoretically represents approximately 4000 beads. 22 mg of one full 4-fold copy of the library were weighed into a 2.0 mL eppendorf tube, suspended in 500 µL pH 7.5 phosphate buffer, and agitated by pipetting the suspension up and down. The suspension was then pipetted into one well of a Falcon 6 well culture plate, and the eppendorf was rinsed with 2 more 500 µL aliquots of buffer to ensure complete transfer of the beads. Bead suspension was agitated side to side and back and forth to spread the beads evenly over the bottom of the well (swirling tended to result in beads aggregating at the center of the well). If possible, floating beads were removed or sunk by pipetting. The

plate was then secured in the XY stage of the fluorescence microscope. With the fluorescent lights in the room turned off (these caused all beads to appear bright) and the microscope UV lamp turned on, the well was prescreened for any beads which were bright prior to addition of DMED. The XY stage was labeled with X and Y increments of 2mm. This divided the screened area into 340 2x2mm squares. The observed field on the computer screen was slightly greater than a 2x3mm rectangle. Therefore there was a slight overlap, which hopefully avoided missing any Section of the well. The grid was labeled A-Q (Y plane) and 1-20 (X plane) and locations of any selected beads were recorded as their XY coordinates. During prescreening, the typical procedure was to scan the entire XY grid and note any visible beads then to return afterwards to remove them with a pipetman set to 5μL. Typically the bead was transferred to a second well containing fresh buffer. Typically 5-15 beads were removed during prescreening plus 10-20 "innocent bystanders."

After prescreening, 1.5 mL 2 μM DMED in assay buffer (with <1% MeOH) were added to the assay well, slowly dripping it down the side so as not to stir up beads to the surface. The plate was agitated and left covered for 2 hours to incubate. After the incubation, the plate was agitated as before to spread the beads evenly and placed on the stage and screened through the XY grid. As before, the coordinates of winning beads and medium intensity beads were recorded, and then selected beads were removed. In this case, the winning beads were first placed into a fresh well with assay buffer, identified, then placed into an individual well of a 96 well plate containing 0.1% TFA in water. Typical time for screening post incubation and transfer of selected beads was 1 hour. Beads remaining in the assay well were transferred to a 10 mL conical tube, and the

supernatant DMED/buffer was removed. The beads were rinsed with 0.1% TFA in water 2 more times then stored suspended in 0.1% TFA in water.

Screening of the complete CPLB library generated a total of 280 beads for further screening. These beads were viewed with the fluorescent microscope, and the bead fluorescence was digitally captured using the XCAP-Lite software to determine the size and baseline fluorescence intensity of the beads in aqueous 0.1% TFA solution. Optical conditions utilized were identical to those described for the initial screening. Each of these beads was transferred to an individual well of a 96 well plate containing 150 μL of pH 7.5 0.1 M phosphate buffer. After equilibration, the fluorescence intensity of the beads was digitally captured to determine the baseline fluorescence of the beads at pH 7.5. Rescreening of the beads with the DMED probe was initiated by the addition of 150 μL of a 2 μM stock of DMEA probe as described above. After 2 hours, the fluorescence intensity of the beads was digitally captured, and the beads were returned to the wells with aqueous 0.1% TFA for storage.

To determine which beads would not undergo successful sequencing, the fluorescence intensity images of the beads in aqueous 0.1% TFA and pH 7.5 buffer were compared. Average fluorescence intensities for the beads were determined using the Image Pro 5.0 software, where the average pixel intensity for a given color channel could be calculated for a selected bead. Beads that could not be successfully sequenced did not show a significant increase in size or decrease in average fluorescence intensity under acidic TFA conditions as compared to pH 7.5 conditions. These beads were set aside and not considered further, leaving 128 beads.

To rank the remaining beads in terms of their ability to bind the DMED probe, average fluorescence intensities for the beads under buffer and DMED conditions were determined. The difference between the average green channel intensity in the presence and absence of DMED provided the relative increase in fluorescence. This value was used to rank the 128 beads. Because of concerns about waning lamp intensity over time, the top 15 from the total CPLB set as well as the best 5 beads from the second half of the screening procedure were selected for further analysis.

CPLB library sequencing.

Sequencing of CPLB beads was performed externally at the Caltech Protein Microanalytical Laboratory or at the University of British Columbia Protein Service Laboratory. Following conversion of the N-terminal chloroacetyl cap to glycine (Section 4.2), individual beads were submitted whole in 1:1 methanol/water, transferred to a filter support, and sequenced on a 476A protein sequenator (Perkin Elmer/Applied Biosystems Inc.) using standard Edman degradation protocols. For the initial CPLB test peptides, the full sequence of the peptides were determined. For the "winners" of the DMED association screen, sequences were only determined up through position 19 (with position 18 being the last amino acid to vary between CPLB peptides). For the unnatural amino acids present at position 14 of the CPLB peptides (Orn, Dab, Dap in addition to Lys), amino acid standards were submitted and analyzed. Under the basic conditions of the glycine conversion chemistry, the CPLB peptides containing Dap at position 14 were known to generate a side product that during sequencing appeared as a truncation peptide starting from position 14. This side

product was not the major species present and did not interfere with the sequence determination.

Circular dichroism studies.

Circular dichroism experiments were performed on an Aviv Model 202 CD Spectrometer. CD samples were prepared by dilution of peptide stocks to concentrations between 2 and 50 μM in degassed 0.01M pH 7.50 phosphate buffer. CD measurements were performed at room temperature in a 0.1 cm or 1.0 path length quartz cell with the optical chamber continually flushed with dry N_2 gas. Scans were collected in the range from 195 nm to 400 nm, with a band width of 1.0 nm, a scan speed of 50 nm/min, a time constant of 0.5 sec, and a scan resolution of 0.25 nm. Three accumulations per sample were performed. CD spectra were analyzed in KaleidaGraph software, version 2.11 and are reported in (deg x cm^2)/(dmol x amide bonds).

Nuclear magnetic resonance studies.

2D NMR experiments were performed on a 600 MHz Brucker Avance spectrometer equipped with a triple resonance probe. NMR samples were prepared in either a 90:10 H_2O/D_2O solution or in a 90:10 solution of 50 mM phosphate buffer/D_2O. In both cases, the pH of the sample was adjusted by the addition of small amounts of a NaOH solution. 0.25 μl dimethyl sulfoxide was added as an internal standard. Peptide concentrations ranged from 1-5 mM. Spectra were acquired by Jennifer Ottesen at 7 °C. Nuclear Overhauser Effects (NOEs) were detected using TOCSY with a 70 ms mixing time or NOESY with a

250 ms mixing time. Water suppression was obtained using the WATERGATE pulse sequence. Spectra were processed and analyzed by Jennifer Ottesen in Felix 95 (Biosym Technologies, San Diego, CA, USA) and NOESY spectra were baseline corrected.

Analytical ultracentrifugation studies.

Equilibrium sedimentation experiments were performed on a Beckman XL-A analytical ultracentrifuge equipped with a AN60TI rotor and an Epon centerpiece with saphire windows. Aqueous samples were prepared with a peptide concentration of 50 or 500 μM and brought to pH 4.50 or 7.50 using a 50 μM solution of acetate or phosphate buffer. These samples were then loaded into the centerpiece. Acquisitions of the absorption spectra were performed in triplicate at °7 C at wavelengths from 291 to 305 nm. A baseline absorption spectrum was collected at 3,000 rpm before samples were spun at 48,000 rpm for 42 h. This duration was sufficient to reach equilibrium, as no significant changes in absorption spectra were observed between 39 and 48 h. Averaged absorption spectra were analyzed in KaleidaGraph software, version 2.11. The plot of radius^2 versus the natural log of absorption was fit to a linear curve to determine the aggregate molecular weight of the peptide species, where the partial specific volume was calculated and the solution density was based on the buffer concentration.[27]

Dap amino acid pK_a determination.

Two-dimensional NMR TOCSY experiments were performed from pH 4.50 to 10.50 using buffered and unbuffered Dap-containing peptides according to the conditions described above. The chemical shifts of the α- and β-protons of the Dap amino acid were determined by first performing a full sequence assignment of the TOCSY spectrum at pH 4.50. The changes in chemical shift were then tracked over increasing pH. Chemical shift data was analyzed in KaleidaGraph software, version 2.11. Data collected under buffering conditions was corrected to account for ion strength effects according to the Debye-Hückel equation.[29] From the fit of the plot of proton chemical shift versus pH, a pKa value for the Dap amino acid was determined.[28]

DMED dissociation constant determination.

Rough determination of the dissociation constant between DMED and bead-associated peptides was performed using an IX50 Inverted System Microscope (Olympus America Inc.) equipped with a digital color camera. Bead-associated peptides of interest were added to 150 μL of pH 7.50 0.1 M phosphate buffer in individual wells of a 96 well plate. DMED stocks dissolved in the same buffer were added to reach a final volume of 300 μL and a final concentration of between 0.01 μM and 100 μM DMED. At regular intervals over several days the fluorescence intensity of the beads was digitally captured using conditions described in the DMED association screen. Captures were continued until the system appeared to reach an intensity equilibrium. For beads in high concentration of DMED, the dynamic detection range was increased by reducing

the camera gain by a consistent increment. Intensity varied linearly as a function of gain, allowing intensities to be normalized to a common gain value. The average fluorescence intensity for each bead was calculated as described above. Fluorescence intensity data was analyzed using KaleidaGraph software, version 2.11. Plotting the average fluorescence intensity of a peptide-bead for a given concentration as a function of time and fitting this plot to a first-order exponential increase allowed the equilibrium fluorescence intensity to be calculated. Plotting the calculated equilibrium fluorescence intensity values as a function of concentration and fitting this plot to a hyperbolic curve provided an estimate of the dissociation constant.

DMED reaction intermediate detection.

Covalent association between DMED dione and CPLB-A2 peptide was detected using a PE Biosystems Mariner TOF electrospray mass spectrometer. Covalent intermediates were generated at 25 °C at pH 7.50 in 0.10 M phosphate buffer with 20 μM CPLB-A2 peptide and 100 μM DMED. After 48 h the crude reaction mixture was analyzed by the mass spectrometer. Except for reducing the nozzle potential of the electrospray source, standard positive ion conditions for peptides were utilized for analysis. Low molecular weight background signal was significant, but did not prevent the detection of the expected dehydration intermediate.

Enamineone formation assay.

Enamineone formation was monitored using a DU 7500 Spectrophotometer equipped with a Peltier temperature regulator. Standard

assays were run at 25 °C at pH 7.50 in 0.10 M phosphate buffer with 20 μM peptide/protein and 100 μM dione compound (either DMED or 2,5-pentanedione). After blanking the instrument against buffer alone, 300 μL of the reaction solution was added to a 1.0 cm quartz cuvette and the UV-Vis absorption spectrum from 250-450 nm was monitored as a function of time. Absorption data was analyzed using KaleidaGraph software, version 2.11. For samples that did not exhibit significant spectral drift above 400 nm (drift was observed for samples where solubility was an issue), the increase in 315 nm absorption was plotted as a function of time. Fitting this increase to a pseudo-first order curve provided equilibrium intensity values and pseudo-first order rate constants.

Aldol condensation product assay

Sensitive monitoring of aldol condensation product generation is possible by an HPLC/UV-Vis assay.

In a typical 100 μl assay, degassed pH 7.50 phosphate buffer (89.6 μl of pH 7.5 0.1M phosphate buffer, μ=0.25 NaCl) were added slowly to a solution of acetone donor (5.0 μl of reagent grade acetone-5% by volume, roughly 300,000,000 pmoles) and p-nitrobenzaldehyde acceptor (2.4 μl of a 0.04M solution of p-nitrobenzaldehyde in DMF, 100,000 pmoles). A solution of aldolase antibody (3.0 μl of a 33.3 μM aqueous stock of 38C2 aldolase antibody) was added and the reaction was capped and stirred in the dark.

Currently the assay has been performed for acceptor/donor pairs as mediated by the aldolase antibody. However, the same procedure was amenable

to either single CPLB library beads or solution phase peptides. For a total volume of 100 µl, a single CPLB library bead (typically containing 50-100 pmoles of peptide) used in place of the aldolase antibody to give a peptide concentration very similar to that of the antibody.

Reactions with aldolase antibody were typically followed over the course of several days. When the reaction was to be analyzed, 100 µl aliquots of the reaction solution were removed, and the constituents were resolved using a System Gold HPLC system (Beckman) equipped with a 5µl x 150 mm C_{18} reverse phase analytical column (Beckman). Water/0.1% TFA and acetonitrile/0.1% TFA were employed as co-solvents. Solvent gradient conditions were optimized for each donor/acceptor pair tested. UV-Vis absorption signals for the starting materials and product were monitored at 228 nm using a 167 scanning detector module (Beckman). A qualitative sense of reaction yield was determined by integration of the absorption signal and comparison to starting materials.

References

(1) Larock, R. C. *Comprehansive Organic Transfornations: A Guide to Functional Group Preparations*; Wiley-VCH: New York, 1999; pp .

(2) Machajewski, T. D.; Wong, C. H. "The catalytic asymmetric aldol reaction," *Angew. Chem. Int. Edit.*, 2000, *39*, 1353-1374.

(3) Stryer, L. *Biochemistry*; W. H. Freeman and Company: New York, 1995; pp .

(4) Humphrey, A. J.; Turner, N. J. Carbon-carbon bond-forming enzymes. In *Enzyme Chemistry: Impact & Application*; Sucking, C. J.; Gibson, C. L. Pitt, A. R.; Blackie Academic & Professional, London, 1998; pp 105-136.

(5) Lawrence, M. C.; Barbosa, J. A. R. G.; Smith, B. J.; Hall, N. E.; Pilling, P. A.; Ooi, H. C.; Marcuccio, S. M. "Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase," *J. Mol. Biol.*, 1997, *266*, 381-399.

(6) Reymond, J. L.; Chen, Y. "Catalytic, enantioselective aldol reaction with an artificial aldolase assembled from a primary amine and an antibody," *J. Org. Chem.*, 1995, *60*, 6970-6979.

(7) Wagner, J.; Lerner, R. A.; Barbas, C. F. "Efficient aldolase catalytic antibodies that use the enamine mechanism of natural enzymes," *Science*, 1995, *270*, 1797-1800.

(8) Hoffmann, T.; Zhong, G.; List, B.; Shabat, D.; Anderson, J.; Gramatikova, S.; Lerner, R. A.; Barbas, C. F. "Aldolase antibodies of remarkable scope," *J. Am. Chem. Soc.*, 1998, *120*, 2768-2779.

(9) Barbas, C. F.; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Bjornestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A.

"Immune versus natural selection: Antibody aldolases with enzymic rates but broader scope," *Science*, 1997, *278*, 2085-2092.

(10) Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Design of a monomeric 23-residue polypeptide with defined tertiary structure," *Science*, 1996, *271*, 342-345.

(11) Struthers, M. D.; Cheng, R. P.; Imperiali, B. "Economy in protein design: Evolution of a metal-independent ββα-motif based on the zinc finger domain," *J. Am. Chem. Soc.*, 1996, *118*, 3073-3081.

(12) Struthers, M.; Ottesen, J. J.; Imperiali, B. "Design and NMR analysis of compact, independently folded ββα-Motifs," *Fold. Des.*, 1998, *3*, 95-103.

(13) Mesecar, A. D.; Stoddard, B. L.; Koshland, D. E. "Orbital steering in the catalytic power of enzymes: Small structural changes with large catalytic consequences," *Science*, 1997, *277*, 202-206.

(14) Michael, S. F.; Kilfoil, V. J.; Schmidt, M. H.; Amann, B. T.; Berg, J. M. "Metal-binding and folding properties of a minimalist Cys2His2 zinc finger peptide," *Proc. Natl. Acad. Sci. USA*, 1992, *89*, 4796-4800.

(15) Jencks, W. P.; Regenstein, J. Ionization constants of acids and bases. In *Handbook of Biochemistry and Molecular Biology*; Fasman, G. D.; CRC Press, Cleveland, 1977; pp 305-350.

(16) Lam, K. S.; Lebl, M.; Krchnak, V. "The "one-bead-one-compound" combinatorial library method," *Chem. Rev.*, 1997, *97*, 411-448.

(17) Renil, M.; Meldal, M. "Synthesis and application of a PEGA polymeric support for high-capacity continuous-flow solid-phase peptide-synthesis," *Tetrahedron Lett.*, 1995, *36*, 4647-4650.

(18) Meldal, M.; Svendsen, I. "Direct visualization of enzyme-inhibitors using a portion mixing inhibitor library containing a quenched fluorogenic peptide substrate.1.Inhibitors for Subtilisin Carlsberg," *J. Chem. Soc. Perk. T 1,* 1995, *12,* 1591-1596.

(19) Boutin, J. A.; Fauchere, A. L. "Combinatorial peptide synthesis: Statistical evaluation of peptide distribution," *Trends Pharmacol. Sci.,* 1996, *17,* 8-12.

(20) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. "Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides," *Anal. Biochem.,* 1970, *34,* 595-598.

(21) Tagaki, W.; Guthrie, J. P.; Westheimer , F. H. "Acetoacetate decarboxylase. Reaction with acetopyruvate," *Biochemistry,* 1968, *7,* 905-913.

(22) Haugland, R. P. *Handbook of Fluorescent Probes*; Spence, M. T. Z.; Molecular Probes, Inc: Eugene, 1996; pp 680.

(23) Ito, A. S.; Turchiello, R. D. F.; Hirata, I. Y.; Cezari, M. H. S.; Meldal, M.; Juliano, L. "Fluorescent properties of amino acids labeled with ortho-aminobenzoic acid," *Biospectroscopy,* 1998, *4,* 395-402.

(24) Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl, K. *Short Protocols in Molecular Biology*; Greene Publishing Associates: Brooklyn, 1992.

(25) Fasman, G. D. *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press: New York, 1996; pp 738.

(26) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986; pp 292.

(27) Harding, S. E.; Rowe, A. J.; Horton, J. C. *Analytical Untracentrifugation in Biochemistry and Polymer Science;* Royal Society of Chemistry: Oxford, 1992.

(28) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman and Company: New York, 1985; pp 475.

(29) Beynon, R. J.; Easterby, J. S. *Buffer Solutions*; Oxford University Press: New York, 1996; pp 88.

(30) Oshea, E. K.; Rutkowski, R.; Stafford, W. F.; Kim, P. S. "Preferential heterodimer formation by leucine zippers from Fos and Jun," *Science*, 1989, *245*, 646-648.

(31) Vis, H.; Dobson, C. M.; Robinson, C. V. "Selective association of protein molecules followed by mass spectrometry," *Protein Sci.*, 1999, *8*, 1368-1370.

Chapter 4: *N*-Terminal Peptide Capping Strategies to Facilitate Peptide Purification and Identification

Introduction

In the course of our efforts to incorporate chemical function into small folded peptides, two new chemical methods were developed. In these methods, amino-terminal capping reagents were employed to expedite the synthesis and characterization of peptides without adversely perturbing their structural and functional properties. Specifically, for the pyridoxamine-containing peptides, a reversible affinity tag capping group strategy was developed to facilitate the purification of *N*-terminal glycolyl-capped peptides.[1] For the aldolase peptides, an α-chloroacetyl capping strategy compatible with Edman sequencing was generated.[2]

4.1. Development of a reversible affinity tag strategy for the purification of glycolyl-capped peptides.

Investigation of a reversible affinity tag capping strategy was motivated by a desire to rapidly purify peptides in parallel, while preserving the amino terminal cap. When studying Pam-containing peptides for transamination, screening was performed serially on individual peptides, necessitating that each individual peptide be isolated and purified (Section 2.3). Standard HPLC purification was suitable for small numbers of peptides but was too time-consuming to use with larger families of peptides. Additionally, Pam-containing peptides often exhibited a significant amount of truncation side products that were difficult to resolve from the desired pure peptide. In general the major reaction impurities are those resulting from incomplete chemical couplings.

A technique such as affinity purification seemed well suited for expediting the purification of large numbers of full-length peptides from truncated reaction side products. In affinity purification techniques, it is possible to selectively cap full-length peptides with the affinity group and thereby specifically remove the full-length peptides from truncation peptides. Additionally, because affinity purification can be performed in parallel, the amount of time required to purify a large number of peptides is significantly reduced. While this technique seemed ideal, the available capping agents did not seem suitable for our purposes.

One common strategy is to use *N*-terminal capping agents where an affinity group, such as biotin, is irreversibly attached to the peptide amino terminus.[3, 4] While this strategy offers the benefit of simplicity, it can be inadequate for accurate characterization of the desired peptide. For example, with the BBA family of peptides we were especially concerned about the potential structural perturbation such a large group would introduce. A more appropriate strategy seemed to be one in which this affinity group could be readily removed from the peptide prior to characterization.

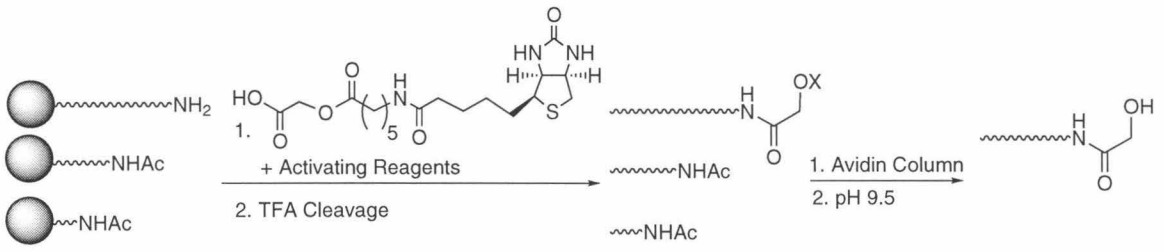
Such reversible *N*-terminal capping agents have also been previously described, where following affinity purification the capping group is removed.[5-15] However, in these methods removal of the capping group results in the exposure of a free amino terminus. For many peptide systems, this seemed problematic. Often it is desirable to have a short peptide approximate an internal stretch of a larger protein, where the sequence is not flanked by charged amino-

and carboxy-termini. Traditionally these shorter peptides are often capped at their amino terminus by a small group, such as an acetyl moiety.

For the BBA family of peptides, the *N*-terminus of the peptide has always been capped with an acetyl group, and it was unclear to what extent a charged amino terminus would perturb the structure and function of Pam-containing BBA peptides. Toward this end, we decided to develop a reversible affinity capping agent that, when removed, would leave the peptide capped by a small, neutral chemical group.

Results and Discussion

To facilitate the rapid purification of peptides modified at the *N*-terminus by a small, neutral capping group, the following reversible affinity purification strategy was developed (Scheme 4.1).[1]



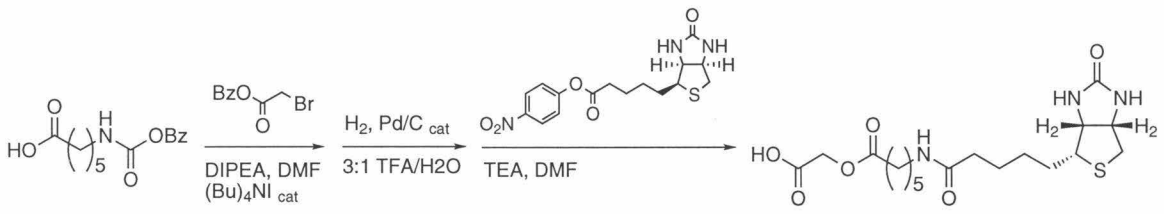
Scheme 4.1. Use of a reversible biotin affinity-capping reagent to purify *N*-terminal glycolyl-capped peptides.

In this scheme, the reversible affinity-tagging agent is made up of a biotin group for tight binding to avidin,[16] a 6-amino-caproic acid linker to maximize biotin-avidin interactions,[17] and a glycolate group, the eventual *N*-terminal capping group. This affinity agent was selectively coupled only to full-length peptides

generated by Fmoc-based solid phase peptide synthesis, since truncation peptides had been acetylated after incomplete coupling cycles. Use of an avidin protein affinity column bound *N*-terminally capped full-length peptides but not the non-tagged truncation products. Mild basic treatment of the bound peptides hydrolyzed the glycolate ester, liberating the pure full-length peptides in their *N*-terminal glycolyl-capped form.

For this strategy to be useful, it was necessary to demonstrate that the affinity-capping group could be readily synthesized and coupled to peptides. Additionally, this group had to show good stability and allow for isolation and recovery of the peptide in high yield and purity.

The biotin affinity-capping reagent was synthesized stepwise from commercially available starting materials (Scheme 4.2).



Scheme 4.2. Synthetic scheme for the reversible biotin affinity tag.

In this scheme, alkylation of the Cbz-*N*-protected 6-amino hexanoic acid linker by the benzyl–protected 2-bromo acetate condensed the amino-hexanoic acid and glycolate fragment of the capping agent. Reductive deprotection of this molecule liberated an amine, which was then linked to biotin, activated as the *p*-nitrophenyl ester, under standard amide coupling conditions in DMF. The crude reaction products were used directly in the subsequent coupling to the peptide.

Delaying attachment of biotin until the final step avoided many of the synthetic difficulties associated with this molecule. Moreover, this approach provided the opportunity to explore other affinity groups in the place of biotin. The reasonable synthetic yields and limited number of chemical steps required to generate the capping group made the utilization of the reagent practical, particularly for use with multiple peptides in parallel.

The biotin affinity tag was effectively coupled to peptides and showed good stability. The biotin-capping group was coupled to the free amino termini of peptides using standard HATU/HOAT coupling chemistry in DMF. Monitoring of the reaction by the well-known Kaiser test showed that the coupling was relatively rapid and quantitative. HPLC and ESI-MS analysis of model peptides BTP1 and BTP2 (sequences FSRS and FSRSDELAKLLRLHAG, respectively) derivatized with the biotin capping group further confirmed that coupling proceeded cleanly to give the desired product. This analysis also demonstrated the stability of the capping group, as it was not perturbed by the peptide cleavage conditions (trifluoroacetic acid/phenol/H_20/thioanisole/ ethanedithiol, 82.5:5.0:5.0:5.0:2.5) or by the aqueous acidic conditions used in HPLC analysis (water/acetonitrile, 0.1% trifluoroacetic acid).

The tagged peptides were isolated using avidin-based affinity column chromatography. Biotin-capped peptides bound to the avidin support within three hours of equilibration in pH 7.0 phosphate buffer. This association appeared to be quite tight, as no model peptide was observed in the eluent under conditions in which avidin-binding sites were in excess of biotin-tagged peptides. The quantity of peptides purified using the column appeared to be

limited only by the amount of avidin support used, although these supports tend to be relatively costly considering the modest number of binding sites they contain.

Release of the peptide was achieved by ester hydrolysis of the affinity tag under basic conditions after the column had been washed to remove any uncapped peptides. Experiments over a range of pH were performed to determine the mildest basic conditions necessary to effect complete ester hydrolysis. At pH 9.5 the biotin-capped model peptides were completely converted to the glycolyl-capped form in 24 hours. These conditions were utilized in subsequent experiments. Under these conditions, peptides were liberated from the avidin support in good yield, with a recovery of 76% relative to the theoretical number of avidin binding sites.

Together, this binding and release resulted in a high degree of purification of peptides. When a crude preparation of capped model peptide BTP2 (Figure 4.1A) was subjected to affinity column purification and release, only the desired glycolyl-capped peptide was observed as determined by HPLC and ESI-MS (Figure 4.1B). After purification peptide solutions were either used directly or stored at –80 °C, as desalting of the peptide solution was neither practical nor necessary.

This technique of affinity purification of peptides by a reversible affinity-capping reagent has now been employed to expedite the purification of two families of Pam-containing peptides each containing more than ten members. Overall, this technique offers a rapid and effective method for the purification of full-length peptides.

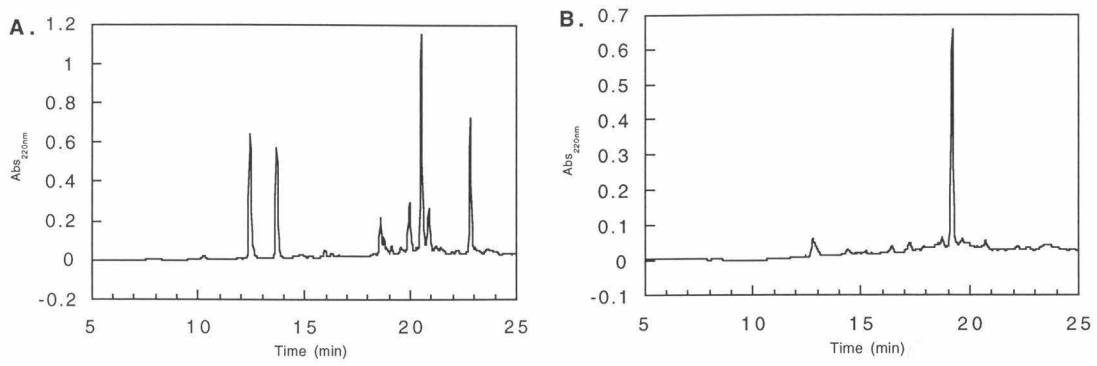


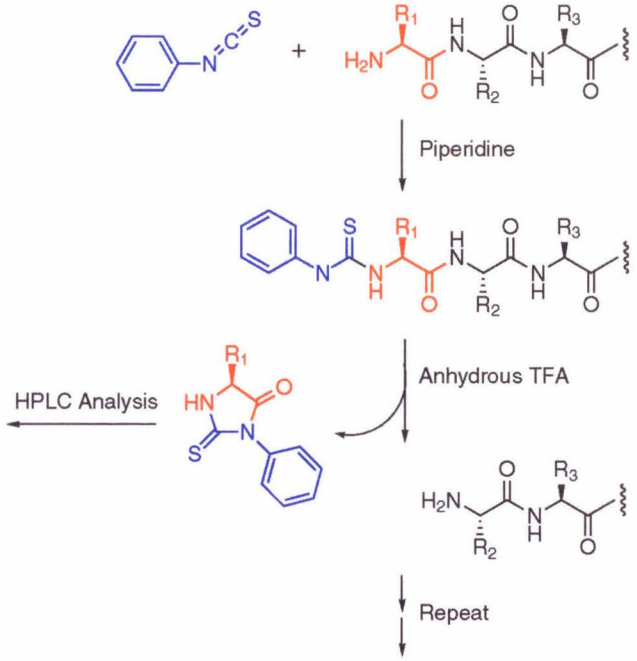
Figure 4.1. HPLC trace of peptide BTP2 A) before and B) after avidin purification. 0-70% gradient of acetonitrile/water/0.1% trifluoroacetic acid.

4.2.Development of an α-chloroacetyl capping strategy compatible with Edman peptide sequencing.

For the generation of peptides capable of mediating aldol condensation chemistry, a need to identify *N*-terminally capped peptides motivated the development of an α-chloroacetyl capping strategy compatible with Edman sequencing. Unlike the largely serial methods employed to study BBA peptides capable of mediating transamination, efforts to generate peptides capable of mediating the aldol condensation focused on a broader parallel approach. In this strategy peptides were generated by split and pool synthesis, allowing over 100,000 different BBA peptide sequences to be synthesized rapidly (Section 3.3). By keeping the peptides covalently attached to their support and performing screening assays directly on the bead, it became possible to screen large fractions of the bead library in parallel (Section 3.2). The split and pool synthetic approach has the benefit of allowing for the rapid generation and screening of a library of

peptides. The drawback, however, is that the ability to track the identity of each individual peptide is lost in the pooling step.

The identity of unknown peptides attached to beads can be determined, and a number of encoding and readout schemes have been devised to facilitate the identification of compounds generated by split and pool chemical synthesis.[18] One of the standard approaches for peptides attached to a bead is the chemical method of Edman degradation sequencing[19] (Scheme 4.3).



Scheme 4.3. Chemical steps involved in peptide sequencing by Edman degradation chemistry.

In this approach, the free amino terminus of the peptide is derivatized by a phenyl isothiocyanate group. Induced rearrangement of the peptide backbone results in the release of the terminal amino acid, which can then be identified, typically by HPLC. The release process also regenerates a free amino terminus, which can undergo further sequencing cycles.

Edman degradation sequencing offers a number of advantages for identifying bead-bound peptides. The technique is straightforward, directly identifying the amino acid residues rather than relying on additional groups meant to encode their identity. The method is also very sensitive; high femtomole amounts of peptides and proteins are routinely analyzed.[19] This sensitivity is more than sufficient for the 10-100 picomole amounts of peptide found on individual PEGA beads used for the BBA library, by this technique. Finally, Edman sequencing is attractive because it is compatible with sequencing directly off beads.[20]

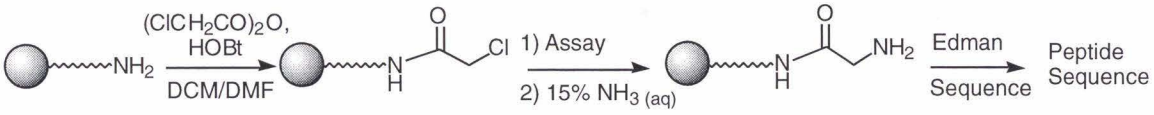
While Edman sequencing offers a number of practical advantages for peptide sequencing, it does suffer from one distinct drawback in this setting. The first step of the analysis process involves the chemical modification of the peptide amino terminus (Scheme 4.1). Thus, peptides compatible with Edman sequencing require a free *N*-terminus. For the BBA peptides, as previously discussed, a free amino terminus is potentially problematic since a charged *N*-terminus could potentially perturb the structure of a BBA peptide. Of even greater concern is the fact that a free *N*-terminus might alter the function of the peptide. In our design of BBA peptides capable of mediating aldol condensation chemistry, a key component of our strategy was to explore the ability of the BBA hydrophobic core to perturb the pK_a of a primary amine downward, thereby making it a more competent nucleophile. Free *N*-terminal amines of peptides are already fairly acidic, with pK_a values of about 8,[21] thus these functional groups could compete with the core amine to participate in chemical reactions. Screening might thereby produce false positive results. Thus, for both structural

and functional reasons, it would be advantageous to cap the amino terminus during screening.

In order to be able to analyze the peptides by Edman sequencing while maintaining the structural and functional integrity during screening, we investigated methods to introduce a small, neutral *N*-terminal capping group which would be present during screening, but which could subsequently be deprotected to generate a free amino terminus.

Results and Discussion

The following strategy was developed to provide a small, neutral *N*-terminal capping group that is stable in common aqueous screening conditions but also compatible with Edman sequencing (Scheme 4.4).[2]



Scheme 4.4. Introduction and use of the α-chloroacetyl capping group.

In this strategy, the free amino termini of full-length peptides were capped with α-chloroacetyl groups. These capped peptides were side chain deprotected, washed, and then screened under desired assay conditions. When sequencing of a particular peptide was desired, the peptide was subjected to aqueous ammonia conditions to convert the terminal capping group into a glycine residue. This

peptide, now with a free amino terminus, could be sequenced under standard Edman conditions.

To demonstrate that this was a useful strategy several requirements had to be met. It had to be shown that the α-chloroacetyl group could be introduced efficiently and that this group was stable to both the acidic conditions of side-chain deprotection and the potentially basic conditions of functional assays. Furthermore, it was necessary to prove that this group could be aminated cleanly and efficiently for subsequent sequencing.

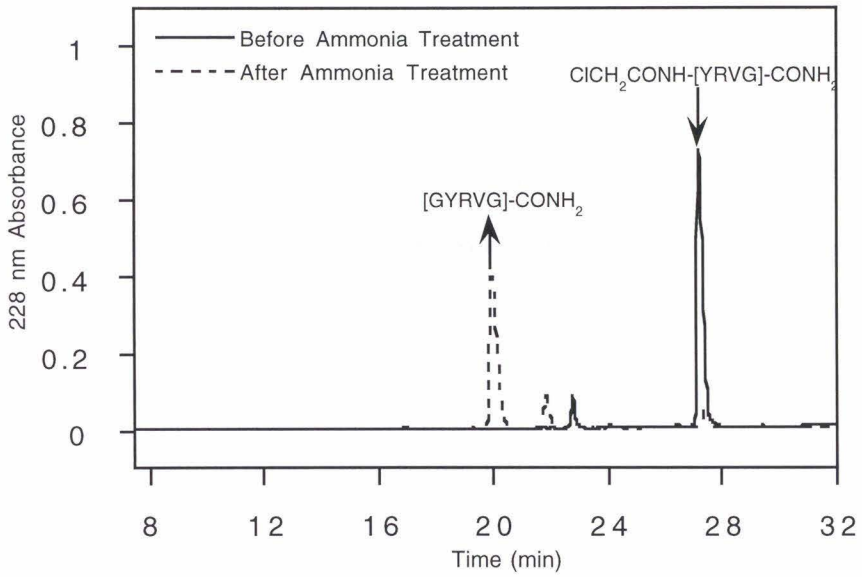


Figure 4.2. Reversed-phase analytical HPLC trace of peptide YRVG A) before and B) after 12 hour treatment with 15% aqueous ammonia.

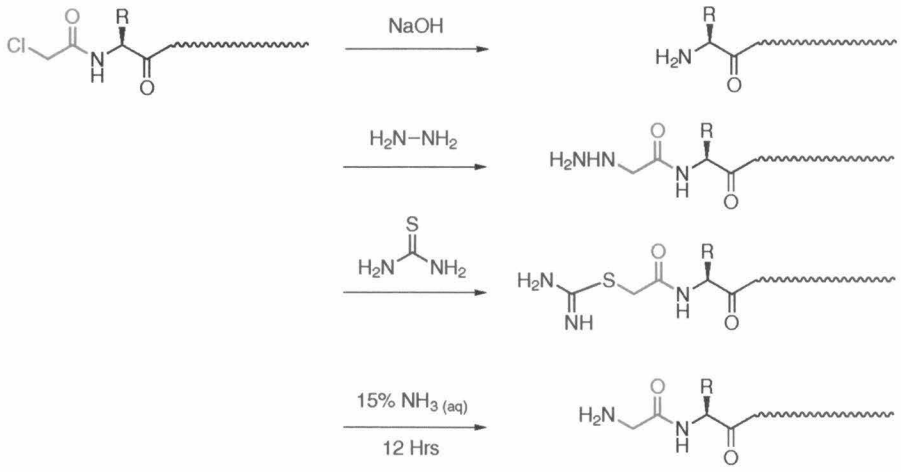
The α-chloroacetyl capping group was coupled to peptides generated by Fmoc-based SPPS efficiently and with high yield and purity. This group was introduced by acetylation of the free amino terminus of peptides with α-

chloroacetic anhydride under standard amide-coupling conditions. Coupling was rapid and efficient, typically going to completion in less than one hour as monitored by the Kaiser test. The yield and purity of this coupling was demonstrated by capping the shorter peptide (YRVG). HPLC and ESI-mass spectral analysis of the crude synthetic products showed that the desired capped peptide was the predominant species (Figure 4.2A) While the di- and tri-chloroacetyl analogs of the α-chloroacetyl capping group were also initially studied, they were not pursued due to the favorable properties of the mono-chloroacetyl group.

α-Chloroacetyl capped peptides also demonstrated the desired stability to both acidic and basic conditions. The acidic stability of α-chloroacetyl capped peptides was apparent from the studies of the capped YRVG peptide. This species remained intact both under the acidic conditions used to side-chain deprotect and cleave the peptide from the resin (trifluoroacetic acid/phenol/ H_20/thioanisole/ethanedithiol, 82.5:5.0:5.0:5.0:2.5) and those used to purify the peptide (acetonitrile/water, 0.1% trifluoroacetic acid). Subjecting the capped YRVG peptide to aqueous basic conditions (pH 10.5, 0.1M CAPSO) for 24 hours resulted in no new products as detectable by HPLC, suggesting good base stability for the capping group. Taken together, these results suggest that the α-chloroacetyl capping group is stable and suitable for assays performed over a wide pH range.

Several strategies were investigated with the capped YRVG peptide in order to generate peptides competent for Edman sequencing (Scheme 4.5). Methods for removing the chloroacetyl-capping group from small organic

molecules were known.[22] However, these techniques were relatively harsh and potentially incompatible with unprotected peptides. It was believed that a simpler and gentler treatment might be possible. Strategies investigated included removal of the capping group by hydroxide treatment and nucleophilic displacement of the halide by hydrazine, thiourea, or ammonia. Reaction with hydroxide at concentrations sufficient to remove the capping group generated a range of side products. After some initial studies sequencing these products, this strategy was not pursued further. Reaction with hydrazine proceeded more cleanly but also generated some side products, so it was not pursued further. Reactions with thiourea and ammonia proceeded cleanly, but only studies with ammonia were carried further.



Scheme 4.5. Reagents tested with the α-chloroacetyl capping group.

Peptides capped with α-chloroacetyl group were effectively converted by ammonia treatment to peptides competent for Edman sequencing. Subjecting peptide YRVG to 15% aqueous ammonia treatment for 12 hours resulted in clean and nearly quantitative conversion of the capping group to glycine as

demonstrated by HPLC and ESI-MS analysis (Figure 4.2). Shorter exposure times and lower ammonia concentrations were insufficient to produce efficient conversion. The efficacy of this approach for sequencing was demonstrated on resin-bound peptides. BBA peptides from the CPLB library of peptides were capped with the α-chloroacetyl group, side chain deprotected, and washed (Section 3.2). These peptides were then subjected to aqueous ammonia treatment and submitted for automated Edman degradation sequencing (Section 3.4). The resulting sequences (Table 4.1) showed that the *N*-terminal capping group had been replaced by glycine (residue 0). At positions with known sequences (2-7, 9-13, 15-16, and 19-23), the expected amino acids were found.

Peptide	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
CPLB-R1	**G**	F	R	T	P	S	Y	D	R	S	R	S	D	E	K	A	K	W	N	R	Q	H	A	G
CPLB-R2	**G**	I	R	T	P	S	Y	D	Y	S	R	S	D	E	Dap[7]	A	K	F	R	R	Q	H	A	G
CPLB-R3	**G**	Y	R	I	P	S	Y	D	R	S	R	S	D	E	K	A	K	W	I	R	Q	H	A	G

Table 4.1. Sequencing results of randomly selected members of the CPLB family of peptides.

At positions where multiple amino acids (1, 8, 14, 17, and 18) are possible, only expected residues were observed. Thus, these results suggest that the chloroacetyl capping group can be introduced and converted cleanly in a manner compatible with most amino acids. As described in the next paragraphs further experiments have clarified some of the issues that can complicate sequence analysis.

Although not problematic for the peptides we were interested in, some analytical ambiguity was found for specific types of amino acid residues. It is

known that aqueous ammonia treatment of amino acids possessing side chain amides (glutamine and asparagine) can result in some degree of conversion of these groups to their corresponding acidic residues (glutamate and aspartate, respectively).[23] This can be problematic when trying to determine the original identity of such residues in an unknown sequence. In our experience these side reactions have been relatively minor, but have been observed. In our library this problem was not relevant, because at positions where the amino acid could be varied the pool of possible amino acids never contained both asparagine and aspartate or glutamine and glutamate. Should this become an issue for future use of this method it should be possible to eliminate this problem. Treatment of peptides with ammonia in anhydrous methanol does not result in hydrolysis of side chain amides[23] but should still be useful for the conversion of the chloroacetyl group to glycine.

An additional complication in the analysis arises from the presence of the Dap amino acid in the peptides. When these peptides were sequenced following ammonia treatment, some amount (usually below 30%) of what appeared to be truncation products were observed just prior to the Dap group. This process appears to depend on the presence of ammonia and Dap. Peptides containing Dap that had not been either terminally capped or treated with ammonia did not exhibit this truncation product upon sequencing. Furthermore, peptides that were not capped but that were treated with ammonia, showed the generation of a side product with a mass identical to that of the original peptide when Dap was present (but not in the presence of Dab, Orn, or Lys). A model consistent with these results is depicted in Figure 4.3.

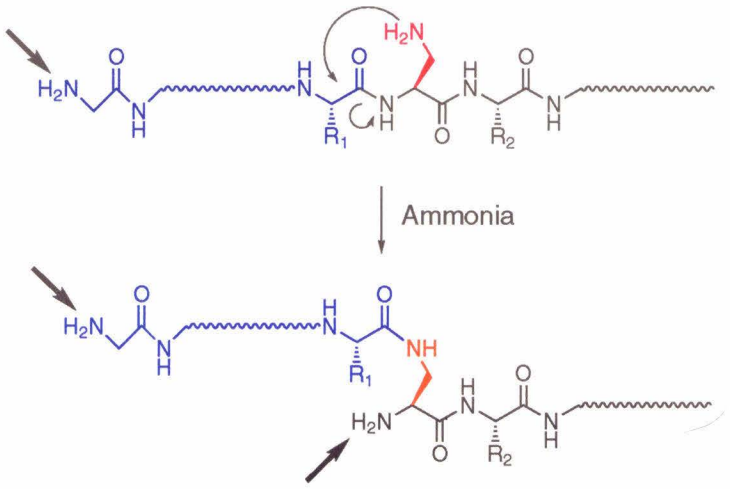


Figure 4.3. Potential mechanism of Dap-dependant sequencing "truncation." Arrows highlight free amino terminal amines.

In this model, the presence of ammonia aids in the attack of the β-amine on the preceding carbonyl center. This migration results in a "truncated" peptide in which the Dap residue has a new side-chain and has become the *N*-terminal residue. Altogether, this rearrangement does not present a problem for sequence analysis. Since the "truncation" product is not the major product, readout of the full sequence is still possible. Additionally, because the Dap residue is not a common amino acid substitution and because the problem is not observed with the more commonly used lysine or ornithine residues, it is not a problem that should be encountered often.

Altogether the α-chloroacetyl capping strategy described offers a useful method for conditionally protecting the amino-terminus of peptides, as it is rapid, effective, and stable. This group can be transformed to a glycine residue rapidly and simply under mild conditions in the presence of a large range of amino acids. Since its initial development, it has been used extensively in the

screening and sequencing of the CPLB peptide library (Chapter 3.4) and has proven itself valuable as a technique for controlling the properties of these peptides.

Conclusion

The reversible biotin-affinity tag and α-chloroacetyl capping reagent provide versatile tools for facilitating the purification and sequencing of *N*-terminally capped peptides, respectively. These groups have been shown to be effective with model peptides and with larger, more complicated peptides, aiding us in our development of the CBP and CPLB family of peptides. Continued use of these reagents offers the potential to expedite future discoveries of functional peptides.

Experimental

Peptide synthesis.

Peptides described above were synthesized by standard solid phase peptide synthesis as described in earlier chapters (Chapters 2 and 3).

Chemical synthesis of *N*-(benzyloxy-carbonyl)6-aminohexanoic acid.

3.0 g (22.80 mmoles) of 6-aminohexanoic acid were dissolved in 50 mL (30.50 mmoles) of a 0.30 M aqueous sodium bicarbonate solution and chilled to 0°C. A clear, colorless solution of 1.9 g (7.62 mmoles) of *N*-benzyloxycarbonyl-

oxy succinimide in 50 mL of dioxane was added dropwise while stirring the clear, colorless aqueous solution to produce a cloudy white mixture, that ultimately became clear and colorless again. The solution was allowed to come back to room temperature and was stirred vigorously for roughly 12 hours. When the reaction was complete the solution was reduced to a minimal volume, diluted with 200 mL water, brought to a pH of about 2, and extracted with 4 x 75 mL of ethyl acetate. The pooled organic layers were dried with sodium sulfate and then removed *in vacuo* leaving a clear colorless syrup. Removal of the remaining solvents by azeotrope resulted in 2.02 g (quantitative yield) of a white solid. R_f=0.33 in 1:1 hexane/ethyl acetate, 1% Acetic Acid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.38-1.70 (m 6H), 2.37 (t 2H), 3.20 (m 2H), 4.81 (broad 1H), 5.13 (s 2H), 7.38 (m 5H) ppm.

Cbz-6-aminohexanoic acid-*O*-glycolate benzylester.

1.5 g (5.65 mmoles) of Cbz-aminohexanoic acid were dissolved in DMF to get a clear, colorless solution. 984 µl (5.65 mmoles) diisopropylethylamine and 880 µl (5.65 mmoles) of benzyl,2-bromoacetate were then added. A spatula tip full of tetrabutylammonium iodide was added to get a faint yellow, clear solution. The solution was brought to 50 °C and stirred approximately 12 h after which time the solution was brought to a minimum volume *in vacuo*. The product was isolated by flash column chromatography with 3:1 hexane/ethyl acetate and rotovapped to a syrup. The product, dried twice by lyophalization from benzene, afforded 1.611 g (69% yield) of a white powder. R_f=0.20 in 3:1 hexane/ethyl acetate, 1% acetic acid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.38-1.76 (m

6H), 2.42 (t 2H), 3.21 (q 2H), 4.65 (s 2H), 4.98 (broad 1H), 5.12 (s 2H), 5.19 (s 2H), 7.38 (m 10H) ppm.

6-Aminohexanoic acid *O*-glycolate.

1.0 g (2.42 mmoles) of the Cbz-6-aminohexanoic acid-*O*-glycolate benzyl ester was dissolved in 3:1 THF/water. 0.129 g (5 mole %) of 10% Pd/C was added. The black mixture was stirred vigorously for approximately 20 h, with a supplemental addition of 0.06 g catalyst after 12 h. The reaction mixture was filtered through celite to get a clear, colorless solution. This solution was then rotovapped to a syrup and lyophalized from benzene to give 0.286 g (62.6%) of a white powder. ^{1}H NMR (300 MHz, D_2O): δ 1.33 (m 2H), 1.58 (m 4H), 2.39 (t 2H), 2.87 (t 2H), 3.34 (s 2H) ppm.

Biotin-6-aminohexanoic acid-*O*-glycolate.

27.2 mg (0.144 mmoles) of 6-aminohexanoic acid-*O*-glycolate were dissolved in 313 µl of water to get a clear, colorless solution. 13.91 µl (0.262 mmoles) of triethylamine were added. 47.9 mg (0.131 mmoles) of nitrophenyl biotin ester were dissolved in 625µl of DMF to give a faint yellow, clear solution. Activated biotin solution was added to the vigorously stirred aqueous solution producing an opaque, yellow mixture, which became a clear, yellow solution over time. The reaction mixture was stirred at room temperature for approximately 12 h. Water was removed *in vacuo*, and the total volume of the crude solution was adjusted to 500 µl with DMF (0.28M theoretical concentration

for desired product). The crude solution was used directly in peptide capping. R_f=0.24 in 9:1:1 chloroform/methanol/acetic acid.

Capping of BTP1 with Biotin-AHA-Glc.

25 mg (0.0044 mmoles, theoretical) of BTP1 resin (Fmoc-F-S-R-S-NH_2) were deprotected by suspending the resin in 500 µl of DCM for 10 min, 500 µl of DMF for 10 min, and finally 500 µl of 20% piperidine in DMF. The resin was washed with 5 x 500 µl of DMF. 33.3 µl (0.0088 mmoles) of 0.28 M Biotin-AHA-Glc crude solution were added followed by 25 µl of a solution of 2.97 mg HATU (0.0132 mmoles) and 1.77 mg HOAT (0.0132 mmoles) dissolved in 25 µl of DMF (sonication required) and a 10 µl solution of 1:6.5 DIEA/DMF (0.0132 mmoles). The reaction mixture was allowed to sit for 3 h. Upon completion, the resin was washed with 5 x 500 µl of DMF and 1 x 500 µl of DCM and was brought to dryness.

Cleavage of Biotin-AHA-Glc-BTP1.

12.5 mg biotin capped peptide (0.0022 mmoles, theoretical) resin were cleaved with reagent K (trifluoroacetic acid/phenol/H_20/thioanisole/ ethanedithiol, 82.5:5.0:5.0:5.0:2.5) for 2 h. The filtered supernatant was reduced to a minimum volume, triturated, washed with 2:1 ether/hexane (6 x 12 mL), and lyophalized from water (1 mL) to give 1.65 mg (84%) of crude peptide.

Avidin affinity column purification of Glc-BTP1.

1 mL of 1 mg/mL (1.2×10^{-4} mmoles binding sites) avidin gel (6% crosslinked to agarose from Pierce) was washed with 10 mL of pH 7.0 20 mM phosphate/500 mM NaCl buffer. 40 µl of a 10 mg/mL solution (4.5×10^{-4} mmoles-excess of peptide to avidin binding sites) of crude biotin-AHA-glc-BTP1 peptide were diluted into 1 mL of buffer and added to the resin. The resin suspension was thoroughly mixed and allowed to equilibrate for 1 hour. Resin was then washed with 5 mL of pH 7 buffer, 5 mL of MilliQ water, 5 mL of 50% acetonitrile/water, and then 10 mL of MilliQ water. The first two washes were collected and stored at –80 °C for reuse. 1 mL of 0.075 M 3-(cyclohexylamino)-2-hydroxypropanesulfonic acids (CAPSO) pH 9.50 buffer was added to the resin and the mixture was agitated and left for 24 h. Collection of the column eluent and a 1.5 mL water wash afforded *N*-terminal glycolyl-capped peptide in 76% yield. Purified peptides were either used directly or stored in solution at –80 °C.

α-Chloroacetyl capping of the CPLB library.

One full CPLB library (0.133 mmoles) was capped with the *N*-terminal α-chloroacetyl cap. The Fmoc-amine protecting group was removed by a one hour treatment with 20% PIP/DMF. After a DMF wash, capping was performed in 10 ml of a 9:1 DMF/DCM solution of monochloroacetic anhydride (0.414 g, 0.30 M) and HOBT (0.513 g, 0.27 M) for six h. A Kaiser test at this point was negative. The resin was then washed with DMF and methanol and then dried *in vacuo*.

Side-chain deprotection and cleavage of the α-chloroacetyl capped CPLB library.

Removal of the side chain-protecting group was accomplished with a two hour treatment of the resin with an excess of reagent K (trifluoroacetic acid/ phenol/H_2O/thioanisole/ethanedithiol, 82.5:5.0:5.0:5.0:2.5). The resin was then thoroughly washed with TFA and methanol and then stored at -80° C.

Conversion of the α-chloroacetyl cap to glycine.

Conversion of the *N*-terminal α-chloroacetyl capping group to glycine was achieved by aqueous ammonia treatment. Resin beads derivatized with peptides capped by the α-chloroacetyl group were washed in pH 7.5 0.1M phosphate buffer and then transferred to 150 µl of water. 150 µl of fresh 30% aqueous ammonia was added to the water. After mixing, the ammonia solution was tightly capped and allowed to sit at 25° C. After 12 h the beads were removed, washed in pH 7.50 0.1M phosphate buffer, and stored in 400 µl of 1:1 methanol/water at 4° C until sequencing was performed.

References

(1) Shogren-Knaak, M. A.; Imperiali, B. "A reversible affinity tag for the purification of N-glycolyl capped peptides," *Tetrahedron Lett.*, **1998**, *39*, 8241-8244.

(2) Shogren-Knaak, M. A.; McDonnell, K. A.; Imperiali, B. "α-Chloroacetyl capping of peptides: An N-terminal capping strategy suitable for Edman sequencing," *Tetrahedron Lett.*, **2000**, *41*, 827-829.

(3) Bianchi, E.; Sollazzo, M.; Tramontano, A.; Pessi, A. "Affinity purification of a difficult-sequence protein," *Int. J. Pept. Prot. Res.*, **1993**, *42*, 93-96.

(4) Thiele, C.; Fahrenholz, F. "Photocleavable biotinylated ligands for affinity chromatography," *Anal. Biochem.*, **1994**, *218*, 330-337.

(5) Kellam, B.; Chan, W. C.; Chhabra, S. R.; Bycroft, B. W. "Transient affinity tags based on the Dde protection/deprotection strategy: Synthesis and application of 2-biotinyl- and 2-hexanoyldimedone," *Tetrahedron Lett.*, **1997**, *38*, 5391-5394.

(6) Roggero, M. A.; Corradin, S. G. "A simple and rapid procedure for the purification of synthetic polypeptides by a combination of affinity chromatography and methionine chemistry," *FEBS Lett.*, **1997**, *408*, 285-288.

(7) Ball, H. L.; Mascagni, P. "Chemical synthesis and purification of proteins: A methodology," *Int. J. Pept. Prot. Res.*, **1996**, *48*, 31-47.

(8) Kazmierski, W. M.; McDermed, J. "Synthesis of carbonic acid benzotriazol-1-yl-ester-(2-biotinylamino)-9H-fluoren-9-ylmethyl ester: A convenient transient-biotinylation reagent for use in affinity chromatography," *Tetrahedron Lett.*, **1995**, *36*, 9097-9100.

(9) Brown, A. R.; Irving, S. L.; Ramage, R. "Affinity purification of synthetic peptides and proteins on porous graphitised carbon," *Tetrahedron Lett.*, **1993**, *34*, 7129-7132.

(10) Funakoshi, S.; Fukuda, H.; Fuji, N. "Chemoselective one-step purification method for peptides synthesized by the solid-phase technique," *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 6981-6985.

(11) Merrifield, R. B.; Bach, A. E. "9-(2-Sulfo)fluorenylmethoxycarbonyl chloride, a new reagent for the purification of synthetic peptides," *J. Org. Chem.*, **1978**, *43*, 4808-4816.

(12) Krieger, D. E.; Erickson, B. W.; Merrifield, R. B. "Affinity purification of synthetic peptides," *Proc. Natl. Acad. Sci. USA*, **1976**, *73*, 3160-3164.

(13) Olejnik, J.; Sonar, S.; Krymanska-Olejnik, E.; Rothschild, K. J. "Photocleavable biotin derivatives: A versatile approach for the isolation of biomolecules," *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 7590-7594.

(14) Canne, L. E.; Winston, R. L.; Kent, S. B. H. "Synthesis of a versatile purification handle for use with Boc chemistry solid phase synthesis," *Tetrahedron Lett.*, **1997**, *38*, 3361-3364.

(15) Garcia-Echeverria, C. "On the use of hydrophobic probes in the chromatographic purification of solid-phase-synthesized peptides," *J. Chem. Soc, Chem. Commun.*, **1995**, *7*, 779-780.

(16) Green, N. M. "Avidin," *Adv. Prot Chem.*, **1975**, *29*, 85-133.

(17) Wilcheck, M.; Bayer, E. A. Avidin-Biotin Technology. In *Methods in Enzymology*; Wilcheck, M. Bayer, E. A.; Academic Press Inc., San Diego, 1990; pp 123-139.

(18) Lam, K. S.; Lebl, M.; Krchnak, V. "The "one-bead-one-compound" combinatorial method," *Chem. Rev.*, **1997**, *97*, 411-448.

(19) Smith, B. J., Ed. *Protein Sequencing Protocols*; Humana Press: Totowa, 1997; pp 375 .

(20) Meldal, M.; Svendsen, I.; Breddam, K.; Auzanneau, F. I. "Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity," *Proc. Natl. Acad. Sci. U.S.A.*, **1994**, *91*, 3314-3318.

(21) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman and Company: New York, 1985; pp 475.

(22) Greene, T. W.; Wuts, P. G. M. In *Protective Groups in Organic Synthesis*; John Wiley & Sons, New York, 1999; pp 555-556.

(23) Novabiochem *1999 Novabiochem Catalog and Peptide Synthesis Handbook*; Novabiochem: San Diego, 1999.