

**Protein Secondary Structure Mimetics:**

**Design, Synthesis and Evaluation**

by

**Zhong Qi Li**

**M.A. & B.A., Harvard University, 1991**

**Submitted in Partial Fulfillment**

**of the Requirements for the**

**Degree of**

**Doctor of Philosophy**

**at the**

**Massachusetts Institute of Technology**

**June, 1996**

**Copyright © Massachusetts Institute of Technology 1996**

**Signature of Author \_\_\_\_\_**

**Department of Chemistry**

**May 23, 1996**

**Certified by \_\_\_\_\_**

**Professor Daniel S. Kemp**

**Thesis Supervisor**

**Accepted by \_\_\_\_\_**

**Professor Dietmar Seyferth**

**Chairman, Departmental Committee on Graduate Students**

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

JUN 12 1996

Science

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor W. H. Orme-Johnson

Chairman

Professor D. S. Kemp

Thesis Supervisor

Professor F. D. Greene

Protein Secondary Structure Mimetics:  
Design, Synthesis and Evaluation

by

Zhong Qi Li

Submitted to the Department of Chemistry at the Massachusetts Institute of Technology in June, 1996 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

A novel tolan-based template was designed to induce  $\beta$ -hairpin formation as well as to report the conformational changes resulting from different  $\alpha$  amino acid derivatives. The synthesis of 2-amino-2'-carboxydiphenylacetylene (the template) and thirty template-peptide conjugates and model compounds are described. The ability of the template to act as a reporting conformational template (RCT) for the antiparallel  $\beta$ -sheet formation was investigated and the results are reported herein.

As a key reaction in the synthesis of the template and the template-peptide derivatives  $\text{AcAA}_1\text{-Taa-AA}_2\text{NHMe}$ , the  $\text{Pd}^0\text{-Cu}^{\text{I}}$  catalyzed coupling of an aryl acetylene with an aryl iodide to yield tolanes was investigated and the scope of this coupling reaction was defined. The commonly-used Sonogahira's procedure was modified, and excellent coupling results were obtained for aryl reactants that are functionalized with peptides at the *ortho* positions.

Conformational studies of these template-peptide conjugates and model compounds by using Nuclear Magnetic Resonance Spectroscopy as well as

Ultraviolet Absorption and Circular Dichroism Spectroscopy indicate that the template is indeed capable of inducing  $\beta$ -sheet formation when linked to peptides. Although the two termini are relatively free, the first hydrogen-bond of the  $\beta$ -sheet is preserved at the tolanoic template among all derivatives examined. At the long wavelengths ( $\lambda_{\text{max}}$  ~250 nm and 315 nm), the ultraviolet absorption spectra and circular dichroism spectra of these template-peptide conjugates display intense bands that have wide range differences in intensities. Both UV and CD spectra are sensitive to the antiparallel  $\beta$ -sheet conformations of the different  $\alpha$  amino acid derivatives.

**Thesis Supervisor: Professor Daniel S. Kemp**

**Title: Professor of Chemistry**



## Acknowledgments

I would like to thank Professor D. S. Kemp for the privilege of working in his laboratories, and for the freedom and independence he gave me in my work.

Financial support from a teaching assistantship (1991-1992) from the Chemistry Department and a research assistantship (1992-1996) from Professor Kemp is appreciated.

I am thankful to the staff of the Chemistry Department Spectroscopy Laboratory for their assistance, and especially to Ms. Li Li for those speedy HRMS data. I am also thankful to the Orme-Johnson group for letting me use their spectrophotometer and to the Biemann group for allowing me to use their delicate balance.

I would like to thank the present and past members of the Kemp group for their friendship and camaraderie. Special thanks do to Dr. Tom Allen and Dr. Gene Hickey for their advice and enduring friendship, and to Janette Lee for proofreading this thesis. My best wishes for a successful future go to Janette, Dr. Kwok Tsang, Evan Powers, Linda Shimizu and Robert Bieganski.

My sincere gratitude goes to all my teachers and friends, especially to Teacher Guan Fengqin of the Tiefeng Elementary School, Teacher Chen Guilan of the Tongdong Elementary School, Teacher Han Jinxiu of the Eighth Middle School, Teacher Wang Guizhi and Principal Jin of the Experimental High School, President Yang Shiqin and Ms. Ao Lifeng of the Harbin Institute of Technology. I am also thankful to Phillips Academy at Andover for the exchange program and to Harvard for offering me the opportunity and privilege of studying there.

I am forever indebted to my parents for their everlasting love and support. I am grateful to my brothers for always wanting to protect their little brother, and to my sister for being my first teacher and for taking care of me when I was young and our mother was at work.

Finally, and most importantly, I want to thank Miyuki for her love, companionship and support, and to C. Li for having been the source of my joy and happiness. To you I dedicate this thesis.

To Miyuki and C. Li

## Table of Contents

<b>1. Introduction</b>	<b>9</b>
1.1 Basic protein structure	11
1.2 Protein folding	18
1.3 <i>De novo</i> design of proteins	20
1.4 $\beta$ -Sheet models	21
1.5 Design consideration of the tolan-based template for $\beta$ -hairpin	27
<b>2. Synthesis</b>	<b>35</b>
2.1 A brief review of the coupling reaction of aryl iodides with acetylenes	40
2.2 The synthesis of 2-amino-2'-carbomethoxytolane	42
2.3 Functionalization of H-Taa-OMe with peptide conjugates	43
2.3.1 Via <i>N</i> -acetylamino acids	45
2.3.2 Via <i>N</i> -( <i>tert</i> -butoxy)carbonylamino acids	46
2.4 An alternative approach to the functionalization of the tolanoic template with peptide conjugates	47
2.5 The synthesis of model compounds	50
2.5.1 The synthesis of AcAib-Taa-AibNHMe	51
2.5.2 The synthesis of other model compounds	55
2.6 The synthesis of tetrapeptide analogues via H-Taa-OH	58
2.7 The synthesis of a cyclic template-peptide system	60
2.8 Nucleophilicity of <i>o</i> -alkynylanilines toward acylation	65
2.9 The scope of the Pd <sup>0</sup> -Cu <sup>I</sup> catalyzed coupling reaction	66
2.10 Summary	69
<b>3. Conformational Properties Studied by NMR</b>	<b>70</b>
3.1 Introduction	71
3.2 Solvent conditions and expectations from NMR	78

3.3 Ac-Taa-NHMe and AcNH-Tln-Me	80
3.4 Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe	83
3.5 AcAib-Taa-AibNHMe	84
3.6 AcGly-Taa-AA <sub>2</sub> NHMe	87
3.7 AcVal-Taa-AlaNHMe	95
3.8 AcVal-Taa-IleNHMe and AcAA <sub>1</sub> -Taa-ValNHMe	97
3.9 AcPro-Taa-AA <sub>2</sub> NHMe	100
3.10 Chemical shift correlation	110
3.11 Cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala]	114
3.12 Summary	119
<b>4. Conformational Properties Studied by UV and CD</b>	<b>121</b>
4.1 Introduction: UV and CD	123
4.2 The tolan chromophore	128
4.3 UV results and discussion	132
4.4 CD results and discussion	152
4.5 Conclusion and future directions	170
<b>5. Experimental Details</b>	<b>172</b>
<b>Appendix A: a mathematical derivation of the molar ellipticity</b>	<b>243</b>
<b>Appendix B: Tabulation for Figure 4.10 and 4.15</b>	<b>253</b>
<b>Bibliography</b>	<b>255</b>
<b>Biographical Note</b>	<b>265</b>

## **Chapter One**

### **Introduction**

:

Proteins, along with nucleic acids, are vital to living organisms. They act as enzymes to catalyze reactions, as hormones to transmit information, as storage and transport for particles across membranes, as protective skins and hairs, as connective tendons, as machinery muscles, as antibodies for defense, and as controls for genes.

The wide range of properties of proteins stem from the diversity of their structures. Nucleic acids encode genetic information for the structures and functions of proteins. The genetic code of nucleic acids encoding for the primary sequences of proteins is well-understood, but knowledge on how the protein primary sequence encodes for its unique three-dimensional fold is still lacking. This so-called “second half of the genetic code” is often referred to as the protein folding problem.

Protein folding and protein design are two different faces of the same coin. With the introduction of protein engineering, whereby a protein of any desired amino acid sequence can be produced in substantial quantities by expressing a cloned or synthetic gene, there is great demand for insight into the function of a protein from just its primary sequence. The function of a protein is closely related to its three-dimensional structure. Therefore, to understand how a primary sequence of amino acids is translated into its unique three-dimensional structure is of paramount importance in science today. We know the basic forces (electrostatic, hydrogen bonding, hydrophobic packing, van der Waals interactions, *etc.*), but how much each of them contributes to the final three-dimensional fold of a protein is not so clear. Therefore, studies toward the

understanding of contributions of different forces to the stability of basic protein structure are in urgent need.

We choose a peptidomimetic approach to isolate elements of the protein secondary structure and study them in detail, whereby we hope to gain deeper insight into the basic forces responsible for protein folding and structure. In order to design efficient and useful peptide mimetics, we must first have a good understanding about the basic structure of proteins.

## 1.1 Basic protein structure

A protein is built from a long-chain polymer of amino acids, called a polypeptide chain. The 20 natural  $\alpha$  amino acids, which are all in the *L* configuration, are the basis for the diverse structures and functions of proteins. They all share the same backbone structure (Figure 1.1), and are classified into three general categories according to the different side chains: nonpolar, polar but uncharged, and charged polar. The nonpolar residues include those with aliphatic hydrocarbon side chains: glycine (Gly), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), proline (Pro), an aromatic group, phenylalanine (Phe), and a "pseudohydrocarbon," methionine (Met). For the polar/neutral group, there are hydroxyl-containing residues serine (Ser) and threonine (Thr), amides asparagine (Asn) and glutamine (Gln), aromatic tyrosine (Tyr) and tryptophan (Trp), and sulfhydryl cysteine (Cys). In the charged polar category are acidic aspartic acid (Asp) and glutamic acid (Glu), basic histidine (His), lysine (Lys), and arginine (Arg). The large nonpolar residues (Val, Leu, Ile, Pro, and Phe) are found usually in the interior of proteins, whereas polar ones on the outside. Gly and Ala are small so that they can be accommodated either in the interior or on the surface of proteins. The hydrocarbons Ala, Val, Leu, and Ile,

and the norleucine-equivalent Met (Smith, 1959; Gellman, 1991) vary in size and shape, and therefore exhibit varying degrees of hydrophobic character (Kauzmann, 1968).

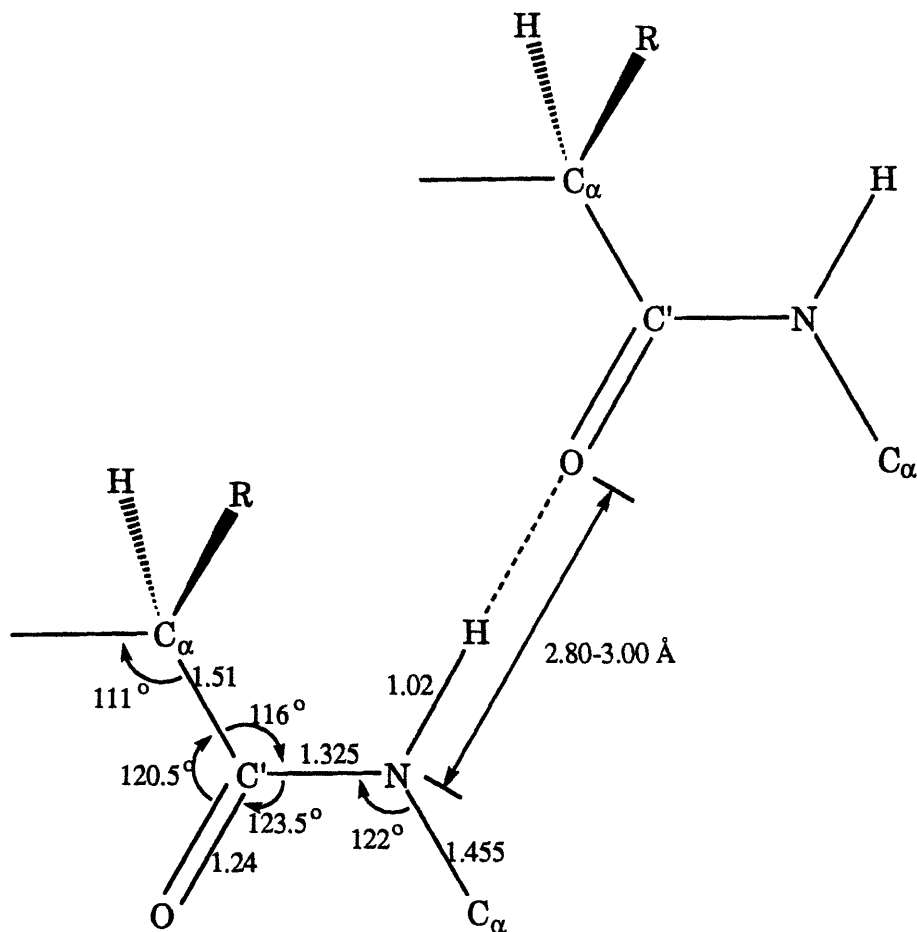


Figure 1.1 Dimensions for a peptide group.

Amino acids are linked together by the  $-CO-NH-$  group, called a peptide bond. Many of the special properties of a polypeptide chain arise from the nature of the backbone chain. The dimensions for a peptide group and the hydrogen-bond distance between two amino acid residues are well known, and





The peptide bond is usually planar ( $\omega = 180^\circ$ ) and has a barrier of about 21 kcal/mole for rotation around this peptide bond. Due to steric hindrance, almost all peptide bonds, except in the case of prolyl residue which has a tertiary amide, are in the *trans* conformation, in which the two flanking  $\alpha$  carbons are at diagonally opposite corners of the amide plane. Proteins are linear polymers, often cross-linked but never branched. These constraints, therefore, greatly limit the ways in which a polypeptide can fold.

Due to steric hindrance of short contacts between the atoms in the adjacent residues, only a limited set of values of  $\phi$  and  $\psi$  are allowed. The permitted values of  $\phi$  and  $\psi$  were first determined by Ramachandran and coworkers, on the basis of stereochemical criteria involving limited contact distances between atoms by using hard-sphere models of the atoms and Pauling-Corey coordinates of the bonds (Ramachandran *et al.*, 1963; Ramakrishnan & Ramachandran, 1965; Ramachandran *et al.*, 1966; Ramachandran & Sasiesekharan, 1968). Ramachandran plots for glycyl residue (top plot) and alanyl residue (bottom) are shown in Figure 1.3. The normally allowed values are shaded and partially allowed ones are enclosed by solid lines. The connecting region, indicated by dashed lines, is permitted if the angle at the  $\alpha$ -carbon atom is varied. As shown in Figure 1.3, 45% of the total area in the  $\phi$ - $\psi$  plane is fully allowed for glycyl residue, and only about 7.5% is allowed for alanyl residue. Residues with longer and larger side chains will have additional restrictions, and the allowed region is even smaller. Because of its rigid five-membered ring, the value of proline  $\phi$  is confined to be around  $-60^\circ$ .

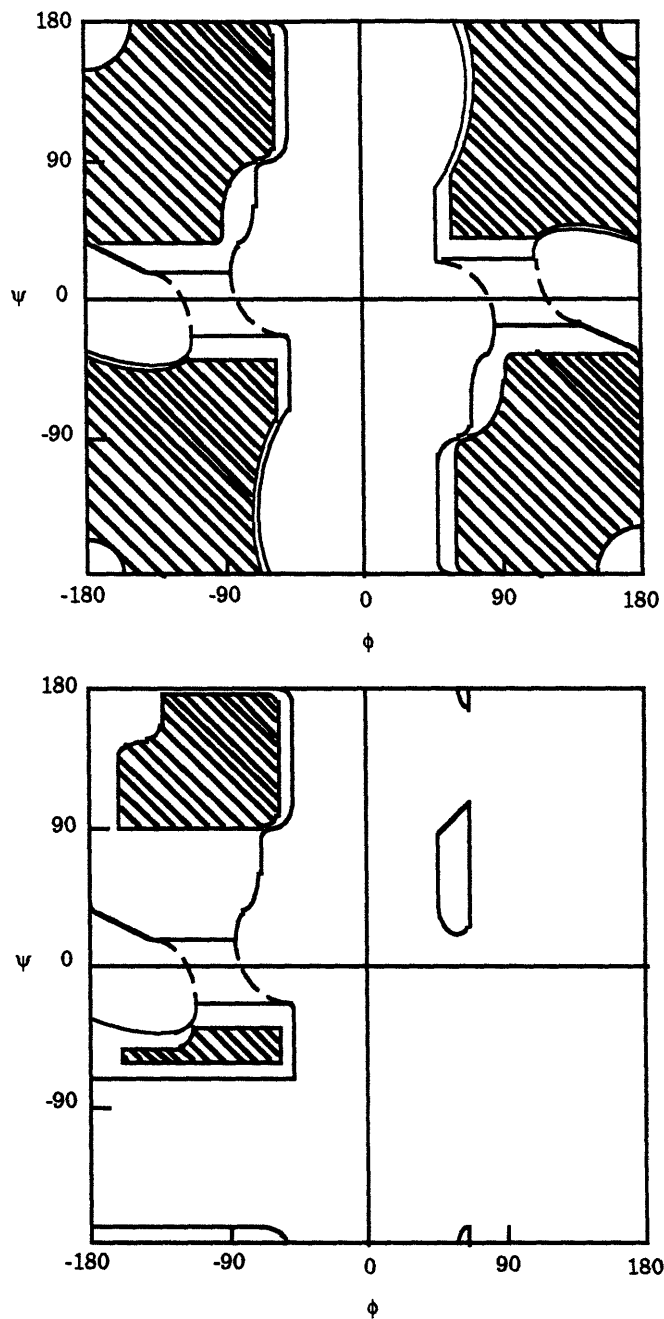


Figure 1.3 Ramachandran Plots for Gly (top) and Ala (bottom) residues. The fully allowed regions are shaded, and the partially allowed regions are enclosed by a solid line. The connecting regions enclosed by the dashed lines are permissible with slight alterations of bond angles.

Well-defined conformations of regular values of  $\phi$  and  $\psi$  angles are observed in proteins and polypeptides. These are called the secondary structure of proteins, and are important building blocks for the three dimensional structures of proteins. The most important secondary structures are  $\alpha$ -helices,  $\beta$ -sheets, and reverse turns. In an  $\alpha$ -helix, the preferred  $\phi$ ,  $\psi$  values are  $-57^\circ$  and  $-47^\circ$ , respectively, which fall in the allowed region of the lower-left quadrant in a Ramachandran plot. The amino acid backbone of an  $\alpha$ -helix curls up in a spiral fashion in which the backbone carbonyl oxygen of each residue hydrogen-bonds to the backbone amide proton of the fourth residue along the chain. These hydrogen bonds are about 2.86 Å long between the non-hydrogen atoms involved and are nearly parallel to the helix axis.

$\beta$ -Sheet is another important class of protein secondary structure. In a  $\beta$ -sheet, all the residues are almost fully extended, in contradistinction to  $\alpha$ -helices. The basic unit for a  $\beta$ -sheet is the  $\beta$ -strand, which is a single chain. A  $\beta$ -strand alone is not stable. It is only when two  $\beta$ -strands come together, with regular hydrogen bonds formed between them, that a stable  $\beta$ -sheet conformation can be formed. Depending on the relative directions of the two strands, two stable  $\beta$ -sheet structures can be formed: parallel or antiparallel. The  $\phi$ ,  $\psi$  values are about  $-119^\circ$  and  $+113^\circ$  for the parallel  $\beta$ -sheet, and about  $-139^\circ$  and  $+135^\circ$  for the antiparallel  $\beta$ -sheet. These values fall in the allowed upper-left quadrant region in a Ramachandran plot.

A third class of important secondary structure is reverse turns that usually change the direction of polypeptide chains. Reverse turns are very

diverse, depending on which types of secondary structure they link and on how many residues are involved in the turn. The best characterized are the  $\beta$ -turns that link adjacent strands in an antiparallel  $\beta$ -sheet. The two central residues in a  $\beta$ -turn are not involved in the hydrogen bonding of the  $\beta$ -sheet; the two residues on either side of the two central residues are hydrogen-bonded and are part of the anti-parallel  $\beta$ -sheet. The conformation of a  $\beta$ -turn can be completely defined with the  $\phi$  and  $\psi$  angles of the two central residues. There are considerable variations of  $\phi$  and  $\psi$  values, and observed  $\beta$ -turns are classified accordingly into several different groups.

A  $\beta$ -hairpin, in which two antiparallel  $\beta$ -sheet strands are linked by a  $\beta$ -turn, is schematically drawn in Figure 1.4. Notice the difference between the residue pair that is hydrogen-bonded and the pair that is not hydrogen-bonded. The side chain groups of the former point away from each other, whereas those of the latter are in close contact.

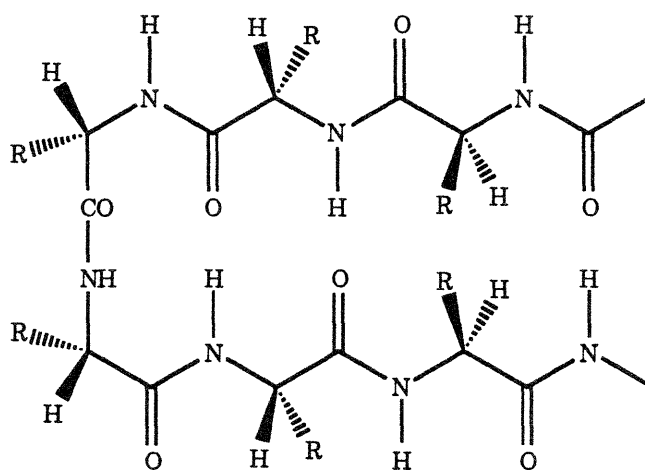


Figure 1.4 Schematic drawing of a  $\beta$ -hairpin.

Studies of basic protein structures, such as the  $\beta$ -hairpin, are of supreme significance because understanding how these structures are formed and what propensities various amino acid residues have for each structure is key to solving two of the most important tasks in biochemistry today: protein folding and the design of novel protein structures.

## 1.2 Protein folding

“Chemists have long understood that a protein’s amino acid sequence determines its three-dimensional structure. Moreover, they have known for some time that proteins are able to carry out their myriad biological functions only when they have folded up into compact three-dimensional structures. Indeed, the late chemist Christian B. Anfinsen received the 1972 Nobel Prize in Chemistry for his work on the relationship between the structural properties of proteins and their biological functions, specifically with regard to the enzyme ribonuclease. Yet one of the great unsolved mysteries of life - and thus an enduring scientific problem - has been a full understanding of the process by which protein folding occurs.” (Jacobs, M. C&EN November 6, 1995, 5.)

The protein folding problem first gained prominence in the 1950s and ‘60s, when it was discovered that ribonuclease could regain its enzymatic activity after it had been fully denatured in high concentrations of urea. This finding along with other evidence led to the conclusion that structure is encoded in sequence - a protein’s amino acid sequence fully determines its three-dimensional structure (Anfinsen, 1973; Anfinsen & Scheraga, 1975).

In addition, Anfinsen proposed that a protein folds into the conformation that represents its lowest possible energy state - a concept referred to as the "thermodynamic model" of protein folding (Anfinsen & Scheraga, 1975). Because of the proposed correspondence between sequence and structure - a relationship sometimes referred to as the second genetic code - it should theoretically be possible to predict structure from sequence (Gierasch & King, 1990). So it might be possible to determine the folded structure of a protein by evaluating the energy of all possible protein conformations and selecting the one with lowest energy. However, the number of possible conformations is astronomically high and it would take centuries of computer time to solve a single folding problem by evaluating all possible conformations (Levinthal, 1968).

Studies of the protein folding problem have gained much progress due to recent advances in new techniques, especially protein/peptide engineering and NMR spectroscopy. A great deal of information can be obtained on the process of protein folding and on the energetics of the native conformation through genetic mutagenesis of proteins or protein domains (point mutations or random mutations and screening), for example, on tailspike protein (King, 1986), tryptophan synthase (Matthews, 1987), bovine pancreatic trypsin inhibitor (Goldenberg *et al.*, 1989; Creighton, 1990; Weisman & Kim, 1991), and on barnase (Fersht *et al.*, 1992).

Advances on the studies of the protein folding problem also came from the initial discovery that the 13-residue C-peptide of ribonuclease A can adopt stable  $\alpha$ -helical secondary structure in solution (Bierzynski *et al.*, 1982; Shoemaker *et al.*, 1987). Also, studies on other small fragments of proteins and synthetic model peptides all implicate a tendency for stable secondary structures in solutions (Dyson *et al.*, 1988; Shortle & Meeker, 1989; Rohl *et al.*, 1992). The fact that a segment of a protein can adopt the same structure as

found in the native protein indicates that the independent development of native-like substructures may be an important aspect of the protein folding process, and it also suggests that the information content of protein structure under native conditions can be investigated independently from the rest of the polypeptide chain (Lecomte & Matthews, 1993). These autonomous folding units are small enough to be probed in detail by techniques such as NMR spectroscopy.

Great strides have been made and substantial progress will surely follow in solving the great mystery of protein folding - the deciphering of the folding code. These advances will be of great importance in the prediction of the three-dimensional structure from a given amino acid sequence. They will also be of great utility in the *de novo* design of polypeptides of defined structures.

### **1.3 *De novo* design of proteins**

The *de novo* design of peptides and proteins has recently emerged as an approach for studying protein structure and function (Bryson *et al.*, 1995). Stable short  $\alpha$  helices have been created by incorporating many helical stabilizing alanine residues (Scholtz *et al.*, 1991), by adding salt bridges (Lyu *et al.*, 1990; Huyghues-Despointes *et al.*, 1993), by covalent macrocyclization (Bracken *et al.*, 1994), and by the ingenious design of templates to initiate helix formation (Kemp *et al.*, 1991). By systematically varying the sequences of these model helices,  $\alpha$ -helical propensities of amino acids can be obtained. *De novo* designs of higher level structures of helices, such as  $\alpha$ -helical coiled coil (Monera *et al.*, 1994; Myszka & Chaiken, 1994; Kuroda *et al.*, 1994) and four-helix bundles (Iwamoto *et al.*, 1994; Nyanguile *et al.*, 1994; Munson *et al.*, 1994;



Olofsson *et al.*, 1995), have been achieved by extending the knowledge obtained from studies of model helical peptides.

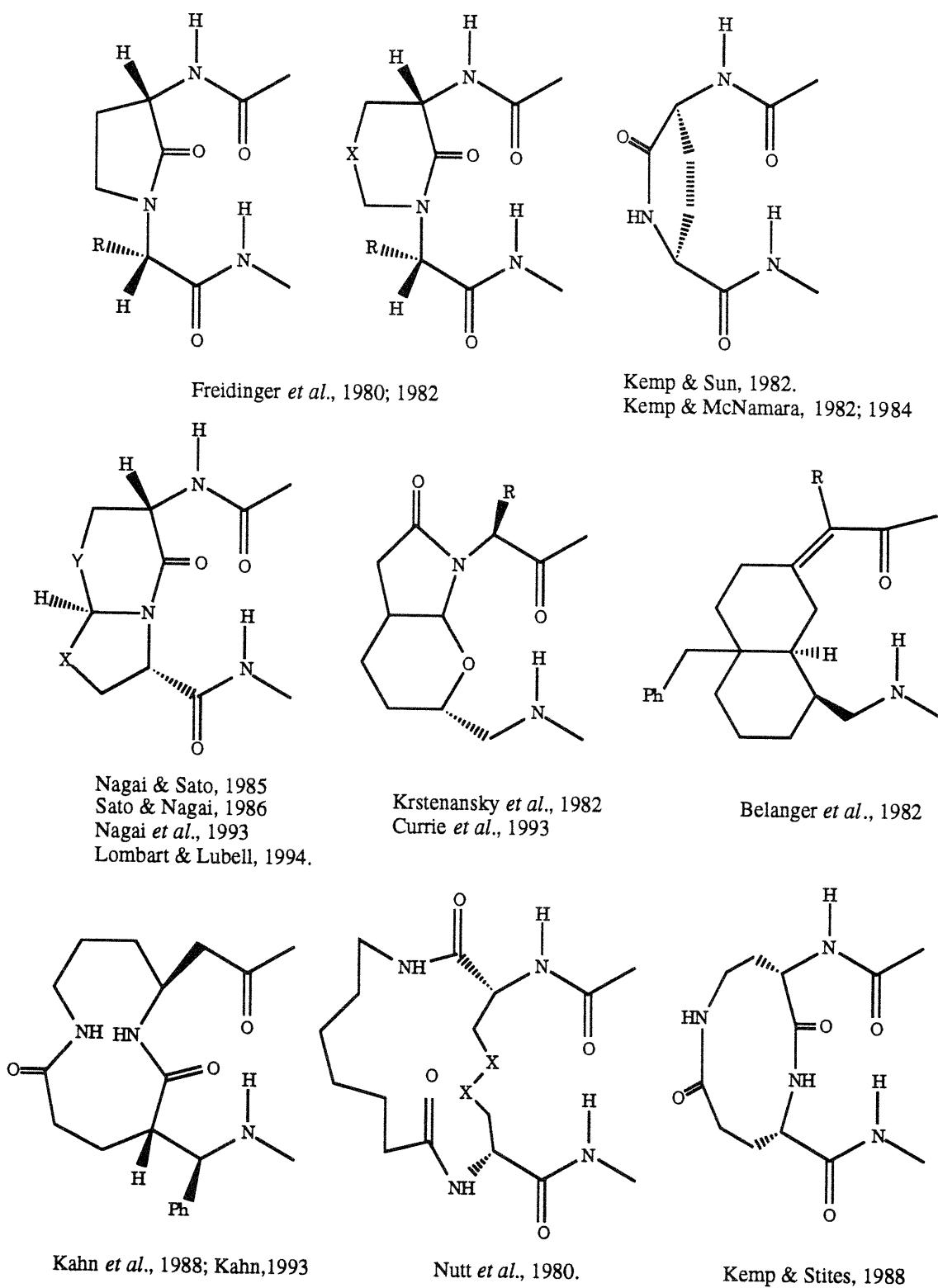
A number of experimental difficulties have beset the development of models for  $\beta$ -sheet formation. Unlike  $\alpha$  helices, where there is a regular succession of hydrogen bonds between amides four residues apart in sequence,  $\beta$  sheets are formed by residues at variable and often distant positions in the sequence. Also, the exposed amides at the edge of  $\beta$  sheets can hydrogen-bond to other sheets, leading to insoluble aggregates. As a result, thermodynamic parameters controlling sheet stability are not as well understood as those for helices, and there exist today fewer *de novo* designed  $\beta$ -sheet structures than  $\alpha$ -helices. Examples of *de novo* designs of  $\beta$ -sheet structures are the beta sandwich “betabellin” (Yan & Erickson, 1994) and its closely related four-stranded antiparallel “betadoublet” (Quinn *et al.*, 1994), six-stranded beta sandwich “minibody” (Bianchi *et al.*, 1994), and  $\alpha$ - $\beta$  barrel proteins (Tanaka *et al.*, 1994).

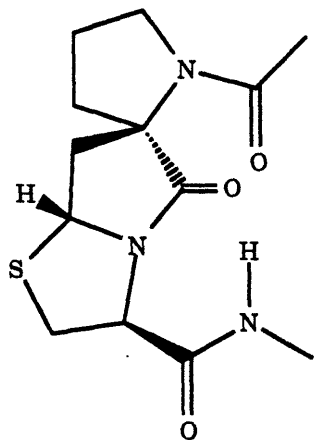
#### 1.4 $\beta$ -Sheet models

The  $\beta$ -hairpin, which is an intramolecular antiparallel  $\beta$ -sheet, as shown previously in Figure 1.4, is not completely a regular structure because the  $\beta$ -turn segment must adopt very different conformations.  $\beta$ -Sheets also have a tendency to form intermolecular aggregates because hydrogen bonds can grow indefinitely along the outer strands. For these reasons, there has been no good model system for studying  $\beta$ -sheet formation, and very little is known about

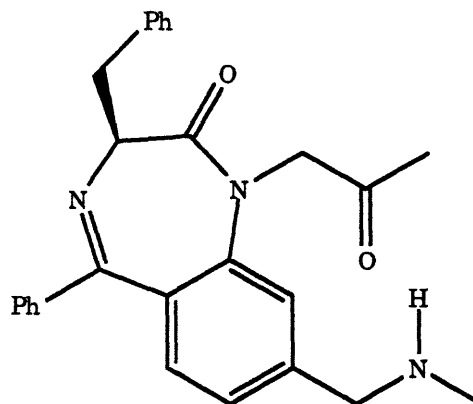
how the various amino acid residues differ in their  $\beta$ -sheet propensities. Many different research groups have set out to circumvent this thorny problem by using an approach that involves the design, synthesis, and study of peptidomimetic model systems.

Some of these earlier works had focused upon the mimicry and induction of  $\beta$ -turns. Selected examples of  $\beta$ -turn mimetics are shown in Figure 1.5a. To stabilize the peptide chain in a  $\beta$ -turn, cyclic and bicyclic dipeptides are most commonly utilized. Examples of modifications which enhance  $\beta$ -turn propensity are the dipeptide lactams that fix the central *trans* peptide bond of the  $\beta$ -turn (Freidinger *et al.*, 1980; 1982), the 2-piperidone that links the  $\alpha$ -carbon atoms of the two  $\beta$ -turn residues (Kemp & Sun, 1982; Kemp & McNamara, 1982; 1984), the bicyclic dipeptide BTD and similar proline-based derivatives (Nagai & Sato; 1985; Sato & Nagai, 1986; Nagai *et al.*, 1993; Lombart & Lubell, 1994), the rigid bicyclic spacer groups that are not based on peptide bonds (Krstenansky *et al.*, 1982; Belanger *et al.*, 1982; Currie *et al.*, 1993), the large lactam macrocycles of the  $\beta$ -turn region (Nutt *et al.*, 1980; Kemp & Stites, 1988; Kahn *et al.*, 1988; Kahn, 1993), the spirolactam (Genin & Johnson, 1992; Genin *et al.*, 1993), and the bicyclic benzodiazepine (Ripka *et al.*, 1993). Almost all of the aforementioned  $\beta$ -turn mimetics were prepared for medicinal chemistry purposes.

Figure 1.5a  $\beta$ -Turn mimetics



Genin & Johnson, 1992  
Genin *et al.*, 1993

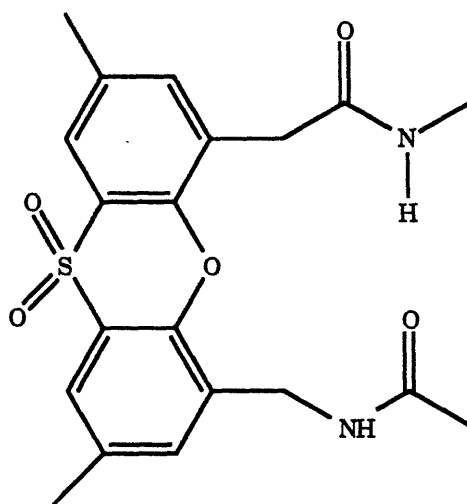


Ripka *et al.*, 1993

Figure 1.5a (cont'd)  $\beta$ -Turn mimetics

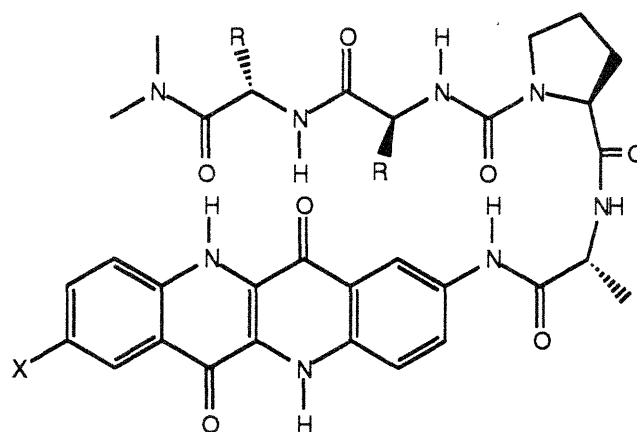
However interest on  $\beta$ -sheet mimetics began to grow after phenoxathiin *S*-dioxide (Figure 1.5b), an organic substitute for a  $\beta$ -turn, was used as a spacer to induce a cyclic antiparallel  $\beta$ -sheet structure (Feigel, 1986). And in 1988, Kemp and coworkers reported the first reporting conformational template (RCT) for antiparallel  $\beta$ -sheet based on the tetracyclic epindolidione (Figure 1.5c), which has been successfully applied in extracting  $\beta$ -sheet propensities of  $\alpha$  amino acid residues (Kemp *et al.*, 1990; Blanchard, 1992). Afterwards, some research groups have developed chemical models for  $\beta$ -hairpins in which a synthetic template holds two peptide strands or a peptide strand and a peptide strand mimic in proximity. Kelly and coworkers had demonstrated that dibenzofuran (Figure 1.5d), a  $\beta$ -turn cap, was able to nucleate  $\beta$ -sheet

structure, and hydrophobic interactions between the template and peptide side chains help stabilize  $\beta$ -sheet formation (Diaz & Kelly, 1991; Diaz *et al.*, 1992; 1993). Recently, Nowick reported that a diurea template with an aminobenzoate  $\beta$ -strand mimic (Figure 1.5d) is also capable of nucleating  $\beta$ -sheet structure (Nowick *et al.*, 1996). Complementary strategies are systems that are wholly peptide-based, including small naturally occurring  $\beta$ -sheet proteins (Kim & Berg, 1993; Minor & Kim, 1994; Smith *et al.*, 1994; Otzen & Fersht, 1995), and *de novo* designed proteins as discussed earlier.



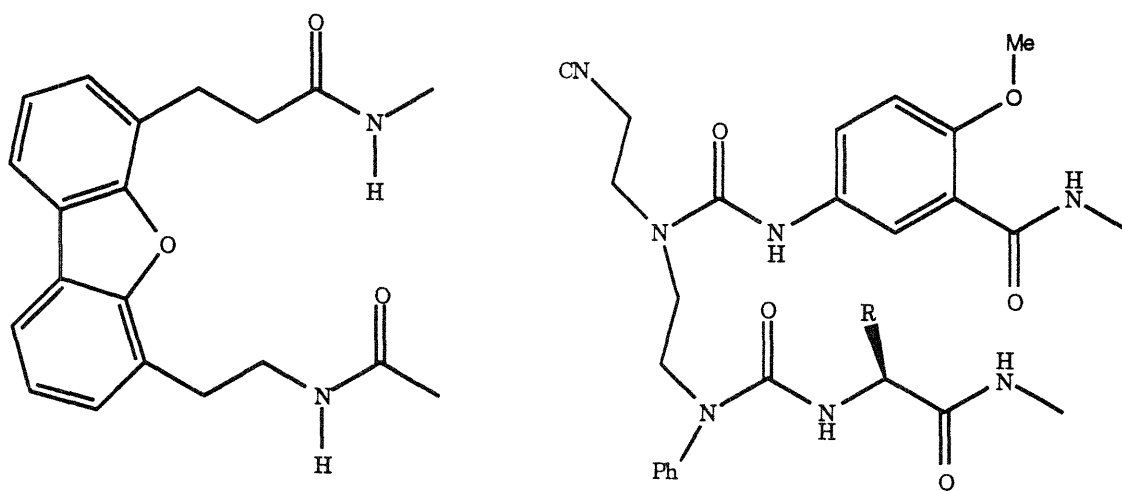
Feigel, 1986.

Figure 1.5b An early  $\beta$ -hairpin mimetic.



Kemp & Bowen, 1988;  
Kemp *et al.*, 1990.

Figure 1.5c The first antiparallel  $\beta$ -sheet RCT.



Diaz & Kelly, 1991  
Diaz *et al.*, 1992; 1993

Nowick *et al.*, 1996

Figure 1.5d Recent examples of  $\beta$ -hairpin mimetics.

At present, none of the template-based strategies have successfully modeled interactions between the side chains of the members of the non-hydrogen-bonded amino acid residue pairs in an antiparallel  $\beta$ -sheet. In addition, except the epindolidione template employed by Kemp *et al.*, none of the aforementioned  $\beta$ -hairpin templates has been shown to possess conformation-reporting properties that could shed light on  $\beta$ -sheet forming propensities of the amino acid residues. Moreover, there are considerable degrees of freedom in the existing  $\beta$ -hairpin templates. Therefore, we would like to devise a new  $\beta$ -hairpin template in which not only is the turn mimetic able to bring two  $\beta$ -strands together to form an antiparallel  $\beta$ -sheet structure, but also at the same time is capable of conveying conformational information.

The new template we are proposing is based on the tolan (diphenylacetylene) moiety, whose unique structure and photochemical properties make it a good candidate for the  $\beta$ -hairpin RCT. Before addressing whether the tolan functionality can serve as a  $\beta$ -hairpin initiator and reporter, a brief review of the structural chemistry of the tolan functionality is appropriate.

### **1.5 Design consideration of the tolan-based template for $\beta$ -hairpin**

Diphenylacetylene, commonly known as tolane, can assume two stable conformations: a planar conformation with  $D_{2h}$  symmetry, in which the two phenyl rings share the same  $p\pi$ -orbitals of the middle triple bond [Figure 1.6, structure (a)]; or an orthogonal conformation with  $D_{2d}$  symmetry, in which the

aromatic rings interact preferentially with different  $p\pi$ -orbitals of the triple bond [structure (b)] and an infinite number of others with  $D_2$  symmetry. The structure of tolane was first examined by Robertson & Woodward (1938) and was later refined by others using a variety of spectroscopic techniques and molecular modeling methods (Mavridis & Moustakali-Mavridis, 1977; Espiritu & White, 1978; Baranovic *et al.*, 1986; Abramnikov *et al.*, 1988; Saebo *et al.*, 1989).

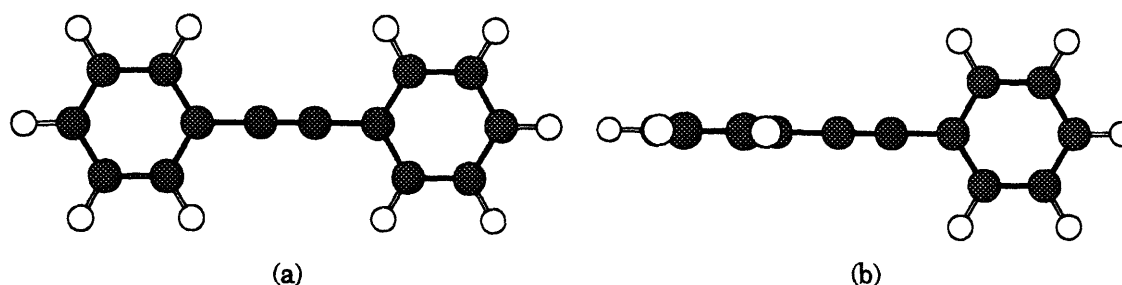


Figure 1.6. The planar (a) and orthogonal (b) conformations of tolane. Carbon atoms are shaded gray and hydrogen atoms are white.

X-ray diffraction data indicate that tolane is in the planar conformation in the crystal, consisting of two non-equivalent molecules, which are only slightly different from each other and from the ideal planarity (Mavridis & Moustakali-Mavridis, 1977; Espiritu & White, 1978). Electronic absorption, Raman, infrared, and electronic diffraction spectroscopic studies indicate that the planar conformation is also the most probable structure in solution and in the gas phase (Suzuki *et al.*, 1982; Okuyama *et al.*, 1984; Baranovic *et al.*, 1986; Abramnikov *et al.*, 1988). However, calculations suggest that there is only a small energy difference between these two conformations, with the planar molecular geometry being around 0.6 kcal/mole more stable than the



orthogonal one (Saebo *et al.*, 1989). Since this energy is on the same order as thermal energy ( $\sim 0.6$  kcal/mole), both conformations should be fairly populated at room temperature. The consensus is that tolane is planar in the solid state, and in solution the most probable geometry is planar with a fraction of non-planar molecules present.

Selected dimensions of tolane, listed in table 1.1, are obtained from the most recent and accurate measurements performed by electron diffraction in the gas phase and by X-ray crystallography for the crystal at 150 °C and at -150 °C, respectively (Abramenkov *et al.*, 1988). The distances are given in ångstrom (Å), and the angle for the triple bond is also included in the table. As expected, the angle  $\angle C-C\equiv C$  is 180°. The C-C triple bond distance (1.212 Å) is not too different from that of acetylene (1.203 Å), and this is believed to support the argument that the presence of the two phenyl rings does not appreciably disturb the character of the triple bond in tolane (Mavridis & Moustakali-Mavridis, 1977).

Table 1.1. Bond distances and the  $\angle C-C\equiv C$  angle for tolane.

	C $\equiv$ C	C-C	C-C $\equiv$	$\angle C-C\equiv C$
Gas	1.215	1.400	1.417	180°
crystal	1.212	1.399	1.426	178°

From these bond distance values, we calculated the C-1 to C-1' and C-2 to C-2' separation distances of tolane to be 4.015 Å and 5.449 Å in the gas phase, and 4.010 Å and 5.463 Å for the crystal, respectively. Therefore, for the planar conformation of tolane, the C-1 to C-1' and the C-2 to C-2' separations are 4.01 Å and  $5.45 \pm 0.01$  Å, respectively, as shown in Figure 1.7(a).

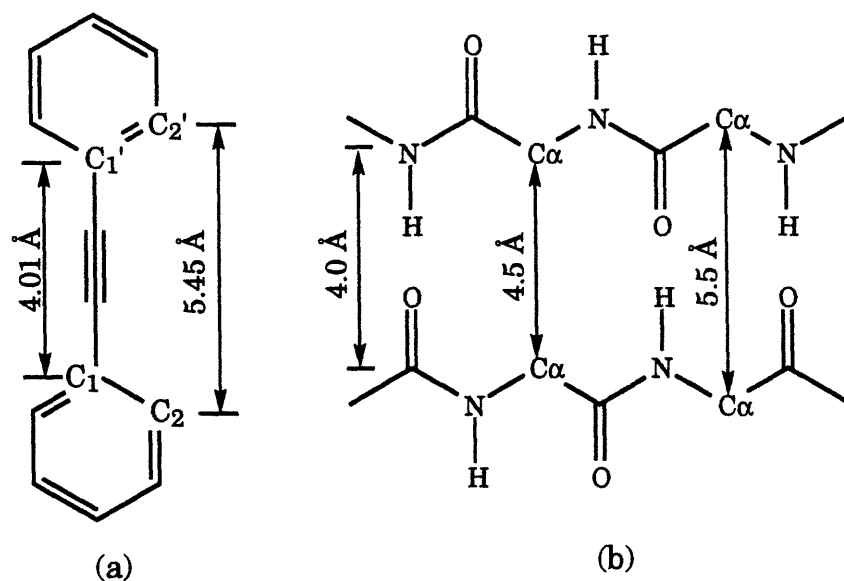


Figure 1.7. C<sub>1</sub> to C<sub>1'</sub> and C<sub>2</sub> to C<sub>2'</sub> inter-atomic distances of tolane (a) versus hydrogen bond and C<sub>α</sub> to C<sub>α</sub> distances in an antiparallel β-sheet (b).

According to the latest analysis of antiparallel β-sheets in the Brookhaven Protein Data Bank, the α-carbon atoms of members of hydrogen-bonded pairs are separated by about 5.5 Å whereas α-carbon atoms for non-hydrogen-bonded pairs are separated by about 4.5 Å as shown in Figure 1.7(b) (Wouters & Curmi, 1995; see also Ashida *et al.*, 1981 and Baker & Hubbard, 1984). The hydrogen bond distance between non-hydrogen atoms in the antiparallel β-sheet is found to be 4.0 Å (Marsh *et al.*, 1955). Therefore the 4.0 Å C-1 to C-1' distance of tolane is in excellent agreement with the interstrand hydrogen-bond distance and is in good agreement with the C<sub>α</sub> to C<sub>α</sub> distance for the non-hydrogen-bonded pair in the antiparallel β-sheet. The 5.45 Å

distance between the C-2 and C-2' carbons in tolane is also in full accord with that of the C- $\alpha$  to C- $\alpha$  interstrand spacing of members of hydrogen-bonded pair in the antiparallel  $\beta$ -sheet.

We envisioned that a tolane-based amino acid residue linked to peptides on both ends will be able to cap two antiparallel  $\beta$ -sheet strands and nucleate  $\beta$ -sheet formation. Residue 2-amino-2'-carboxydiphenylacetylene, the proposed template, is abbreviated as -Taa- ("Taa" stands for tolanoic amino acid) and its peptide conjugates are named accordingly as -AA<sub>1</sub>-Taa-AA<sub>2</sub>-, where AA<sub>1</sub> and AA<sub>2</sub> represent  $\alpha$  amino acid residues. The template -Taa- and the portion of the  $\beta$ -sheet that it mimics are enclosed in boxes (Figure 1.8).  $\alpha$ -Amino acid side chains of the four residues defining the  $\beta$ -turn are omitted for clarity. Notice how the template -Taa- mimics not only the  $\beta$ -turn residues but also the two  $\beta$ -sheet residues flanking the  $\beta$ -turn. Therefore we call this tolanoic amino acid a  $\beta$ -hairpin template. The two  $\alpha$  amino acid residues linked to the template are wide-spaced non-hydrogen-bonded pair. As a result their side chains are in close contact in the antiparallel  $\beta$ -sheet conformation.

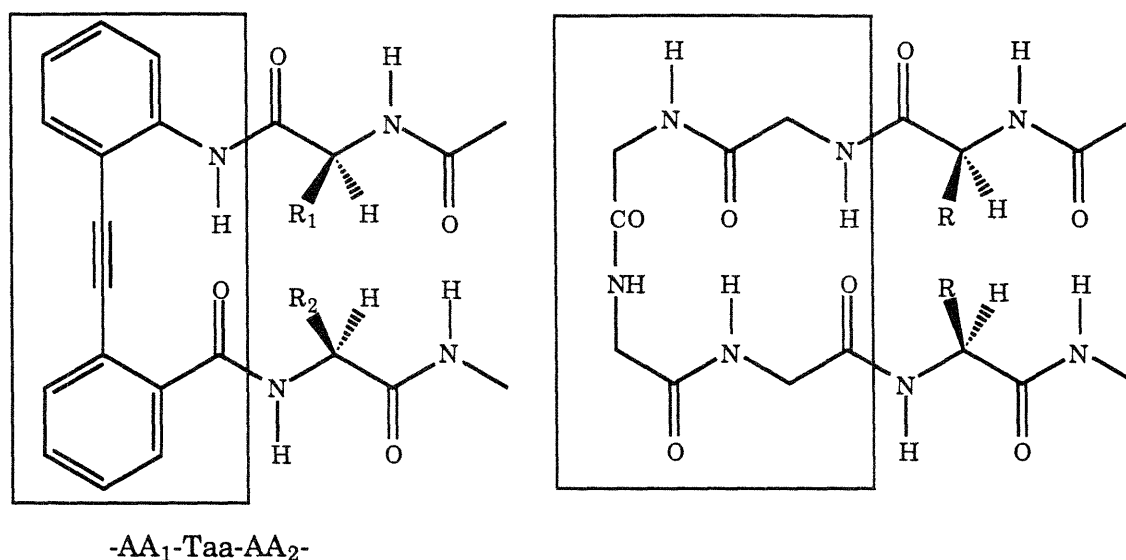


Figure 1.8 Schematic presentation of the proposed  $\beta$ -sheet template functionalized with two  $\alpha$  amino acid residues ( $-AA_1-Taa-AA_2-$ , left structure) and a  $\beta$ -hairpin structure (right). The template and the mimicked region of the  $\beta$ -hairpin are enclosed in boxes for comparison. The side chains of those  $\alpha$  amino acids in this region of the  $\beta$ -hairpin are omitted for clarity.

We are attracted to the tolanoic template for its geometric fit with the antiparallel  $\beta$ -sheet ( $\beta$ -hairpin) as well as potential conformation-reporting properties that the tolanoic chromophore may provide. There are three degrees of rotational freedom at the tolanoic chromophore: the interplanar angle  $\alpha$  about the triple bond formed between the two phenyl planes, dihedral angle  $\beta$  about the bond formed between the aromatic carbon atom C-2' and the aromatic carbonyl carbon atom C', and the dihedral angle  $\gamma$  about the bond formed between aromatic carbon atom C-2 and the aryl nitrogen atom. The

definitions of these angles are shown in Figure 1.9. Angle  $\alpha$  defines the planarity of the two phenyl planes, thus it determines the degree of the extended conjugation of the tolan system, whereas angles  $\beta$  and  $\gamma$  relate to what effect the substitutions at aryl carbon atoms C-2 and C-2' may have on the tolan chromophore. Different degrees of representation of  $\beta$ -sheet conformation in the template-peptide derivatives will induce changes in the values of these angles, which may correlate with their electronic absorption and circular dichroism spectra in a systematic fashion.

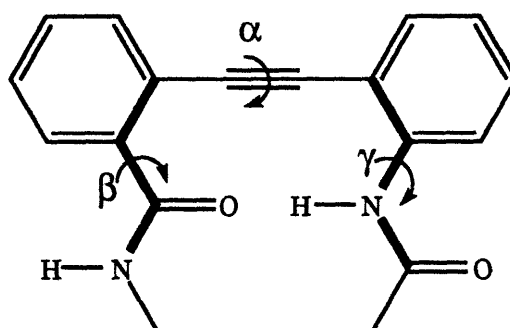


Figure 1.9 Schematic presentation of the interplanar angle  $\alpha$ , dihedral angles  $\beta$  and  $\gamma$ . Bonds that define the dihedral angles  $\beta$  and  $\gamma$  are highlighted as well as the two aromatic bonds that define the interplanar angle  $\alpha$ .

Previously, Kemp and colleagues had pioneered the concept of reporting conformational template and had amply demonstrated the importance of these RCT's in measuring propensities of  $\alpha$  amino acids for  $\alpha$ -helices (Kemp *et al.*, 1991 & 1995; Cammers-Goodwin *et al.*, 1996; Groebke *et al.*, 1996) and for  $\beta$ -sheets (Kemp & Bowen, 1988 & 1990; Kemp *et al.*, 1992). However, the NMR

techniques used for conformational quantitation is not as sensitive as UV or CD techniques. The former requires millimolar concentrations of sample whereas in the latter, samples of micromolar concentration usually suffice for measurement.

In this thesis we want to address the following two main questions: whether the tolanoic template -Taa- can be used to initiate and stabilize  $\beta$ -sheet formation and whether it has conformation-reporting capabilities in UV and/or CD.

In order to address these two issues, we must first synthesize the template, derivatize it with peptides and prepare model compounds that can serve as controls. Once the syntheses of these template-peptide derivatives and model compounds are completed, conformational studies by using a variety of spectroscopic techniques will commence to evaluate the template in terms of its ability to promote  $\beta$ -sheet formation and its capability of reporting subtle conformational changes induced by different  $\alpha$  amino acid residues. Therefore the rest of this thesis is organized in the following order: the syntheses of the template ester H-Taa-OMe and the template-peptide conjugates AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe and model compounds are described in Chapter 2; the results of conformational studies by using Nuclear Magnetic Resonance spectroscopy are summarized in Chapter 3, whereas the results of conformational studies by Ultraviolet (UV) Electronic Absorption and Circular Dichroism (CD) techniques are presented in Chapter 4.

## **Chapter Two**

### **Synthesis**

In the last chapter we inferred that a tolanoic amino acid might be a good template for  $\beta$ -sheet structure when it is linked to  $\alpha$  amino acid residues (Figure 2.1, I). In this chapter we will describe the synthesis of this tolanoic template, its derivatization with peptide conjugates and the synthesis of model compounds as controls.

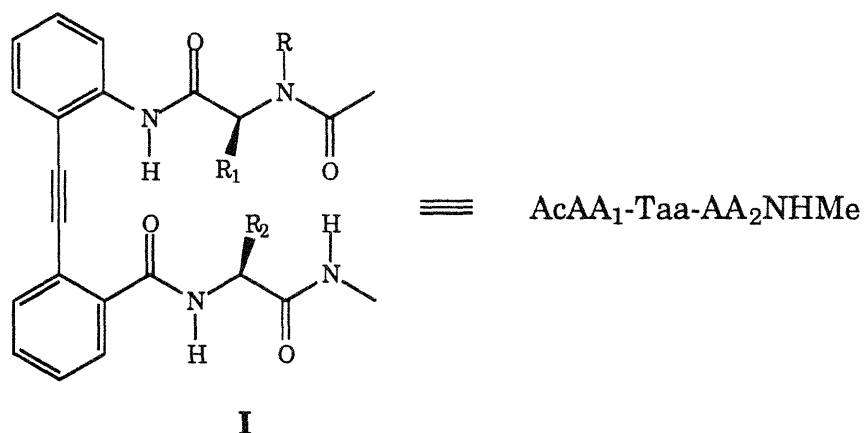


Figure 2.1 Template-peptide conjugates

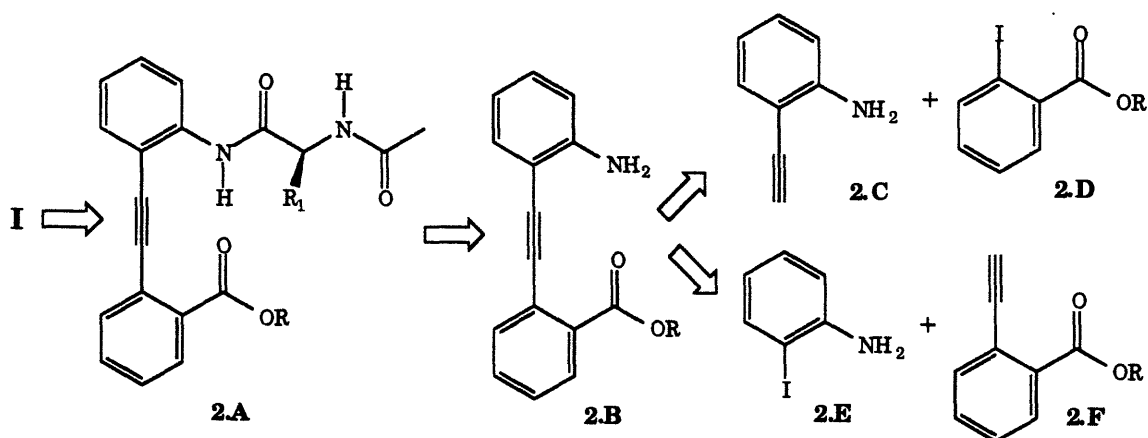
The objectives of the synthesis are the following: 1) to synthesize the tolanoic amino ester, 2-amino-2'-carbomethoxy-tolane; 2) to find an efficient method to derivatize the template with peptides; 3) to define the scope of the coupling reaction of aryl acetylenes with aryl iodides catalyzed by palladium (0) and copper (I) via the modified Sonogashira's procedure; 4) to synthesize necessary model analogues as aids to conformational evaluation of the tolanoic template.



To reach this template-peptide target AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (**I**), we explored several retrosynthetic pathways (Corey & Cheng, 1989) as shown in Scheme 1a-c. The retrosynthetic analysis we first explored involves saponification and peptide coupling of the tolanoic ester **2.A** (Scheme 1a). The dipeptidyl ester **2.A** can be obtained from the tolanoic amino ester H-Taa-OR (**2.B**), where R is an acid protecting group, such as a methyl group. However, these two steps may be problematic when the synthesis is carried out. The issues in question are the nucleophilicity of the tolanoic amino group in **2.B**, the hydrolysis of the tolanoic ester **2.A** (R=Me, for example) and the coupling of the tolanoic acid **2.A** (R=H) with an  $\alpha$  amino acid derivative. Certain anilines are difficult to acylate due to their reduced nucleophilicity. The tolanoic amino group in **2.B** is *ortho* to an electronically deficient triple bond and it may be hydrogen-bonded intramolecularly with the carbonyl oxygen across the chain. As a result, its ability of being acylated with an  $\alpha$  amino acid derivative may be drastically reduced as compared with aniline. Hydrolysis of the aryl ester **2.A** may not be a problem, but it may be a problem to obtain the tolanoic acid **2.A** (R=H) in pure form -- High Pressure Liquid Chromatography (HPLC) may be needed, which would make scaling-up of the reaction somewhat difficult. In addition, the peptide coupling step of **2.A** (R=H) with an  $\alpha$  amino acid may present a synthetic challenge. On the one hand, the tolanoic acid **2.A** (R=H) is a dipeptide and peptidyl acids usually give poor acylating yields; on the other hand, it is an aromatic acid whose acylating power is lesser than non-aromatic acids. Although these steps seemed problematic, we proceeded with the synthesis because of the need to prepare the tolanoic template ester H-Taa-OMe (**2.B**). The ester **2.B** (R=Me) may be accessed via Sonogashira's procedure from Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling of the aryl acetylene **2.C** or **2.F** with the aryl iodide **2.D** or **2.E**, respectively (Sonogashira *et al.*, 1975). However, Sonogashira's procedure has been mainly applied to relatively simple aryl acetylenes and aryl iodides, and to our best knowledge, it has never been attempted on complex systems that have functionalized substituents at *ortho*

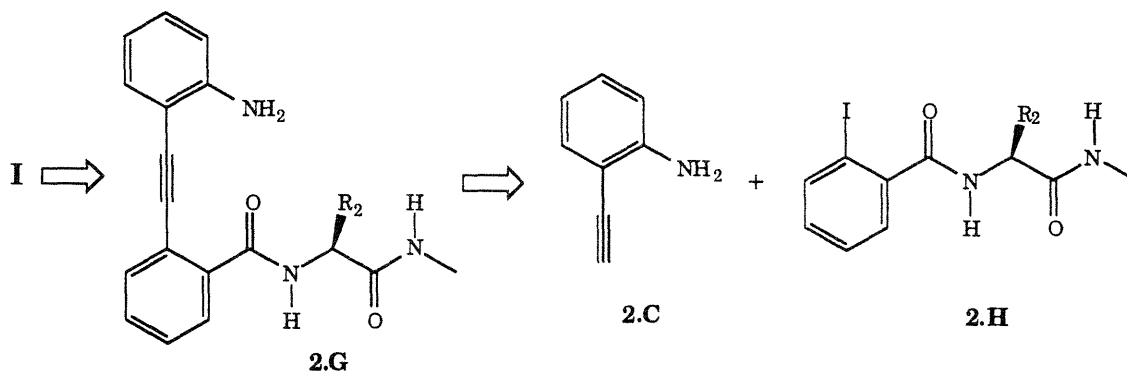
positions of both reactants. The anilinyne acetylene **2.C** is a known compound from the literature (Arcadi *et al.*, 1989). 2-Iodoaniline **2.E** is commercially available and the ester **2.D** can be easily obtained from its acid. The alkynyl benzoate ester **2.F** may be synthesized from **2.D** via Sonogashira's procedure.

### Scheme 1a. Retrosynthetic analysis.



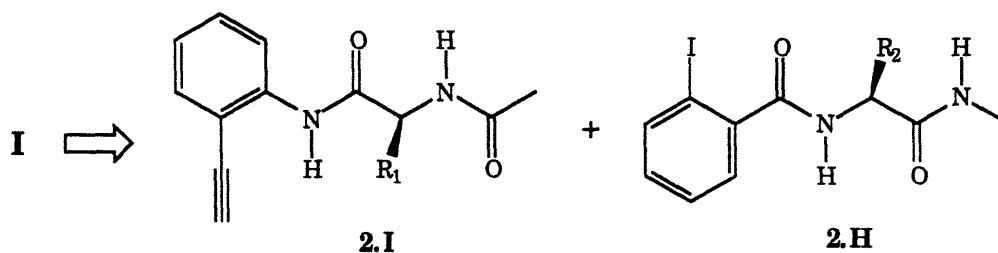
The retrosynthetic analysis we next explored starts with the acylation of the amino group of a tolanoic dipeptide H-Taa-AA<sub>2</sub>NHMe (**2.G**) followed by a coupling reaction between the aryl acetylene **2.C** and an aryl iodide **2.H** already functionalized with an  $\alpha$  amino acid derivative (Scheme 1b). Although this pathway still has to cope with the acylation of a tolanoic amino group, it is highly attractive because it is more convergent than Scheme 1a. In addition, because it is already derivatized with the required  $\alpha$  amino acid at site-2, it avoids the potentially problematic tolanoic acid AcAA<sub>1</sub>-Taa-OH (**2.A**, R=H) that has been a concern in Scheme 1a. Another attractive feature of this scheme is that a large number of template-peptide derivatives may be obtained from **2.G**, therefore it may greatly facilitate the conformational studies of these template-peptide conjugates.

**Scheme 1b. Retrosynthetic analysis.**



Scheme 1c is appealing because it is highly convergent. Both the aryl iodide **2.H** and the aryl acetylene **2.I** have already been derivatized with the required  $\alpha$  amino acid residues before the  $\text{Pd}^0\text{-Cu}^{\text{I}}$  coupling reaction is applied. **2.I** may be obtained either from the acylation of 2-ethynylaniline **2.C** with an  $\alpha$  amino acid derivative or it may be obtained from the coupling between an aryl iodide and acetylene. This pathway is appealing also because the potential application of this coupling reaction on two peptide fragments. Peptide coupling is generally not chemically selective and usually difficult to achieve between peptide fragments. As a result, there have been growing efforts in finding chemo-selective reactions for peptide coupling (Kemp & Carey, 1993; Liu & Tam, 1994; Dawson *et al.*, 1994). The  $\text{Pd}^0\text{-Cu}^{\text{I}}$  coupling reaction is chemically very selective and can tolerate many different functional groups found in peptides. It would be desirable if a fragment of a targeted peptide or protein can be replaced by the tolanoic template without affecting the overall conformation and function of the molecule. This may have profound implication for protein and peptide designs. In any event, an investigation of this type of coupling is highly valuable.

**Scheme 1c. Retrosynthetic analysis.**

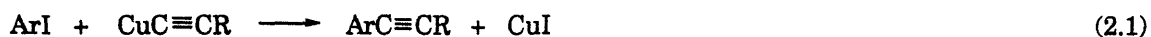


All these retrosynthetic pathways rely on a common key reaction, namely the carbon-carbon bond formation between an aromatic acetylene and an aromatic iodide. Before we proceed with the synthesis, a brief review of this type of coupling reaction will be helpful.

Following the review on the Pd<sup>0</sup>-Cu<sup>I</sup> coupling chemistry (Section 2.1) are descriptions of the synthesis of H-Taa-OMe (Scheme 1a and Section 2.2), the functionalization of H-Taa-OMe (Section 2.3), an improved synthesis of the template-peptide conjugates AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (Scheme 1b and Section 2.4), the synthesis of model compounds (Section 2.5), the synthesis of the template-peptide conjugates via the tolanoic acid H-Taa-OH (Section 2.6), and the synthesis of a cyclic template-peptide system (Section 2.7), followed by discussions on the reactivity of *o*-alkynylanilines toward acylation (Section 2.8) and on the scope of the Pd<sup>0</sup>-Cu<sup>I</sup> coupling chemistry (Scheme 1c and Section 2.9), and a conclusion (Section 2.10).

**2.1 A brief review of the coupling reaction of aryl iodides with acetylenes**

The reaction of cuprous acetylides with aryl iodides (Equation 2.1) was



first discovered in 1963 by Stephens and Castro, and henceforth has been named the Stephens-Castro reaction (Stephens & Castro, 1963). Simple tolanes are obtained in good yields by exposing aryl iodides to cuprous acetylides in refluxing pyridine (Castro & Stephens, 1963; Stephens & Castro, 1963). However, when *ortho* nucleophilic substituents such as amino and carboxyl groups are present, heterocyclic compounds, such as indoles and isocoumarins, are often obtained instead of tolanes. Even in cases for which tolanes are obtained, the yields are usually highly unpredictable (Castro *et al.*, 1969). In addition to requiring relatively harsh reaction conditions, the Stephens-Castro reaction also requires pre-formed cuprous acetylide, which is sometimes difficult to obtain in pure form. Later on, several research groups sought to find a catalytic system that would allow the formation of tolanes from phenyl iodides. Most noteworthy are the studies conducted by the research groups of Cassar, Heck, and Sonogashira. All of the catalytic reactions investigated employed palladium complexes. In addition to a Pd<sup>0</sup> catalyst, Cassar's procedure requires the use of a base such as sodium methoxide or sodium phenoxide (Cassar, 1975). In Heck's procedure, a Pd<sup>II</sup> catalyst is used and heating in a sealed reaction vessel at 100 °C is necessary. Although milder bases (excess triethylamine or piperidine is used as the solvent as well as the base) are used by Heck, the reactions do not work well with aryl iodides that bear strongly electron-donating substituents (Dieck & Heck, 1975). Shortly after Cassar and Heck published their results, Sonogashira and coworkers reported a similar procedure as Heck's, except that a co-catalyst, cuprous iodide, was employed, and the reaction could now be run successfully under a much milder condition with good results (Sonogashira *et al.*, 1975). The reaction condition employed in Sonogashira's procedure is the following: in addition to acetylene (excess) and a phenyl iodide (10 mmol), CuI (0.05 mmol), (Ph<sub>3</sub>P)<sub>2</sub>PdCl<sub>2</sub> (0.1 mmol), and diethylamine (60 ml) at room temperature for 6 hours. This procedure was later successfully applied to the

synthesis of indoles, in which simple 2-aminotolanes were isolated as intermediates (Arcadi *et al.*, 1989).

These results are indeed very encouraging for our proposed synthetic plan to prepare H-Taa-OMe (**2.B**, R=Me), H-Taa-AA<sub>2</sub>NHMe (**2.G**), and AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (**I**). The application of Sonogashira's procedure to the preparation of H-Taa-OMe will be reported in the following section, and the modification of this procedure for the preparation of even more complex systems such as H-Taa-AA<sub>2</sub>NHMe and BocAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe will be described in Section 2.4 and Section 2.9, respectively. Also in Section 2.9, the scope of the modified Sonogashira's procedure will be described as well.

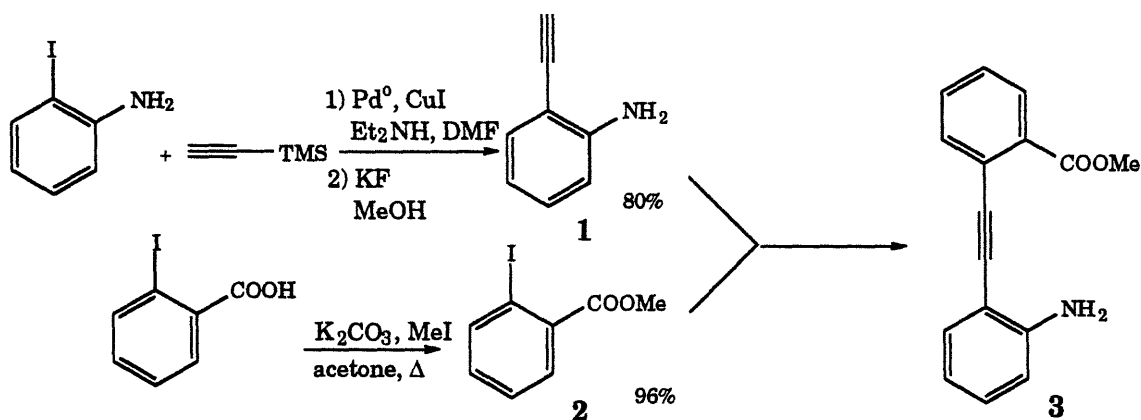
## 2.2 The synthesis of 2-amino-2'-carbomethoxytolane.

For the synthesis of the tolanoic template H-Taa-OMe (Scheme 2, **3**), we decided to follow the top pathway in Scheme 1a because 2-ethynylaniline is a known compound, and a carboxyl substituent at the *ortho* position of the aryl iodide has been found to be a beneficial coordination site for palladium and its ligand (Castro *et al.*, 1969). Methyl 2-iodobenzoate (**2**) was obtained almost quantitatively from esterification of 2-iodobenzoic acid with iodomethane. Following a published procedure (Arcadi *et al.*, 1989), 2-ethynylaniline (**1**) was synthesized in two steps by Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling reaction between trimethylsilylacetylene and 2-iodoaniline followed by desilylation of the intermediate with KF in methanol. Compound **1** is quite volatile and great care must be given when drying *in vacuo* to avoid losses.

After 2-ethynylaniline (**1**) and carbomethoxyphenyl iodide (**2**) were treated with catalysts Pd(PPh<sub>3</sub>)<sub>4</sub> and CuI in diethylamine and DMF as co-solvents at room temperature overnight, tolanoic ester **3** was obtained in 84% isolated yield (Scheme 2). Sonogashira's procedure proved to be applicable to

the preparation of H-Taa-OMe (**3**), but whether it will still be viable for even more complex substrates will be explored later in this chapter. Once the template was in hand, we began to derivatize it with peptide conjugates according to our retrosynthetic plan (Scheme 1a).

**Scheme 2. Synthesis of 2-amino-2'-carbomethoxytolane:  
the tolanoic ester H-Taa-OMe (**3**).**

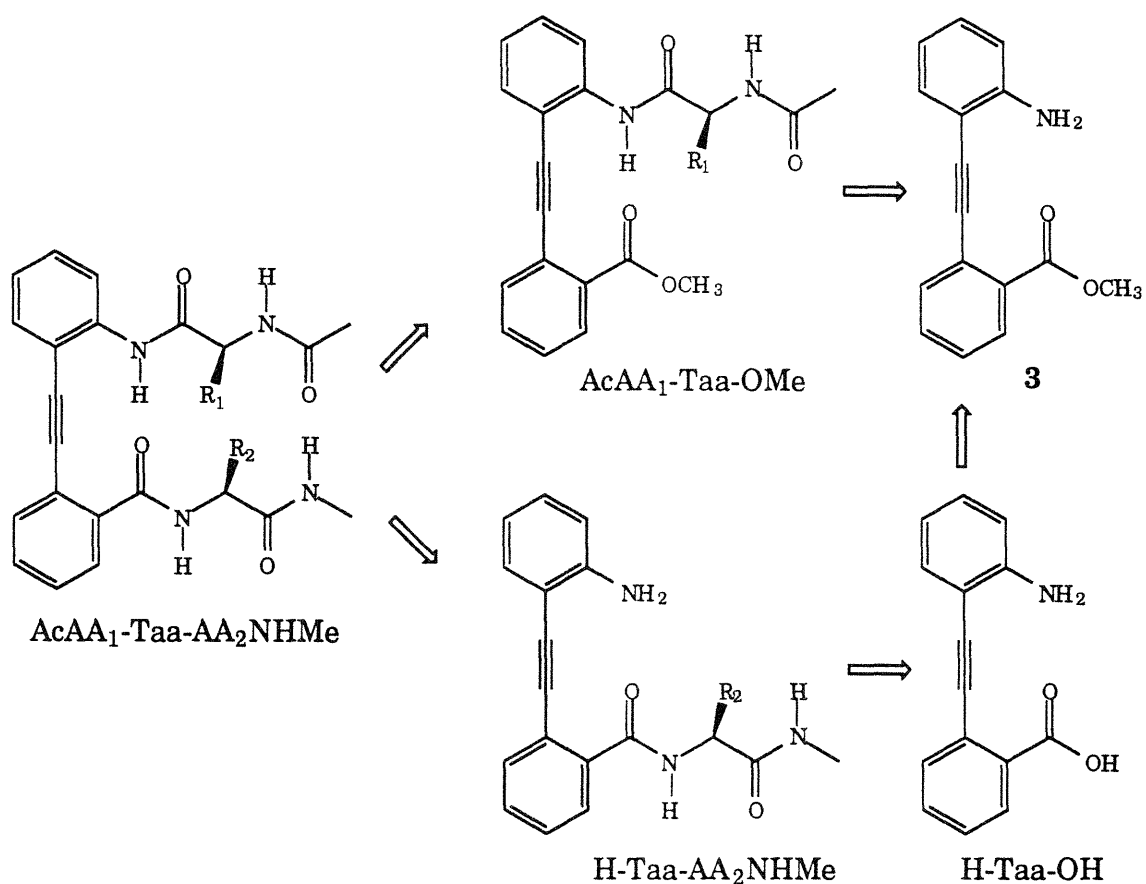


### 2.3 Functionalization of H-Taa-OMe with peptide conjugates.

There are two different ways to derivatize H-Taa-OMe with  $\alpha$  amino acid derivatives: direct acylation of **3** with an activated *N*-protected amino acid followed by hydrolysis of the intermediate acid ester AcAA<sub>1</sub>-Taa-OMe and coupling of the acid intermediate with the free amino group of an amino acid ester or amide (Scheme 3, top pathway). Alternatively, **3** can be hydrolyzed to its acid H-Taa-OH. Coupling of the acid with an amino acid derivative may afford a dipeptide intermediate H-Taa-AA<sub>2</sub>NHMe, which is then acylated with another amino acid derivative to give the target tripeptide AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (Scheme 3, lower pathway). Both methods rely on the successful acylation of an aryl amine of low nucleophilicity with an amino acid derivative.

However, the lower pathway will rest heavily on the reactivity difference between the aryl amino group of the tolanoic amino acid H-Taa-OH and that of an  $\alpha$  amino acid derivative under competitive environment. The former strategy will be explored first, and the latter will be examined later in Section 2.6.

**Scheme 3. Ways to functionalize H-Taa-OMe (3).**

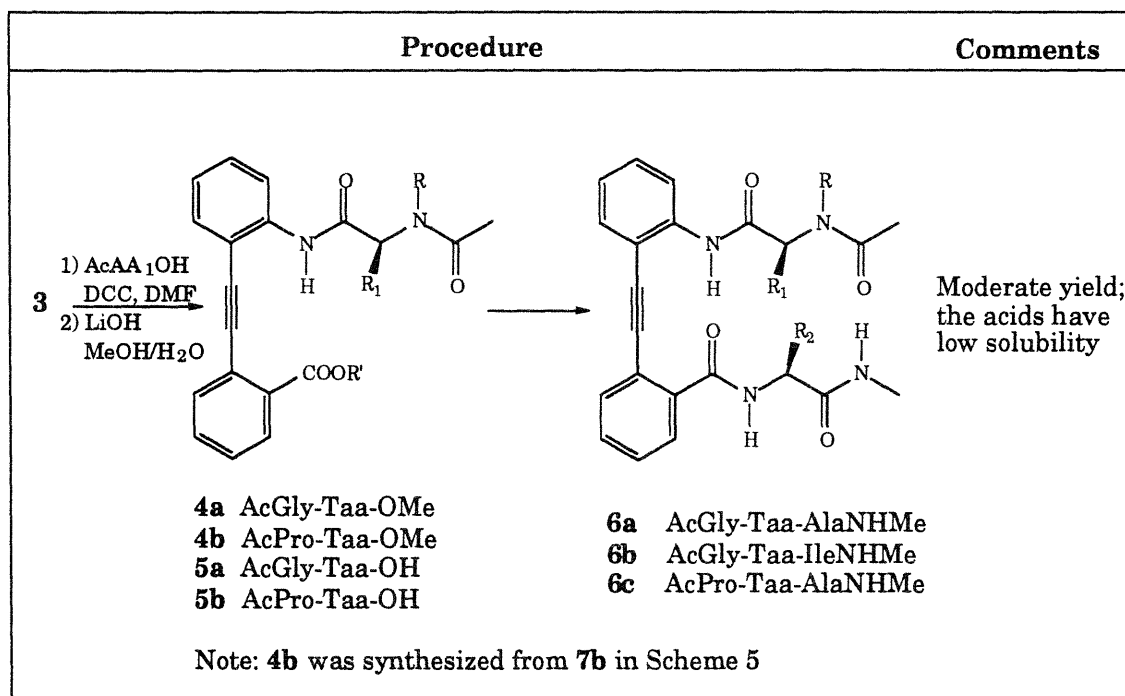




### 2.3.1 Via *N*-acetylamino acids.

For the direct derivatization of H-Taa-OMe (**3**) with  $\alpha$  amino acid derivatives, we first employed *N*-acetylamino acids directly since an acetylamide is desired as the *N*-terminus of these template-peptide conjugates, and this seems to be the shortest path to the target (Scheme 4). Treatment of H-Taa-OMe and AcGlyOH with the peptide coupling reagent dicyclohexylcarbodiimide (DCC) afforded AcGly-Taa-OMe (**4a**), although in a disappointingly low yield (Scheme 4).

#### Scheme 4. Synthesis of Ac-AA<sub>1</sub>-Taa-AA<sub>2</sub>-NHMe (**6a-c**) via AcAA<sub>1</sub>OH.



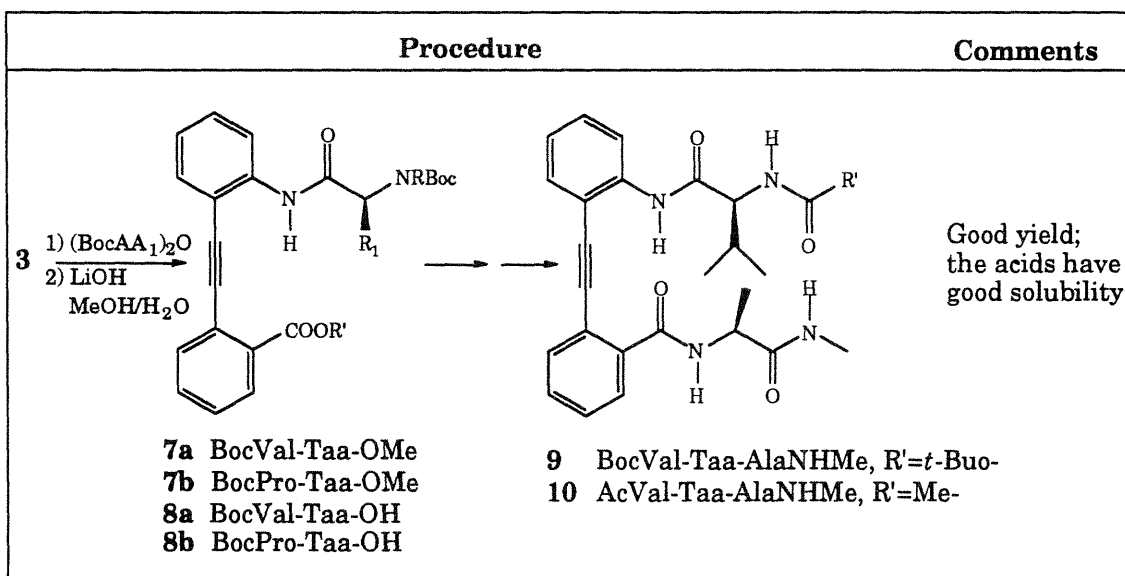
The intermediate ester AcGly-Taa-OMe (**4a**) was hydrolyzed to the acid AcGly-Taa-OH (**5a**). Coupling of the free acid AcGly-Taa-OH (**5a**) with  $\alpha$  amino acid derivatives H-Ala-NHMe and H-Ile-NHMe in the presence of 1-hydroxybenzotriazole (HOBt) and diisopropylcarbodiimide (DIC) yielded the

desired final products AcGly-Taa-AlaNHMe (**6a**) and AcGly-Taa-IleNHMe (**6b**), again, in low yields (28% and 32%, respectively, in two steps). However, the coupling of AcPro-Taa-OH (**5b**), which was obtained by a separate route, with H-Ala-NHMe under the same condition afforded the pure AcPro-Taa-AlaNHMe (**6c**) in a 73% overall yield (in two steps) after purification.

### 2.3.2 Via *N*-(*tert*-butoxy)carbonylamino acids.

In an attempt to find a more practical route to peptide-functionalized template conjugates, *t*-butoxycarbonylamino acids were used to explore the use of symmetrical anhydrides in the coupling with H-Taa-OMe **3** (Scheme 5).

**Scheme 5. Synthesis of AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (10) via BocAA<sub>1</sub>OH.**



As shown in Scheme 5, for example, the symmetrical anhydride (BocValO)<sub>2</sub>O was pre-formed with DCC in dichloromethane at 0 °C and was treated with H-Taa-OMe (**3**) to give the desired BocVal-Taa-OMe (**7a**) in 83% yield. Alternatively, the symmetrical anhydrides can be formed *in situ* and still

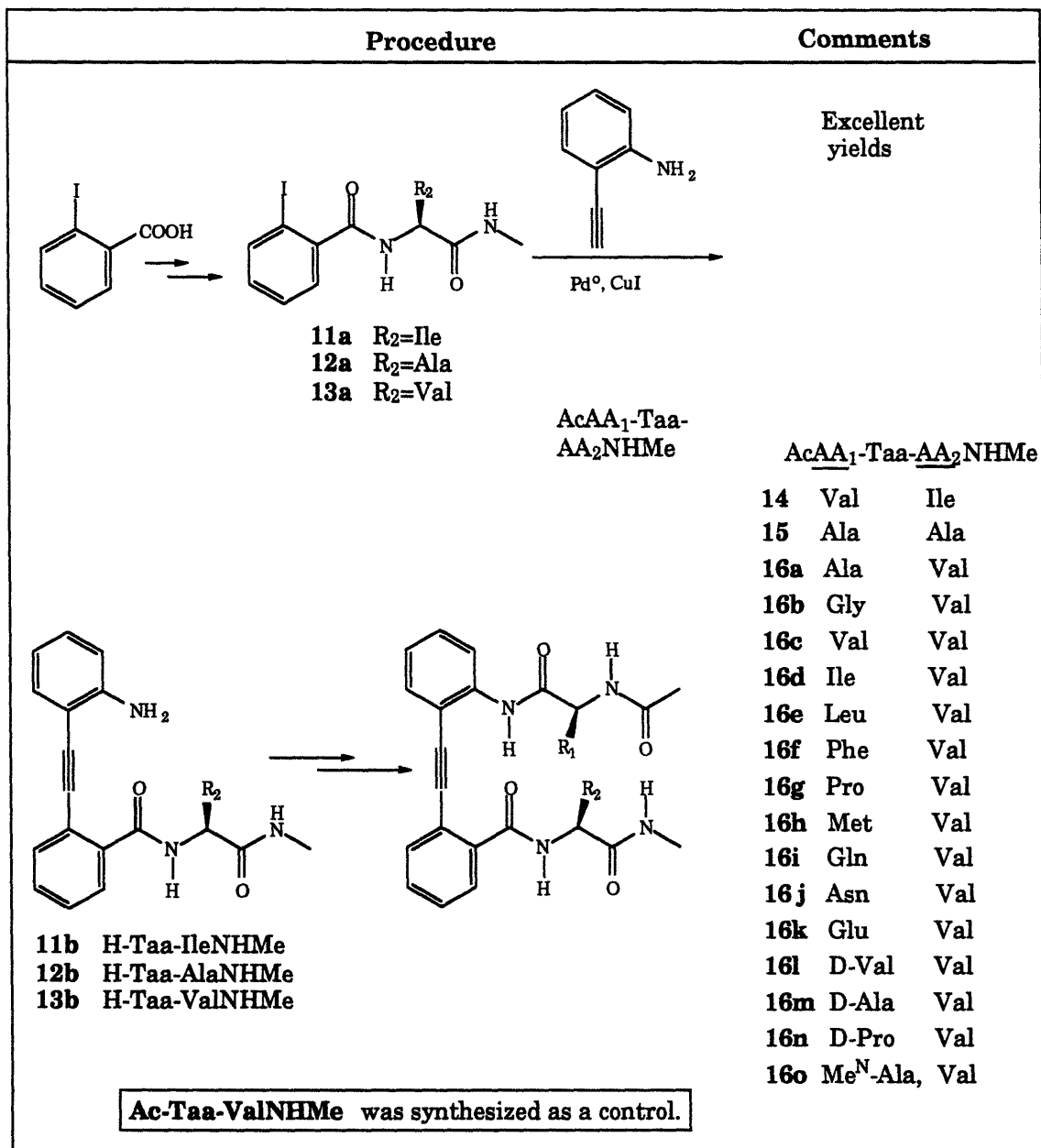
give the desired coupling in good yields. For the *in situ* formation of symmetrical anhydrides, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) was often used with good results (for the advantage of using the water soluble carbodiimide EDC as a peptide coupling reagent, see Sheehan *et al.*, 1961). In this procedure, for example, a solution of **3** and BocProOH in dry THF at 0 °C was treated with EDC to give 93% yield of the desired BocPro-Taa-OMe (**7b**) after the reaction was allowed to go overnight at ambient temperature.

Saponification of **7a** followed by coupling with H-AlaNHMe afforded BocVal-Taa-AlaNHMe (**9**) in a 43% yield. Acetylation with acetic anhydride following deprotection of **9** with trifluoroacetic acid in dichloromethane yielded 81% of the final product AcVal-Taa-AlaNHMe (**10**). The sequence of saponification followed by coupling with an  $\alpha$  amino acid methylamide is still problematic (as exemplified by the relatively low yield for **9**), and we would like to explore alternative ways to functionalize the tolanoic moiety with peptide conjugates.

#### **2.4 An alternative approach to the functionalization of the tolanoic template with peptide conjugates.**

An alternative approach is clearly needed for the rapid derivatization of the tolanoic template with  $\alpha$  amino acid derivatives. The route we eventually applied to most template-peptide derivatives with great efficiency was outlined at the beginning of this chapter (Scheme 1b). The key feature of this pathway is that it avoids the problematic saponification-followed-by-coupling sequence, and makes it possible to synthesize, in a short time, a large number of different template-peptide conjugates that share the same  $\alpha$  amino acid residue at site-2.

**Scheme 6. An improved general approach to the synthesis of template-peptide conjugates: AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe.**



As shown in Scheme 1b and outlined in Scheme 6, the synthetic plan we employed for this purpose relies on the success of Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling

between an aryl iodide, functionalized with an  $\alpha$  amino acid derivative, and 2-aminophenyl acetylene (**1**). As shown earlier, Sonogashira's procedure worked in the case of **1** and 2-iodobenzoic acid methyl ester (**2**). The same conditions were explored here in the hope that more complex tolanic systems would be attained.

The free amino groups of  $\alpha$  amino acid methyl amides HCl•H-IleNHMe, HCl•H-AlaNHMe, or HOBt•H-ValNHMe, were acylated with the 2-iodobenzoyl group, either by using 2-iodobenzoyl chloride (cases **11a** and **12a**), or by applying the symmetrical anhydride of 2-iodobenzoic acid, pre-formed with DCC in THF (case **13a**).

Coupling of the amino-acid-derivatized aryl iodide **11a** with 2-aminophenyl acetylene (**1**) under Sonogashira's condition at room temperature was only completed after 5 days. Crude H-Taa-IleNHMe (**11b**) was used directly in the following step.

During some of these Pd<sup>0</sup>-Cu<sup>I</sup> coupling reactions following Sonogashira's procedure and using a constant flow of nitrogen or argon gas, oxidative coupling products -- diacetylenes -- were recovered at the end of these reactions, which made it necessary to use excess alkynyl starting materials.

The lengthened reaction time and increased side products are probably due to the reaction conditions used and the complexity of functionalization in these aryl iodide reagents (**11a-13a**). We set out to explore ways to shorten the reaction time and to increase yields. The best condition found for these coupling reactions is under an inert atmosphere in a thick-walled Pyrex reaction tube with a self-sealing Teflon screw cap (referred to as the "sealed tube" condition hereon). We also discovered that purer products and higher yields are obtained in a shorter reaction time if diethylamine is replaced with triethylamine and the sealed reaction tube is heated below 100 °C in an oil bath

for a few hours. By running the reactions in a sealed tube, side reactions such as oxidative homo-coupling of the acetylene substrate were reduced. By controlled heating, the desired coupling proceeded faster with little side reactions. A general procedure follows: after specified quantities of aryl iodide (1.0 equiv.), aryl acetylene (1.2 equiv.), Pd(PPh<sub>3</sub>)<sub>4</sub> (1 mole %) and CuI (3 mole%) in excess triethylamine are gently heated below 100 °C in a sealed tube under argon for about 6 hours, the crude tolanoic amino intermediates (**12b** and **13b** in Scheme 6) obtained are generally pure enough to be used in the next step without further purification. Alternatively, they can be further purified by flash column chromatography. H-Taa-ValNHMe (**13b**) was obtained almost quantitatively as a light yellow powder, and was used directly without further purification. H-Taa-AlaNHMe (**12b**) was further purified by flash chromatography to give a 85% isolated yield.

The tolanoic amino intermediates H-Taa-AA<sub>2</sub>NHMe (**11b-13b**) were then acylated with *t*-butoxycarbonylamino acids via their symmetrical anhydrides either with DCC or with EDC to give tripeptide BocAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe, which were usually used directly in the next step after standard workup but a few of which were purified by flash chromatography. Deprotection by using equal volume amounts of trifluoroacetic acid and dichloromethane followed by acetylation of these tripeptidyl TFA salts with either acetic anhydride or acetyl chloride (labeled with carbon-13 at the carbonyl carbon) afforded the final products **14-16**. After they were purified by reversed phase preparative HPLC or by a combination of HPLC and flash chromatography, excellent overall yields (80-100% in three steps) were obtained for most derivatives.

## 2.5 The synthesis of model compounds.

In order to understand the conformational properties of these template-peptide conjugates, various model compounds that can serve as controls are desired. In this section, syntheses of these model compounds will be described.

### 2.5.1 The synthesis of AcAib-Taa-AibNHMe.

As part of our effort to elucidate the conformational properties of these template-peptide conjugates, we want to synthesize a model compound such as AcAib-Taa-AibNHMe (**21**), in which the bulky  $\alpha$ -aminoisobutyric acid (Aib) is placed at both sites of the tolanoic template. With reduced reactivity of the tolanoic amino group, however, this very steric bulkiness of Aib may also present a synthetic challenge, since *t*-butoxycarbonylamino acid derivatives with  $\alpha,\alpha$ -dialkylation are known to be extremely difficult to incorporate into peptides (Frerot *et al.*, 1991; Spencer *et al.*, 1992). Coste and co-workers found that PyBOP<sup>®</sup> was the best reagent for the coupling of Aib derivative with the free amino group of amino acids to form peptides (Frerot *et al.*, 1991). PyBOP<sup>®</sup>, which stands for benzotriazolyl-oxy-tris[pyrrolidino]-phosphonium hexafluorophosphate, is a reagent for introducing the activated HOBT ester of  $\alpha$  amino acids in peptide coupling reactions with very high yields (Coste *et al.*, 1990). However, PyBOP<sup>®</sup> failed to achieve the desired coupling between the acid Boc-AibOH and the aryl amino groups of 2-ethynylaniline and iodoaniline (Table 2.1, entries 7 and 8, respectively). After prolonged reaction time (2 days), only the activated ester BocAibOBt (Figure 2.2) was isolated after purification by flash chromatography.

Figure 2.2 Structure of BocAibOBt

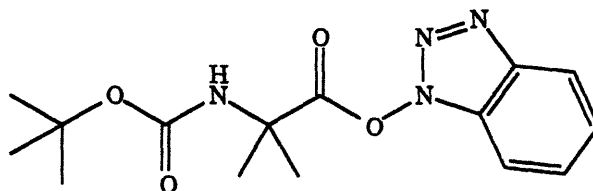
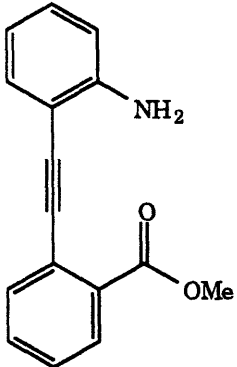
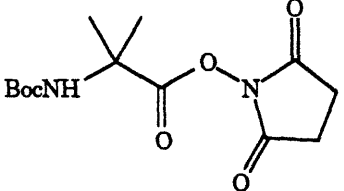
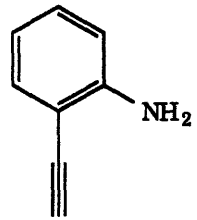
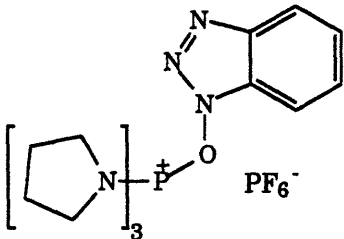
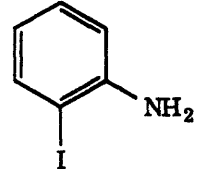


Table 2.1 Failed couplings of aryl amines with derivatives of BocAib.

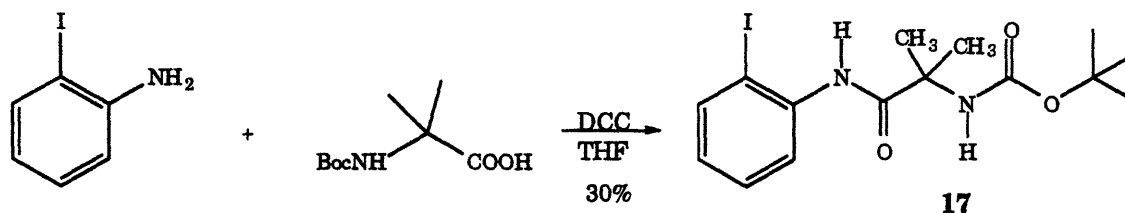
Aryl amine	Coupling method
	<p>1) (BocAib)<sub>2</sub>O</p> <p>2) BocAibOSuc</p>  <p>3) Boc AibOH + <i>N</i>-HOSuc + DCC</p> <p>4) Boc AibOH + HOBt + DCC</p> <p>5) BocAibOPdn</p>
	<p>6) BocAibOSuc</p> <p>7) BocAibOH + PyBOP</p> 
	<p>8) BocAibOH + PyBOP</p>



Under a variety of coupling conditions, BocAib was unable to be incorporated with either H-Taa-OMe (**3**) or 2-ethynylaniline or 2-iodoaniline (Table 2.1). As shown earlier, H-Taa-OMe (**3**) was acylated in high yields with symmetrical anhydrides of *t*-butoxycarbonylamino acids. However, attempts to couple **3** with the symmetrical anhydride (BocAib)<sub>2</sub>O proved futile (Table 2.1, entry 1). Neither the use of different activated esters (Table 2.1, entries 2 and 5) nor the addition of different coupling additives (Table 2.1, entries 3 and 4) yielded any better results. As mentioned earlier, even changing the nature of the aryl amine (Table 2.1, entries 6-8) or employing the hot additive PyBOP<sup>®</sup> (Table 2.1, entries 7 and 8) did not improve the situation.

Fortuitously, a desired intermediate, *N*-BocAib-(2-iodo)anilide (**17**) was obtained after 2-iodoaniline and the acid BocAibOH were treated with DCC in dry THF without any additive, such as HOSuc or HOBT, which are commonly used in peptide coupling reactions to give better yields. Thus, treatment of BocAibOH with DCC in the presence of excess 2-iodoaniline at 0 °C for 4 hours and then at ambient temperature for 20 hours afforded BocAib-(2-iodo)anilide in 30% isolated yield after workup and flash chromatography (Scheme 7).

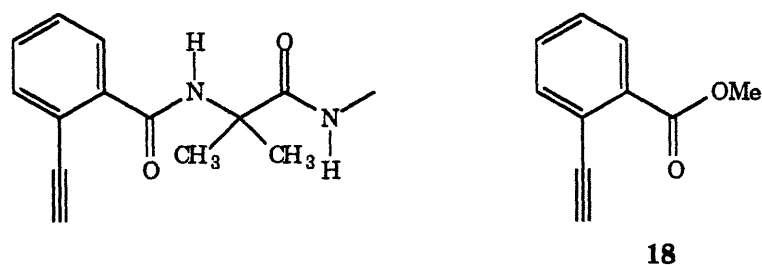
**Scheme 7. Synthesis of BocAib-(2-iodo)anilide (17).**



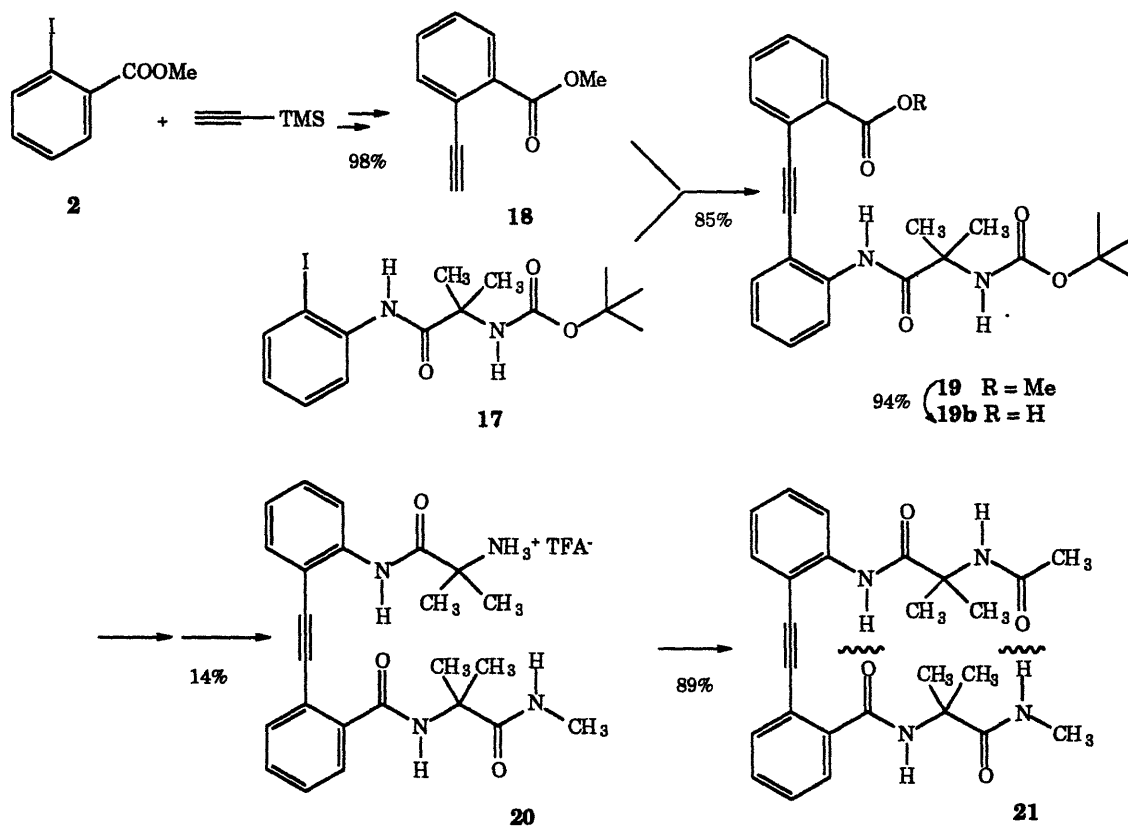
Once the aniline is acylated by BocAib, we are ready to derivatize it with an acetylene. We can either couple **17** with a fully functionalized acetylene (Figure 2.3, structure on the left), or with a simpler precursor (Figure 2.3, **18**).

Since the yield of coupling the aryl iodide **17** with a fully functionalized aryl acetylene might give lower yield compared with a simpler system, we decided to explore first the use of methyl 2-ethynylbenzoate (**18**).

Figure 2.3 Possible acetylene reactants



Scheme 8. Synthesis of AcAib-Taa-AibNHMe (**21**).



2-Carbomethoxyphenylacetylene **18** was obtained in 98% yield from 2-carbomethoxyphenyl iodide (**2**) in two steps by following Sonogashira's procedure (Scheme 8). By coupling the aryl iodide **17** with the aryl acetylene **18** under sealed tube conditions with heating, BocAib-Taa-OMe (**19**) was obtained in 85% isolated yield. Hydrolysis of the methyl ester **19** afforded 94% of the acid BocAib-Taa-OH (**19b**). Derivatization of the acid **19b** with H-AibNHMe in the presence DIC and HOBT in DMF gave the tripeptide BocAib-Taa-AibNHMe, which was deprotected with trifluoroacetic acid. The crude TFA salt (**20**) was then purified with preparative HPLC (14%, 3 steps from **19**). The final product AcAib-Taa-AibNHMe (**21**) was obtained in 89% yield after **20** was acetylated with acetic anhydride in a basic aqueous THF solution.

### 2.5.2 The synthesis of other model compounds.

In Table 2.2 are listed other model compounds that have been synthesized. Since the syntheses of these derivatives are similar to those described earlier, only a brief description will be given below.

Attempts to synthesize Ac-Taa-NHMe (**24a**) and Ac-Taa-N(Me)<sub>2</sub> (**24b**) through the coupling of Ac-Taa-OH with H<sub>2</sub>NMe or HN(Me)<sub>2</sub> were unsuccessful; Ac-Taa-OH was obtained from H-Taa-OMe by acetylation with acetic anhydride followed by saponification. We then opted for applying the Pd<sup>0</sup>-Cu<sup>I</sup> chemistry to aryl acetylene **22** and aryl iodides **23a-b** that were already derivatized with the required functional groups. The synthesis is outlined in Scheme 9. Following Sonogashira's procedure, Ac-Taa-NHMe (**24a**) and Ac-Taa-N(Me)<sub>2</sub> (**24b**) were obtained in 76% and 74% yield, respectively.

2-Carbomethoxyphenylacetylene **18** was obtained in 98% yield from 2-carbomethoxyphenyl iodide (**2**) in two steps by following Sonogashira's procedure (Scheme 8). By coupling the aryl iodide **17** with the aryl acetylene **18** under sealed tube conditions with heating, BocAib-Taa-OMe (**19**) was obtained in 85% isolated yield. Hydrolysis of the methyl ester **19** afforded 94% of the acid BocAib-Taa-OH (**19b**). Derivatization of the acid **19b** with H-AibNHMe in the presence DIC and HOBT in DMF gave the tripeptide BocAib-Taa-AibNHMe, which was deprotected with trifluoroacetic acid. The crude TFA salt (**20**) was then purified with preparative HPLC (14%, 3 steps from **19**). The final product AcAib-Taa-AibNHMe (**21**) was obtained in 89% yield after **20** was acetylated with acetic anhydride in a basic aqueous THF solution.

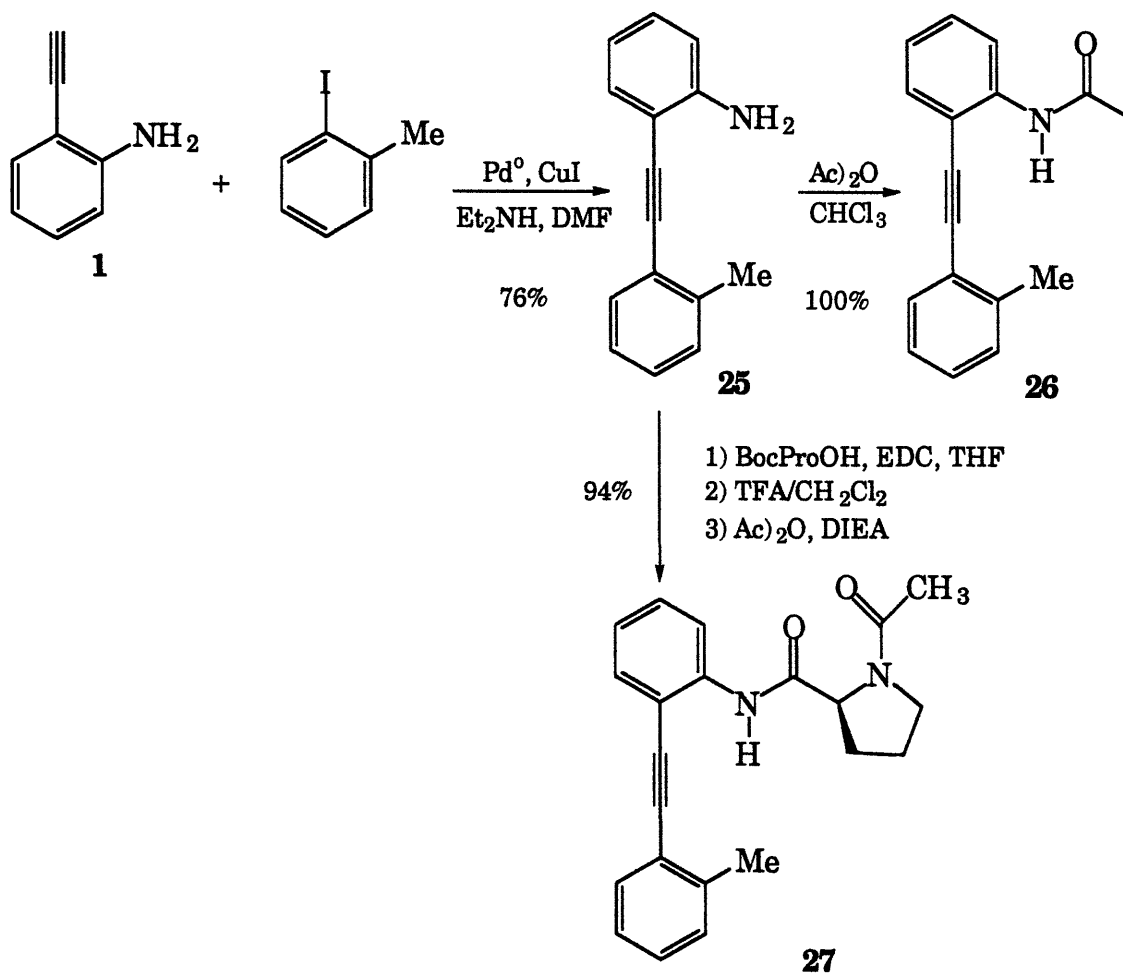
### 2.5.2 The synthesis of other model compounds.

In Table 2.2 are listed other model compounds that have been synthesized. Since the syntheses of these derivatives are similar to those described earlier, only a brief description will be given below.

Attempts to synthesize Ac-Taa-NHMe (**24a**) and Ac-Taa-N(Me)<sub>2</sub> (**24b**) through the coupling of Ac-Taa-OH with H<sub>2</sub>NMe or HN(Me)<sub>2</sub> were unsuccessful; Ac-Taa-OH was obtained from H-Taa-OMe by acetylation with acetic anhydride followed by saponification. We then opted for applying the Pd<sup>0</sup>-Cu<sup>I</sup> chemistry to aryl acetylene **22** and aryl iodides **23a-b** that were already derivatized with the required functional groups. The synthesis is outlined in Scheme 9. Following Sonogashira's procedure, Ac-Taa-NHMe (**24a**) and Ac-Taa-N(Me)<sub>2</sub> (**24b**) were obtained in 76% and 74% yield, respectively.

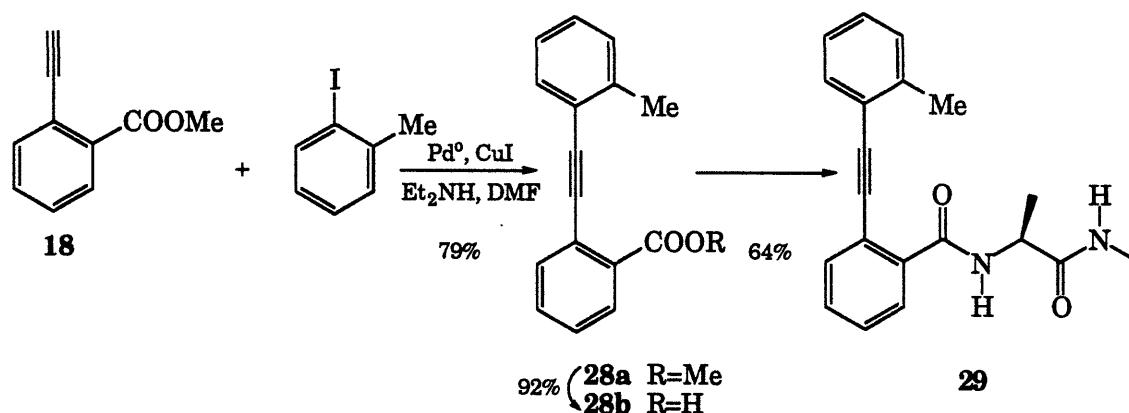
2-Amino-2'-methyltolane (**25**) was obtained in a 76% yield by applying Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling of 2-aminophenylacetylene with 2-methylphenyl iodide under Sonogashira's condition. Derivatization of **25** with acetic anhydride afforded AcNH-Tln-Me (**26**) in quantitative yield. Treatment of **25** with *t*-butyloxycarbonylproline under the water soluble carbodiimide EDC coupling condition yielded the intermediate BocProNH-Tln-Me, which was deprotected with TFA and acetylated with acetic anhydride to give the desired control compound **27** (Scheme 10).

**Scheme 10. Synthesis of AcNH-Tln-Me (26) and AcProNH-Tln-Me (27).**



Control compound Me-Tln(C=O)AlaNHMe (**29**) was obtained similarly, and its synthesis is outlined in Scheme 11. Coupling of 2-iodotoluene with methyl 2-ethynylbenzoate (**18**) under Sonogashira's condition gave rise to 2-carbomethoxy-2'-methyltolane (**28a**) in a 79% yield. Saponification of **28a** followed by peptide coupling with H-AlaNHMe in the presence of HOBt and EDC afforded the desired compound **29**.

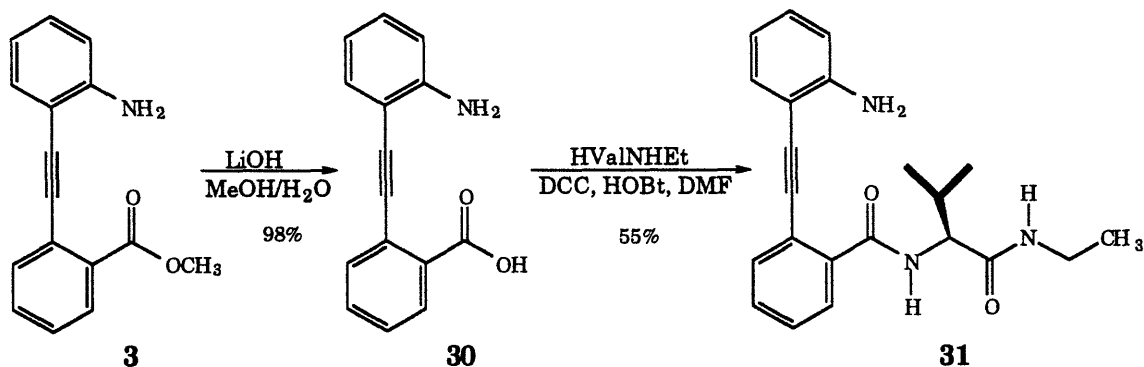
**Scheme 11. Synthesis of Me-Tln-(C=O)AlaNHMe (**29**).**



## 2.6 The synthesis of tetrapeptide analogues via H-Taa-OH.

Because of the much-reduced nucleophilicity of the tolanoic amino group, we had proposed earlier in Section 2.3 that it might be possible to couple the acid H-Taa-OH with the amino group of an  $\alpha$  amino acid without protection for the tolanoic amine (Scheme 3, bottom pathway). Here is the report on such a competitive study (Scheme 12).

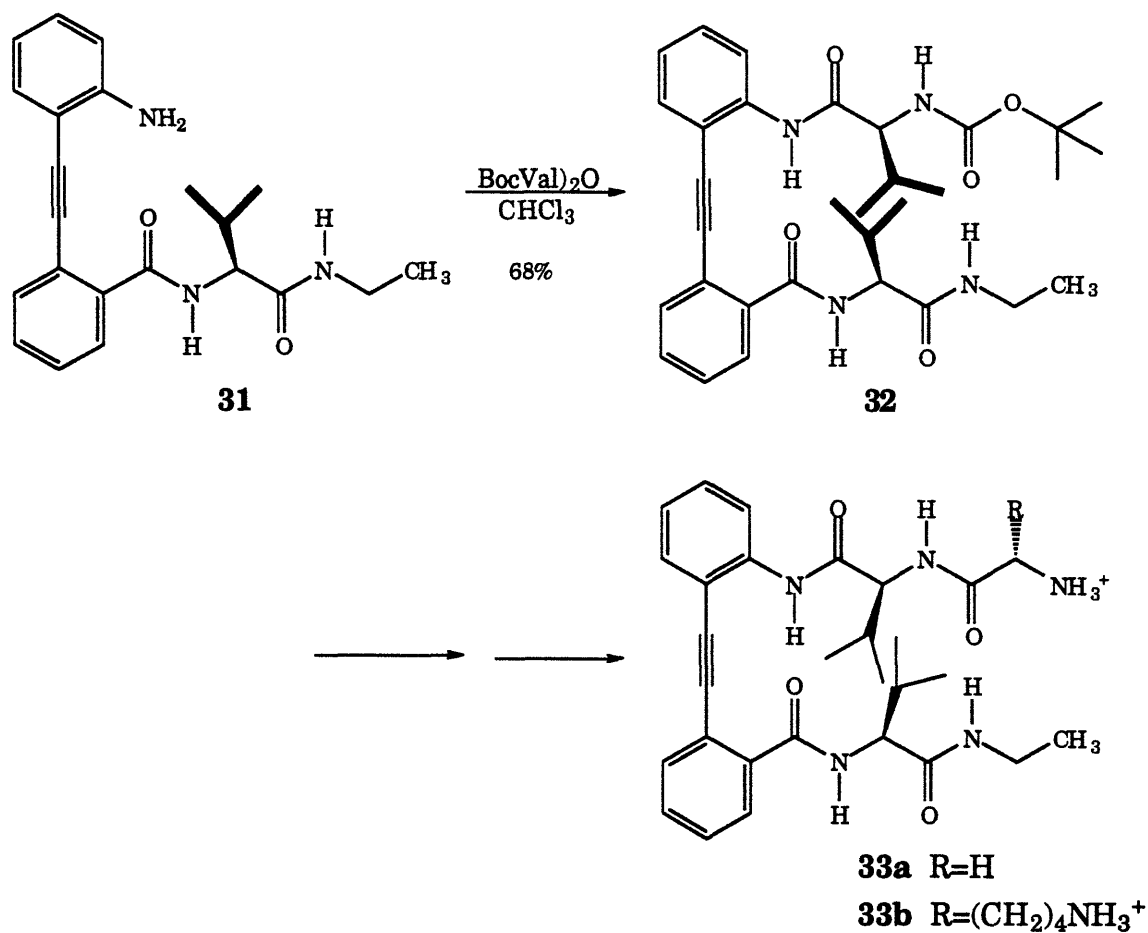
### Scheme 12. Chemistry of H-Taa-OH



As shown in Scheme 12, the tolanoic ester H-Taa-OMe (**3**) was hydrolyzed cleanly to give the free tolanoic amino acid H-Taa-OH (**30**). Coupling of **30** with 2.8 equivalent H-ValNHEt•HCl salt in the presence of 3.3 equivalent DIEA, 1.2 equivalent HOBt, and 1.1 equivalent DCC, at room temperature in DMF afforded dipeptide H-Taa-ValNHMe (**31**) in a 55% yield after purification by preparative TLC. Although this seemed like a promising route to functionalize the tolanoic template, no attempt was made to optimize reaction conditions and yields. The unique chemistry of tolanoic acid H-Taa-OH may prove useful in preparation of peptides containing the tolanoic template.

Further functionalization of **31** by methods described earlier for the preparation of template-peptide conjugates afforded tolanoic tetrapeptides shown in Scheme 13. Tetrapeptides Gly-Val-Taa-ValNHEt (**33a**) and Lys-Val-Taa-ValNHEt (**33b**) were obtained first as their respective TFA salts after they were purified by preparative HPLC, and were then ion-exchanged with chloride by aqueous HCl solution.

**Scheme 13. Synthesis of charged  
Gly-Val-Taa-ValNHEt (33a) and Lys-Val-Taa-ValNHEt (33b).**

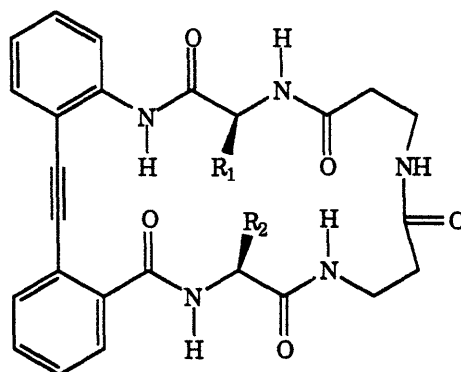


**2.7 The synthesis of a cyclic template-peptide system.**

A cyclic template-peptide conjugate of the structure **39** in Figure 2.4 is desired for comparative conformational analysis, in which the two ends are linked by two  $\beta$ -alanine residues. The synthesis of **39** is described below.



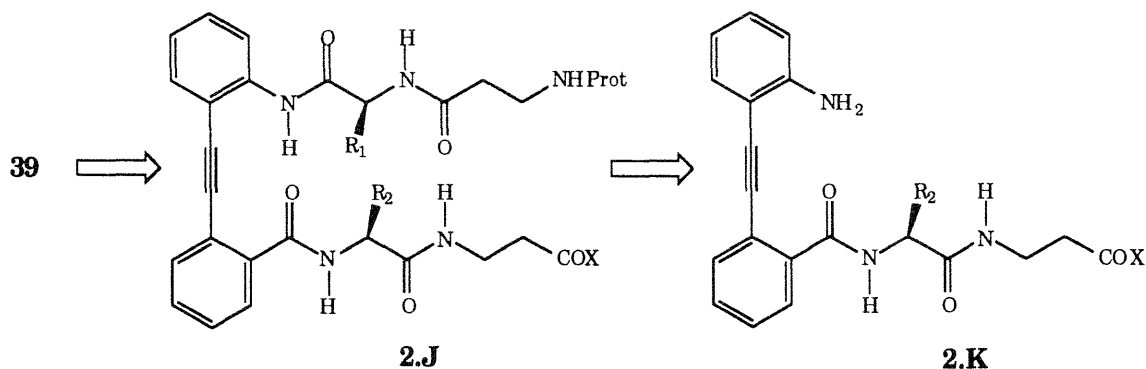
Figure 2.4 Cyclic tolanoic pentapeptide

**39**

The retrosynthetic analysis was envisioned as shown in Scheme 14. There are many possible ways to open the ring: at the triple bond carbon to the phenyl ring carbon junctions, or at any one of the peptide bonds. Although it appears attractive to use the Pd<sup>0</sup>-Cu<sup>I</sup> coupling chemistry to achieve the final ring closure, we feel compelled to open the ring at the β-alanyl to β-alanyl amide junction. If such a macrocyclization can be achieved under relatively mild condition, it may be the result of β-sheet conformation present in the penultimate pentapeptide. It will facilitate the intramolecular cyclization if the ends are held close together, but it will not if strain exists at the cyclization site. The tolanoic pentapeptide intermediate **2.J** can be obtained from **2.K** by stepwise coupling with activated Boc-protected amino acid derivatives, the chemistry of which have been worked out earlier. Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling under the sealed-tube condition, described earlier, is used to generate the requisite tolanoic tripeptide **2.K**.

The synthesis of **39** (R<sub>1</sub> = R<sub>2</sub> = -CH<sub>3</sub>), where the two center amino acid residues are both alanine, is outlined in Scheme 15.

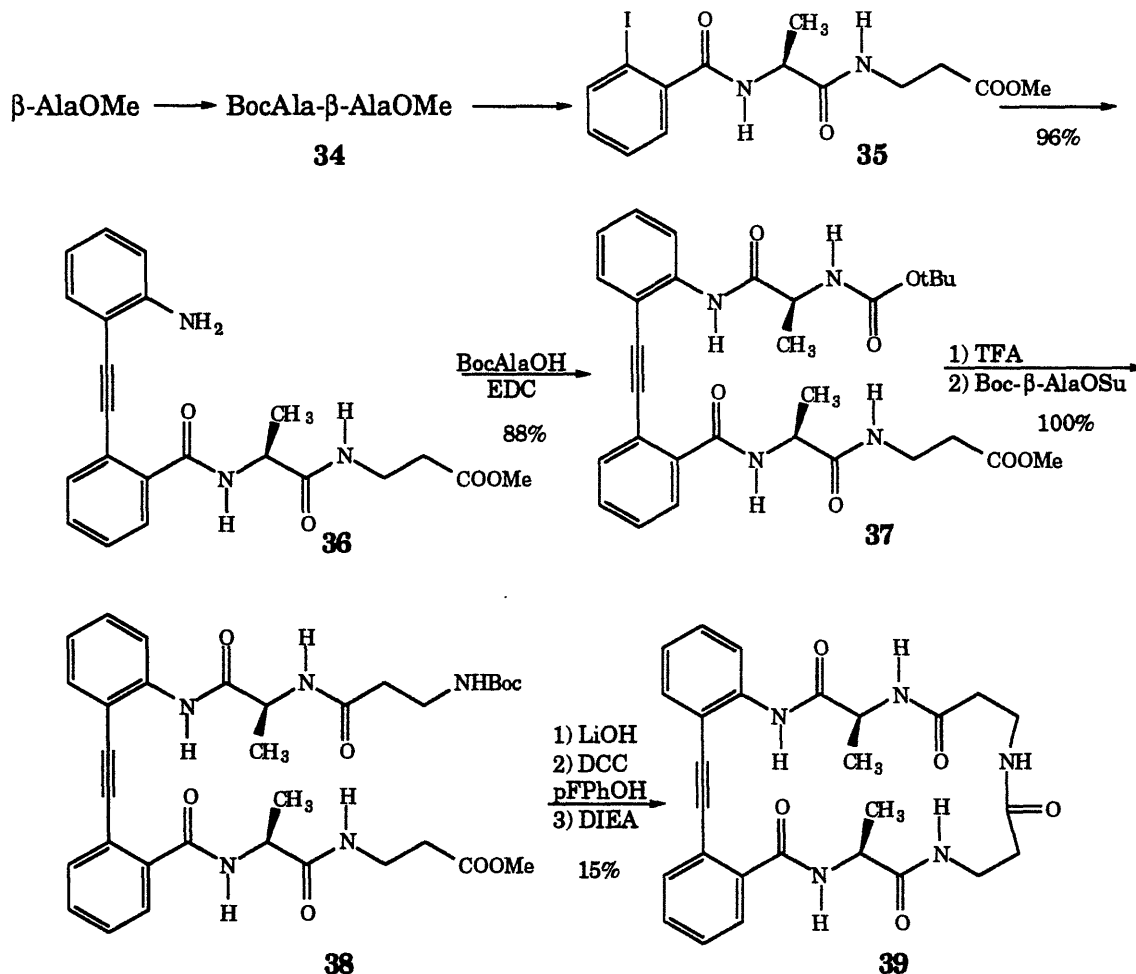
**Scheme 14. Retroanalysis of the cyclic target.**



Dipeptide BocAla- $\beta$ -AlaOMe (**34**) was easily obtained from BocAlaOH and  $\beta$ -AlaOMe hydrogen chloride salt. Deprotection of **34** with TFA followed by acylation of the intermediate with 2-iodobenzoyl chloride afforded 2-iodobenzoylalanyl- $\beta$ -alanyl methyl ester **35**. Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling of aryl iodide **35** with 2-aminophenylacetylene (**1**) under the sealed-tube condition gave the tripeptide **36** in 96% yield, which was then acylated with the symmetrical anhydride of BocAlaOH formed *in situ* to yield the tetrapeptide **37** (88%). Deprotection of **37** with TFA followed by acylation with the activated ester Boc- $\beta$ -AlaOSuc yielded quantitatively the requisite pentapeptide precursor Boc- $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -AlaOMe (**38**).

Now that all the required amino acid residues are in their respective places, a strategy for the final cyclization is required. To reduce yield loss, all subsequent steps were undertaken with only a few simple workup procedures and all intermediates were used directly without purification. Although there is a large literature on macrocyclization conditions, we chose a simple strategy: hydrolysis of the  $\beta$ -alanyl methyl ester **38** followed by activation of the acid with a suitable reagent, and intramolecular amide formation following deprotection of the *N*-terminal *t*-butoxycarbonyl protecting group.

**Scheme 15. Synthesis of the Cyclic [ $\beta$ -Ala-Ala-Tolan-Ala- $\beta$ -Ala].**

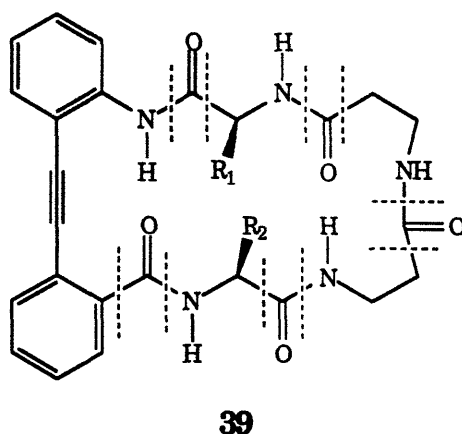


**38** was saponified to the acid quantitatively with LiOH in 3:1:1 THF/MeOH/H<sub>2</sub>O. Perfluorophenol was chosen for the active ester precursor because any excess amount used in preparing the activated ester can be evaporated due to its low boiling point. Therefore, in the presence of 1 equivalent DCC and 4 equivalent perfluorophenol, the acid Boc- $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -AlaOH was converted into its perfluorophenol ester, which was

concentrated, deprotected with TFA to yield a TFA salt of the precursor H- $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -AlaOpfP.

For the cyclization step, we chose high dilution conditions to minimize competing intermolecular coupling. The perfluorophenol ester was dissolved in dry THF, and the solution was slowly injected over a period of four hours into a mixture of DIEA and  $K_2CO_3$  in 800 ml of dry THF (final concentration 0.2 mM). Standard workup followed by multiple steps of purification by HPLC and TLC eventually afforded the final compound cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] (**39**) in 15% overall yield, which was confirmed by NMR studies, by the detection of its molecular mass by HRMS, and by its GC mass spectrum. The GC mass spectrum was noteworthy in that it shows ten well-resolved peaks containing the molecular ion 503 atomic mass, which correspond to ten different stable fragments formed from cleavage of the C-C and C-N bonds at the amide carbonyl groups, as shown in Figure 2.5.

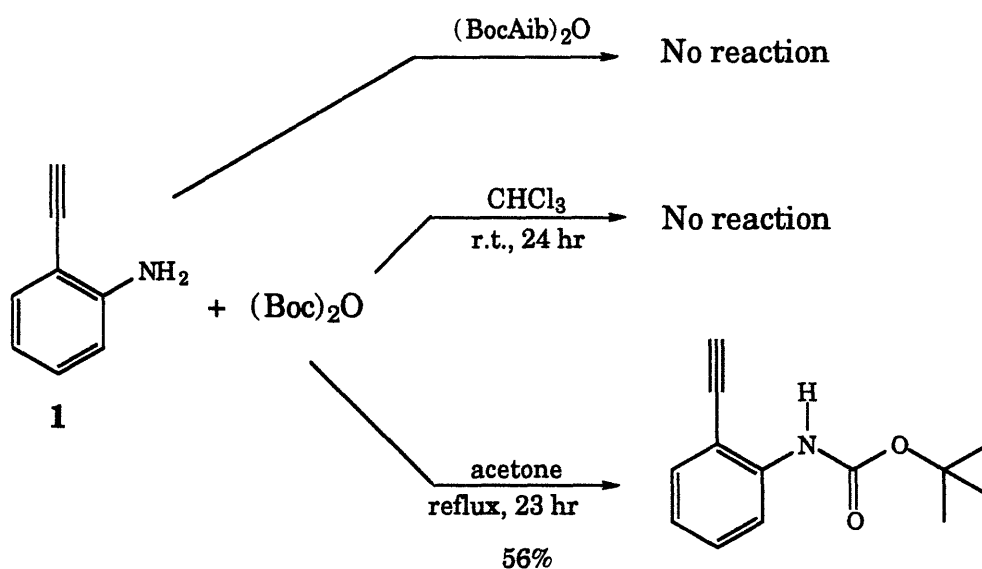
Figure 2.5 MS fragmentation pattern



## 2.8 Nucleophilicity of *o*-alkynyl-anilines toward acylation.

During the synthesis of the AcAib-Taa-AibNHMe (**21**) we were unable to couple the tolanoic amino groups of H-Taa-OMe (**3**) and 2-ethynylaniline (**1**) directly with the symmetrical anhydride of  $\alpha$ -aminoisobutyric acid (BocAib)<sub>2</sub>O. Also, in Section 2.6 we described a successful coupling between the tolanoic acid H-Taa-OH (**30**) and a valine derivative. It is evident that the reduced reactivity of the tolanoic amino group contributed to the failure in the former case and the success in the latter.

**Scheme 16. Nucleophilicity of alkynylanilines.**



Di(*t*-butyl) dicarbonate (or Boc anhydride) is a commonly used protecting reagent for amines. Protection for amino groups of  $\alpha$  amino acids by using Boc anhydride is usually complete within an hour. However, when 2-ethynylaniline (**1**) was treated with Boc anhydride in chloroform or in acetone,

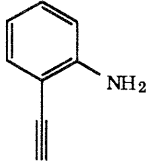
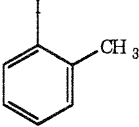
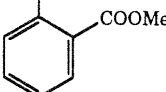
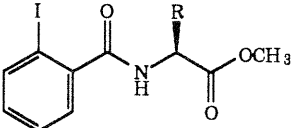
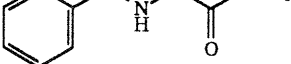
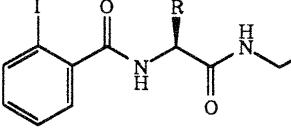
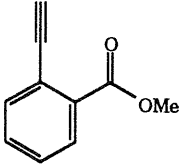
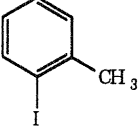
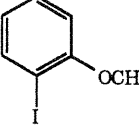
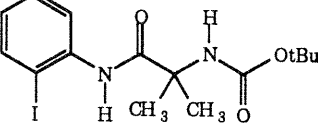
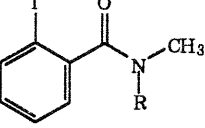
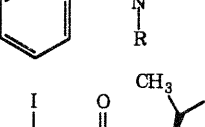
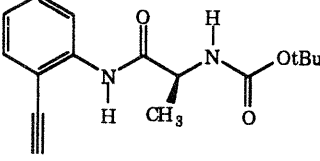
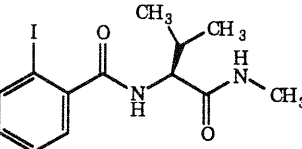
no detectable amount of the acylated product was observed after 24 hours at room temperature. Only after an extended period of refluxing in acetone was **1** eventually acylated by the Boc group, however, in a 56% yield (Scheme 16). On the other hand, acetylation of the same aryl amino group by acetic anhydride at room temperature afforded cleanly acetyl 2-ethynylanilide (**22**) as shown in Scheme 9.

## 2.9 The scope of the Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling reaction.

As mentioned earlier in this chapter, we would like to explore the scope of the coupling reaction between an aryl acetylene and an aryl iodide. Successfully linking two such substrates that are fully functionalized with amino acid or peptide derivatives, as shown in Scheme 1c, may offer a novel method in designing new protein and peptide architecture. In the previous sections, we described the synthesis of the template ester H-Taa-OMe and the template-peptide conjugates by a variety of conditions, where the powerful utility of the Pd<sup>0</sup>-Cu<sup>I</sup> coupling reaction of aryl acetylenes with aryl iodides is clearly evident. We also discovered that by modifying Sonogashira's reaction conditions, yields could be improved and reaction time could be shortened for the case in which the aryl iodide was functionalized with an  $\alpha$  amino acid derivative.

Also mentioned at the beginning of this chapter, the scope of this type of Pd<sup>0</sup>-Cu<sup>I</sup> coupling reaction has not been fully defined in the literature, and the successful coupling of two *ortho*-functionalized reactants such as **1** and **2** was unprecedented. And in Section 2.4, even more exciting and encouraging is the successful coupling result of **1** with aryl iodides that are fully functionalized with peptides such as **11a-13a** in Scheme 6. We would like to expand these findings and apply this Pd<sup>0</sup>-Cu<sup>I</sup> coupling chemistry to an even more complex system.

Table 2.3 Preparation of 2,2'-disubstituted tolanses via Pd<sup>0</sup>-Cu<sup>I</sup> catalysis.

Acetylene	Iodide	Condition	Time	Yield
		a	2 days	76%
		a	overnight	84%
		R=Ala b	6 hrs	90%
		R=Val b	7 hrs	91%
		R=Ala b	9 hrs	96%
		a	2 days	79%
		a	2 days	91%
		b	2.5 hrs	85%
		R=H a	6 days	76%
		R=Me a	4 days	74%
		b	13 hrs	63%

a: r.t., DEA, DMF; b: sealed tube, ~ 60-100 °C bath temperature, TEA, DMF.

The goal of this section is two-fold. Firstly, we want to find out to what degree of functionalizations on both aryl acetylenes and aryl iodides this Pd<sup>0</sup>-Cu<sup>I</sup> coupling reaction can tolerate. Secondly, we want to know whether or not we can apply this very chemo-selective coupling to acetylenes and iodides that are both functionalized with peptides.

In Table 2.3 are tabulated aryl acetylenes and aryl iodides that had been successfully coupled to give the corresponding 2,2'-disubstituted tolanes by applying the Pd<sup>0</sup>-Cu<sup>I</sup> methodology under either the Sonogashira's condition or the modified version.

There were many new findings in this investigation. We found that the Pd<sup>0</sup>-Cu<sup>I</sup> chemistry is quite general. Excellent yields of couplings are achieved even with the so-called difficult cases under the Heck's condition, in which the aryl iodides bear strongly electron-donating substituents. For example, *o*-methoxyphenyl iodide reacted with *o*-carbomethoxyphenyl acetylene under our modified Sonogashira's condition to give the coupled product, 2-carbomethoxy-2'-methoxytolane in a surprisingly high 91% yield (entry 7, Table 2.3). We also discovered that it is beneficial for some of these more complex systems to run this reaction in a sealed tube under argon, use excess triethylamine as a co-solvent of DMF, and apply controlled heating (< 100 °C). When the couplings are run in a sealed tube under argon with heating, the reaction time can be shortened from several days to a few hours, side reactions are kept to a minimum, and yields are usually high. For example, the coupling between 2-(*t*-butoxycarbonyl-alanyl)amidophenyl acetylene and (*o*-iodobenzoyl-valyl)methyl amide failed after many days at room temperature and under a continuous flow of argon gas, but a 63% yield of the coupled product was obtained after only 13 hours of heating under argon in a sealed reaction tube (last entry, Table 2.3). Although only one such case was tested, the moderate yield of the tripeptide BocAla-Taa-ValNHMe led us to believe that under the modified



conditions, Sonogashira's procedure indeed works well with cases in which both aryl reactants have fully functionalized peptides at *ortho* positions.

We have shown that the Pd<sup>0</sup>-Cu<sup>I</sup> catalyzed coupling reaction between aryl acetylenes and aryl iodides is quite general and can afford good to excellent yields for a variety of *ortho*-substituting functional groups under the modified condition. Its excellent chemo-selectivity and mild reaction condition offers a novel method to facilitate the coupling between two peptides.

## 2.10 Summary

We have accomplished the following objectives outlined at the beginning of this chapter. An efficient methodology was found for the preparation of the tolanoic template H-Taa-OMe and its peptide conjugates AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe. From the syntheses of these conjugates and other control compounds, we obtained a much better understanding about the unique acylation chemistry involving alkynylanilines. The scope of the Pd<sup>0</sup>-Cu<sup>I</sup> coupling chemistry was defined in terms functionalization tolerance at *ortho* positions of both aryl reactants; and by modifying Sonogashira's procedure, a variety of tolanoic peptide derivatives, including the tripeptide BocAla-Taa-ValNHMe, were obtained in good to excellent yields. After the successful synthesis of cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] (**39**), AcAib-Taa-AibNHMe (**21**), 7 control compounds (**4b**, **24a-b**, **26**, **27**, **29**, and Ac-Taa-ValNHMe), and 23 template-peptide conjugates (**6a-c**, **10**, **14**, **15**, **16a-o**, and **33a-b**), we are ready to examine the conformational properties of the new template.

## **Chapter Three**

### **Conformational Properties Studied by NMR**

Next to X-ray Crystallography, modern NMR spectroscopy is the most powerful tool for acquiring structural information on biomolecules. The purpose of this chapter is to apply this tool to the derivatives synthesized in the previous chapter. The results of NMR studies on the conformational properties of template-peptide conjugates and model compounds are described. From these results, we hope to resolve the question whether the newly synthesized template can induce  $\beta$ -sheet structure when linked to peptides.

Compared with  $\alpha$ -helices,  $\beta$ -sheets are less characterizable by NMR. The available NMR techniques commonly used for acquiring conformational information are chemical shifts, coupling constants and NOEs. Before commencing NMR studies on the tolanoic template and its peptide derivatives, a brief review of the techniques employed is in order.

### 3.1 Introduction

Nuclear Magnetic Resonance Spectroscopy (NMR) gives information about the number of each type of hydrogen atoms in a molecule, which provide information regarding the nature of the immediate environment of each of these types of hydrogen atoms. It is one of the best tools available for determining structures and conformations of a molecule.

NMR utilizes spin properties of nuclei. An atomic nucleus has a quantized spin angular momentum and a magnetic moment if either its mass or its atomic number is odd, such as  $^1_1\text{H}$  and  $^{13}_6\text{C}$ . When an appropriate

magnetic pulse is applied, spin transition (resonance) occurs and the populations of different atomic nuclear spin states are perturbed from the equilibrium. A NMR spectrum records these spin transitions at their respective resonance frequencies. These resonance frequencies are called chemical shifts,  $\delta$ , and are recorded in the units of parts per million (ppm) with respect to the carrier frequency, which is a property of the spectrometer. The chemical shifts are usually different for different types of protons depending on their magnetic environment, and the area under a resonance peak corresponds to the number of protons present at that chemical shift. There are also coupling constants (in Hz) that give information for the through-bond neighboring environment, and nuclear Overhauser effects (NOEs) that provide information about spatial interaction between different types of protons. The relaxation times  $T_1$  and  $T_2$  can shed light on the dynamics of a molecule as well. Therefore NMR can provide very rich information on molecular structures and conformations, and thus has been a preferred technique for conformational studies of peptides in solution.

Methods of determining the conformation of a molecule by NMR spectroscopy include analysis of chemical shift (such as amide NH chemical shifts) as a function of solvent or temperature, scalar couplings (including vicinal, and geminal couplings of diastereotopic methylene protons), unusual chemical shifts due to anisotropy effects, and dipolar couplings through space manifested in NOEs. These methods provide information on the conformational properties of a molecule, such as orientation of an amide proton (hydrogen bonded or exposed to the solvent), dihedral angles of the backbone and side chains, orientation of a particular functional group that gives rise to unusual anisotropy effects, and intramolecular proton to proton distances.

Solvent-exposed amide protons respond to temperature changes more emphatically than amide protons that are sequestered or intramolecularly

hydrogen bonded and the most numerous and meaningful results have been obtained for solutions in DMSO (Kopple *et al.*, 1969; Ovchinnikov & Ivanov, 1974; 1975; Ballardin *et al.*, 1978; Kessler & Kondor, 1979; Pease *et al.*, 1979; Kessler & Hölzemann, 1981). In DMSO, amide protons with temperature gradients in excess of 4 ppb/K (ppb stands for parts per billion) are generally considered evidence of external NH-orientation, while those under 3 ppb/K point to intramolecular hydrogen bonding and solvent shielding (Llinas & Klein, 1975; Bara *et al.*, 1978). In addition, solvents of different hydrogen-bonding strengths will induce very different chemical shifts on amide NH protons that are exposed to solvent. Replacing a strongly hydrogen-bonding solvent, such as DMSO, with a less polar solvent, such as chloroform or dichloromethane, will greatly shift these externally-facing amide NH protons to the higher field (lower chemical shifts), assuming no conformational change occurs.

The vicinal H-NC<sub>α</sub>-H proton coupling constant is another factor in conformational analysis of peptides. The dependence of the vicinal coupling constant, <sup>3</sup>J, on dihedral angle θ between the H-N-C<sub>α</sub> and N-C<sub>α</sub>-H planes is expressed by the Karplus equation (Karplus, 1959; 1963)

$${}^3J = A \cos^2 \theta - B \cos \theta + C. \quad (3.1)$$

Vicinal H-NC<sub>α</sub>-H proton coupling constants for peptides are found to be in the range between 0 and 10.7 Hz, and the average value is shown empirically to be related to angle θ by the following Karplus equation

$${}^3J = 9.4 \cos^2 \theta - 1.1 \cos \theta + 0.4, \quad (3.2)$$

which is plotted in Figure 3.1 (dashed line) expressed as a function of  $\theta$  and in Figure 3.2 (dashed line) as a function of  $\phi$ ,  $\theta = |\phi - 60^\circ|$  (Bystrov, 1976). This expression was later modified as

$${}^3J = 6.4\cos^2\theta - 1.4\cos\theta + 1.9, \quad (3.3)$$

and is shown in Figure 3.1 and 3.2 (solid lines). Equation 3.3 was obtained from detailed studies of the globular protein basic pancreatic trypsin inhibitor (BPTI) (Pardi *et al.*, 1984).

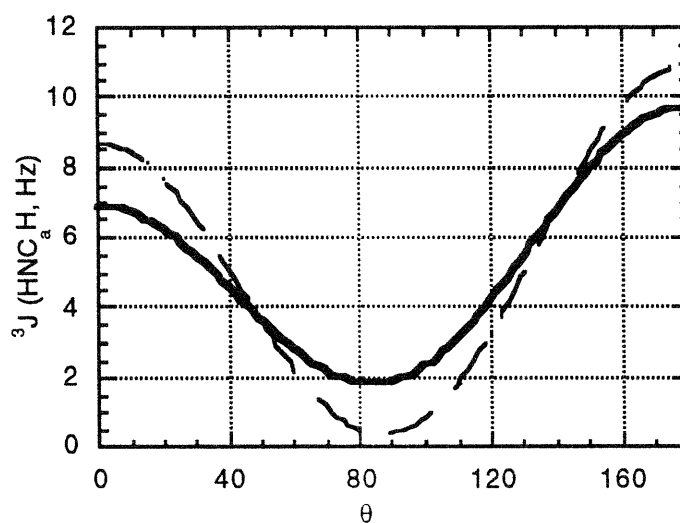


Figure 3.1 Dependence of the vicinal H-NC $_{\alpha}$ -H proton coupling constant on the dihedral angle  $\theta$ , derived from peptides (broken line, Bystrov) and BPTI (solid line, Pardi *et al.*).

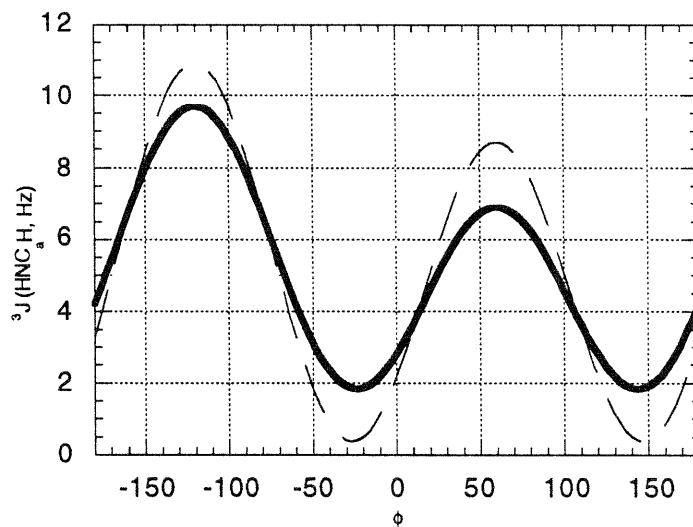


Figure 3.2 Dependence of the vicinal H-NC $_{\alpha}$ -H proton coupling constant on the dihedral angle  $\phi$ , derived from peptides (broken line, Bystrov) and BPTI (solid line, Pardi *et al.*).

However, vicinal H-NC $_{\alpha}$ -H proton coupling constants alone are not definitive in determining molecular conformations and must be used with caution. Since the vicinal H-NC $_{\alpha}$ -H proton coupling constants for  $\alpha$  amino acid residues in the  $\beta$ -sheet conformation are similar to those in an extended conformation, these values do not necessarily indicate the presence of hydrogen-bond interactions between two  $\beta$ -sheet strands.

For a glycol residue, a different Karplus function is used. It is found that the averaged value of its H-NC $_{\alpha}$ -H proton coupling constants correlates to

$${}^3J = -9.4 \cos^2 \phi - 1.1 \cos \phi + 14.9, \quad (3.4)$$

which is plotted in Figure 3.3.

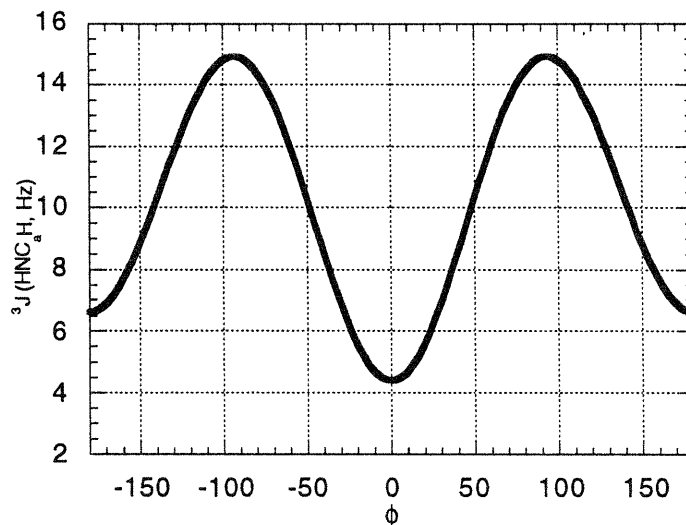


Figure 3.3 Dependence of the averaged glycylic vicinal H-NC<sub>α</sub>-H proton coupling constant on the dihedral angle  $\phi$  (Bystrov).

Vicinal H-NC<sub>α</sub>-H proton coupling constants can be measured from the multiplet splittings of the C<sub>α</sub>H signals whose components are narrower than those of the NH signals in simple peptides. The NH signals, on the other hand, are relatively broad due to the quadrupolar relaxation of the <sup>14</sup>N nucleus as well as proton exchange. As a result, the components of the NH doublet (or triplet in the case of a glycylic residue) can overlap significantly, so an appropriate correction factor may need to be applied when the NH components are used for the measurement of vicinal H-NC<sub>α</sub>-H constants (Bystrov, 1976). In DMSO, the doublets of most amide NH protons are sharp enough and no correction is necessary.



When a series of compounds of similar structures and conformations are available, very meaningful results can generally be obtained from chemical shift differences (Bowen, 1988; Arico-Muendel, 1992; Blanchard, 1992). Anisotropy effects from neighboring carbonyl groups or aromatic groups are by far the most useful source of information on molecular geometry. For example, the orientation of a proton in the plane of an adjacent carbonyl or aromatic groups often leads to considerable shift to lower field (Dale & Titlestad, 1969 and 1970; Rich & Bhatnagar, 1978), whereas location inside the cone of an electromagnetic field generated by these functional groups results in shift to higher field (Pavia *et al.*, 1979).

NOE interactions provide completely independent evidence for conformational preferences. NOEs are caused by dipolar cross-relaxation processes, and the efficiency of cross-relaxation has a distance dependence of  $r^{-6}$ , therefore there is a correlation between a NOE observed and the distance between the protons involved (Noggle & Schirmer, 1971; Macura *et al.*, 1982). NOEs can be observed only if the proton to proton distance is no greater than 3-4 Å. Therefore a NOE is observed when two protons (or proton types) are close in space. The NOE is induced by transient or steady-state perturbation of a spin, and is observed as an intensity difference or crosspeak at the affected chemical shift position by 1D difference NOE and 2D NOESY NMR (Jeener *et al.*, 1979; Kumar *et al.*, 1981; Macura *et al.*, 1981), respectively. Since a NOE results from cross-relaxation that is modulated by molecular motion, its sign and magnitude are dependent on the molecular size, temperature, solvent viscosity, and other environments. Slowly tumbling macromolecules such as proteins exhibit negative NOEs, whereas positive NOEs are observed in small molecules. However, systems of intermediate size (200-1000 amu) may not show any NOEs of measurable intensities. In such a situation, the rotating-frame NOE (ROESY) experiment (Bothner-By *et al.*,

1984; Bax & Davis, 1985; Kessler *et al.*, 1987) is employed and all cross-relaxation processes display positive crosspeaks in a ROESY plot.

For complex systems, the COSY (Aue *et al.*, 1976; Nagayama *et al.*, 1980; Bax, 1982) technique is commonly used for NMR peak assignment. For the conformational analysis of the template-peptide derivatives, COSY, NOESY and ROESY were utilized along with 1D NMR techniques, and the results are described in the following sections.

### **3.2 Solvent conditions and expectations from NMR**

Although the final choice of the solvent for any biomolecular mimetics is water, initial studies are invariably done in organic solvents such as  $\text{CDCl}_3$  and DMSO. The primary reasons are that there is more structure (less frayed) in these solvents than in water and that it is easier to apply NMR in these solvents in which amide protons are exchanged slowly on the NMR time-scale. In the structure-inducing solvents such as  $\text{CD}_2\text{Cl}_2$  and  $\text{CDCl}_3$ , all hydrogen bonding requirements are usually satisfied and the most structural information is observed in these solvents. In contrast, the strongly hydrogen-bond accepting solvent DMSO is a much better model for the aqueous environment. One drawback is that although more hydrogen bonded structures may be obtained in the same solvent, the hydrophobic interaction, which is an important driving force for structures in water, may be reduced in DMSO. In this chapter, DMSO was used as the NMR solvent for all derivatives, and in a few cases, NMR spectra were recorded in  $\text{CD}_2\text{Cl}_2$  and  $\text{CDCl}_3$  for comparison with those in DMSO.

In the  $\beta$ -sheet conformation, the tolanoic template-peptide conjugate, shown in Figure 3.4, assumes a near-planar conformation concurrent with two hydrogen bonds.

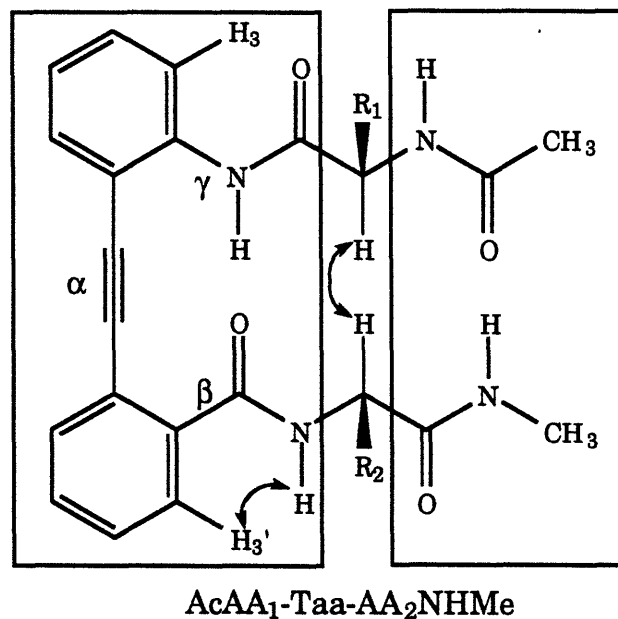


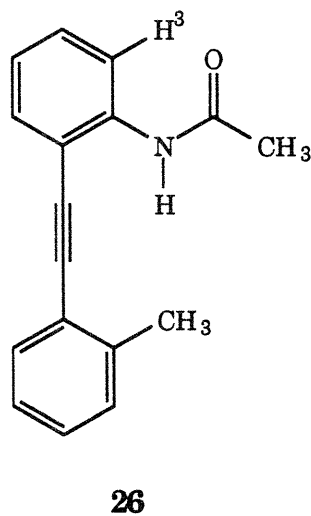
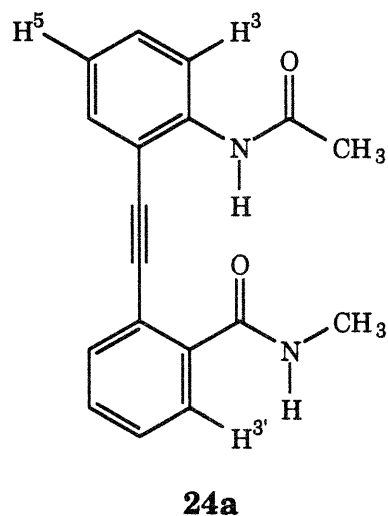
Figure 3.4 Dissection of regions of the template-peptide conjugate that are indicative of  $\beta$ -sheet conformation.

When the first hydrogen bond is formed between the tolanoic amide proton and its carbonyl oxygen (left box, Figure 3.4), the three angles  $\alpha$ ,  $\beta$ , and  $\gamma$  are fixed. As a result, we expect to see intense NOE interactions between proton H-3' and its adjacent amide proton, whereas no or very little NOE is expected for the tolanoic amide proton and proton H-3. In addition, significant changes in the chemical shifts of protons H-3 and H-3' are expected when compared with control compounds in which the hydrogen bond is weaker or

absent. With the formation of the second hydrogen bond, the two protons at the  $\alpha$  carbons of the  $\alpha$  amino acid residues are positioned close to each other (*ca.* 2.1 Å). Therefore, substantial NOE crosspeaks are expected for these two  $\alpha$ -carbon protons in the  $\beta$ -sheet conformation. The three tolanic angles  $\alpha$ ,  $\beta$  and  $\gamma$  will also be further restricted by the second hydrogen bond. In addition to these NOEs and chemical shift changes, the dihedral angles obtained from vicinal H-NC $_{\alpha}$ -H coupling constants should indicate an extended conformation for the two strands.

The results of the NMR conformational studies on the template-peptide derivatives and model compounds are reported below.

### 3.3 Ac-Taa-NHMe and AcNH-Tln-Me



Model compounds Ac-Taa-NHMe (**24a**) and AcNH-Tln-Me (**26**) were examined first. In **24a** the tolanoic template is capped with an acetyl group and a methyl amide group. By analyzing these compounds, we hope to ascertain whether the first hydrogen bond can be formed. The model compound **26** is a truncated version of **24a**, in which the methyl amide moiety is replaced by a methyl group. By examining **26** in comparison with **24a**, the mobility of the tolanoic amide group at C-2 may be determined.

In the  $^1\text{H}$  NMR spectrum for Ac-Taa-NHMe, chemical shifts for the amide NH protons and two methyl groups are easily identifiable. Chemical shift assignment of the phenyl rings was aided by using COSY NMR. There is a large downfield shift for H-3 as compared with H-5, which are 8.40 ppm and 7.10 ppm, respectively, most likely the result of a large anisotropy effect on H-3 from its proximity to the acetyl carbonyl group.

According to the results of 1D difference NOE study on Ac-Taa-NHMe, the tolanoic NH group has only a weak NOE interaction with the aryl H-3 proton at C-3 position, methylamide NH has a very strong NOE interaction with aryl H-3', and there are weak to medium interactions between the acetyl methyl and amide methyl groups as well as a weak NOE between aryl NH and N-Me across the strands. The temperature gradients of the two amide NH's are quite different: for the aryl NH, the  $\Delta\delta/\Delta T$  is -1.95 ppb/K, indicative of solvent-shielding; for the methylamide NH,  $\Delta\delta/\Delta T$  is -4.93 ppb/K, indicating that it is externally oriented and exposed to the solvent.

Molecular modeling with NOE constraints using Quanta resulted in a hydrogen-bonded structure as shown in Figure 3.5. In this conformation, however, the two phenyl planes are rotated about  $\alpha = 13^\circ$  from co-planarity, the plane containing the methylamide group is tilted from its linked phenyl

plane by about  $\beta = \pm 30^\circ$ , while on the other hand, the acetylamide plane is only slightly off co-planarity with its phenyl ring ( $\gamma = \pm 13^\circ$ ).

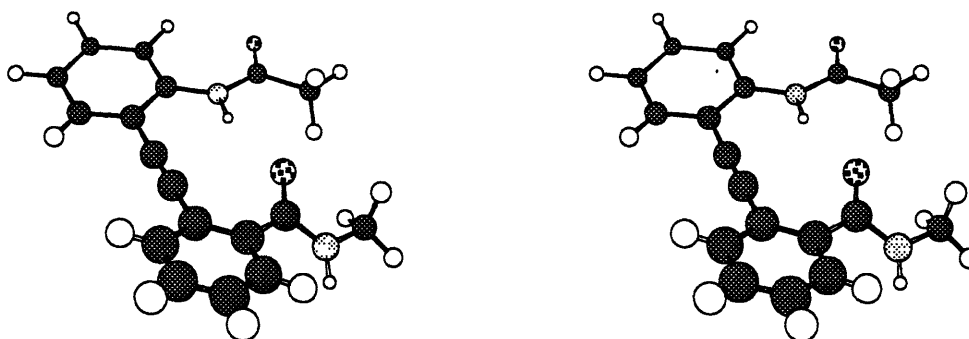


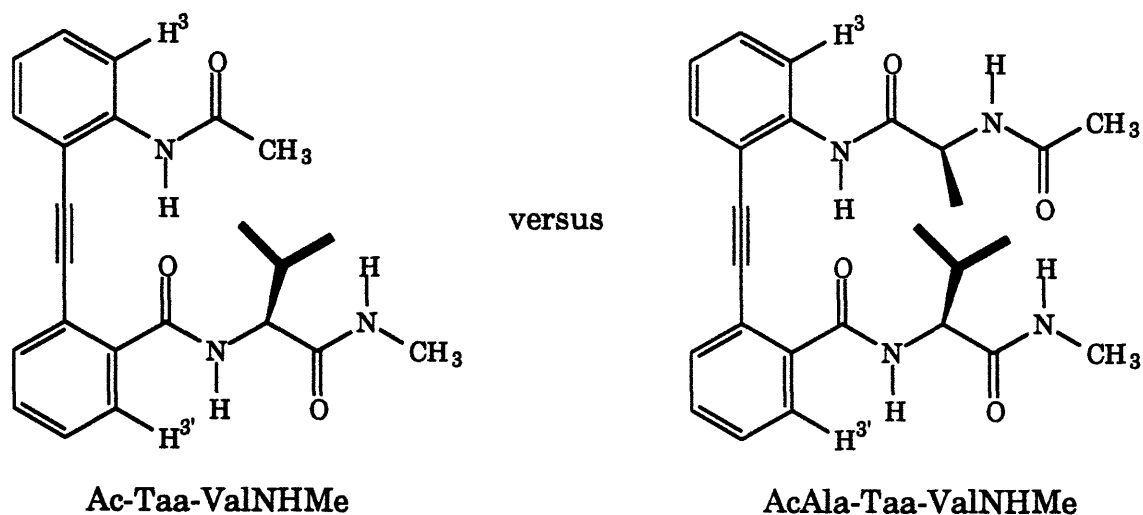
Figure 3.5 Stereo view of Ac-Taa-NHMe ( only one mirror image is shown).

Evidence for coplanarity of the acetylamide group with its connected phenyl ring also comes from the 1D difference NOE study on the model compound AcNH-Tln-Me (**26**), in which a methyl group has replaced the methylamidocarboxyl group in Ac-Taa-NHMe. Surprisingly, the NOE interaction between the NH and aromatic H-3 is also very weak, which indicates the rigidity of the planar conformation for the top half of Ac-Taa-NHMe (Figure 3.5). Favorable resonance interactions between the acetylamide group and the aromatic ring system, and weak hydrogen-bonding between the acetylamide NH and its adjacent triple bond may contribute to the overall coplanar conformation of the tolanoic amide with the phenyl ring. Similar types of weak hydrogen bonds have been reported in the literature (Prey & Berbalk, 1951; Schleyer *et al.*, 1958; Oki & Iwamura, 1959; West, 1959; Ferstandig, 1962; Allerhand & Schleyer, 1963; Yoshida & Osawa, 1965).

These results indicate that the tolanoic hydrogen bond is formed in the absence of peptide linkages, and the tolanoic amide functionality has a limited degree of rotation with respect to its attached phenyl ring. In a short summary, the tolanoic template is indeed capable of inducing the turn conformation.

### 3.4 Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe

The structures of Ac-Taa-ValNHMe and AcAla-Taa-ValNHMe are shown below. Ac-Taa-ValNHMe is truncated in comparison with AcAla-Taa-ValNHMe and cannot form the second hydrogen bond.



A comparison of the NMR spectra for Ac-Taa-ValNHMe and AcAla-Taa-ValNHMe indicates there are significant interactions between the two chains in the latter (Table 3.1). The chemical shifts of the two diastereotopic

valinyl methyl groups fall almost on top of each other ( $\Delta\delta$  0.004 ppm) for Ac-Taa-ValNHMe; as for AcAla-Taa-ValNHMe the two methyl peaks are well-separated ( $\Delta\delta$  0.051 ppm). The valinyl  $\beta$ -methine proton and the amide methyl protons in the former NMR spectrum are shifted downfield by 0.021 ppm and 0.074 ppm, respectively; the valinyl  $\alpha$  proton is modestly shifted to higher field (-0.015 ppm); a large upfield shift of -0.109 ppm for the methylamide proton is also observed.

Table 3.1 Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe in DMSO.

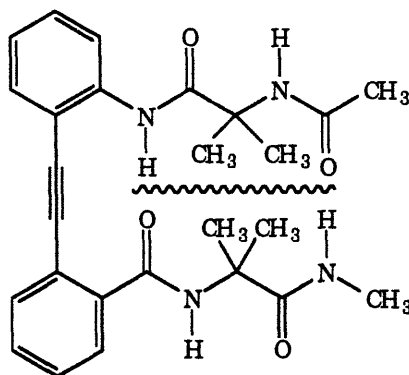
	Ac-Taa-ValNHMe	AcAla-Taa-ValNHMe
Aryl NH	9.240 ppm	9.357 ppm
H3	8.304 ppm	8.316 ppm
Val NH	8.512 ppm	8.684 ppm
MeNH	7.932 ppm	8.041 ppm
Val $\alpha$ H	4.263 ppm	4.278 ppm
Val $\beta$ H	2.067 ppm	2.046 ppm
NMe	2.557 ppm	2.483 ppm
$\Delta\delta$ (Val Me's)	0.004 ppm (0.880, 0.884)	0.051 ppm (0.914, 0.863)

These changes in chemical shifts for Ac-Taa-ValNHMe as compared with AcAla-Taa-ValNHMe support the sheet structure in which the methylamide proton is hydrogen-bonded to the acetyl carbonyl oxygen.

### 3.5 AcAib-Taa-AibNHMe

The structure of AcAib-Taa-AibNHMe (**21**) is depicted below, in which the bulky  $\alpha$ -aminoisobutyric acid (Aib) residue is linked to the tolanoic template at both sites.





21

The added steric bulkiness of the  $\alpha$ -substitution ought to deter the two Aib residues from coming together to form a  $\beta$ -sheet structure, in which the two ends, the acetyl oxygen and the methyl amide proton, can hydrogen bond. Such a compound will provide useful spectral information on the tolan moiety that is functionalized so that it cannot form a sheet structure.

Proton NMR peak assignment was aided by NOESY studies. Observed NOE interactions are tabulated in Table 3.2 as a 10 x 10 matrix. Entries S, M, and W in the matrix stand for strong, medium and weak NOE interactions, respectively. Noticeably missing are inter-chain NOE interactions in the matrix, which are observed in other template-peptide systems. Clearly these chains are no longer extended due to the disubstitution at the  $\alpha$ -carbon. The chemical shift of acetylmethyl group is shifted upfield to 1.55 ppm, whereas in other template-peptides systems it is usually between 1.8 and 2.1 ppm. This acetylmethyl chemical shift did not change much when DMSO was replaced by a less polar solvent  $\text{CDCl}_3$  (1.54 ppm). Also in DMSO, the aryl NH has a very low temperature gradient (-0.70 ppb/K), Aib NH-1 and MeNH have border-line

temperature gradients (-3.82 and -3.96 ppb/K, respectively), and only Aib NH-2 has a clearly solvent-exposed temperature gradient (-4.83 ppb/K). The conformation obtained from molecular mechanics modeling (Figure 3.6) is consistent with those NOEs observed in the NOESY spectrum and the temperature gradients obtained. In this structure, the two strands are far apart, each strand forms either a  $\gamma$ - or a helical-turn-like conformation, with the two phenyl planes almost perpendicular to each other (angle  $\alpha \approx 106^\circ$ ). The chemical shifts of the tolanoic amide proton and the methyl amide proton are shifted downfield (0.3 ppm and 0.4-0.6 ppm, respectively) compared with other template-peptide conjugates, as well as the aryl H-3 proton, which is also shifted downfield by 0.1 ppm (Table 3.3).

Table 3.2 NOEs observed for AcAib-Taa-AibNHMe in DMSO.

	AcMe	NH-1	Me-1	aryl NH	H-3	H-3'	NH-2	Me-2	MeNH
AcMe		S							
NH-1			S	M					
Me-1				M					
aryl NH									
H-3									
H-3'							M		
NH-2								S	
Me-2									S

Table 3.3 A comparison of chemical shifts (ppm) with other derivatives.

	aryl NH	MeNH	H-3
AcAib-Taa-AibNHMe	9.02	7.49	8.18
AcGly-Taa-Ala-NHMe	9.33	7.86	8.32
AcAla-Taa-ValNHMe	9.36	8.04	8.32

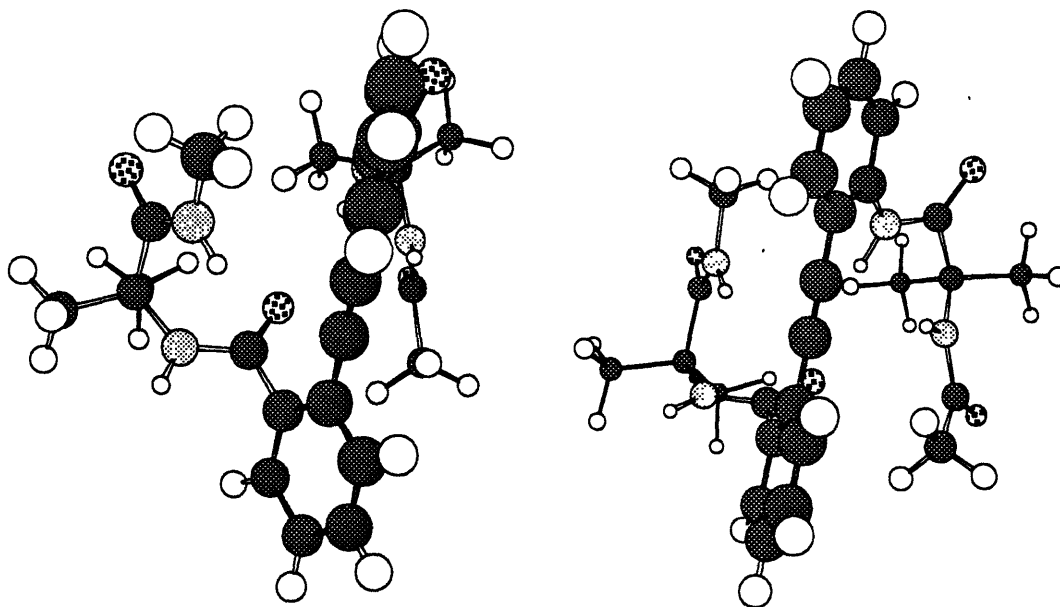
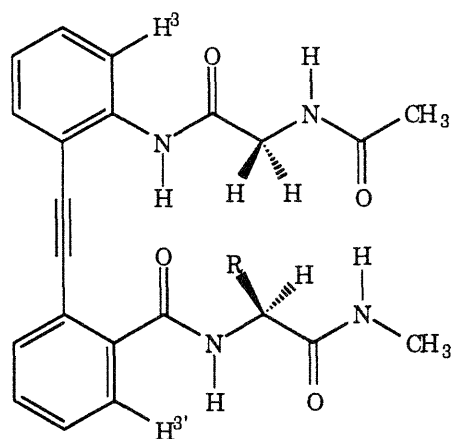


Figure 3.6 Two different views of AcAib-Taa-AibNHMe

It is clearly evident that the Aib residues not only prevent the formation of the second hydrogen bond but also destroy the tolanoic hydrogen bond.

### 3.6 AcGly-Taa-AA<sub>2</sub>NHMe

Three template-peptide derivatives AcGly-Taa-AlaNHMe (**6a**), AcGly-Taa-IleNHMe (**6b**), and AcGly-Taa-ValNHMe (**16b**), all of which contain a *N*-acetylglycyl residue at site-1, are examined. The structures of these compounds, along with selected numbering of the aromatic protons, are pictured below.



**6a** R = -CH<sub>3</sub> (Ala)

**6b** R = -CH(CH<sub>3</sub>)CH<sub>2</sub>CH<sub>3</sub> (Ile)

**16b** R = -CH(CH<sub>3</sub>)<sub>2</sub> (Val)

The results of the NOESY NMR of AcGly-Taa-AlaNHMe (**6a**) in the structure-inducing solvent CD<sub>2</sub>Cl<sub>2</sub> are tabulated in Table 3.4 as a 10 x 10 matrix. In the matrix, NOE interactions near the diagonal represent intra-strand NOE crosspeaks, and the ones that are far off diagonal are inter-strand NOE interactions. The intra-strand NOE crosspeaks indicate an extended conformation for the amino acid residues, whereas the inter-strand crosspeaks point to close proximity between proton types of different amino acid residues across the strands, and are indicative of the presence of a sheet-like structure. As shown in the matrix, all the intra-chain NOE interactions indicative of an extended conformation are present. The strong NOE between glycylic and alanyl  $\alpha$ -carbon protons, the weak NOE between glycylic  $\alpha$ -carbon protons and alanyl  $\beta$ -methyl protons and that between glycylic  $\alpha$ -carbon protons and methylamide NH are strong indications of a sheet-like conformation for AcGly-Taa-AlaNHMe.

Table 3.4 NOEs observed for AcGly-Taa-AlaNHMe in CD<sub>2</sub>Cl<sub>2</sub>.

	AcMe	Gly NH	Gly H	aryl NH	H-3	H-3'	Ala NH	Ala H	Ala Me	MeNH
AcMe		S								
Gly NH			M							
Gly H				S				S	W	W
aryl NH										
H-3										
H-3'							S			
Ala NH								W	W	
Ala H										S
Ala Me										
MeNH										

In DMSO, the NOE interactions for AcGly-Taa-AlaNHMe (Table 3.5) are very similar to those in CD<sub>2</sub>Cl<sub>2</sub>, suggesting that conformations in these two different solvents are similar. However the inter-chain NOEs are weaker, which indicate a more frayed or random structure in DMSO than that in CD<sub>2</sub>Cl<sub>2</sub>. ROESY NMR spectra of template-peptide conjugates that do not contain glyceryl residue, such as AcVal-Taa-AlaNHMe (**10**) and AcVal-Taa-IleNHMe (**14**) in DMSO show a similar pattern of NOE interactions as the one seen in the spectra of AcGly-Taa-AlaNHMe. In these ROESY spectra, all the intra-chain NOEs indicative of extended conformation for the two strands are observed, as well as significant cross-chain NOEs, such as medium to strong NOE interactions between C- $\alpha$  protons, weak NOEs between Val  $\alpha$ -carbon proton and methylamide NH, between Ala  $\alpha$ -H and Val methyl groups, in AcVal-Taa-AlaNHMe, and weak NOEs between Ile  $\alpha$ -H and Val methyl groups, between Val  $\alpha$ -H and methylamide NH, in AcVal-Taa-IleNHMe. More detailed information on these different derivatives will be described later. The

similarity in observed NOE patterns indicates that they also share a similar conformational motif with these glycyI-containing template conjugates.

Table 3.5 NOEs observed for AcGly-Taa-AlaNHMe in DMSO.

	AcMe	Gly NH	Gly H	aryl NH	H-3	H-3'	Ala NH	Ala H	Ala Me	MeNH
AcMe		S								
Gly NH			M							
Gly H				S				M	VW	W
aryl NH					W			VW		
H-3										
H-3'							S			
Ala NH								W	M	
Ala H										S
Ala Me										W
MeNH										

The very similar pattern of NOE interactions in the NOESY spectra of AcGly-Taa-AlaNHMe in DMSO and in CD<sub>2</sub>Cl<sub>2</sub> strongly supports the hypothesis that conformations in these different solvents are very similar, if not the same.

The two glycyI  $\alpha$  protons in these template-peptide conjugates exhibit non-equivalent chemical shifts, indicating the influence exerted by the amino acid residue on the opposite side of the chain. They are split into an ABX spin system, where A and B stand for glycyI  $\alpha$  protons, X stands for glycyI amide NH, and the distances between the letters indicate their closeness (or distance) in chemical shifts. These differences in chemical shifts, along with their

geminal coupling constants, are listed below for the three derivatives (Table 3.6).

Table 3.6 Chemical shift non-equivalence of glycylic methylene protons

derivative	$\Delta\delta$ of Gly $\alpha$ H's	$^2J$ (Hz)
AcGly-Taa-AlaNHMe	0.072 ppm	17.1
AcGly-Taa-ValNHMe	0.175 ppm	17.1
AcGly-Taa-IleNHMe	0.169 ppm	17.1

Chemical shift non-equivalence of the glycylic methylene protons is a strong support for a compact conformation, in which a sheet-like structure with two inter-chain hydrogen bonds is formed. The differences among the  $\Delta\delta$ 's are noteworthy, because not only do they implicate sheet structure but also indicate different degrees of influence that the amino acid residue on the opposing strand has on the glycylic methylene protons in the sheet conformation. If a simple two state equilibrium between non-sheet and sheet conformations could be assumed, the non-equivalence in  $\Delta\delta$  reflects a free energy difference of at least 0.5 kcal/mole.

A geminal proton coupling is manifested in the NMR spectra of glycylic-containing peptides when the  $C_\alpha H_2$  methylene protons are magnetically non-equivalent (ABX systems), as observed in these glycylic tolanoic derivatives. The value of the H- $C_\alpha$ -H coupling constant ranges from 14.5 to 19 Hz (Dale & Titlestad, 1970; Titlestad *et al.*, 1973) with an average value of about 15 Hz (Bystrov, 1976). For the three tolanoic derivatives containing glycylic residue the geminal H- $C_\alpha$ -H coupling constant is 17.1 Hz (AcGly-Taa-AlaNHMe, AcGly-Taa-ValNHMe and AcGly-Taa-Ile-NHMe), which is well above the average value. The geminal H- $C_\alpha$ -H coupling is believed to arise from different orientations of the  $\pi$ -electron orbital of the glycylic carbonyl bond (Siemion,

1971; Titlestad *et al.*, 1973). According to a prescribed relationship between the H-C<sub>α</sub>-H coupling constant and  $\psi$  angle, the glycylic  $\psi$  angle in these tolanoic derivatives was calculated to be around either  $\pm 20$  or  $\pm 165^\circ$ , the latter value is consistent with the molecular modeling results in which  $\psi$  was calculated to be  $\pm 175^\circ$ . The “+” and “-” signs correlate with different sheet conformations in these tolanoic derivatives according to the molecular mechanics modeling.

The amide NH temperature gradients of chemical shifts ( $\Delta\delta/\Delta T$ 's) indicate that the aryl amide NH (- 2.1 ppb/K) is strongly solvent-shielded and the other amides are exposed to DMSO. The chemical shifts for all four amide NH's correlate strictly in a linear relationship with temperature (Figure 3.7). When DMSO is replaced by CD<sub>2</sub>Cl<sub>2</sub> or CDCl<sub>3</sub>, the glycylic and alanyl amide NH's are shifted upfield by about 1.5 ppm, whereas the methanamide NH and aryl amide NH are shifted, in contradistinction, to lower field by 0.03-0.04 ppm and 0.3-0.4 ppm, respectively (Table 3.7). These findings are strong evidence for a sheet-like structure, in which the aryl amide NH and methanamide NH are intramolecularly hydrogen-bonded, whereas the two amino acid amide NH's are exposed to the solvent.

Table 3.7 Chemical shifts of AcGly-Taa-AlaNHMe in different solvents

	aryl NH	Ala NH	Gly NH	MeNH	H-3
in DMSO	9.328	8.783	8.292	7.856	7.783
in CD <sub>2</sub> Cl <sub>2</sub>	9.660	7.273	6.772	7.884	8.606
in CDCl <sub>3</sub>	9.729	7.261	6.751	7.918	8.604



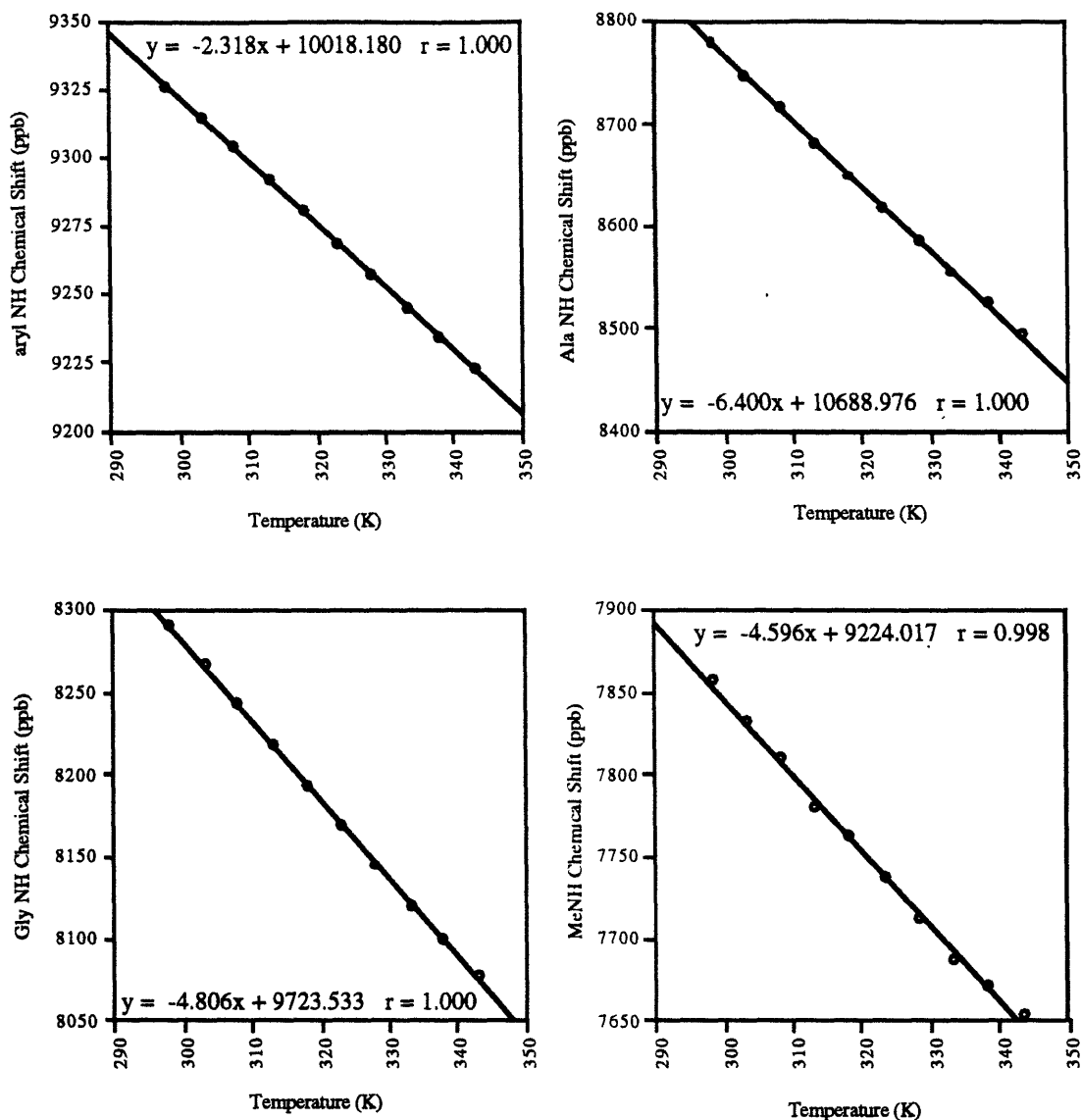


Figure 3.7 Chemical shift versus temperature for AcGly-Taa-AlaNHMe

Molecular mechanics modeling resulted in two distinctively different  $\beta$ -sheet conformations for these derivatives. In the first sheet conformation, the interplanar angle  $\alpha$  between the two phenyl planes is negative, with the aromatic carbonyl group being angled above its phenyl ring as shown in Figure

3.8. In the second conformation, signs of these angles are reversed. The two conformations are very similar energetically with the “+” conformer slightly favored over the “-” conformer. The  $\phi$  values of the valinyl residue in the molecular modeling results, for example, are similar for the two conformations ( $\phi = -159^\circ$  for “-” and  $-151^\circ$  for “+” conformers) and are consistent with the value calculated from the vicinal H-NC <sub>$\alpha$</sub> -H coupling constant (around  $-150^\circ$ ).

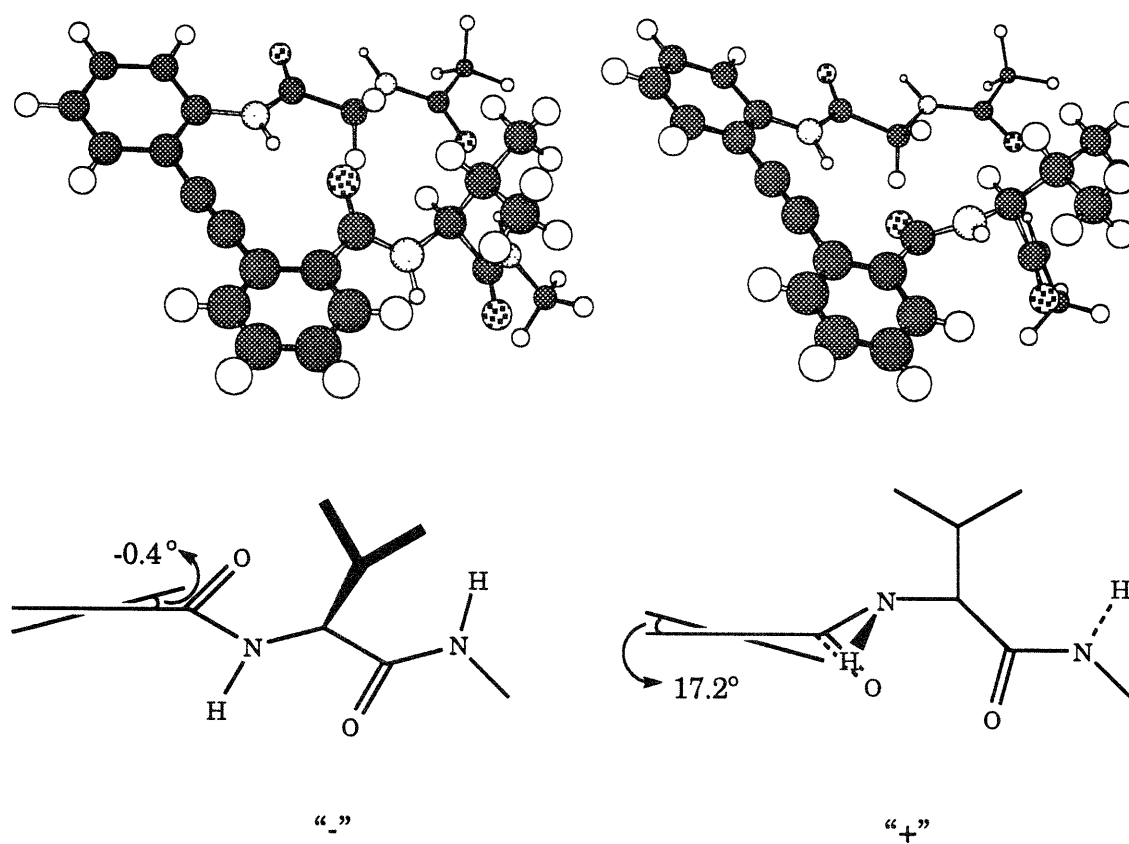


Figure 3.8 Conformations of AcGly-Taa-ValNHMe



There is little difference between the chemical shifts in DMSO and those in CDCl<sub>3</sub> for the aryl amide NH and methylamide NH, but again, large changes in chemical shifts are observed for the two  $\alpha$  amino acid amide NH's (Table 3.9). These different behaviors of amide NH's support the presence of sheet structure, in which the aryl amide NH and methylamide NH are hydrogen-bonded to carbonyl groups, and the other two amide protons are exposed to the solvent.

Table 3.9 Chemical shifts (ppm) of AcVal-Taa-AlaNHMe in different solvents

	aryl NH	Ala NH	Val NH	MeNH	H-3
in DMSO	9.461	8.771	8.212	7.903	8.231
in CDCl <sub>3</sub>	9.323	7.059	6.566	7.877	8.616

The dependence of the chemical shifts on temperature for each amide is again strictly linear and the temperature gradients ( $\Delta\delta/\Delta T$ ) (Table 3.10) again implicate a solvent-shielded aryl amide NH, while all the other three amide NH's are somewhat exposed to the solvent.

Table 3.10 Temperature gradients of AcVal-Taa-AlaNHMe

	aryl NH	AlaNH	ValNH	MeNH
$-\Delta\delta/\Delta T$ (ppb)	3.06	6.81	5.69	4.80

The vicinal H-NC <sub>$\alpha$</sub> -H coupling constants for the valinyl and the alanyl residues are 8.75 and 7.53 Hz, respectively. According to the Karplus equations, these values correlate to the dihedral angle value  $\phi = -153^\circ$  for the alanyl residue and  $\phi = -140^\circ$  for the valinyl residue, which are consistent with

the molecular modeling results of  $\phi = -161^\circ$  for alanyl residue and  $\phi = -136^\circ$  for the valinyl residue.

In contrast to the cases of AcGly-Taa-AA<sub>2</sub>NHMe, molecular mechanics modeling on AcVal-Taa-AlaNHMe generated only one preferred conformation -- the “-” conformer with the aromatic carbonyl being above the phenyl ring, which is shown in Figure 3.9. This is hardly surprising. When there are considerable interactions between the side chains of the two amino acid residues across the strands, their allowable conformational space will be severely restricted. Thus the modeling result is another piece of supporting evidence for the sheet structure.

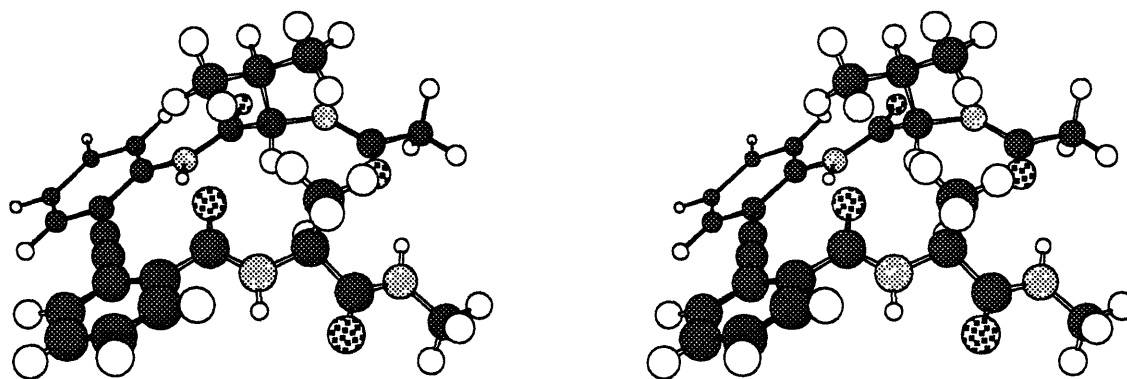


Figure 3.9 Stereoview of AcVal-Taa-AlaNHMe

### 3.8 AcVal-Taa-IleNHMe and AcAA<sub>1</sub>-Taa-ValNHMe

NMR spectra of AcVal-Taa-IleNHMe (14) and AcAA<sub>1</sub>-Taa-ValNHMe (16a, 16c-k) are similar to those observed for AcGly-Taa-AA<sub>2</sub>NHMe (6a-b and 16b) and AcVal-Taa-AlaNHMe (10). The temperature gradients of amide

proton chemical shifts for these derivatives in DMSO follow the same trend as those observed for **10**. They indicate a solvent-shielded aryl amide proton and three externally oriented amide protons. Although evidence so far supports the existence of sheet structure, the methylamide proton is still inevitably exposed to solvent due to an end-group effect. Therefore a solvent-exposed methylamide proton in no way precludes the presence of the sheet structure. The fact that it is exposed to the solvent to a lesser degree compared with the other two amide protons of the  $\alpha$  amino acid residues also supports the presence of the  $\beta$ -sheet structure.

In the ROESY spectra of these template-peptide derivatives in DMSO, all sequential NOEs indicating an extended conformation are observed. Also observed are interstrand NOEs characteristic of a sheet conformation, for example, between MeNH and Val  $\alpha$  H (weak), between Val  $\alpha$  H and Ile  $\alpha$  H (medium strong), between Ile  $\alpha$  H and Val Me (weak), for AcVal-Taa-IleNHMe (**14**). NOEs of different strengths observed for **14** are shown in Figure 3.10.

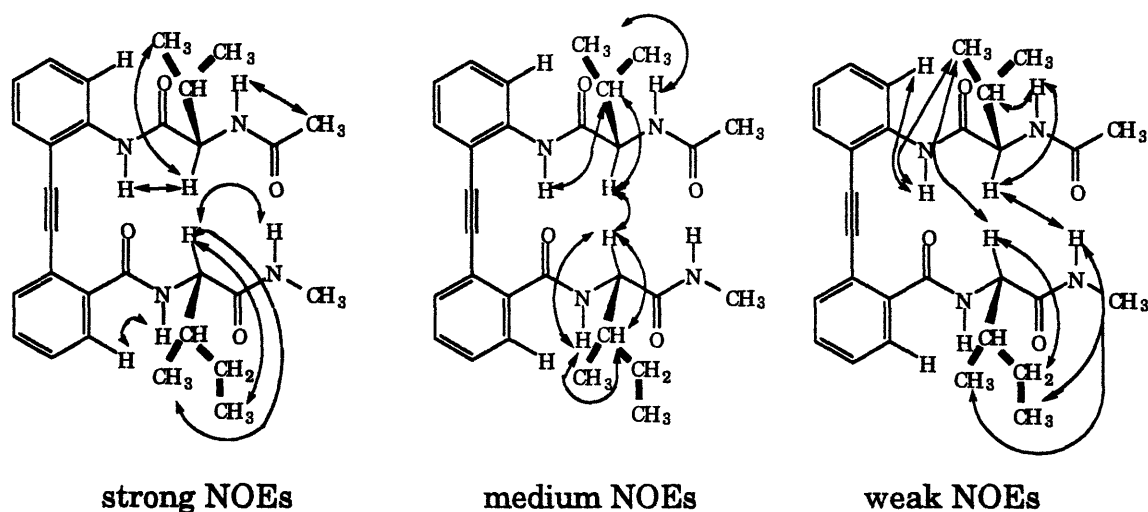


Figure 3.10 NOEs observed for AcVal-Taa-IleNHMe in DMSO

The ROESY spectrum of a charged tolanoic tetrapeptide analogue Cl<sup>-</sup> H<sub>3</sub>N<sup>+</sup>Gly-Val-Taa-ValNH<sub>2</sub> (**33a**) in DMSO also shows similar NOE crosspeaks that are characteristic of a sheet conformation.

Molecular mechanics modeling on these template-peptide derivatives gave results that are consistent with NOE interactions observed. Values of the dihedral angles  $\phi$  generated by molecular mechanics modeling are close to those calculated from vicinal H-NC $\alpha$ -H proton coupling constants observed in their NMR spectra. All these results indicate that these template-peptide derivatives share a similar sheet conformation with AcVal-Taa-AlaNHMe discussed earlier. The structures of AcVal-Taa-IleNHMe **14** (Figure 3.11) and of AcAla-Taa-ValNHMe **16a** (Figure 3.12) generated by Quanta are shown below.

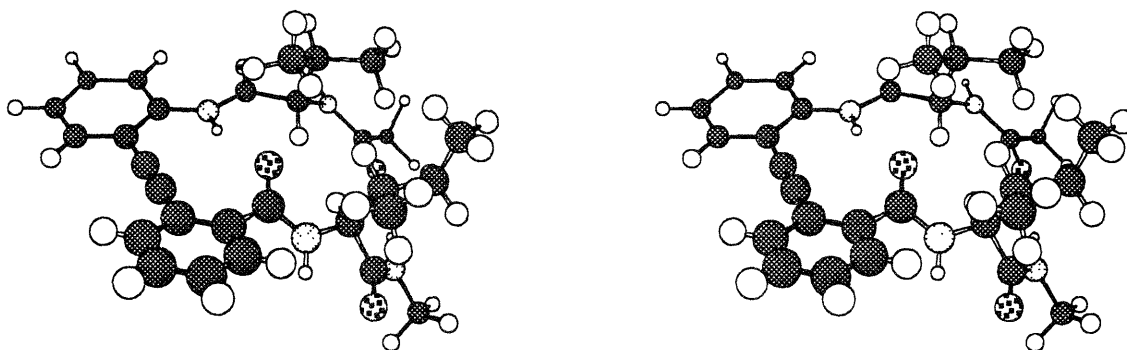


Figure 3.11 Stereoview of AcVal-Taa-IleNHMe

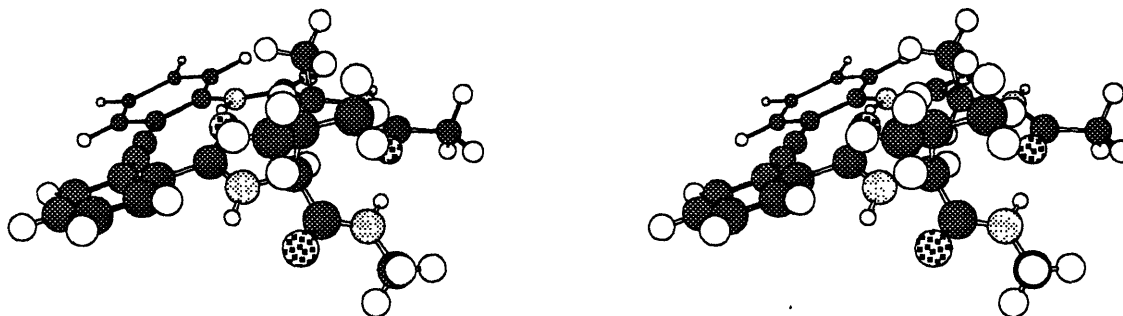


Figure 3.12 Stereoview of AcAla-Taa-ValNHMe

### 3.9 AcPro-Taa-AA<sub>2</sub>NHMe

Owing to the unique conformational rigidity of proline and its importance in the specific construction of peptide and protein structures (Schultz & Schirmer, 1978; Kopple *et al.*, 1975), the properties of proline-containing tolanoic derivatives are discussed separately in this section.

A prolyl residue is different from other  $\alpha$  amino acids in that its side chain loops back and forms a five-membered ring with its amino group, restricting the backbone dihedral angle  $\phi$ . As a result, the conformation of the prolyl residue depends mainly on the value of  $\psi$ . For an isolated prolyl residue,  $\psi$  value is found to be around  $-55^\circ$  and  $+145^\circ$  (Schimmel & Flory, 1968), which falls in the helix (or turns) and sheet regions on a Ramachandran plot, respectively. About 11% of *trans*-prolyl residues are found to occur in  $\beta$ -sheets,



of which 72% are in antiparallel  $\beta$ -sheet structure with the average values  $\phi$ ,  $\psi$  =  $-65^\circ$ ,  $150^\circ$  (MacArthur & Thornton, 1991).

Because it is tertiary, the peptidyl amide bond of proline can assume substantial *cis* isomer (Figure 3.13) that has eluded all the other  $\alpha$  amino acid residues. About 35% of prolyl residues found in small peptides and 8% in proteins are *cis* isomers (Milner-White *et al.*, 1992). The intrinsic probability of a prolyl residue having the *cis* isomer of the preceding peptide bond is about 0.1 to 0.3 (Brandts *et al.*, 1975), and the energy barrier for *cis-trans* isomerization is about 13 kcal/mole (Schultz & Schirmer, 1978). A recent study using empirical energy functions and *ab initio* calculations puts that barrier for *N*-acetylproline methylamide at  $\sim 15$  kcal/mole (Fischer *et al.*, 1994). Therefore the chemical process of prolyl *cis/trans* isomerization is slow on the NMR time-scale and two distinctive sets of NMR peaks may be observed for the two isomers (Kessler, 1970; Binsch & Kessler, 1980; Kessler, 1982).  $^{13}\text{C}$  NMR has been mostly used to obtain the prolyl *trans/cis* ratio because the peaks in  $^1\text{H}$  NMR are not always well separated. Intensities of  $^{13}\text{C}$  peaks are used according to the procedure prescribed by Grathwohl & Wüthrich (1976) because  $^{13}\text{C}$  peak area measurements are generally unreliable and can lead to quite conflicting results compared with those obtained from  $^1\text{H}$  NMR (Luthman & Hacksell, 1993). NMR Studies of proline-containing peptides indicate that the ratio of *cis/trans* isomers is characteristic for the amino acid sequence in the immediate environment of the prolyl residue, and this equilibrium is sensitive to the overall molecular conformation (Wüthrich & Grathwohl, 1974; Grathwohl & Wüthrich, 1976).

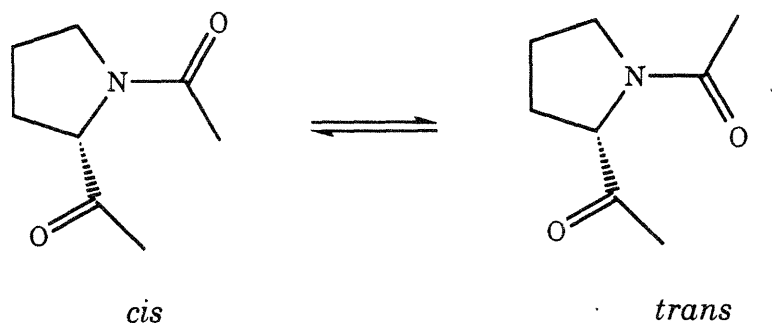


Figure 3.13 Acetylprolyl *cis/trans* isomerization.

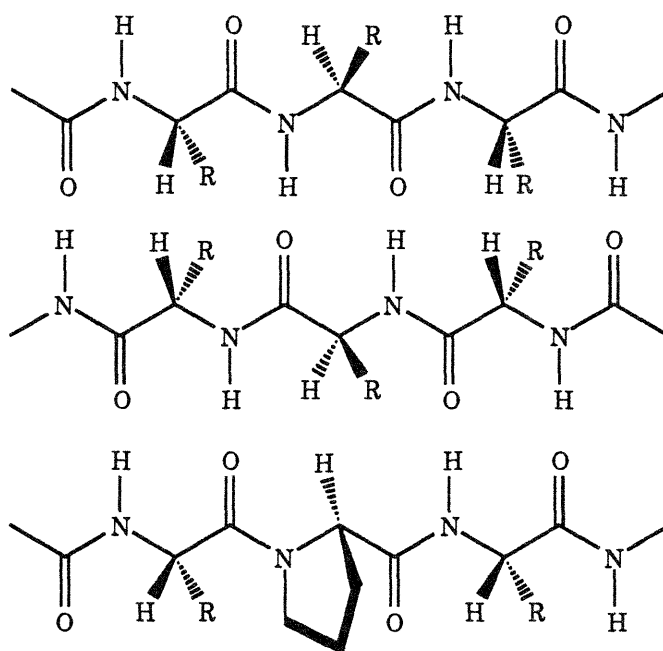
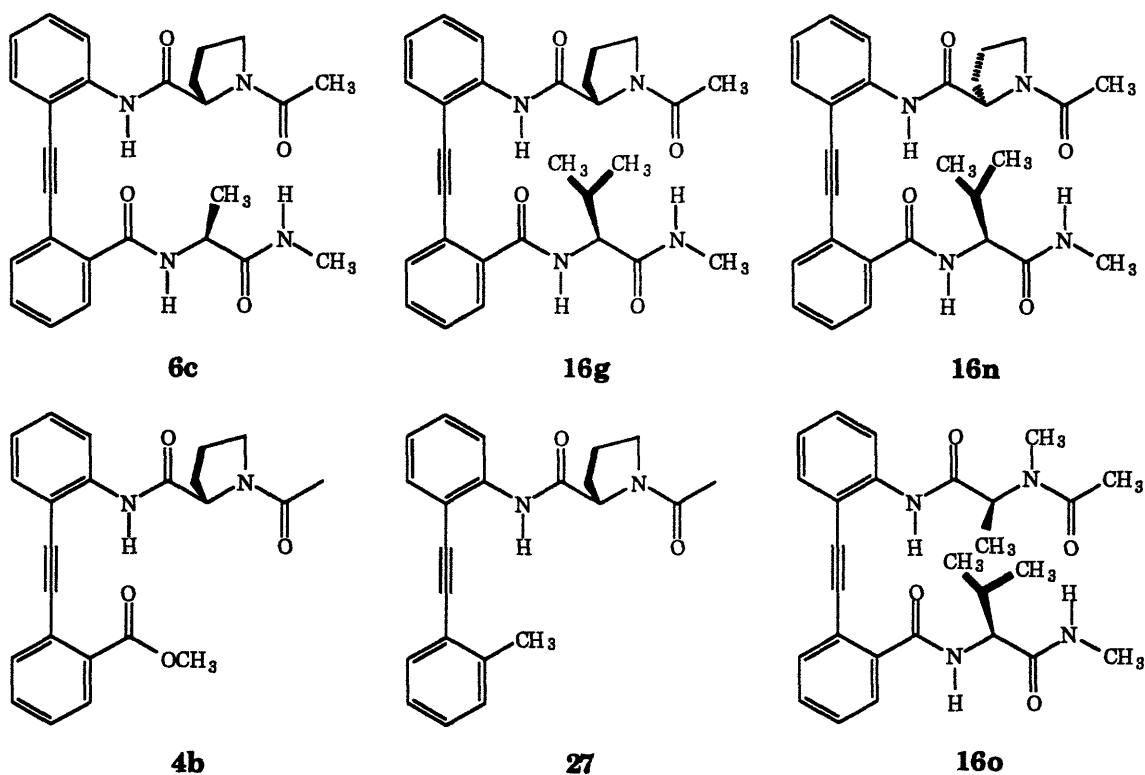


Figure 3.14 Preferred location for prolyl residue found in  $\beta$ -sheets

The lack of amide hydrogens renders a proline residue unable to form a hydrogen bond at its amide site in a peptide. As a result, almost all the prolyl residues found in  $\beta$ -sheet structure occur on the edge strand of the sheet, and between the widely spaced hydrogen-bond pairs with the carbonyl oxygen making no contribution to the hydrogen-bonding network as pictured in the above Figure 3.14 (MacArthur & Thornton, 1991). Due to its restrictive  $\phi$  values, prolyl residue has the effect of intensifying the strand twist (Chothia & Janin, 1981; 1982).

A total of five proline-containing derivatives plus a *N*-methylalanyl derivative were synthesized, as shown below.



These derivatives are AcPro-Taa-AlaNHMe (**6c**), and AcPro-Taa-ValNHMe (**16g**), Ac-D-Pro-Taa-ValNHMe (**16n**), AcPro-Taa-OMe (**4b**), AcPro-NHTln-Me (**27**), and Ac-Me<sup>N</sup>-Ala-Taa-ValNHMe (**16o**). All were purified by HPLC and/or flash chromatography before NMR and CD spectra were taken. Derivatives **4b** and **27** are controls without amino acid residues on the opposing strand of the proline residue, thus excluding the possibility of a sheet-like structure. The prolyl residue in **16n** is a mis-matched D-proline, and as a result, its acetylamide should exhibit a quite different *trans/cis* ratio if the opposing strand exerts any effect on the equilibrium of *trans/cis* isomers at all.

The reason that Ala and Val are chosen for this initial study is that they are both non-polar, and the only difference between them is their different sizes. All 5 derivatives have well separated peaks in the <sup>1</sup>H spectra in DMSO (amide NH, aromatic, aliphatic). For the two controls in DMSO, the ratios obtained from <sup>13</sup>C NMR agree very well with those from <sup>1</sup>H NMR. Since in CD<sub>2</sub>Cl<sub>2</sub> the two sets of peaks of Ac-Pro-Taa-OMe <sup>1</sup>H NMR overlap significantly, <sup>13</sup>C NMR is used exclusively for calculation in this case. Our unique template-peptide system is superior in the sense that the <sup>1</sup>H NMR spectra of these tripeptidyl derivatives possess 5 or 6 sets of well separated peaks available for more precise calculation of the prolyl *trans/cis* ratio, and 4 sets of prolyl ring <sup>13</sup>C NMR peaks serve as a check for the calculated ratio from proton NMR area integration. In our studies, the <sup>13</sup>C measurements all agreed very well with the <sup>1</sup>H results.

The *trans/cis* isomers were distinguished by a combination of NOESY/ROESY and <sup>13</sup>C NMR spectra. Assignments of *trans/cis* isomers using <sup>13</sup>C NMR follows the prescribed procedure in the literature by recognizing distinctive peak locations (Grathwohl & Wüthrich, 1976). The <sup>13</sup>C chemical

shifts of  $C_\beta$  and  $C_\gamma$  in proline rings are indicative of the stereochemistry of *cis* and *trans* prolines. The  $^{13}\text{C}$  chemical shifts of  $C_\beta$  and  $C_\gamma$  are around 31.3 and 22.5 ppm respectively for a proline in a *cis* conformation, and 29.5 and 24.2 for the one in a *trans* conformation (Dormann & Bovey, 1973; Deber *et al.*, 1976). In the prolyltolanoic derivatives the  $^{13}\text{C}$  chemical shifts of  $C_\beta$  and  $C_\gamma$  in DMSO are found to be around 32.1 and 22.5 ppm respectively for the *cis* conformation, and 29.5 and 24.5 ppm for the *trans* conformation. In addition,  $C_\alpha$  and  $C_\delta$  are also well-separated and are around 60.9 and 46.3 ppm for the *cis* conformation, and 59.9 and 47.6 ppm for the *trans* conformation. The assignment of *cis* and *trans* acetylprolyl residue in our tolan systems is in full agreement with that obtained from 2D NMR studies.

The NOEs in the ROESY spectra are very similar to their NOESY counterparts. The NOESY and ROESY spectra in DMSO for the major conformation are very similar to those in  $\text{CD}_2\text{Cl}_2$ . The following NOESY spectrum of AcPro-Taa-AlaNHMe in  $\text{CD}_2\text{Cl}_2$  is typical among these proline-containing derivatives. Observed NOEs of interests are the following:

- 1) Aryl NH -- Pro  $\alpha$  H: very strong, indicating extended conformation;
- 2) Aryl NH -- Pro  $\beta$  H: very weak, indicating extended conformation;
- 3) MeNH -- Ala  $\alpha$  H: very strong, indicating extended conformation;
- 4) MeNH -- Pro  $\alpha$  H: very weak, indicating sheet-like conformation;
- 5) MeNH -- Ala Me: very weak, indicating extended conformation;
- 6) Ala NH -- Ala  $\alpha$  H: medium weak, indicating extended conformation;
- 7) Pro  $\alpha$  H-- Ala  $\alpha$  H: strong, indicating sheet-like conformation;
- 8) Pro  $\alpha$  H-- Ala Me: very weak, indicating sheet-like conformation;
- 9) Pro  $\delta_1$  -- Pro  $\gamma_1$ : very strong;
- 10) Pro  $\delta_1$  -- Pro  $\gamma_2$ : weak;
- 11) Pro  $\delta_1$  -- AcMe: strong (inferred, see below);
- 12) Pro  $\delta_2$  -- Pro  $\gamma_1$ : weak (inferred, see below);
- 13) Pro  $\delta_2$  -- Pro  $\gamma_2$ : very strong;

14) Pro  $\delta_2$  -- AcMe: strong (inferred, see below);

The proof that prolyl *trans* conformer is contributing to the sheet-like conformation comes from two separate facts. Firstly from the  $^{13}\text{C}$  NMR spectra, the major conformer is assigned to the *trans* isomer according to the chemical shift patterns observed in the literature (Grathwohl & Wüthrich, 1976). The second evidence comes from NMR analysis of both the 1D  $^1\text{H}$  NMR spectra and the 2D NOESY data. In the 1D  $^1\text{H}$  NMR spectrum in DMSO, the  $\alpha$ -carbon proton for the *trans* acetylprolyl isomer is shifted to lower field compared with the *cis* isomer because in the *trans* isomer the  $\alpha$ -carbon proton is deshielded significantly by the acetyl carbonyl group. As for the NOESY spectrum, the NOEs between the prolyl  $\delta$  protons and those of the acetylmethyl group should be strong for the *trans* isomer and weak for the *cis* isomer. However, the assignment of these NOE crosspeaks was not straightforward. The complication stems from the fact that the  $^1\text{H}$  NMR peak of one of the two prolyl  $\gamma$  protons overlaps with that of the acetylmethyl group. However, an inference can be made, with a fair degree of confidence, from the asymmetry and the different intensities of the apparent NOEs observed between the prolyl  $\delta$  protons and the prolyl  $\gamma$  protons plus between the prolyl  $\delta$  protons and the acetylmethyl protons, as shown in Figure 3.15a-b. In addition, the NOEs between the acetylmethyl protons and the prolyl  $\alpha$ -carbon proton are absent for the major isomer.

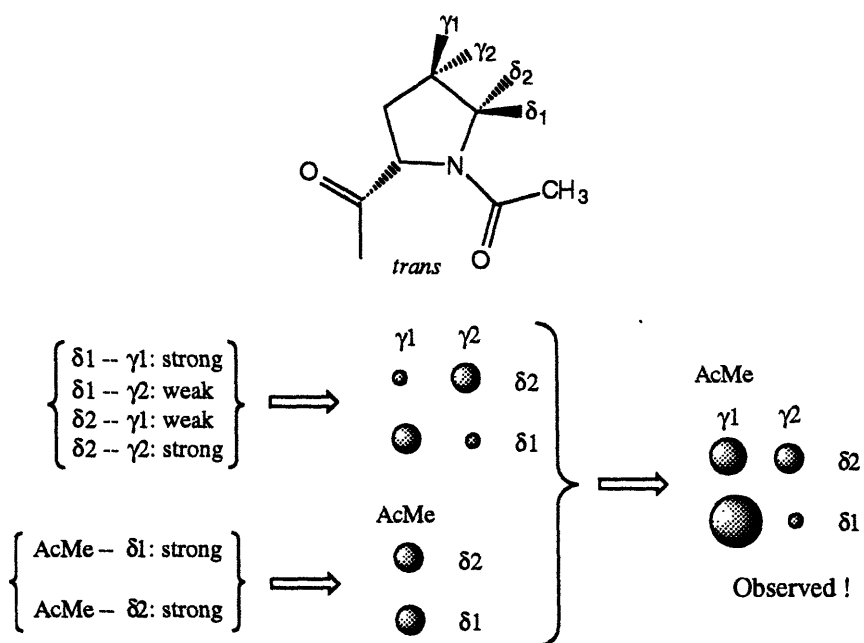


Figure 3.15a Analysis of the NOE pattern observed.

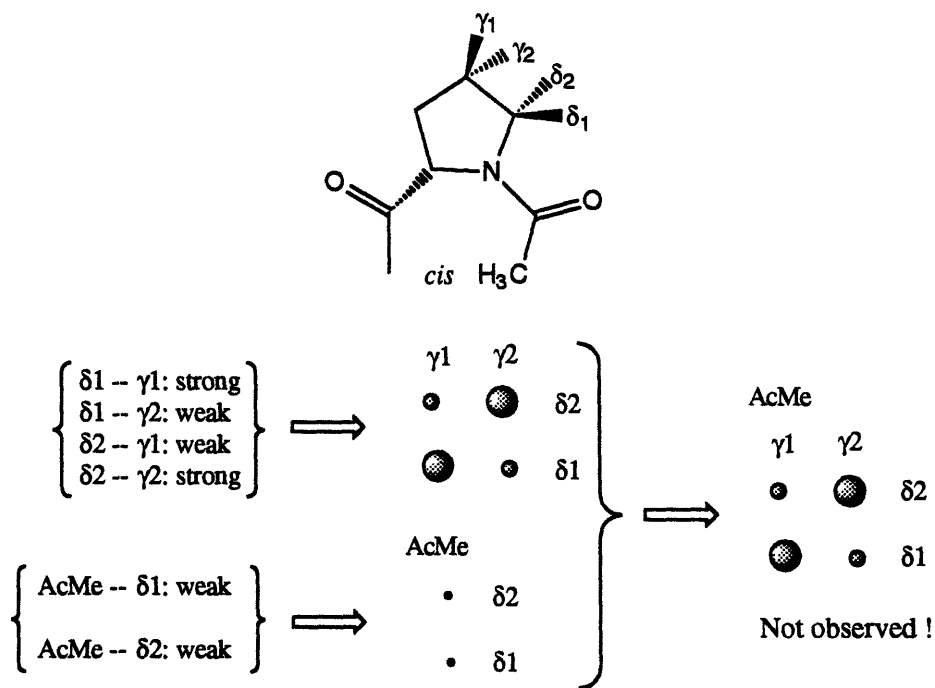


Figure 3.15b Analysis of the NOE pattern not observed.

Variable temperature studies of AcPro-Taa-AlaNHMe and AcPro-Taa-ValNHMe gave similar temperature gradients (Table 3.11) as those observed for other template-peptide derivatives.

Table 3.11 Dependence of chemical shift on temperature ( $\Delta\delta/\Delta T$ )

	AcPro-Taa-AlaNHMe	AcPro-Taa-ValNHMe
major aryl NH	-1.77 ppb/K	-1.43 ppb/K
minor aryl NH	-2.45 ppb/K	-2.93 ppb/K
major Ala NH	-6.79 ppb/K	-8.19 ppb/K
minor Ala NH	-5.96 ppb/K	-6.45 ppb/K
major MeNH	-5.19 ppb/K	-5.22 ppb/K
minor MeNH	-4.50 ppb/K	-4.67 ppb/K

Table 3.12 *Trans/cis* ratio of prolyl derivatives at 25 °C.

Derivatives	<i>trans/cis</i> ratio in DMSO	<i>trans/cis</i> ratio in CD <sub>2</sub> Cl <sub>2</sub>
Ac-Pro-Taa-Ala-NHMe	3.46 ± 0.07	96 ± 1% <i>trans</i>
Ac-Pro-Taa-Val-NHMe	5.4 ± 0.2	100% <i>trans</i>
Ac-Pro-Taa-OMe	3.6 ± 0.2	2.8 ± 0.1 (from <sup>13</sup> C)
Ac-Pro-NHTIn-Me	3.2 ± 0.2	8.2 ± 0.5
Ac-D-Pro-Taa-ValNHMe	3.1 ± 0.3	not measured
Ac-Me <sup>N</sup> -Ala-Taa-ValNHMe	5.0 ± 0.3	96 ± 2% <i>trans</i> (CDCl <sub>3</sub> )

As mentioned earlier, the *trans/cis* ratio of the acetylprolyl amide isomers is sensitive to its molecular conformation. In non-polar solvents, such as CD<sub>2</sub>Cl<sub>2</sub> and CDCl<sub>3</sub>, the *trans* isomer predominates: more than 95% for AcPro-Taa-AlaNHMe, and almost no *cis* isomer was detected for AcPro-Taa-ValNHMe. In comparison, the *trans/cis* ratio is much lower for the two prolyl analogues: about 8:1 for AcPro-NHTIn-Me (27) and only about 3:1 for AcPro-



Taa-OMe (**4b**), indicating drastic conformational differences between AcPro-Taa-AA<sub>2</sub>NHMe (AA<sub>2</sub> = Ala and Val) and these two model compounds. The difference in *trans/cis* ratios observed for **4b** and **27** is noteworthy. This discrepancy is undoubtedly the result of a competing hydrogen bonded conformation in **4b**, whereas in **27** it is absent (Figure 3.16). The value observed for **27** is almost the same as that for *N*-acetylprolyl methylamide measured by Gellman (Liang et al., 1992).

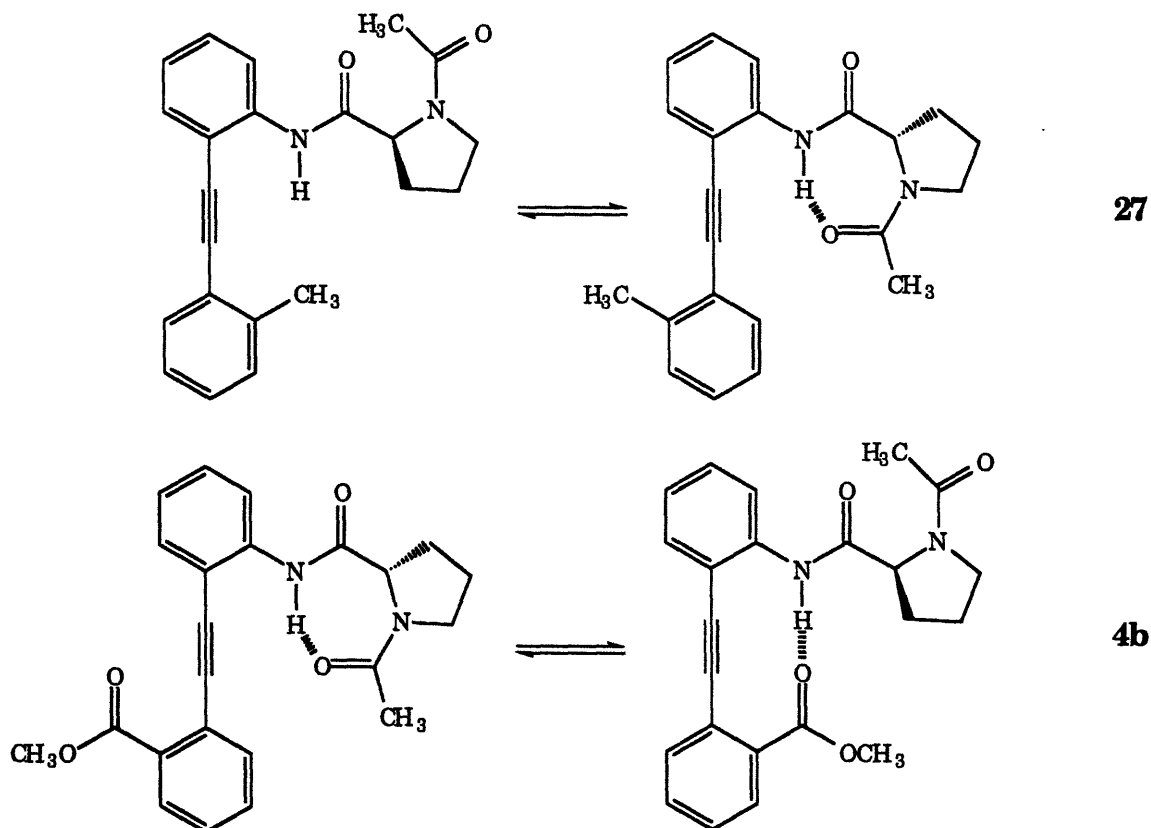


Figure 3.16 Proposed conformational equilibrium for AcPro-NHTln-Me (**27**) and AcPro-Taa-OMe (**4b**)

For AcPro-NHTIn-Me (**27**), the *trans* isomer is selectively stabilized by intramolecular hydrogen bonding ( $\gamma$ -turn, as shown in Figure 3.16), whereas for AcPro-Taa-OMe (**4b**), a second carbonyl group is competing for the same hydrogen-bond donor -- the aryl amide proton (Figure 3.16, bottom equilibrium). As a strongly hydrogen-bonding solvent, DMSO has a leveling effect and the trend is reversed: *trans/cis* ratio is slightly higher for **4b** (3.6:1) than that for **27** (3.2:1), and the difference of these two values becomes much smaller. As expected, the mis-matched Ac-D-Pro-Taa-ValNHMe (**16n**) has the lowest *trans/cis* ratio 3.1:1 in DMSO among all the derivatives examined, whereas AcPro-Taa-ValNHMe has the highest value 5.4:1, being close to the ratio 5.0:1 for the *N*-methylalanyl derivative Ac-Me<sup>N</sup>-Ala-Taa-ValNHMe (**16o**). It is very intriguing that the ratio for AcPro-Taa-AlaNHMe approaches that observed for AcPro-Taa-OMe (3.4-3.5 versus 3.4-3.8), which seems to indicate that the energy stabilization from prolyl-alanyl pairwise interaction is very small.

In this study we have shown that the prolyl *trans/cis* ratio is affected by the overall molecular conformation of AcPro-Taa-AA<sub>2</sub>NHMe derivatives (**6c** and **16g**), and that the amino acid residue (Ala or Val) on the opposite strand has a strong influence on the equilibrium of the prolyl *cis/trans* isomers.

### 3.10 Chemical shift correlation

If the conformational states for these template-peptide derivatives can be approximated as an equilibrium between three well-defined states, the chemical shifts for the tolanoic amide proton and the methylamide proton should follow a linear relation as shown in Figure 3.17.

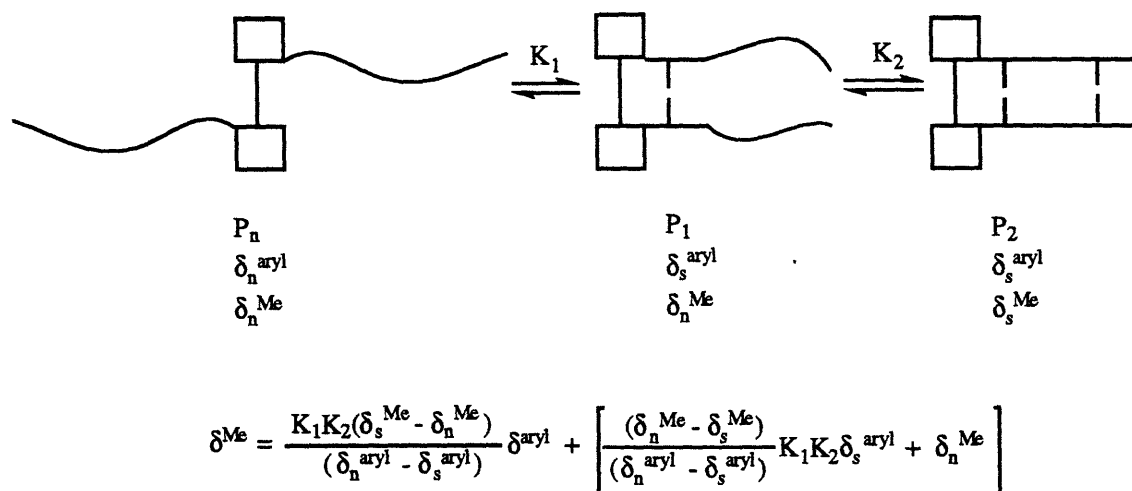


Figure 3.17 Correlation of aryl amide H chemical shift with methylamide H.  $P_n$ ,  $P_1$ , and  $P_2$  represent mole fraction of non-sheet, first hydrogen bond, and sheet conformation, respectively, and their sum = 1. Aryl amide H and methylamide H chemical shifts for each state are designated as shown. Except for  $\delta^{\text{aryl}}$  and  $\delta^{\text{Me}}$ , all the other  $\delta$  values are constant numbers.

The chemical shifts of the tolanoic amide protons are plotted against the methylamide protons in Figure 3.18 for the twelve template-peptide derivatives  $\text{AcAA}_1\text{-Taa-AA}_2\text{NHMe}$ , in which  $\text{AA}_2$  is either a valinyl or an isoleucyl residue. Except for the phenylalanyl derivative  $\text{AcPhe-Taa-ValNHMe}$ , all the other points fall in a close range of the line. This correlation is consistent with the three-state approximation as discussed above. There is also positive correlation between the chemical shifts of the tolanoic amide proton and those of the phenyl ring H-3 proton, and between the chemical shifts of the methylamide proton and the phenyl ring H-3' proton (Figure 3.19).

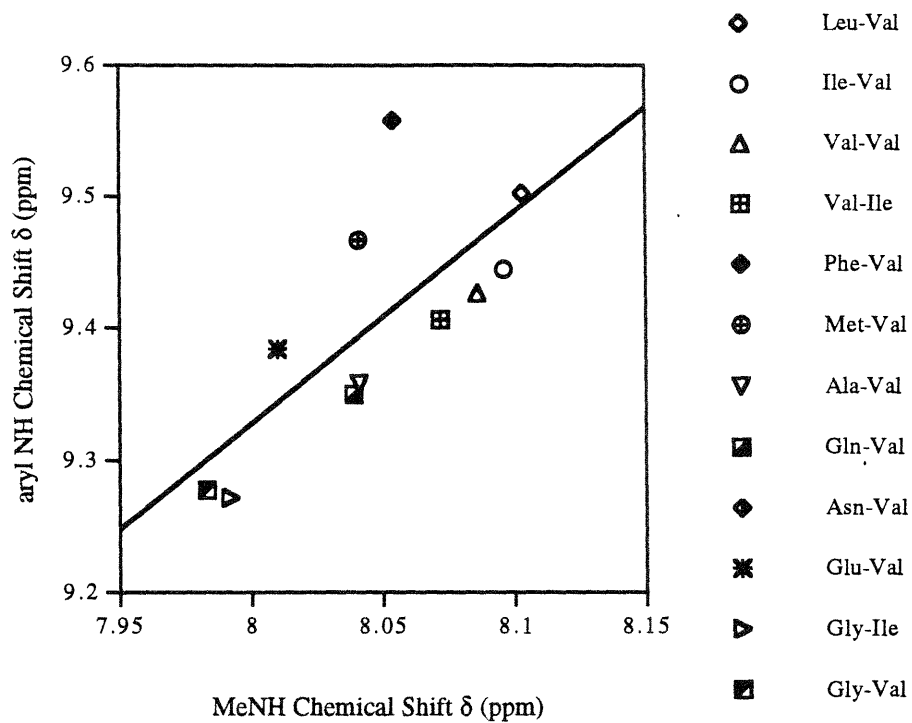


Figure 3.18 Chemical shift of aryl amide proton versus chemical shift of methylamide proton for twelve template-peptide derivatives  $\text{AcAA}_1\text{-Taa-AA}_2\text{NHMe}$ , where  $\text{AA}_2$  is either a valinyl or an isoleucyl residue.

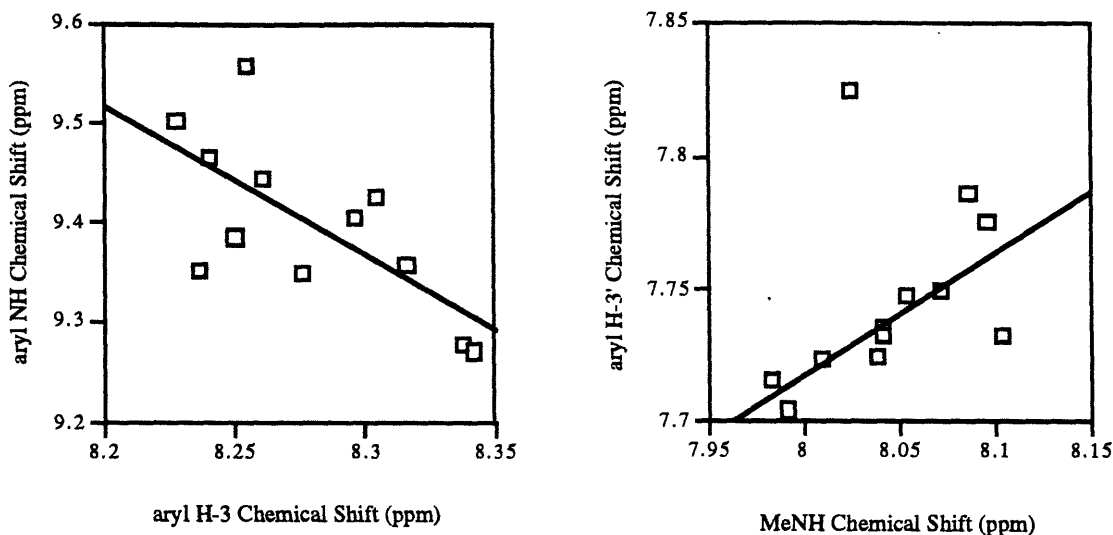
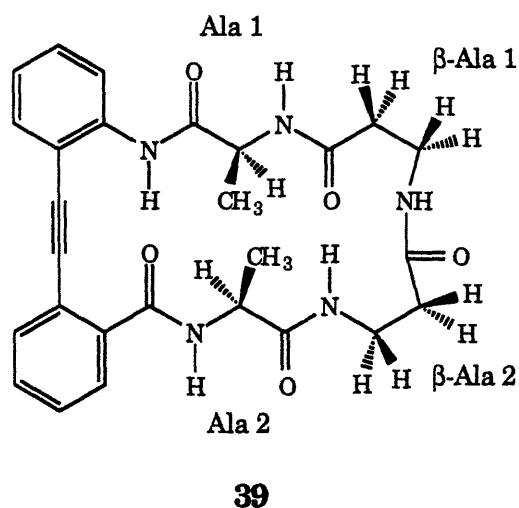


Figure 3.19 Chemical shift correlation between the tolanoic amide protons and the aryl H-3 protons (left), and between the aryl H-3' protons and the methylamide protons (right) for the twelve template-peptide derivatives AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe, where AA<sub>2</sub> is either a valinyl or an isoleucyl residue.

These weak positive correlations in chemical shifts between the distant protons of the template-peptides conjugates suggest that functionalities at the two ends of the chain are related to the tolanoic template in a non-random fashion, which in turn supports the idea that the tolanoic turn induces sheet structure in the  $\alpha$  amino acid derivatives.

### 3.11 Cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala]

The structure for cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] (**39**) is depicted below. Amino acid residues, including the tolanoic moiety, are numbered as such: start with the amino terminus of the top  $\beta$ -alanyl residue as residue number 1, then trace down the chain to rank, sequentially, alanyl residue number 1, tolanoic residue, alanyl residue number 2, and  $\beta$ -alanyl residue number 2.



**39** is a key compound in our conformational analysis because it helps in addressing the major question whether the two ends of  $\beta$ -hairpin conformation observed in these template-peptide conjugates are close. In this cyclic compound, the two ends are forced to be in such a position.

According to molecular mechanics modeling, the tolanoic carbonyl amide functionality is significantly tilted from its attached phenyl ring for all tolanoic peptide derivatives examined. This is perhaps due to the electronic repulsion between the  $p\pi$ -orbitals of the triple bond and the non-bonded  $p$ -orbital of the

carbonyl oxygen. A cyclic tolanoic derivative is useful in determining whether the additional constraints from the cyclization will force the carbonyl group into coplanarity with the ring.

Molecular mechanics modeling indicates there are two viable sheet conformations for cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] (39) resulting from a flipping motion of the central  $\beta$ -alanyl to  $\beta$ -alanyl amide bond (Figure 3.20): an "up" conformation in which the  $\beta$ -alanyl to  $\beta$ -alanyl amide NH is positioned slightly above the pseudo-plane formed by the two  $\beta$ -alanyl residues and forms a bifurcated hydrogen bond with the carbonyl of its  $\beta$ -alanyl residue, shown in Figure 3.21a, and a second "down" conformation in which the  $\beta$ -alanyl to  $\beta$ -alanyl amide NH points into the concave face of the molecule, as shown in Figure 3.21b.

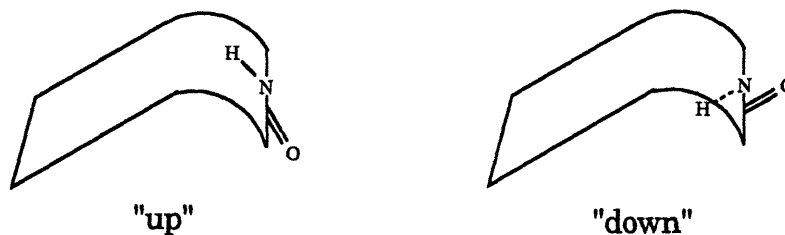


Figure 3.20 Cartoon drawing depicting the "up" and "down" conformers

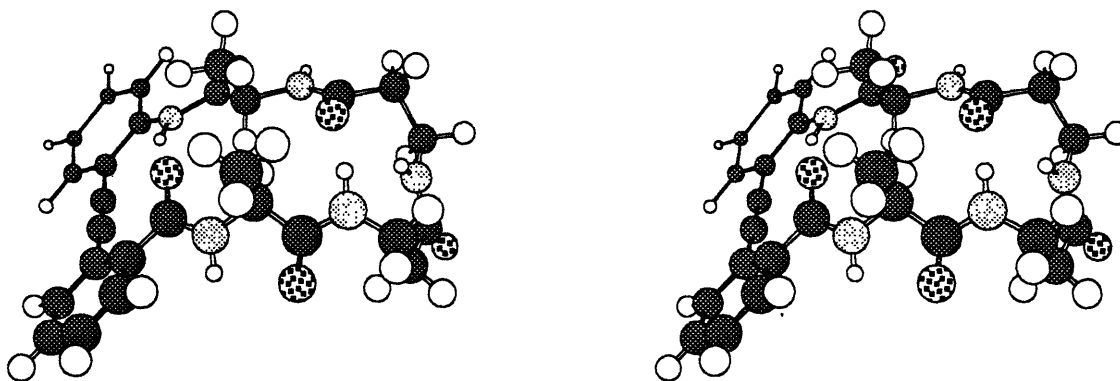


Figure 3.21a The "up" conformation

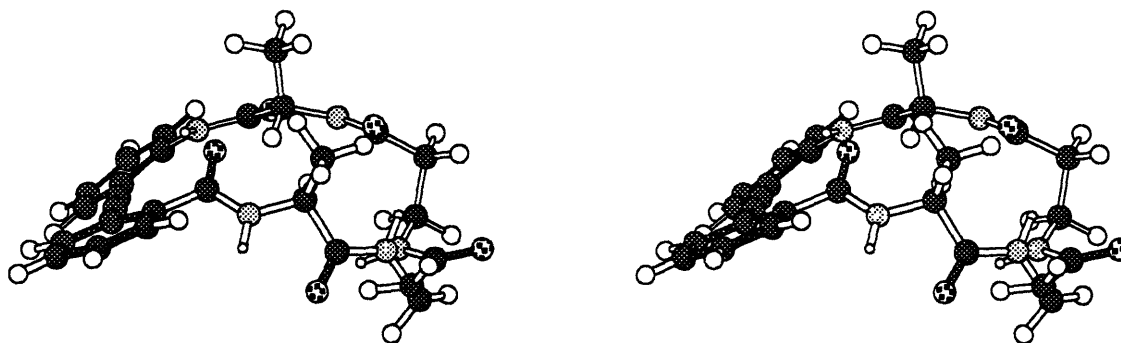


Figure 3.21b The "down" conformation

The tolanoic conjugation system in the "up" conformation is more planar than the "down" conformation ( $\alpha$  angle  $2.2^\circ$  as compared with  $7.8^\circ$ ). In both conformations, the aromatic amide group is substantially rotated off its phenyl plane ( $\gamma = -23.9^\circ$  in the "up" and  $-29.2^\circ$  in the "down" conformation), so is the



aromatic carbonyl group ( $\beta = 35.2^\circ$  in the “up” and  $40.7^\circ$  in the “down” conformation). Both conformations exhibit significant twist along the chain.

Two slowly-equilibrating conformations are in fact detected by NMR. In  $\text{CDCl}_3$ , one conformation dominates the  $^1\text{H}$  NMR spectrum (>90%), whereas in DMSO two conformations are significantly represented (7:1). One noticeable feature of the 1D NMR spectrum in  $\text{CDCl}_3$  is the widely dispersed chemical shifts for methylene protons of the  $\beta$ -alanine residues. The two  $\alpha$ -carbon (covalently bonded to the  $\beta$ -alanyl carbonyl) protons of  $\beta$ -alanine 1 experience different magnetic environments and are diastereotopic. They are separated from each other by 0.95 ppm. The  $\beta$ -carbon methylene protons of  $\beta$ -alanine 1, on the other hand, are split by an even larger difference of 1.19 ppm. The  $\beta$ -carbon methylene protons of  $\beta$ -alanine 2 are also no longer magnetically equivalent and are distanced from each other by 1.21 ppm. However, chemical shifts for the two  $\alpha$ -carbon protons of  $\beta$ -alanine 2 fall on top of each other and are virtually coupled to each other, which reflect their similar electromagnetic environments. Since the magnetic ring currents generated by the tolanic moiety are far-removed from these methylene protons according to molecular modeling, these large anisotropy effects (or its absence in the case of  $\beta$ -Ala 2's  $\alpha$  protons) are undoubtedly due to their proximity to the amide carbonyl groups. Anisotropy effects due to proximity of the carbonyl group are usually subjected to substantial solvent effects. Differences in chemical shifts of the methylene protons are considerably smaller in DMSO as compared with those in  $\text{CDCl}_3$ . These large differences in chemical shifts of methylene protons are solid indications for conformational homogeneity and rigidity in  $\text{CDCl}_3$ .

The amide NH of alanyl residue 2, however, is a conspicuous singlet in both DMSO and  $\text{CDCl}_3$ , and its chemical shift remains essentially unchanged

going from  $\text{CDCl}_3$  to DMSO ( $\Delta\delta = 0.25$  ppm) as compared with the large downfield shift experienced by the amide NH of Ala 1 ( $\Delta\delta = 2.59$  ppm). An examination of the models indicates that this amide NH is pointed into the concave face of the molecule, making it less accessible to the solvent. However, the molecular mechanics modeling could not explain the appearance of the singlet alanyl amide NH. From the modeling result, the dihedral angle  $\theta$  of  $\text{H-NC}_\alpha\text{-H}$  is  $139^\circ$  and  $143^\circ$  for the “up” and “down” conformations, respectively, which should give rise to non-zero vicinal  $\text{H-NC}_\alpha\text{-H}$  coupling constants (calculated to be 6.6 and 7.1 Hz, respectively). At the present time, we cannot distinguish whether the absence of this scalar coupling is due to the dihedral angle  $\theta \approx 90^\circ$  that is not accurately reflected in the modeling or to the possibility that Karplus relations do not apply in this case where the amino acid residue (alanine) is covalently linked to a planar aromatic ring system.

Although an unambiguous assignment of the two conformations is difficult to achieve, we are led to believe that the “up” conformation is the dominant one. The  $\alpha$ -carbon proton of Ala-2 is shifted significantly downfield for the major conformation from that of the minor conformation in both DMSO and in  $\text{CDCl}_3$ . The tolanoic amide proton of the major conformation is also further downfield in contrast with the minor conformation (0.5 ppm).

A comparison of the 1D  $^1\text{H}$  NMR spectrum of cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] with its acyclic analogue AcAla-Taa-AlaNHMe in DMSO indicates two regions of significant differences in chemical shifts. The  $\alpha$ -carbon proton of Ala-1 in the cyclic derivative has an upfield shift of -0.67 ppm, whereas that of Ala-2 has a downfield shift of 0.45 ppm. These changes in chemical shifts can be explained by the rigidity of the “up” conformation, and the proximity of Ala-2

$\alpha$ -H to both the aromatic carbonyl and the  $\beta$ -Ala 1 carbonyl groups, as compared with the open-chain case of AcAla-Taa-AlaNHMe. The second chemical shift region of substantial change is the aromatic H-3' proton, which has a large upfield shift of -1.14 ppm. This is most likely due to the result of the cyclization that confines the amide carbonyl in a more or less fixed position relative the phenyl ring, and H-3' coincides inside the cone of the electromagnetic ring current generated by the carbonyl group. In contradistinction to the relative coplanarity of the tolanoic amide functional group with its phenyl ring, the relative inertness to solvent changes and the singlet nature of the Ala-2 amide NH as well as the unusually large chemical shift anisotropy experienced by the aromatic H-3' proton indicate that the tolanoic carboxamide functional group has a significant value of the dihedral angle  $\beta$  with its attached phenyl ring.

These results indicate that the tolanoic carboxamide to the tolan linkage is significantly different in the cyclic case from the open case, suggesting there is a certain degree of strains experienced by the tolanoic carboxamide functionality when the two ends of the  $\beta$ -hairpin are brought together. Therefore, although it can induce the  $\beta$ -sheet conformation in the attached peptide derivatives, the tolanoic template is not an ideal  $\beta$ -hairpin template.

### 3.12 Summary

The results of NMR studies indicate that the tolanoic template is indeed capable of inducing  $\beta$ -sheet formation. There is considerable structure present at the tolanoic moiety due to the persistent hydrogen bonding between the tolanoic amide proton and the tolanoic carbonyl oxygen. However, the two

ends of the template-peptide conjugates are relatively free. This is hardly surprising, since theoretical studies on the  $\beta$ -sheet to random coil transition indicate that the N-terminal and C-terminal residues should be disordered (Mattice, 1989; Bradley *et al.*, 1990; Miick *et al.*, 1991; Yapa *et al.*, 1992). However, in our case, the major deterrent for the two ends being close together seems to come from the fact that there are considerable constraints against a coplanar arrangement of the tolanoic carboxamide functionality with the phenyl ring.

The methylamide proton is exposed to a lesser degree to the solvent than the two  $\alpha$  amino acid amide protons, suggesting the presence of a sheet conformation. The positive correlation between the tolanoic amide proton and the methylamide proton also suggests a sheet conformation. The NMR and molecular mechanics modeling studies also indicate that the tolanoic amide functionality is conformationally constrained and stays in near co-planarity with its attached phenyl ring, whereas the tolanoic carboxamide group is tilted from the phenyl plane.

## **Chapter Four**

### **Conformational Properties Studied by UV and CD**

From the previous chapter, we concluded that the tolanoic template can induce the formation of a  $\beta$ -hairpin. In this chapter, we want to address the second question -- whether the tolanoic template can act as a UV/CD reporting conformational template. More precisely, the following two questions will be addressed:

- 1) Is the UV spectrum of a template-peptide derivative sensitive to the conformation of interest?
- 2) Is the CD spectrum of a template-peptide derivative sensitive to the conformation of interest?

In addition, we want to apply the UV and CD tools to further define the conformational properties of the template-peptide derivatives, since analysis of molecular conformations can be best accomplished by combining different spectroscopic methods. In the composite approach, the results obtained by each method are used to answer the questions they are best fitted to deal with, and thus permit the overall problem to be solved in the most efficient manner.

Before we embark on the actual analysis of the UV and CD results, a brief introduction of the tools employed is appropriate.

#### 4.1 Introduction: UV and CD

Molecular absorption in the ultraviolet and visible region of the spectrum is dependent on the electronic structure of the molecule. The energy absorbed by a molecule is quantized, promoting electrons from orbitals in the ground states to higher-energy orbitals in an excited state. The total energy  $E$  of a molecule in an energy state can be expressed as

$$E = E_{el} + E_{vib} + E_{rot},$$

where  $E_{el}$ ,  $E_{vib}$ , and  $E_{rot}$  represent the electronic, vibrational, and rotational energies, respectively. These three types of energies are all quantized and can only exist at discrete energy states.

The allowed vibrational energies for a diatomic molecule, for example, are simply

$$E_{vib} = E_v = (v + \frac{1}{2})h\nu_e \quad v = 0, 1, 2, \dots$$

where  $h$  is the Planck constant,  $\nu_e$  is the fundamental vibrational frequency for the electronic state, given by  $(1/2\pi)(k_e/\mu)^{1/2}$ ,  $k_e$  and  $\mu$  being the force constant for the electronic state and the reduced mass of the molecule, respectively, and  $v$  is the vibrational quantum number, which can only assume integral values as shown. The energy separations of successive vibrational levels of relatively low quantum numbers are on the order of 1-10 kcal/mole.

The rotational energy is expressed approximately as

$$E_{rot} = J(J+1)hB_v,$$

where  $J$  is the rotational quantum number, and  $B_v$  is the rotational constant for the  $v$ th vibrational state, given by  $h/(8\pi^2I_v)$ ,  $I_v$  being the moment of inertia of the molecule. The energy separations between two neighboring rotational states are much smaller compared with those of vibrational transitions and on the order of  $10^{-3}$  to  $10^{-2}$  kcal/mole.

The energy separations of electronic levels, in contrast, are much greater because they are the energy differences between different molecular orbitals. As the simplest example, consider a hypothetical single-electron molecule (or atom). Its electronic energy can be approximately expressed as

$$E_n = -\frac{e^2}{2a_0} \left(\frac{1}{n^2}\right) \quad n = 1, 2, 3, \dots$$

where  $e$  is the charge of an electron,  $a_0$  the Bohr radius, and  $n$  the electronic quantum number. The energy separations of electronic states are on the order of 100 kcal/mole, much greater than those for vibrational transitions.

The number of molecules in each of the energy states in a thermal equilibrium has a direct proportional relation to the so-called Boltzmann factor  $e^{-E/kT}$ , where  $k$  is the Boltzmann constant and  $T$  is the absolute temperature. At ambient temperature ( $\sim 300\text{K}$ ),  $kT$  is about 0.6 kcal/mole in energy. Therefore most molecules are in a large number of rotational states associated with very low vibrational levels belonging to the lowest electronic state.

When continuous electromagnetic radiation passes through a material (either neat or in solution), a fraction of the radiation may be absorbed and a spectrum with gaps in it may be obtained when the residual radiation is



detected. Such a spectrum is called an absorption spectrum. The electromagnetic radiation absorbed by the material is spent to promote the atoms or molecules from a state of low energy to one of higher energy. When atoms or molecules in an excited state relax back to a state of lower energy, they lose energy to thermal motions, but a portion of that energy may emit radiation and thus give rise to an emission spectrum. The energy change  $\Delta E$  in the atom or molecule and the frequency  $\nu$  of the radiation absorbed or emitted follows the Bohr condition:  $|\Delta E| = |E_f - E_i| = h\nu$ , in which the subscripts  $f$  and  $i$  denote final and initial states respectively. The absolute value of  $\Delta E$  is referred to as the transition energy, and  $\Delta E$  itself is positive for an absorption process and negative for emission. The frequency  $\nu$  is inversely related to the wavelength  $\lambda$  by  $\nu = c/\lambda$ , where the constant  $c$  is the speed of light.

Molecular absorption spectra, i.e., the absorption spectra of molecules, are classified naturally into three different types: rotation, vibration, and electronic.

Due to the low energy of a rotational transition, the rotation spectrum of a molecule usually occurs in the far infrared and the microwave region, at wavelengths on the order of  $\mu\text{m}$  and  $\text{cm}$ , respectively. The vibration or vibration-rotation spectrum of a molecule, on the other hand, is mostly in the near and middle infrared region (1-30  $\mu\text{m}$ ).

Since electronic spectra correspond to molecular transitions between electronic states, they are of much higher energies than the rotational and vibrational transitions. They usually occur in the ultraviolet (below 380 nm) and visible (380-780 nm) regions. The ultraviolet region is again subdivided into two different regions: far or vacuum UV (below 200 nm) and near or quartz

UV (200-380 nm) regions. The vacuum UV region is not readily accessible, and consequently, for many electronic structures, the absorption is more readily measured in quartz UV or visible regions which are accessible by commercial spectrophotometers. An electronic transition is often accompanied by simultaneous changes in the vibrational and rotational states. Therefore, an electronic absorption usually encompasses a range of wavelengths corresponding to a spectrum with a large number of closely spaced lines, whose fine structure cannot be resolved most of the time. In liquids, solutions, and solids, the fine structure of rotational transitions cannot be resolved; but sometimes that of vibrational transitions are discernible. As a result, the electronic spectra are of relatively broad bands (envelopes). However, when available, fine structures can give valuable information on the conformational properties of a molecule.

An absorption spectrum is described by the absorption intensity as a function of the wavelength, and is plotted conventionally with the absorption intensity as the ordinate and the wavelength as the abscissa. Wavelength  $\lambda$  is usually expressed in the nanometer unit ( $1 \text{ nm} = 10^{-9} \text{ m}$ ). Sometimes, the wave number  $\nu$ , which is defined as the reciprocal of the wavelength, i.e.,  $\nu = 1/\lambda$ , is used. The conventional unit for the wave number  $\nu$  is  $\text{cm}^{-1}$ . The wavelength is directly related to energy by the following relationship

$$E \text{ (kcal/mole)} = hc/\lambda = 28,591.2/\lambda \text{ (nm)}.$$

According to the Beer-Lambert Law on the process of light absorption, we have the following expression for the light intensity  $I$ ,

$$A = \log(I_0/I) = \epsilon cl,$$

where  $I_0$  is the intensity of the incident light, the quantity  $\log(I_0/I)$  is called the absorbance ( $A$ ) or optical density ( $OD$ ), and  $I_0/I$  the transmittance ( $T$ ). Intensity  $I$  is proportional to the intensity of its electronic wave  $E^2$ , [ $I/I_0 \propto (E/E_0)^2$ ], in which  $E$  is the amplitude of the electronic wave and can be measured directly. The quantity  $\epsilon$  is called the molar absorptivity or the molar absorption coefficient or molar extinction coefficient. The concentration  $c$  of the sample and the length  $l$  of the sample cell are expressed in units of moles per liter (M) and in centimeters (cm), respectively. The molar absorption coefficient,  $\epsilon$ , is usually below  $10^5$ .

The following definitions will be used throughout the text. A bathochromic shift (or red shift) refers to the shift of absorption to a longer wavelength due to a substitution or solvent effect; a hypsochromic shift (or blue shift) is the shift of absorption to a shorter wavelength; hyperchromic effect means an increase in absorption intensity; and hypochromic effect is an decrease in absorption intensity.

Circular Dichroism (CD) is closely related to UV absorption spectroscopy. However, since CD measures the absorption difference between left-circularly and right-circularly polarized light, it is a very sensitive photospectroscopic technique for detecting minute conformational changes in asymmetrical molecules. Only an optically active medium can give rise to a CD spectrum. The optical activity arises from the fact that most molecules are chiral, i.e., they have handedness and are not superimposable on their mirror images. An optically active substance could be an enantiomeric molecule, a chiral conformation, or an achiral chromophore in a chiral environment. CD is most commonly expressed by the quantity  $[\theta]$ , called molar ellipticity. The following relations hold:

$$\Delta\varepsilon = \varepsilon_l - \varepsilon_r$$

$$[\theta] = 3298.21\Delta\varepsilon$$

in which  $\varepsilon_l$  and  $\varepsilon_r$  are the molar absorption coefficients for the left-circularly polarized and right-circularly polarized light, respectively. The quantity  $\Delta\varepsilon$  is called molar circular dichroism or circular dichroism. The molar ellipticity will be used throughout the text and occasionally, the molar CD will be used to illustrate certain points.

The most likely sources of error come from the uncertainty in weighing and the fractional impurities that may be present in the samples. The MicroGram-Atic balance used in weighing the samples has an accuracy of 0.01 mg. The samples had been subjected to an extended period of drying *in vacuo* and checked for purity by NMR to be above 95% before they were weighed. The error limits are estimated to be 5% or less in the measurements of the molar absorption coefficients and molar ellipticities in this study.

## 4.2 The tolan chromophore

Tolane is the simplest analogue for the tolanoic template chromophore. The correlation of the molecular conformation of tolane with its electronic absorption spectra had been elucidated in connection with studies on stilbenes, which are closely related to tolane in structure (Schlubach & Franzen, 1951; Suzuki, 1960; Suzuki *et al.*, 1982). According to these studies, the absorption bands in the ultraviolet absorption spectra of tolane are classified into three

groups: A-band centered around 280 nm with significant fine structural detail; B-band around 220 nm, which is actually a group of peaks off the C-band shoulder with some fine structure, and at longer wavelengths in a valley at 235 nm; and C-band at 197 nm. Listed in Table 4.1 are the major peaks in the absorption spectrum of tolane in *n*-heptane. The A-band is a conjugation band due to the extended conjugation of tolane and is polarized along the long axis. Although the B-band is affected by the conjugation, it is considered to be localized in the phenyl rings and is polarized along the short axis (Dale, 1957; Suzuki, 1960). According to Suzuki (1960), transitions at both bands are affected by a conjugation factor  $\rho$ , and as a result, both peak intensities and locations increase with an increased coplanarity. The fine structure observed in the UV spectrum of tolane has been attributed to the stretching frequency of the triple bond (Beale & Roe, 1953; Suzuki, 1960). It has also been proposed that the appearance of the fine structure might be related to the coplanarity of the molecular configuration or to its rigidity (Wiegand & Merkel, 1947; Merkel & Wiegand, 1947 and 1948; Kortüm & Dreesen, 1951).

Table 4.1. Ultraviolet absorption spectrum of tolane in *n*-heptane\*

A-band $\lambda_{\max}$ ( $\epsilon$ )	B-band $\lambda_{\max}$ ( $\epsilon$ )	C-band $\lambda_{\max}$ ( $\epsilon$ )
297 (27500)	237 (6930)	197 (38000)
289 (20800)	232 (6520)	
280 (30800)	222 (16500)	
273 (22400)	217 (18500)	
265 (20600)	212 (19000)	

\* Taken from Suzuki, 1960.

For the tolanoic template-peptide derivatives, the pattern of the three principal absorption bands is expected to remain in their UV spectra, although

the *ortho* substituents may induce changes in wavelength positions and in absorption intensities.

From the conformational analyses of the template-peptide derivatives and model compounds using NMR and molecular mechanics modeling techniques, we concluded in the last chapter that there are three degrees of rotational freedom at the tolan moiety, namely, the interplanar angle  $\alpha$  along the triple bond, the dihedral angle  $\beta$  about the tolanoic carboxyamide group and its attached phenyl ring, and the dihedral angle  $\gamma$  about the tolanoic amide functionality and its attached phenyl ring. The values of these angles obtained from molecular mechanics modeling for a few representative template-peptide derivatives are tabulated below (Table 4.2). Although these values should not be over-interpreted, they do indicate different influences the side chains exert on the tolanoic chromophore.

Table 4.2 Molecular mechanics modeling results (in vacuum).

	GV+	GV-	VA	VI	AV	LV	MV
$\alpha$	17.2	-0.4	-11.8	-10.1	-4.5	-12.2	-12.3
$\beta$	-43.5	29.7	38.3	38.5	33.9	38.8	45.2
$\gamma$	6.7	-10.6	-4.6	-5.3	-11.0	-4.1	-4.6

The UV absorptions in the A-band and the B-band regions of these template-peptide derivatives will be affected by the coplanarity of the tolan moiety. As a result, different absorption intensities and wavelengths may be induced by the variations of these angles. For an isolated amide chromophore, the  $n \rightarrow \pi^*$  transition is weak and the absorption is usually below 220 nm ( $\lambda_{\max}$  210 nm and  $\epsilon < 100$  for hydrogen bonding solvents) and only an  $\alpha, \beta$ -

unsaturated amide shows absorptions between 200-220 nm with appreciable  $\epsilon_{\max} < 10,000$ . For the template-peptide derivatives, an UV absorption above 220 nm is undoubtedly due to the tolanoic chromophore. Although the  $\alpha$  amino acid amide functionalities may contribute some absorption below 220 nm ( $\lambda_{\max}$  190-200 nm,  $\epsilon$  7000-9000), the more intense absorption of the tolanoic chromophore is expected to be predominant in this region as well.

Since CD is a chiroptical phenomenon and because our tolan chromophore has a plane of symmetry, only those electronic absorptions of the tolanoic chromophore that reflect a chiral environment will be selected in the CD spectra. Because of the so-called "one-electron effect", the  $n \rightarrow \pi^*$  transition of an amide can borrow its intensity from the more intense  $\pi \rightarrow \pi^*$  transition and may show appreciable CD bands above 220 nm. The characteristic features of the CD spectra for polypeptides in the  $\beta$ -sheet conformation are a negative band near 216 nm ( $[\theta] \sim -2 \times 10^4$ ) and a positive band between 195 and 200 nm ( $[\theta] \sim 3 \times 10^4$ ) (Sarkar & Doty, 1966; Brahms *et al.*, 1977). Although the wavelength positions of these two bands may vary considerably from case to case, these CDs almost never exceed 250 nm. Therefore a CD above 250 nm is undoubtedly due to the tolanoic chromophore. In addition, the CD contributions from the amide groups of the  $\alpha$  amino acid residues are expected to be minimum because the two ends are relatively free for such short peptides (NMR results in chapter 3, and see Ref. Mattice, 1989; Bradley *et al.*, 1990; Miick *et al.*, 1991; Yapa *et al.*, 1992). Therefore, for CDs that appear below 250 nm, those from the tolanoic template are expected to be dominant for the open chain cases of the template-peptide conjugates. However, in the cyclic template-peptide derivative, CD contributions from the

amide groups of the  $\alpha$  (and  $\beta$ ) amino acid residues may be significant ( $[\theta]$  on the order of  $10^4$ ) and cannot be ignored in interpreting the CD spectrum.

Since the conjugation band (A-band) is invariably polarized perpendicular to the  $\beta$ -hairpin, a negative CD Cotton effect is expected in this region. However, the situation at the B-band is more complex. Different signs of the interplanar angle  $\alpha$  and the dihedral angles  $\beta$  and  $\gamma$  may give rise to different asymmetries at the tolanoic chromophore and as a result, this region may display different CD bands. If the molecular mechanics modeling results are correct that there exist two conformations with opposite signs for these angles, we might expect to see two CD bands with opposite signs in the B-band region as well. However, if the two phenyl planes are near perfect coplanarity, a completely different CD spectrum may be expected for such a conformation. If this is true, the CDs in the B-band region for the cyclic case may become insignificant. In contrast, CDs resulting from the amide groups may become substantial in this region, since the cyclization fixes the molecule in a  $\beta$ -hairpin conformation.

### 4.3 UV results and discussion

Shown in Figure 4.1 is an example of the absorption spectra for the tolanoic peptide derivatives AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe. The basic pattern of the three principal bands is apparent. Listed in Table 4.3 are the major peaks in the absorption spectrum of AcLeu-Taa-ValNHMe (**16e**). The three major bands are: two broad peaks at 315 nm ( $\epsilon$  12400) and 292 nm ( $\epsilon$  16500) in the A-band region, two sharper peaks of similar intensities at 254 nm ( $\epsilon$  30500)



and 247 nm ( $\epsilon$  31300) in the B-band region, and a sharp band at 204 ( $\epsilon$  39100) in the C-band region. The four intense peaks at long wavelengths are undoubtedly due to the tolanoic amino acid residue, and the tolanoic chromophore probably dominates the sharp band at low wavelength as well (204 nm and  $\epsilon$  39100 vs. 197 nm and  $\epsilon$  38000 for tolane).

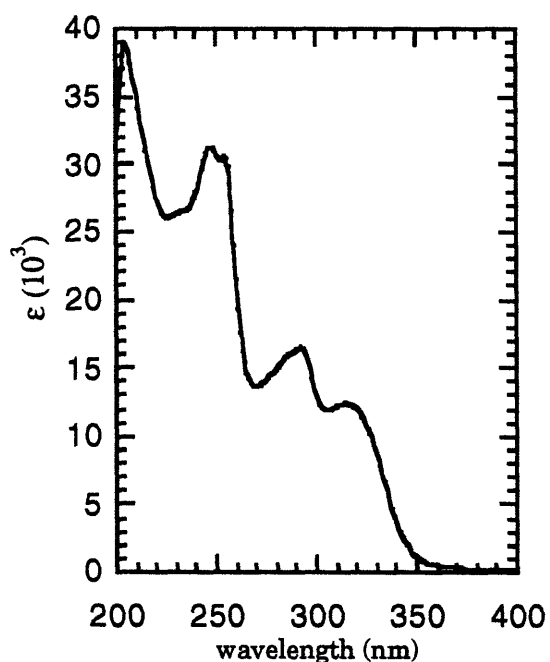


Figure 4.1 UV spectrum of AcLeu-Taa-ValNHMe in EtOH

Table 4.3 Absorption spectrum of AcLeu-Taa-ValNHMe in EtOH at 25 °C.

	A-band $\lambda_{\max}$ ( $\epsilon$ )	B-band $\lambda_{\max}$ ( $\epsilon$ )	C-band $\lambda_{\max}$ ( $\epsilon$ )
$\alpha$	315 (12400)	254 (30500)	204 (39100)
$\beta$	292 (16500)	247 (31300)	

All three bands have bathochromic shifts compared with tolane: the A-band has 12-18 nm red shifts accompanied with large hypochromic effects ( $\Delta\epsilon$  -14300 to -15100); the B-band has 20-30 nm red shifts concurrent with large hyperchromic effects ( $\Delta\epsilon$  12000 to 12300); and the C-band has a relatively small bathochromic shift (~ 7 nm) without much change in the absorption intensity. These values are most certainly  $\pi \rightarrow \pi^*$  transitions from the substituted tolan system. The large hypochromic effects at the A-band region accompanied by the large hyperchromic effects in the B-band region suggests that the A-band is due to the extended tolan conjugation system, whereas the B-band is localized to the aromatic ring systems because increased interactions of the phenyl rings with substitutions at C-2 and C-2' positions will inevitably occur at the cost of reducing their interactions with the acetylene  $\pi$  bonds, unless the two substituents exert a strongly favorable and stable "push-pull" interaction through the triple bond.

The appearance of doublet peaks in the A-band and the B-band regions is very intriguing. At this point there are two plausible explanations. The first hypothesis is that they are the result of two distinctively different and yet very similar conformations of the tolanoic chromophore. The second hypothesis is that these absorption fine structures are due to the stretching motions of the tolanoic chromophore. The energies separating the doublets are 7.15 kcal/mole and 3.19 kcal/mole for the A-band and the B-band, respectively. Although these energy differences are on the same order as vibrational transitions, the first hypothesis cannot be dismissed. If the above doublets observed in the spectrum of AcLeu-Taa-ValNHMe are indeed vibrational fine structures, a near-planar conformation for the tolan-moiety must be assumed. However, it is a well-known fact that a strong polar protic solvent such as EtOH can dilute or completely destroy fine structures. In addition, NMR results indicate that

there is some twist between the two phenyl planes of the tolan chromophore. These pieces of evidence seem to support the first hypothesis. However, further experimental studies are needed to dismiss one of them.

Since the A-band is considered to be a conjugation band resulting from the extended conjugation of the tolan system, the large bathochromic shift compared with tolane are undoubtedly due to *ortho* substitution effects. However, the large hypochromic effect in the A-band region of the template-peptide conjugates is most likely the result of influences from both *ortho* substitutions and from the reduced coplanarity in the tolan moiety as compared with tolane. This seems to support the first hypothesis as well.

The A-band region in the absorption spectrum of Me-Tln(C=O)AlaNHMe (**29**) is quite similar to that of AcNH-Tln-Me (**26**) with only a small decrease in intensity (~ 15%) (Figure 4.2). However, the B-band region is very much different: it disappeared altogether with the fine structure into a valley concurrent with a large decrease in intensity (30%). These changes also support the conclusion that the A-band is the electronic absorption due to the extended conjugation network of the tolan system, which is modulated by the amide substitution at C-2 and the carboxamide at C-2'. They also suggest that the fine structure and most of the absorption intensities in the B-band region come from the interaction of the aromatic amide functional group with the tolan system. This is consistent with NMR results in which the tolanoic carboxamide is considerably off coplanarity with its phenyl plane and enjoys a greater degree of rotational freedom than the tolanoic amide functionality. The similarity between AcNH-Tln-Me and Me-Tln(C=O)AlaNHMe in the A-band region also indicates that substituents at C-2 and C-2' exert their influences on the tolan chromophore through a more direct electron-donating (or accepting) effect rather than an extended resonance between the tolanoic amide and the

tolanoic carboxyamide groups via the central triple bond. Therefore resonance interactions of the substituents with the tolan moiety are considered small.

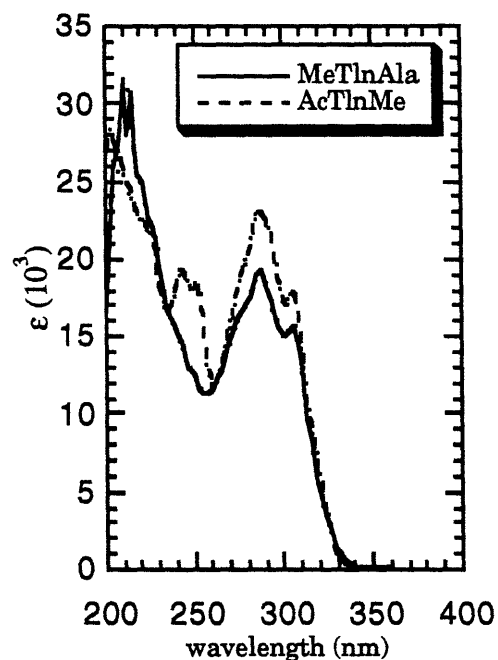


Figure 4.2 AcNH-Tln-Me versus Me-Tln(C=O)AlaNHMe.

Large hypsochromic shifts (3-9 nm) in the A-band region and smaller hypsochromic shifts (2-3 nm) in the B-band region of the absorption spectrum for AcNH-Tln-Me (**26**) were observed in comparison with Ac-Taa-ValNHMe (Figure 4.3). In addition, there is a large hyperchromic effect (60-70% increase) in the A-band region, and in stark contrast, a modest hypochromic effect (20% decrease) in the B-band region (Figure 4.3). These opposite changes in intensities for the A-band and the B-band regions are strong evidence for the reproducibility of the data. In addition these changes are in full accord with the assignment of the A-band as the conjugation band and the B-band as the more

localized transitions of the phenyl rings. A substitution on the phenyl ring increases transition probabilities in the phenyl ring but at a cost to the conjugation with the central triple bond.

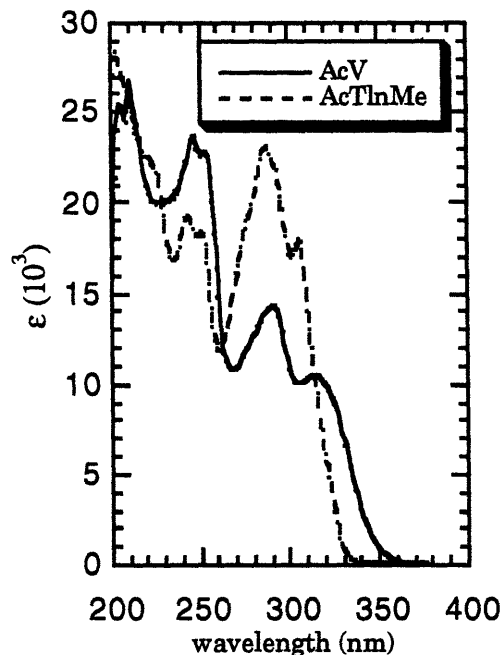


Figure 4.3 AcNH-Tln-Me versus Ac-Taa-ValNHMe.

The added steric bulkiness of the  $\alpha$ -substitution in the  $\alpha$ -aminoisobutyric acid residue ought to deter the two Aib residues from coming together to form a sheet-like structure. According to the NMR studies and the molecular mechanics modeling results, AcAib-Taa-AibNHMe (**21**) did not display a  $\beta$ -hairpin conformation but had a completely different conformation in which the tolanoic turn hydrogen bond was nonexistent and the two phenyl planes were almost perpendicular to each other with no interactions across the chains whatsoever. Therefore a derivative such as **21** should provide useful

spectral information on the tolan moiety in the absence of the  $\beta$ -hairpin structure. It is remarkable that the derivative AcAib-Taa-AibNHMe (**21**) has a very uninteresting UV absorption spectrum, shown in Figure 4.4 against AcAla-Taa-ValNHMe, in which none of the characteristic absorption bands for the tolanoic  $\beta$ -hairpin is visible. This almost certainly suggests that when the  $\beta$ -hairpin and especially, when the first hydrogen bond in the  $\beta$ -hairpin is destroyed, the tolanoic chromophore loses its characteristic UV absorption bands. It also suggests that a geometry close to the coplanarity for the tolan moiety, although not necessarily a prerequisite, is important for the characteristic absorptions. Therefore, the interplanar angle  $\alpha$  cannot be too large in the  $\beta$ -hairpin conformation.

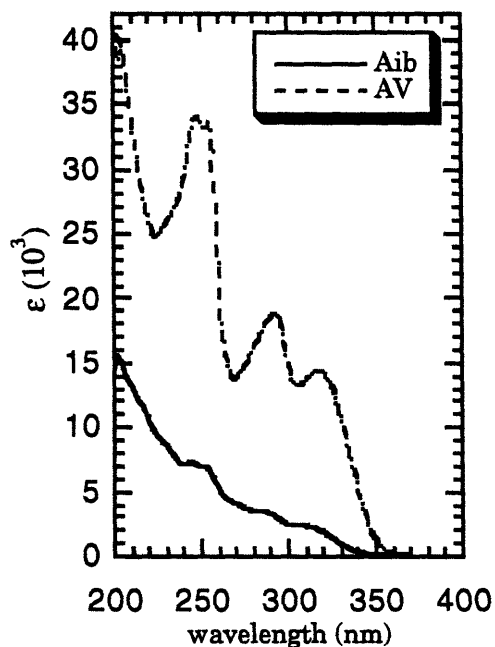


Figure 4.4 AcAib-Taa-AibNHMe versus AcAla-Taa-ValNHMe in EtOH

The absorption spectrum for Ac-Taa-ValNHMe is quite similar to that for AcAla-Taa-ValNHMe (Figure 4.5). These two derivatives are different in that as compared with the latter derivative, the former compound is truncated and can only form one sheet-like hydrogen bond.

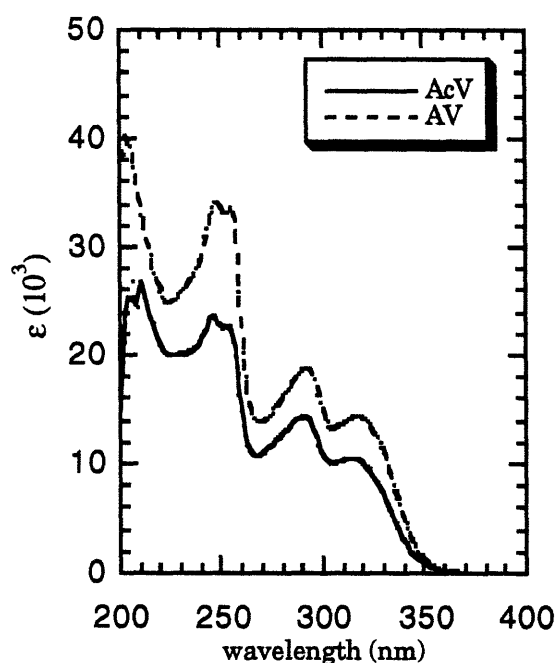


Figure 4.5 Ac-Taa-ValNHMe  
versus AcAla-Taa-ValNHMe in EtOH

The overall shape of the spectrum is conserved. However, when compared with that for AcAla-Taa-ValNHMe, the spectrum for Ac-Taa-ValNHMe exhibits noticeable hypsochromic shifts (2-3 nm for the A band and 2 nm for the B-band) and significant hypochromic effects in both the A-band (20-30%) and B-band (30%) regions (Table 4.3). These results indicate that

these characteristic absorption peaks (A-band and B-band) are dominated by the tolanoic chromophore, and the amino acids residues linked to the template do affect the UV spectra. They also indicate that the first hydrogen bond alone is enough to give the UV spectral shape that is characteristic of the tolanoic  $\beta$ -hairpin, and the formation of the second hydrogen bond between the methylamide NH and the acetylcarbonyl oxygen reinforces the  $\beta$ -hairpin structure.

Table 4.3 Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe in EtOH.

Tolan derivatives	A-band			B-band		
	$\lambda_{\max}$	$\epsilon$	ratio	$\lambda_{\max}$	$\epsilon$	ratio
AcAla-Taa-ValNHMe	317	14400	1.00	255	33600	1.00
	292	18900	1.00	248	34100	1.00
Ac-Taa-ValNHMe	314	10500	0.73	253	22800	0.68
	290	14300	0.76	246	23700	0.69

The absorption spectra for AcGly-Taa-AlaNHMe (**6a**), AcGly-Taa-IleNHMe (**6b**), and AcGly-Taa-ValNHMe (**16b**) are shown in an overlay plot in Figure 4.6. All three spectra have well-defined absorption peaks similar to those of AcLeu-Taa-ValNHMe. The trend in the absorption intensities for these three derivatives follow that observed in the chemical shift nonequivalences of the glycy methylene protons observed in the last chapter: namely, it follows in an ascending order from AcGly-Taa-AlaNHMe to AcGly-Taa-IleNHMe to AcGly-Taa-ValNHMe. The ratios calculated from absorption intensities using AcGly-Taa-AlaNHMe as a reference are 1.2 and 1.3 for AcGly-Taa-IleNHMe and AcGly-Taa-ValNHMe (Table 4.4), whereas the ratios from the NMR nonequivalences of the methylene protons are the same (2.4) for these two derivatives. Although the difference between the ratios 1.2 and



1.3 approach experimental error limits, it does seem to indicate that the UV spectra are more sensitive to the molecular conformations than the NMR spectra.

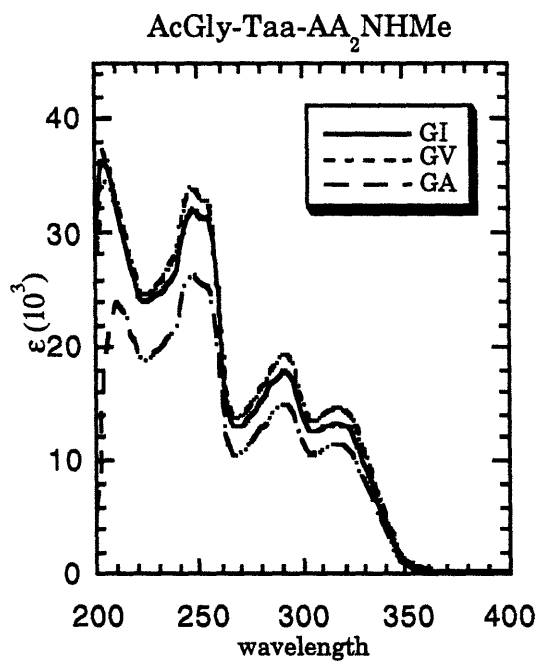


Figure 4.6 Overlay UV spectra of AcGly-Taa-AlaNHMe, AcGly-Taa-IleNHMe, and AcGly-Taa-ValNHMe in EtOH.

Table 4.4 Selected absorption data for AcGly-Taa-AA<sub>2</sub>NHMe in EtOH.

Tolan derivatives	A-band			B-band		
	$\lambda_{\max}$	$\epsilon$	ratio	$\lambda_{\max}$	$\epsilon$	ratio
AcGly-Taa-AlaNHMe	316	11400	1.00	254	25500	1.00
	292	15000	1.00	248	26300	1.00
AcGly-Taa-IleNHMe	316	13300	1.16	254	31400	1.23
	292	17900	1.19	247	32000	1.22
AcGly-Taa-ValNHMe	317	14700	1.29	254	33000	1.30
	292	19400	1.29	246	34000	1.29

The UV absorption spectra for the glyceryl tolanoic derivatives AcGly-Taa-AA<sub>2</sub>NHMe are very similar to the other template-peptide derivatives AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe as well. Shown in Figure 4.7 is an overlay plot of the UV spectra for AcGly-Taa-ValNHMe and for AcAla-Taa-ValNHMe (**16a**). Notice the striking similarity between these two spectra -- they are almost superimposable.

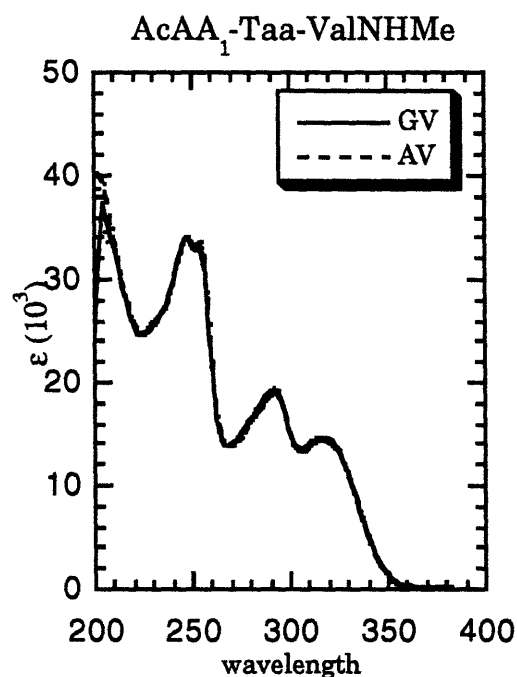


Figure 4.7 Comparison of AcGly-Taa-ValNHMe with AcAla-Taa-ValNHMe in EtOH

A closer examination (Table 4.5) of these two spectra only revealed minute changes. In the B-band region, there is a small but nonetheless visible bathochromic shift of 1-2 nm for AcAla-Taa-ValNHMe as compared with its

glycyl counterpart. Perhaps this is a reflection of the subtle difference in conformation between these two derivatives as revealed through molecular mechanics modeling. However, the overall shape of these two spectra is the same and their molar absorption coefficients are the same within the experimental errors.

Table 4.5 Comparison of absorption data for AcAA<sub>1</sub>-Taa-ValNHMe in EtOH.

Tolan derivatives	A-band			B-band		
	$\lambda_{\max}$	$\epsilon$	ratio	$\lambda_{\max}$	$\epsilon$	ratio
AcGly-Taa-ValNHMe	317	14700	1.02	254	33000	0.98
	292	19400	1.03	246	34000	1.00
AcAla-Taa-ValNHMe	317	14400	1.00	255	33600	1.00
	292	18900	1.00	248	34100	1.00

In the two absorption spectra (Figure 4.8) for AcAla-Taa-ValNHMe (**16a**) and AcVal-Taa-AlaNHMe (**10**), where the two amino acid residues Ala and Val are swapped, the overall spectral shape is the same. Compared with AcAla-Taa-ValNHMe, **10** has only a small hypsochromic shift (1 nm) (Table 4.6). This information confers the same conclusion reached from NMR analysis: the molecular conformations for the two derivatives are similar. However, the large differences in absorption intensities are very revealing. The peak intensities in the absorption spectrum of **10** are only 60% of those observed for AcAla-Taa-ValNHMe (**16a**). This almost certainly reflects the different degrees of the structural conformation that is present in the two derivatives, i.e., AcAla-Taa-ValNHMe is more structured than AcVal-Taa-AlaNHMe. It seems to be the consensus between the NMR analysis and UV study that it is beneficial to have a good sheet-forming amino acid residue at

site-2 of these template-peptide derivatives, due to the observation that the angle  $\beta$  is the most flexible, and thus is the most decisive factor for the degree of overall sheet conformation in these template-peptide conjugates.

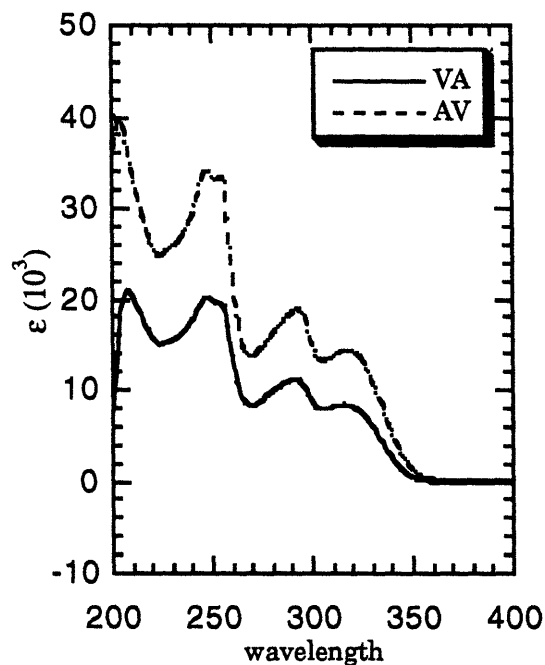


Figure 4.8 AcAla-Taa-ValNHMe versus AcVal-Taa-AlaNHMe in EtOH

Table 4.6 AcAla-Taa-ValNHMe versus AcVal-Taa-AlaNHMe in EtOH.

Tolan derivatives	A-band			B-band		
	$\lambda_{\max}$	$\epsilon$	ratio	$\lambda_{\max}$	$\epsilon$	ratio
AcAla-Taa-ValNHMe	317	14400	1.00	255	33600	1.00
	292	18900	1.00	248	34100	1.00
AcVal-Taa-AlaNHMe	316	8400	0.58	254	19700	0.59
	292	11100	0.59	248	20200	0.59

From the NMR studies on the acetylprolyl tolanoic derivatives, we concluded that the prolyl residue can form sheet structure, but perhaps with a larger molecular twist due to the restricted prolyl angle  $\phi$ . The difference between a sheet-contributing conformation and a non-sheet conformation is reflected in the *trans/cis* ratio of the acetylprolyl amide isomers. The above examples indicate that UV absorption bands, especially in the A-band and B-band regions, are sensitive to the molecular conformations of the template-peptide derivatives. This should be true with the proline-containing tolanoic derivatives as well.

Shown in Figure 4.9 and in Table 4.7 are the UV absorption spectra and peak listings for AcPro-Taa-AlaNHMe (**6c**) and AcPro-Taa-ValNHMe (**16g**) compared against AcAla-Taa-ValNHMe (**16a**). There is little change in  $\lambda_{\max}$  (only one 2 nm hypsochromic shift for AcPro-Taa-ValNHMe compared with AcAla-Taa-ValNHMe), and the overall spectral shape is preserved compared with the absorption spectrum of AcAla-Taa-ValNHMe. However, absorption intensities at A-band and B-band regions are considerably higher as compared with those in the spectrum of AcVal-Taa-AlaNHMe (Tables 4.6 and 4.7). The large twist in the sheet-like structure due to the presence of the prolyl residue undoubtedly affects the conformation of these prolyl derivatives. It is noteworthy that the peak intensities in these UV spectra follow the same trend as those observed for the *trans/cis* ratios by using NMR, although the ratio of the peak intensities is quite different from the ratio of the *trans/cis* ratios observed by NMR in DMSO (1.2 versus 1.5) for these two prolyl derivatives. Again, the UV spectrum is sensitive to the changes of the molecular conformations in the prolyl series.

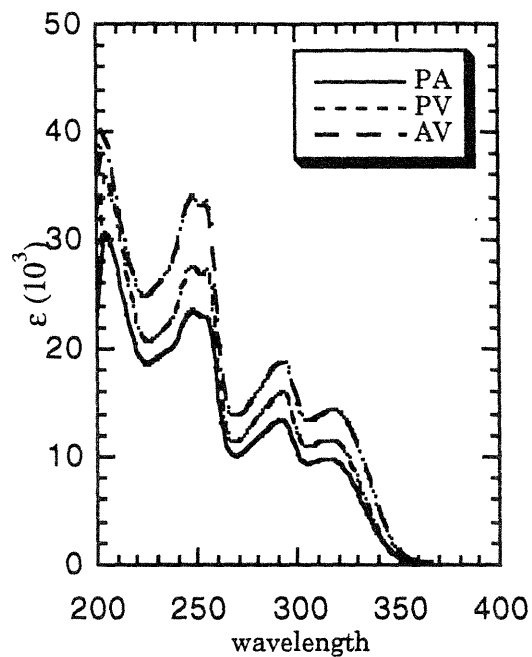


Figure 4.9 AcPro-Taa-AA<sub>2</sub>NHMe  
versus AcAla-Taa-ValNHMe in EtOH

Table 4.7 AcPro-Taa-AA<sub>2</sub>NHMe versus AcAla-Taa-ValNHMe in EtOH.

Tolan derivatives	A-band			B-band		
	$\lambda_{\max}$	$\epsilon$	ratio	$\lambda_{\max}$	$\epsilon$	ratio
AcAla-Taa-ValNHMe	317	14400	1.00	255	33600	1.00
	292	18900	1.00	248	34100	1.00
AcPro-Taa-ValNHMe	315	11600	0.80	255	27300	0.81
	292	16000	0.85	248	27600	0.81
AcPro-Taa-AlaNHMe	315	9800	0.68	255	23100	0.69
	291	13500	0.71	247	23600	0.69

All the tolanoic tripeptides  $\text{AcAA}_1\text{-Taa-ValNHMe}$  display the characteristic A-band and B-band in their absorption spectra (with shifts of only 1-2 nm), but considerable changes occur in the absorption intensities (Figure 4.10).

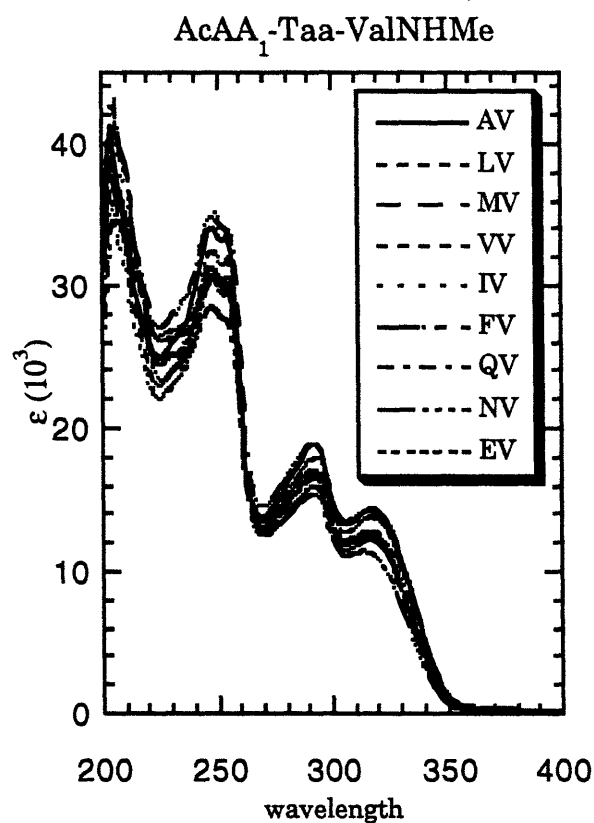
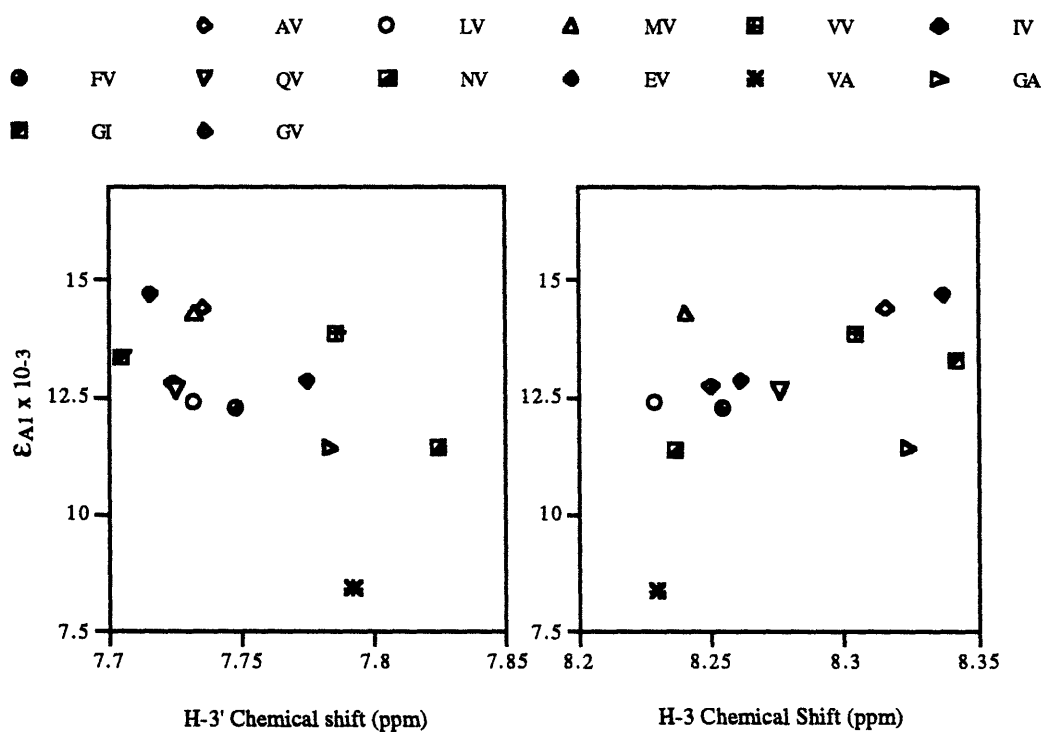


Figure 4.10 Overlay UV spectra of selected valinyl series  $\text{AcAA}_1\text{-Taa-ValNHMe}$  in EtOH.

Plotting the molar absorption coefficients against the chemical shifts of either H-3 or H-3' revealed some interesting patterns. Although the points are scattered about in the plots (Figure 4.11), the two plots follow quite different orders. On the one hand, the correlation plot with the H-3' chemical shifts

reveals a negative trend -- the UV absorption intensity decreases with an increasing H-3' chemical shift (Figure 4.11, left plot). On the other hand, the UV absorption intensity increases with an increasing H-3 chemical shift (Figure 4.11, right plot). These trends are consistent with the  $\beta$ -hairpin structure, in which the rotation of the interplanar angle  $\alpha$  is correlated with the dihedral angles  $\beta$  and  $\gamma$  at the C-2 and C-2' substitution sites.



**Figure 4.11** Correlation of UV absorption with NMR chemical shift. Absorptivity versus H-3' (left) and versus H-3 chemical shift (right).



If the hypothesis is correct that the intensities of these absorption bands directly relate to the degree of the coplanarity of the tolan moiety (and the overall sheet structure), derivatives containing alanyl residue at site-2 (PA, GA, and VA) are far less planar than other derivatives containing valinyl residue at site-2. Furthermore, prolyl residue also decreases the coplanarity, and asparagine is less efficient at inducing coplanarity than glutamine or glutamic acid. However, a more rigorous correlation between the UV absorption spectra and the molecular conformations of tolanoic derivatives cannot be established at this point.

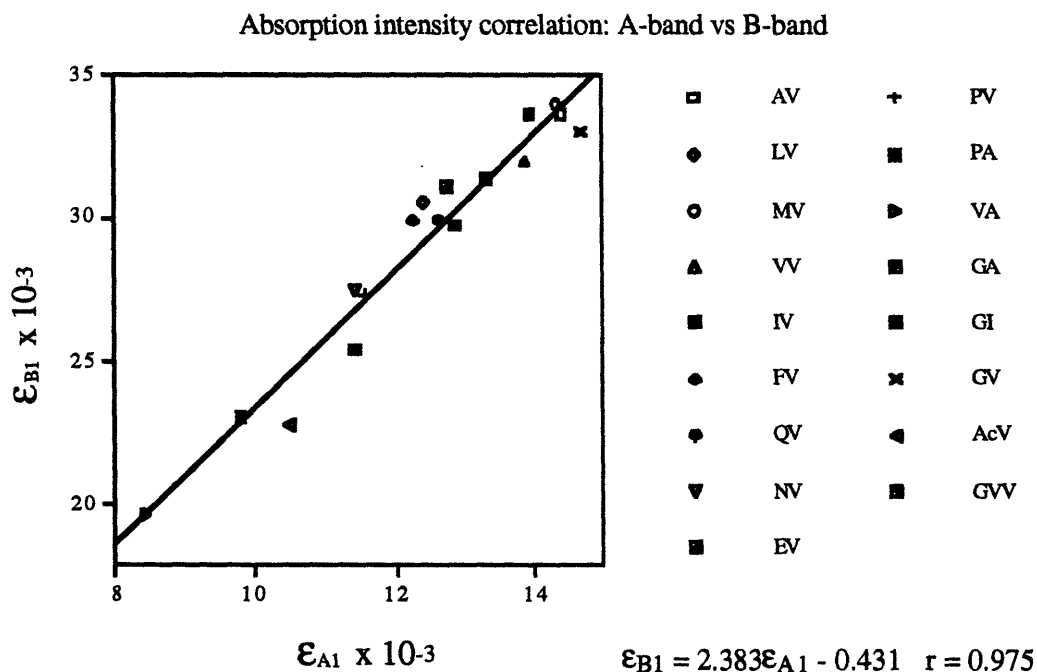


Figure 4.12 Correlation between the A-band and the B-band regions for the tolanoic peptides AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (Ac-Taa-ValNHMe and Gly-Val-Taa-ValNHMe) are included for comparison).

It is noteworthy that for all the tolanoic peptide derivatives examined, the absorption intensities in the A-band region have a non-random correlation with those in the B-band region. In Figure 4.12 plotted are the absorption intensities in the B-band region  $\lambda_{\max} \approx 255$  nm against those in the A-band region  $\lambda_{\max} \approx 316$  nm for seventeen template-peptide derivatives (model compounds Ac-Taa-ValNHMe and Gly-Val-Taa-ValNHMe are included for comparison). All the points fall near a straight line (with a correlation  $r = 0.98$ ). The ratio of the molar absorption coefficients  $\epsilon_{B1}/\epsilon_{A1}$  remains, for all derivatives, a near constant number of 2.4.

The UV spectrum for the cyclic template derivative cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] is quite different from those for the open-chain derivatives AcAA<sub>1</sub>-Taa-AA<sub>2</sub>NHMe (Table 4.8 and Figure 4.13). The fine structures observed previously in the A-band and the B-band regions are no longer there, and in their place, single smooth absorption bands appear. This is a strong piece of evidence in support of the first hypothesis that the fine structure in these UV spectra is the result of two closely related and yet distinctively different conformations. It can be speculated that these conformations might be the “+” and the “-” conformations generated by the molecular mechanics modeling. Since these are conformations of different handedness, the chiroptical CD method might be able to distinguish between them.

Table 4.8 Cyclo [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala]  
versus AcVal-Taa-AlaNHMe in EtOH.

	A-band		B-band		A/B ratio
	$\lambda_{\max}$	$\epsilon$	$\lambda_{\max}$	$\epsilon$	
Tolan derivatives					
cyclic derivative	317	8300	254	14700	0.57
AcVal-Taa-AlaNHMe	316	8400	254	19700	0.43

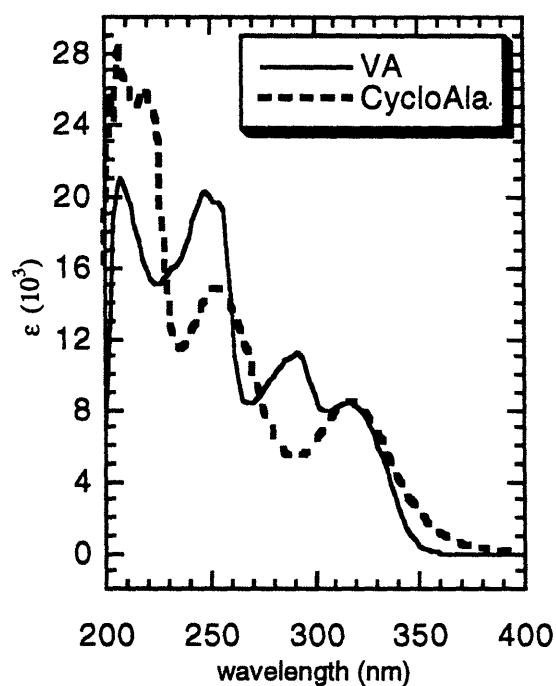


Figure 4.13 UV spectra of the cyclic derivative and AcVal-Taa-AlaNHMe in EtOH.

Since the two ends of the  $\beta$ -hairpin are now linked in the cyclic derivative, the coplanarity at the tolan chromophore is expected to increase, and as a result the conjugation band A-band should have an increase in the absorption intensity. In contrast, as the result of an imperfect  $\beta$ -hairpin template at the tolanoic carboxamide and the phenyl ring junction, the coplanarity of the tolanoic amide with its attached phenyl ring (angle  $\gamma$ ) is expected to decrease and as a result, the intensity of the localized B-band

should decrease as well. This is indeed the case observed. The ratio of the intensities of the A-band over the B-band for the cyclic derivative is considerably higher (~ 30%) than that for the open-chain case AcVal-Taa-AlaNHMe. Although the intensities of the longer wavelengths are used for the doublets in the A-band and the B-band regions for AcVal-Taa-AlaNHMe, the result should be the same if the other two peaks are chosen. In addition, all the open-chain template-peptide derivatives have a similar ratio for these two bands (Figure 4.12). Therefore the conformation for the cyclic case may be quite different from the open chain derivatives.

#### 4.4 CD results and discussion.

The electronic absorption spectrum (Figure 4.14, top plot) for AcAla-Taa-ValNHMe (**16a**) is compared with its CD spectrum (bottom plot). There are three regions of intense UV absorptions: two broad peaks at 317 nm ( $\epsilon$  12400) and at 292 nm ( $\epsilon$  18900), two sharper peaks at 255 nm ( $\epsilon$  33600) and at 248 nm ( $\epsilon$  34100), and a sharp peak at 202 nm ( $\epsilon$  41800). In these corresponding regions, the CD spectrum of **16a** exhibits four major intense Cotton effects (CE):  $\lambda_{\max}$  319 nm ( $\Delta\epsilon$  -5.5),  $\lambda_{\max}$  252 nm ( $\Delta\epsilon$  + 40.3),  $\lambda_{\max}$  229 nm ( $\Delta\epsilon$  -34.5), and  $\lambda_{\max}$  205 nm ( $\Delta\epsilon$  +12.7). Besides these four major intense CD CEs, there is a negative shoulder CD near 290 nm and off the 319 nm band. The negative CD band and the shoulder peak at the long wavelengths correspond to the A-band region of the UV spectrum, whereas the large positive CD at 252 nm and its negative couplet at 229 nm correspond to the B-band region of the UV spectrum, and the small positive CD at 205 nm corresponds to the C-band of the UV spectrum.

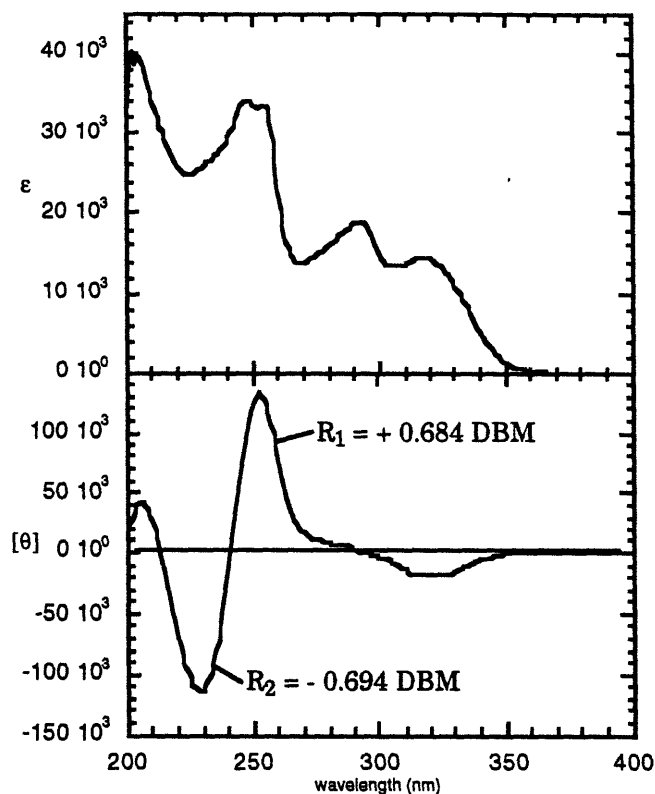


Figure 4.14 Electronic absorption spectrum (top plot) and CD spectrum (bottom) for AcAla-Taa-ValNHMe in EtOH.

The rotational strength values calculated for the two most intense CD bands indicate that the large negative CD at  $\lambda_{\max}$  229 nm and the large positive CD at  $\lambda_{\max}$  252 nm might be coupled. The rotational strength  $R_1$  and  $R_2$  derived from the integrated areas of the CEs at 252 nm and at 229 nm are almost identical: +0.68 Debye Bohr magnetons (DBM) and -0.69 DBM, respectively. They also have much bigger anisotropy factors:  $g$   $1.2 \times 10^{-3}$  and  $g$   $-1.3 \times 10^{-3}$  for  $\lambda$  252 nm and  $\lambda$  229 nm, respectively, compared with  $g$   $-3.9 \times 10^{-4}$  at  $\lambda$  319 nm.

This CD couplet has a positive amplitude of +75. The molar ellipticities of these two CD bands are large compared with polypeptides:  $[\theta]$   $1.1 \times 10^5$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  at 229 nm and  $[\theta]$   $1.3 \times 10^5$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  at 252 nm for AcAla-Taa-ValNHMe. Although the amide chromophores in these derivatives may also contribute to the CD band at 229 nm, it is almost certain that it will be dominated by the tolanoic chromophore. Therefore, these major CD Cotton effects are related to the tolanoic chromophore. As predicted, the CDs for the conjugation band are negative, whereas for the B-band region both a negative and a positive CD are observed.

The CD spectra for the eighteen derivatives, including two controls AcTaa-ValNHMe and Gly-Val-Taa-ValNHMe, three glycylic derivatives AcGly-Taa-AA<sub>2</sub>NHMe (AA<sub>2</sub> = Ala, Ile, and Val), two prolyl derivatives AcPro-Taa-AA<sub>2</sub>NHMe (AA<sub>2</sub> = Ala and Val), one AcVal-Taa-AlaNHMe, one AcVal-Taa-IleNHMe, and nine valinyl derivatives AcAA<sub>1</sub>-Taa-ValNHMe (AA<sub>1</sub> = Ala, Met, Leu, Val, Ile, Phe, Asn, Gln, and Glu), are shown in Figure 4.15.

The CDs at the long wavelengths corresponding to the A-band in the UV spectra are indeed invariably negative for all derivatives, whereas the CDs in the B-band region all appear as a couplet. The negative shoulder peak at 290 nm is more visible in this overlay plot. One noticeable feature of these spectra is that they are all concentrated in the same region with considerable variation in intensities, indicating similar conformations among these derivatives. For example, the CD intensities at 252 nm range from a high  $[\theta] = 2 \times 10^5$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcPhe-Taa-ValNHMe and AcMet-Taa-ValNHMe to a low  $[\theta] = 8 \times 10^3$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcGly-Taa-AlaNHMe.

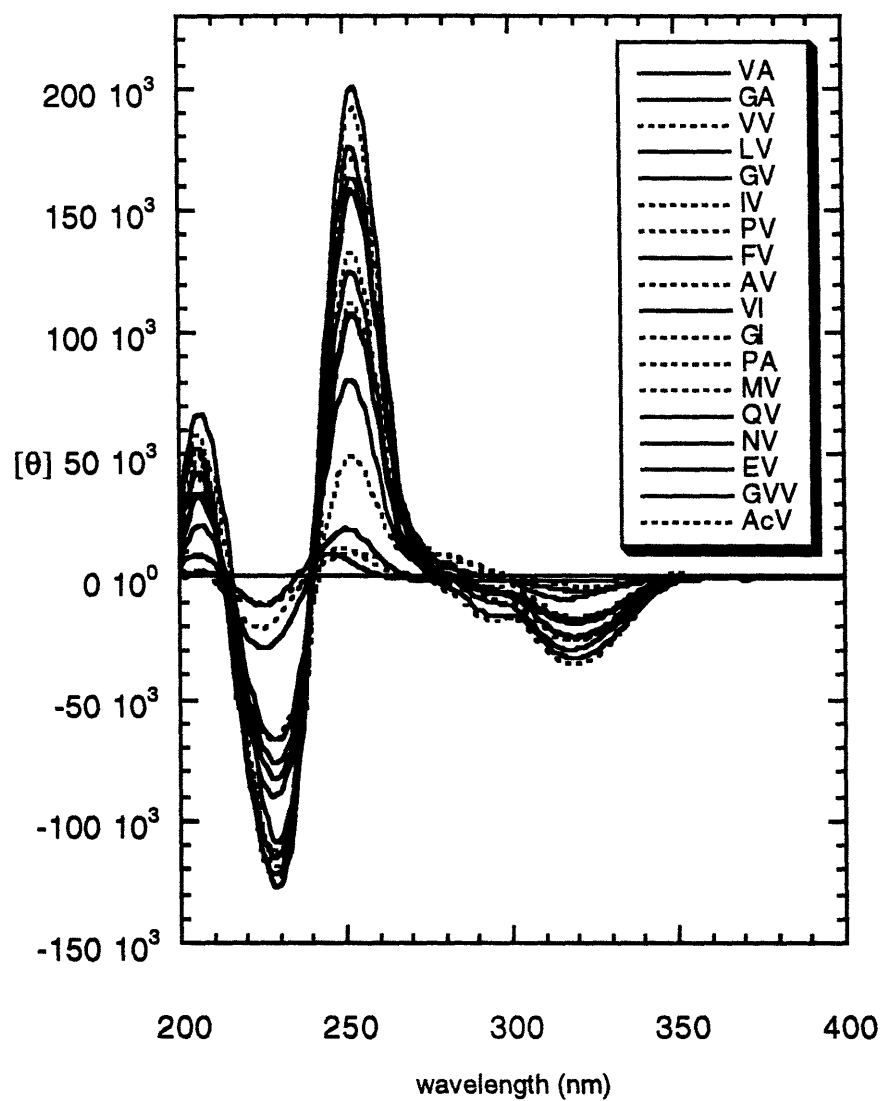


Figure 4.15 CD spectra of template-peptide derivatives in EtOH at 25°.

There are only small changes in wavelength positions of the CD Cotton effects for all derivatives. For the first negative CD, it ranges from 321 nm for the three glycyI derivatives to 320 nm for AcPro-Taa-ValNHMe to 319-318

nm for most template-peptides to 317 nm for AcVal-Taa-AlaNHMe to 316 nm for AcPro-Taa-AlaNHMe and AcAsn-Taa-ValNHMe. The  $[\theta]$  value ranges from -1400 degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcGly-Taa-AlaNHMe to -35700 degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcMet-Taa-ValNHMe. The large positive CD spans from 248 nm for AcGly-Taa-IleNHMe to 252 or 253 nm for most derivatives. The third major CD (negative CE) changes from 224 nm for AcGly-Taa-IleNHMe to 228 or 229 nm for most derivatives. The  $[\theta]$  values are as large as those seen for the intense positive CD CE and they range from  $-1.2 \times 10^4$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcGly-Taa-AlaNHMe to  $-1.3 \times 10^5$  degrees  $\text{cm}^2 \text{dmol}^{-1}$  for AcLeu-Taa-ValNHMe. The variation in the positions of CD CEs is no more than 5 nm, but intensity can change by a factor as large as 26. The correlation between these CD bands and their relation to the  $\beta$ -hairpin conformation will be examined.

The CD spectra of the glycine-containing tolanoic derivatives are visibly different from the other template-peptide derivatives (Figure 4.16 and Table 4.9).

A comparison of AcGly-Taa-ValNHMe and AcAla-Taa-ValNHMe reveals large differences between these two spectra. The three major intense CD Cotton effects for AcGly-Taa-ValNHMe all display certain  $\lambda_{\text{max}}$  shifts accompanied by large decreases in intensities compared with AcAla-Taa-ValNHMe: a bathochromic shift of 2 nm with a hypochromic effect ( $\Delta[\theta]/[\theta]_{\text{AV}}$ ) of -69% near 320 nm, hypsochromic shifts of 2 nm and 3 nm with hypochromic effects of  $\Delta[\theta]/[\theta]_{\text{AV}}$  -85% and -74% near 252 nm and 229 nm, respectively. This is contrary to the UV results in which these two derivatives have nearly identical UV spectra. The findings strongly indicate that the CD Cotton effects are indeed more sensitive to the conformational changes than the UV absorption method.



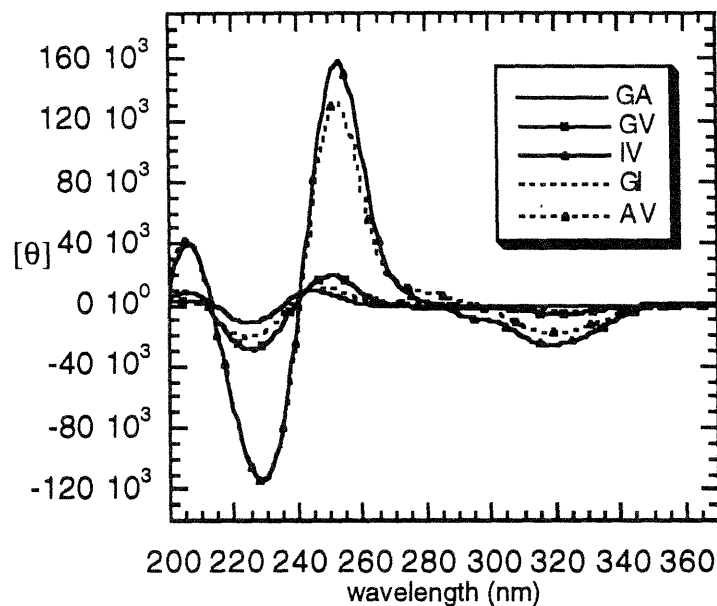


Figure 4.16 CD spectra of glycyll tolanoic derivatives versus non-glycyll tolanoic derivatives in EtOH

Table 4.9 Comparison of CD intensities of glycyll template-peptides with other template-peptide derivatives in EtOH

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-[\theta]$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-[\theta]_2$	$-[\theta]_1/[\theta]_2$
AcAla-Taa-ValNHMe	319	18200	252	133000	229	113700	1.17
AcIle-Taa-ValNHMe	318	26000	253	159000	229	114000	1.39
AcGly-Taa-AlaNHMe	321	1400	249	7900	225	11700	0.68
AcGly-Taa-IleNHMe	321	3600	248	11100	224	20700	0.54
AcGly-Taa-ValNHMe	321	5700	250	19500	226	29200	0.67

The ratios of the peak intensities of the coupled positive CE CD versus the negative CD CE revealed a pattern in which the values are very different for the glycine-containing derivatives compared with the other template-peptide conjugates. All of the three ratios for AcGly-Taa-AA<sub>2</sub>NHMe (AA<sub>2</sub> = Ala, Ile, Val) are significantly lower than unity (range 0.5-0.7), whereas the ratios for AcAA<sub>1</sub>-Taa-ValNHMe (AA<sub>1</sub> = Ala, Ile, for example) are conspicuously above 1.00 (1.2 and 1.4 respectively). These findings are consistent with the two-conformations hypothesis. The two CD CEs of opposite signs in the B-band region correlate with two rotational conformers in which the signs of the angle  $\alpha$  are reversed.

If the hypothesis is correct that the coupled CD CEs are directly related to the two  $\beta$ -hairpin conformations that have opposite signs of the angle  $\alpha$ , a further proof should come from the CD spectrum for the control compound Ac-Taa-ValNHMe. If Ac-Taa-ValNHMe exhibits any CD CEs, the ratio of the coupled CD CEs should approach unity.

The CD spectrum for Ac-Taa-ValNHMe indeed displays non-vanishing CD Cotton effects (Figure 4.17 and Table 4.10). Just as predicted, the ratio of the split CD CEs is 0.94. It shows that the two conformers are equally probable in this case. In addition, these CD CEs are far less intense than either AcAla-Taa-ValNHMe or AcGly-Taa-ValNHMe. For example, the large positive CE is similar to AcGly-Taa-ValNHMe in shape but is only 52% as intense, and is less than 8% as that of AcAla-Taa-ValNHMe

The results indicate that the CD spectrum is indeed sensitive to the conformations of interest. Moreover, it strongly suggests that the negative CD band near 229 nm originates from the tolanoic chromophore and not from the amide functionalities.

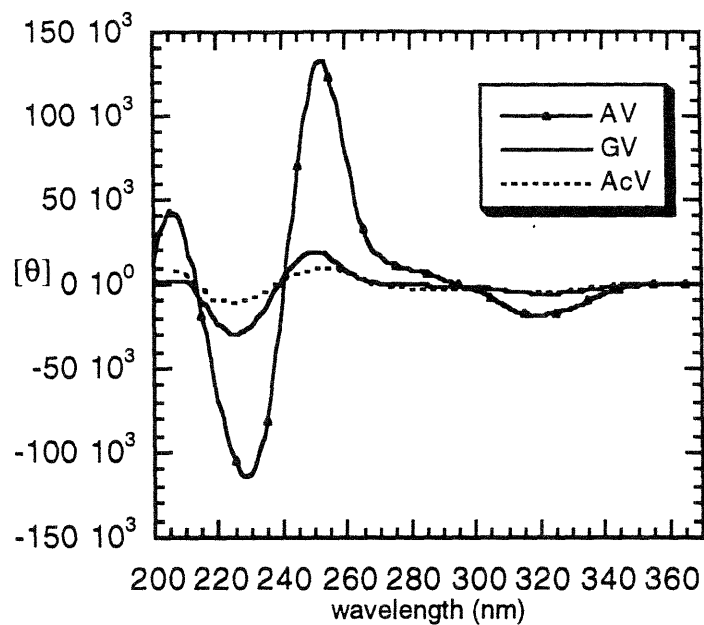


Figure 4.17 Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe and AcGly-Taa-ValNHMe.

Table 4.10 Comparison of CD intensities of Ac-Taa-ValNHMe versus AcAla-Taa-ValNHMe and AcGly-Taa-ValNHMe

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-[\theta]$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-[\theta]_2$	$-[\theta]_1/[\theta]_2$
AcAla-Taa-ValNHMe	319	18200	252	132900	229	113700	1.17
AcGly-Taa-ValNHMe	321	5700	250	19500	226	29200	0.67
Ac-Taa-ValNHMe	317	4900	253	10200	225	10900	0.94

The CD Cotton effects for the prolyl template-peptides are substantially more intense than the glycyll template-peptides, but much less so than other template-peptides (Figure 4.18). For example AcPro-Taa-ValNHMe is 3-6 times as intense as AcGly-Taa-ValNHMe, and has about 80-100% of the intensity for AcAla-Taa-ValNHMe (Table 4.11).

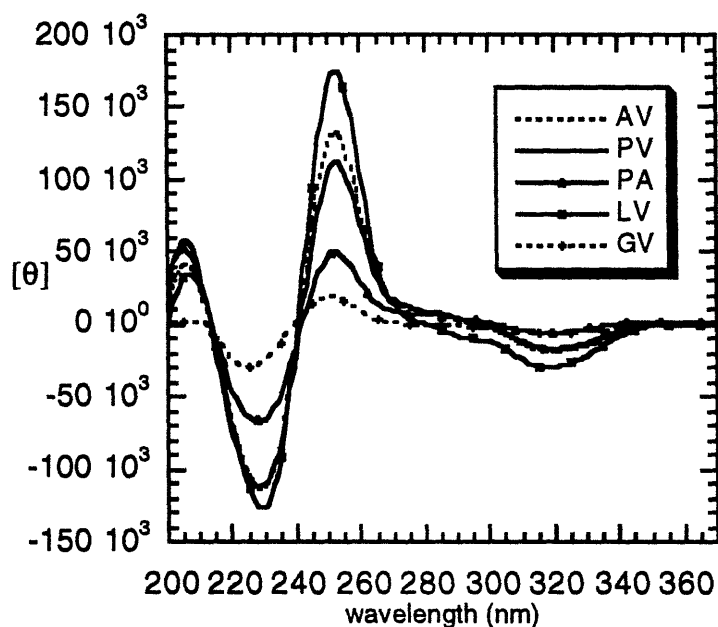


Figure 4.18 CD spectra of prolyl tolanoic peptides versus other tolanoic peptides in EtOH

The intensity ratios of the coupled CD CEs are about 1.00 and 0.74 for AcPro-Taa-ValNHMe and AcPro-Taa-AlaNHMe, respectively, indicating that the conformational behavior of AcPro-Taa-AlaNHMe is perhaps similar to the glycyll derivatives and that AcPro-Taa-ValNHMe is a border-line case and it behaves conformationally similar to AcAla-Taa-ValNHMe. This is consistent with NMR studies in DMSO, which indicate that the prolyl-alanyl derivative is much less structured than its valinyl analogue, in a ratio of 1.5:1 favoring the

valinyl derivative. The ratio of the values ( $-\theta_1/\theta_2$ ) for AcPro-Taa-ValNHMe and for AcPro-Taa-AlaNHMe is 1.3:1. Although these two ratios measure different quantities, they are certainly consistent with each other (according to the models, the former NMR ratio corresponds to the sum of “+” and “-” conformations over the non-sheet conformation, and the latter CD ratio corresponds to the “-” conformer over “+” conformer).

Table 4.11 Comparison of CD intensities of prolyl template-peptides with other template-peptide derivatives in EtOH

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-\theta$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-\theta_2$	$-\theta_1/\theta_2$
AcAla-Taa-ValNHMe	319	18200	252	133000	229	113700	1.17
AcLeu-Taa-ValNHMe	318	29400	252	175400	229	126800	1.38
AcGly-Taa-ValNHMe	321	5700	250	19500	226	29200	0.67
AcPro-Taa-ValNHMe	320	16400	252	112000	228	112500	1.00
AcPro-Taa-AlaNHMe	316	5400	252	49800	228	67000	0.74

In contrast to the UV results, the CD intensities for AcAla-Taa-ValNHMe and AcVal-Taa-AlaNHMe are very similar (Figure 4.19 and Table 4.12).

Table 4.12 Comparison of CD intensities of AcVal-Taa-AlaNHMe and AcAla-Taa-ValNHMe

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-\theta$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-\theta_2$	$-\theta_1/\theta_2$
AcAla-Taa-ValNHMe	319	18200	252	132900	229	113700	1.17
AcVal-Taa-AlaNHMe	317	16800	253	106900	228	90200	1.19

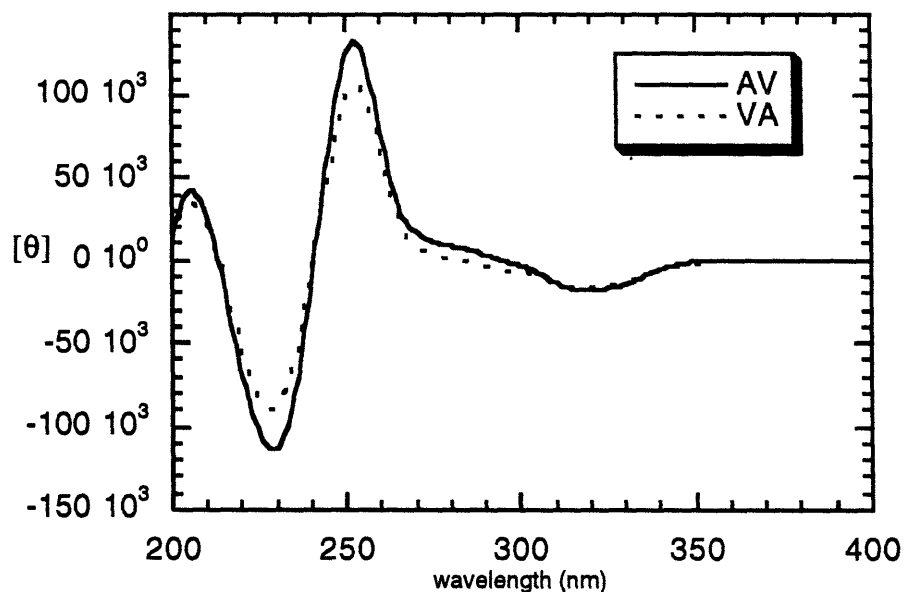


Figure 4.19 AcAla-Taa-ValNHMe versus AcVal-Taa-AlaNHMe

A visual inspection of the CD spectra for the template-peptides that contain polar residues capable of donating a side chain protic hydrogen for hydrogen bonding at site-1 indicates a non-random trend among the three residues glutamine, glutamic acid and asparagine -- they follow a descending order of intensities, especially in the coupled CD region (Figure 4.20 and Table 4.13). All three derivatives have ratios in this region bigger than AcAla-Taa-ValNHMe, and both the glutamine and glutamic acid derivatives have bigger ratios than the asparagine derivative AcAsn-Taa-ValNHMe. This might be attributable to an additional stabilization of the favored  $\beta$ -hairpin conformation by hydrogen bonding of the side chain protic hydrogen with the tolanic carbonyl oxygen across the chain. The length of the side chains of Glu and Gln is perhaps better suited for such an interaction than that of Asn.

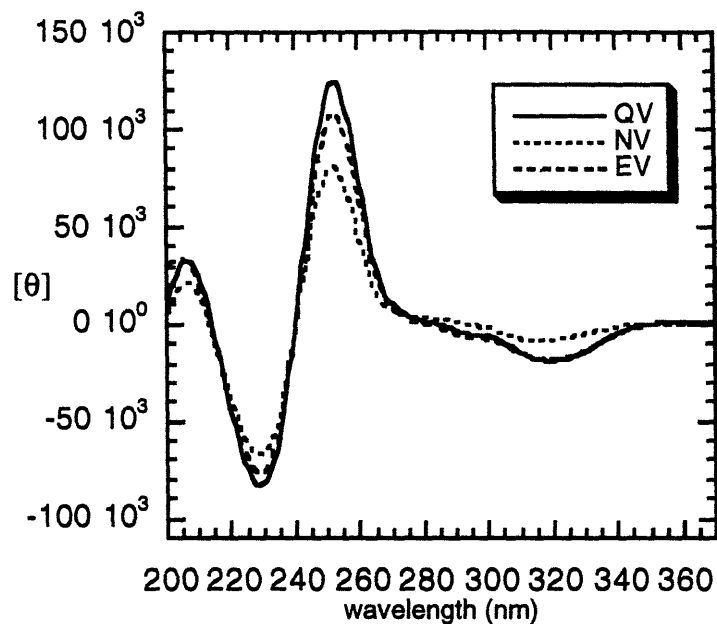


Figure 4.20 CD spectra of AcAA<sub>1</sub>-Taa-ValNHMe (AA<sub>1</sub> = Gln, Asn, Glu), in which the side chain of AA<sub>1</sub> is capable of hydrogen bonding.

Table 4.13 Comparison of CD intensities of AcAA<sub>1</sub>-Taa-ValNHMe (AA<sub>1</sub> = Gln, Glu, Asn)

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-[\theta]$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-[\theta]_2$	$-[\theta]_1/[\theta]_2$
AcGln-Taa-ValNHMe	318	18700	252	124600	229	83100	1.50
AcGlu-Taa-ValNHMe	318	18000	252	108000	229	76100	1.42
AcAsn-Taa-ValNHMe	316	8900	252	80600	229	67100	1.20

An examination of the CD spectra for AcAA<sub>1</sub>-Taa-ValNHMe (AA<sub>1</sub> = Ile, Leu and Met) revealed a pattern. The degree of the side chain branching is decreased by changing the residues. The trend indicates that there is an increasing stabilization of the preferred  $\beta$ -hairpin conformation by the longer side chain of methionine. The order of the preferred pairwise interaction with valinyl residue is Met > Leu  $\approx$  Ile (Figure 4.21 and Table 4.14).

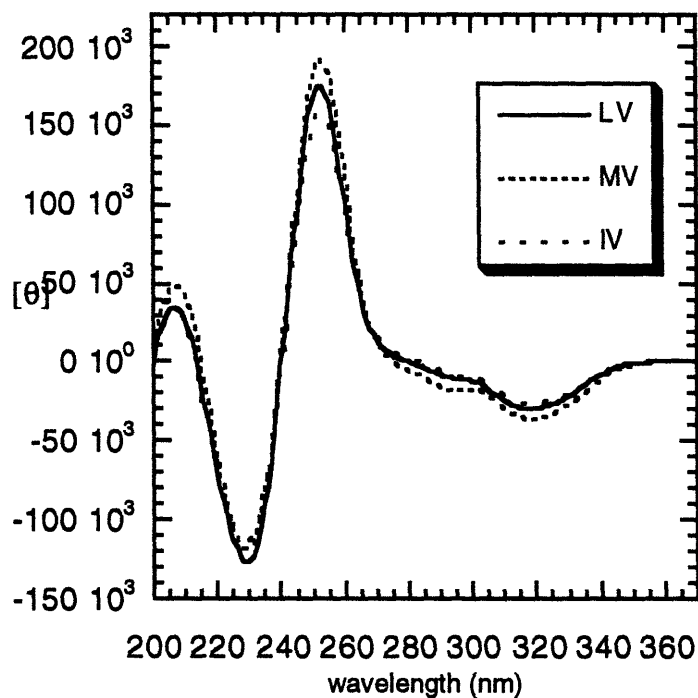


Figure 4.21 Comparison of CD CEs of "homologous" side chains.



Table 4.14 Comparison of CD intensities of  
template-peptides with "homologous" side chains

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	$-[\theta]$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	$-[\theta]_2$	$-[\theta]_1/[\theta]_2$
AcMet-Taa-ValNHMe	318	35700	253	192000	229	118800	1.62
AcLeu-Taa-ValNHMe	318	29400	252	175400	229	126800	1.38
AcIle-Taa-ValNHMe	318	26000	253	158900	229	114000	1.39

In Figure 4.22, the CD CE near 320 nm is plotted against the ratio ( $-[\theta]_1/[\theta]_2$ ) of the coupled CD CEs. The two sloped lines illustrate that there is a non-random correlation between the CD Cotton effect of the conjugation band with those two coupled CD Cotton effects observed in the corresponding B-band region.

The plot exhibits several prominent features: 1) non-polar residues including glycyl and prolyl residues comprise one group of template-peptide conjugates, which all correlate on the same line; 2) a second group of the template-peptide conjugates consists of phenylalanine and polar residues that are capable of donating to a hydrogen bond; it falls on a separate line that is moved almost parallel to the right of the first sloped line; 3) the points in the vicinity of the vertically-drawn line represent the border-line cases in which the two  $\beta$ -hairpin conformers are almost equally populated; 4) points located to the left of the vertical line are glycine- or proline-containing derivatives that favor one  $\beta$ -hairpin conformation; 5) points falling to the right side of vertical line comprise all the other template-peptides that favor the other  $\beta$ -hairpin conformation.

	◆	VV	✕	PV	◁	FV
	●	VI	▷	PA	■	QV
○	MV	▽	▣	GV	◆	EV
△	LV	■	◆	GF	+	NV
▣	IV	◆	✕	GA	⊕	AcV

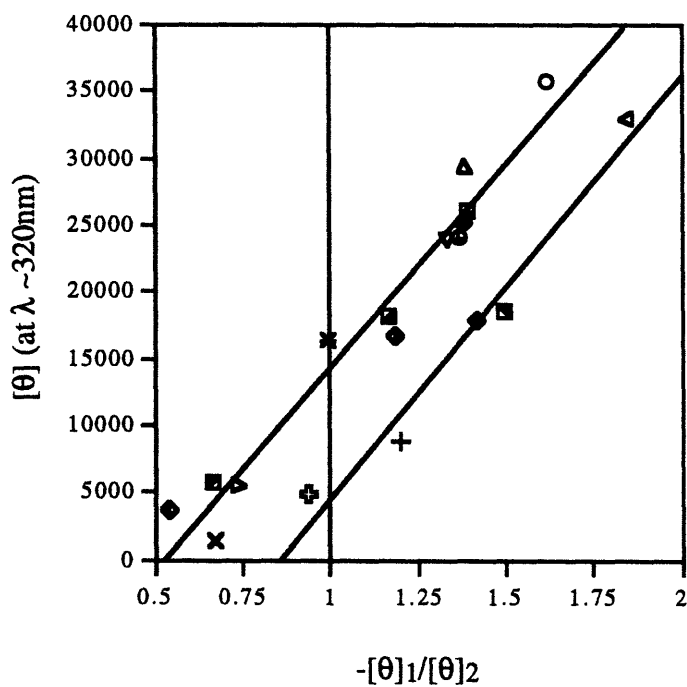


Figure 4.22 CD CEs near  $\lambda_{\max}$  318 nm versus the ratio of CEs near  $\lambda_{\max}$  252 nm and near  $\lambda_{\max}$  229 nm

The appearance of the second parallel line for the three residues Gln, Glu and Asn is attributable in part to favorable and stabilizing interactions between the side chain functionalities and the tolanic carbonyl group across

the chain. The abnormally high ratio ( $-\theta_1/[\theta]_2$ ) observed for AcPhe-Taa-ValNHMe may also be the result of the phenyl side chain having non-vanishing CD Cotton effects in the coupled CD region.

The ratio ( $-\theta_1/[\theta]_2$ ) actually is, according to our analysis, an equilibrium constant corresponding to the interconversion between two  $\beta$ -hairpin conformations. Free energies  $\Delta G^\circ$  obtained from these ratios are listed in Table 4.15.

Table 4.15 The residue pairs are listed in order of decreasing value of the ratio ( $K = -\theta_1/[\theta]_2$ ) of the coupled CD CEs and the corresponding  $\Delta G^\circ$  (kcal/mol).

derivative	$K = -\theta_1/[\theta]_2$	$\Delta G^\circ$ (kcal/mol)
AcPhe-Taa-ValNHMe	1.84	- 0.37
AcMet-Taa-ValNHMe	1.62	- 0.29
AcGln-Taa-ValNHMe	1.50	- 0.249
AcGlu-Taa-ValNHMe	1.42	- 0.21
AcIle-Taa-ValNHMe	1.39	- 0.20
AcLeu-Taa-ValNHMe	1.38	- 0.20
AcVal-Taa-ValNHMe	1.38	- 0.19
AcVal-Taa-IleNHMe	1.37	- 0.19
Gly-Val-Taa-ValNHMe	1.33	- 0.17
AcAsn-Taa-ValNHMe	1.20	- 0.11
AcVal-Taa-AlaNHMe	1.19	- 0.10
AcAla-Taa-ValNHMe	1.17	- 0.09
AcPro-Taa-ValNHMe	1.00	0.00
Ac-Taa-ValNHMe	0.94	0.04
AcPro-Taa-AlaNHMe	0.74	0.18
AcGly-Taa-AlaNHMe	0.68	0.24
AcGly-Taa-ValNHMe	0.67	0.24
AcGly-Taa-IleNHMe	0.54	0.37

The preferred pairwise interactions for an amino acid residue at site-1 with the valinyl residue at site-2 in the tolanoic peptide derivatives (AcAA<sub>1</sub>-Taa-ValNHMe) follow a descending order of Phe, Met, Gln, Glu, Ile, Leu, Val, Asn, Ala, Pro, and Gly.

The ranking of Phe-Val pair before Ile-Val pair is consistent with that obtained from statistical studies (Wouters & Curmi, 1995) and from protein mutation studies (Smith & Regan, 1995). The interactions of valine with methionine or glutamine are very favorable, which is consistent with the assignment by Chou and Fasman (1974) that methionine and glutamine are considered strong sheet-former and sheet-former, respectively.

Shown in Figure 4.23 and Table 4.16 are the CD spectra for cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] and for AcGly-Taa-AlaNHMe and AcVal-Taa-AlaNHMe in EtOH. The CD spectrum for cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] is indeed dramatically different from the open chain cases. The long wavelength CD Cotton effect has a significant bathochromic shift of 3 nm and 7 nm from AcGly-Taa-AlaNHMe and AcVal-Taa-AlaNHMe, respectively. The two CD bands corresponding to the B-band region of the electronic absorption spectrum has hypsochromic shifts as large as 9 nm (compared with AcVal-Taa-AlaNHMe). Compared with AcGly-Taa-AlaNHMe, the signs of the two most intense CD bands are reversed for the cyclic derivative.

These changes almost certainly indicate that the conformation observed for cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] cannot be the same as those observed for the open chain derivatives. It is conceivable that in the cyclic derivative the two rotational conformations of non-vanishing values for the interplanar angle  $\alpha$  are replaced by a  $\beta$ -hairpin conformation that has a coplanar tolan chromophore and as a result, the CD bands corresponding to the former

conformers are no longer significant. However, the CD bands observed in this region may be accountable by the exciton coupling between the planar tolan chromophore and the  $\beta$ -hairpin amide groups.

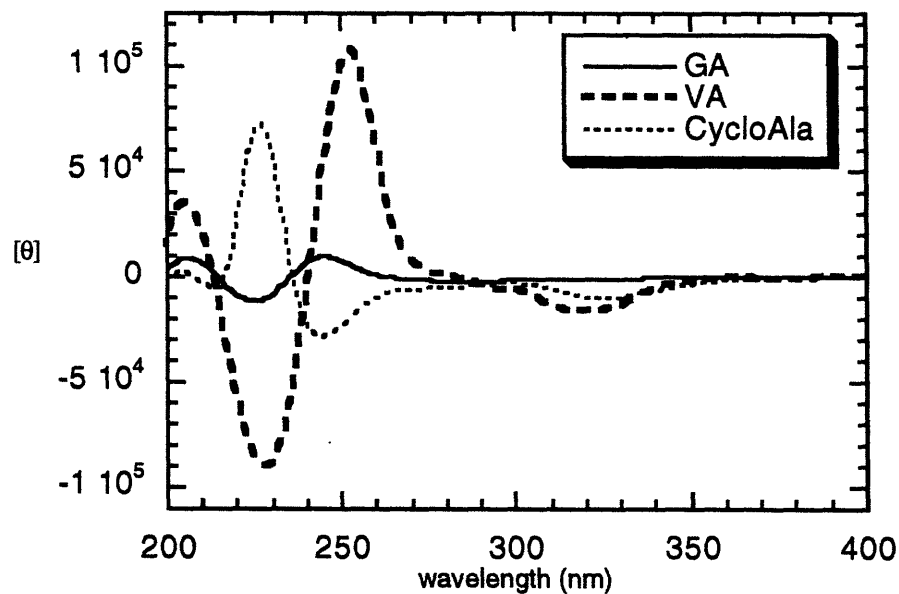


Figure 4.22 The CD spectrum of cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] versus AcGly-Taa-AlaNHMe and AcVal-Taa-AlaNHMe in EtOH.

Table 4.16 Comparison of cyclo[ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] with AcGly-Taa-AlaNHMe and AcVal-Taa-AlaNHMe

Tolan derivative	A-band		B-band				
	$\lambda_{max}$	$-[\theta]$	$\lambda_{max}$	$[\theta]_1$	$\lambda_{max}$	$-[\theta]_2$	$-[\theta]_1/[\theta]_2$
cyclic derivative	324	10200	244	-28100	227	-72300	0.39
AcGly-Taa-AlaNHMe	321	1400	249	7900	225	11700	0.68
AcVal-Taa-AlaNHMe	317	16800	253	106900	228	90200	1.19

These results suggest that the tolanoic amino acid residue is not a perfect template for the  $\beta$ -hairpin structure.

#### 4.5 Conclusion and future directions

The UV and CD results are consistent with those obtained from the NMR studies. In addition, they offer further refinement for the molecular conformations of the template-peptide derivatives. These results are summarized as follows.

- 1) The two substituents have very little resonance interactions with the tolan chromophore. They exert their influences through the more direct electron-donating or electron-withdrawing effects. However, different values of angles  $\beta$  and  $\gamma$  affect their UV and CD spectra.
- 2) The first hydrogen bond at the tolanoic turn is essential for the  $\beta$ -hairpin formation. Its disruption will destroy the  $\beta$ -hairpin conformation .
- 3) The dihedral angle  $\beta$  may be a decisive factor in determining the degree of the  $\beta$ -hairpin conformation, and it is important to have a good sheet-forming residue at site-2.
- 4) The interplanar angle  $\alpha$  cannot be too large. A value close to  $90^\circ$  will certainly disrupt the  $\beta$ -hairpin conformation.
- 5) There exist two  $\beta$ -hairpin conformations that have opposite signs for the angle  $\alpha$ , and can be detected by both UV and CD spectra.

6) Both UV and CD spectra are sensitive to the content of the  $\beta$ -hairpin conformations, but CD is more sensitive than UV in this regard.

7) The tolanoic amino acid is a good  $\beta$ -hairpin RCT but not so good at inducing and stabilizing the  $\beta$ -hairpin structure.

This thesis represents the first attempt in establishing the feasibility of using the tolanoic amino acid as a  $\beta$ -hairpin template, as well as the first in utilizing UV and CD as conformation-reporting techniques for a template-based protein secondary structure. Although the template is not ideal for the  $\beta$ -hairpin, it did point out a new direction for designing RCTs. Preliminary experiments on the syntheses of more constrained tolanoic analogues such as **4.A-C** have been commenced. By restricting the tolanoic carboxamide functional group either through covalent linkages (**4.A** and **4.B**) or through hydrogen bonds (**4.C**), we hope to obtain better  $\beta$ -hairpin templates.

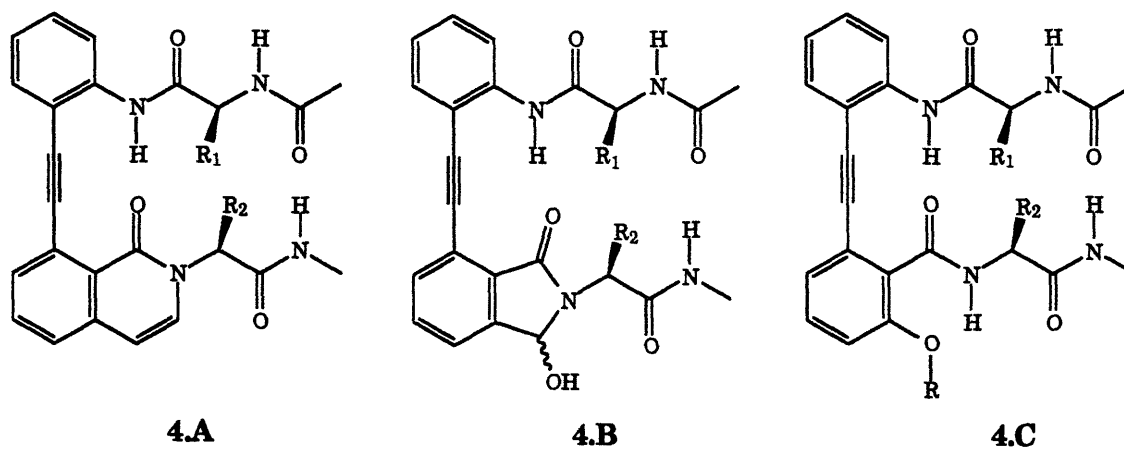


Figure 4.23 Candidates for  $\beta$ -hairpin template.

## **Chapter Five**

### **Experimental Details**



## General

All reactions were conducted in oven-dried (200 °C) or flame-dried glassware under dry argon or nitrogen. Evaporation of solvents was performed under reduced pressure (water aspirator pressure, ~ 10-15 torr) or *in vacuo* (oil pump pressure, ~ 0.1-0.5 torr) by using a Büchi rotary evaporator.

Solvents, when referred to as dry, were distilled as follows. Tetrahydrofuran (THF) and 1,4-dioxane were distilled from sodium benzophenone ketyl. Dimethylformamide (DMF) was distilled from  $\text{CaH}_2$  *in vacuo*. Dimethylsulfoxide (DMSO) was distilled *in vacuo*. DMF was obtained from Burdick & Jackson as high purity solvent, DMSO from Aldrich and absolute ethanol from Pharmco Products. Dichloromethane ( $\text{CH}_2\text{Cl}_2$ ), ethyl acetate (EtOAc), hexane, diethyl ether ( $\text{Et}_2\text{O}$ ), chloroform, methanol, and other solvents were reagent or spectrophotometric grade from either Mallinckrodt or OmniSolv® (E.M. Science) and were used as received. Baker HPLC grade acetonitrile ( $\text{CH}_3\text{CN}$ ) and de-ionized distilled water filtered through a Millipore Milli-Q water system were used for HPLC separations.

Deuterium labeled DMSO- $\text{d}_6$ ,  $\text{CDCl}_3$ ,  $\text{CD}_2\text{Cl}_2$ ,  $\text{CD}_3\text{OD}$ , acetone- $\text{d}_6$ ,  $\text{D}_2\text{O}$ , DMF- $\text{d}_7$ , and other solvents were obtained from either MSD Isotopes or Cambridge Isotope Laboratories.

Triethylamine (TEA), diethylamine (DEA), diisopropylethylamine (DIEA), and trifluoroacetic acid (TFA) were high purity reagents obtained from Aldrich or Pierce and were used as received. Acetic anhydride ( $\text{Ac}_2\text{O}$ ) was redistilled before use. Peptide coupling reagents dicyclohexylcarbodiimide (DCC), diisopropylcarbodiimide (DIC), and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) as well as 1-hydroxybenzotriazole and N-hydroxysuccinimide were purchased from Aldrich. Simple amino acid starting materials were purchased from either Bachem Bioscience or Advanced ChemTech. Tetrakis(triphenylphosphine)-palladium (0) [ $\text{Pd}(\text{PPh}_3)_4$ ] was obtained from Aldrich free of contamination (a yellow color is preferred, and a green color often indicates it is from a contaminated batch). All other reagents were either obtained from Aldrich or from Fluka.

Amino acids and derivatives are abbreviated in accord with the IUPAC-IUB Commission on Biochemical Nomenclature.

## Chromatography

Analytical thin-layer chromatography (TLC) was performed by using plates coated with a 0.25 mm layer of silica gel (E. M. Science). Compounds were visualized by ultraviolet (UV) lamp (at short and long wavelength) directly or by exposing the developed TLC plates to iodine vapor in a chamber or by staining (and charring) of the TLC plates with ninhydrin or potassium permanganate ( $\text{KMnO}_4$ ) or ceric ammonium molybdate solutions. Preparative TLC was performed by using 20 cm x 20 cm plates coated with a 0.25, 0.5, 1.0 or 2.0 mm thick layer of silica gel (Analtech G or GF Uniplates®). Flash column chromatography was typically performed as described by Still using E. Merck silica gel 60 (230-400 mesh) (Still *et al.*, 1978).

Analytical high pressure liquid chromatography (HPLC) separations were performed by using a Waters Associates system equipped with a model U6/K injector, a model 660 solvent gradient programmer, two model 501 pumps, a model 441 single channel UV absorbance detector (214 or 254 or 313 nm) or a model 441 programmable multi-wavelength detector, a model 740 or 746 data module, and a Vydac 218TP54 reverse phase C<sub>18</sub> column or a Waters Associates RCM 8 x 10 reverse phase C<sub>18</sub> column, with the flow rate set at 1 mL/min and 2 mL/min, respectively. Preparative HPLC was performed on a system consisting of an Autochrom™ OPG/S pre-pump solvent mixer, a Rheodyne injector, a Waters Associates model 590 pump fitted with preparative heads, a Waters Associates model 484 tunable absorbance detector, a Houston Instrument OmniScribe recorder, and a Vydac 218TP1022 reverse phase C<sub>18</sub> column or a Waters Associates RCM 25 x 10 reverse phase C<sub>18</sub> column, with flow rates set at constant 4 mL/min and variable 2-6 mL/min, respectively.

Unless otherwise noted, all compounds purified by chromatography are sufficiently pure (>95% by <sup>1</sup>H NMR analysis) for use in subsequent reactions.

### **Nuclear Magnetic Resonance Spectroscopy**

All <sup>1</sup>H (proton) and <sup>13</sup>C (carbon-13) nuclear magnetic resonance (NMR) spectra were recorded on a Varian XL-300 (300 MHz), Varian Unity-300 or Varian VXR-500 (500 MHz) spectrometers maintained by the Spectrometry Laboratory at the Chemistry Department of Massachusetts Institute of Technology (MIT).

Routine one-dimensional (1D)  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded at 300 MHz and 75 MHz, respectively, with either a Varian XL-300 or the Unity 300 spectrometer, without temperature regulation.

High resolution 1D  $^1\text{H}$  spectra, homonuclear decoupling experiments, 1D nuclear Overhauser effects (NOE) difference spectra, two-dimensional (2D) correlated (COSY) spectra, relayed coherence transfer (Relayed-COSY) spectra, 2D nuclear Overhauser and exchange (NOESY) spectra, and 2D rotating-frame transient NOE (ROESY) spectra were all taken on one of the two Varian VXR-500 spectrometers, which were optimized by probe tuning and shimming, and under strict temperature control, before data acquisition started. Shimming followed the method described in the literature (Conover, 1984).

All 2D NMR experiments were performed without sample spinning. Data were acquired at one of the two VXR-500 spectrometers, saved, and processed later with Varian Instruments VNMR software at an off-site Sun Microsystems Sparc 2 workstation or on a Silicon Graphics Personal Iris Indigo workstation. During acquisition, weak homospoil pulses (sspul=y, hs=nn) were often employed in the relaxation delay period. Phase-sensitive detection was achieved with the method of States (States *et al*, 1982). During data-processing, a third-order polynomial baseline correction was generally applied into the  $t_2$  Fourier transformation. The first point in  $t_1$  was baseline-corrected by performing the "cfpmult" correction.

Chemical shifts are reported as  $\delta$  values in parts per million (ppm) relative to tetramethylsilane (TMS).  $^1\text{H}$  NMR coupling constants (J values) are reported in hertz. Multiplicity (splitting pattern) is indicated as follows: s=singlet; d=doublet; t=triplet; q=quartet; dd=doublet of doublets; dt=doublet of triplets; td=triplet of doublets; quin=quintet; m=multiplet; br=broad; etc.

All multiplicities and coupling constants are apparent values without correction.

### **Electronic Absorption and Circular Dichroism Spectroscopy**

Ultraviolet and visible (UV-vis) spectra were taken on a Hewlett Packard model HP8452A spectrophotometer, with a resolution of 2 nm, courtesy of Professor Orme-Johnson's laboratory at MIT, and on a Hewlett Packard model HP 8453 UV-vis spectrophotometer, with a higher resolution of 1 nm. All spectra were acquired either in the wavelength or spectrum/peak mode, scanning from 190 or 200 nm to 400 nm. A baseline of the pure solvent was obtained before sample spectra were recorded. All spectra were taken in absolute ethanol and in 1 cm strain-free quartz cuvettes unless otherwise noted.

Circular dichroism (CD) experiments were performed on an Aviv model 62-DS circular dichroism spectropolarimeter, and temperature was controlled using a jacketed cell and a Neslab Coolflow CFT-33 constant temperature circulating bath. CD spectra were obtained in wavelength mode, with constant temperature control set at 25 °C, step size and bandwidth both set at 1.0 nm, averaging time at 1.5 sec, and scanning region adjusted from 400 to 200 nm, with 5 or 8 repeats of scan for each data point. A baseline of the pure solvent was obtained before sample spectra were recorded. All spectra were taken in absolute ethanol and in 1 cm strain-free quartz cuvettes unless otherwise noted.

Samples for UV and CD studies were all purified by reverse phase HPLC or by a combination of HPLC and preparative TLC (or flash chromatography), checked for purity by NMR to be above 95%, and were subjected to an extended period of drying *in vacuo* before sample solutions

for UV and CD studies were prepared. A small amount of the sample (~0.5-2 mg) was then weighed in a tiny cylinder, made from aluminum foil that had been crimped about the end of a ball-point pen, by using Fisher Scientific MicroGram-Atic balance (accuracy of 0.01 mg), courtesy of Professor Biemann's laboratory at MIT, before it was dissolved in a specified solvent in a volumetric flask marked with the desired volume. All UV and CD spectra were taken in 1 cm strain-free quartz cuvettes.

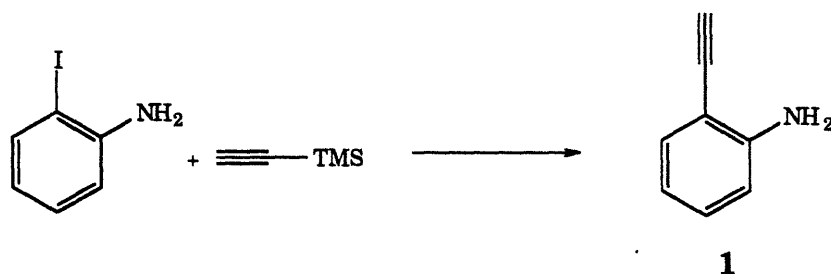
### **Molecular mechanics modeling**

Energy minimizations on derivatives of the tolanoic template were performed on a Silicon Graphics Personal Iris Indigo workstation using the Quanta 4.0 version of the CHARMM molecular modeling software package from Molecular Simulations Inc. The general protocol for the energy minimizations involved the determination of low energy conformations by means of a Steepest Descents algorithm, followed sequentially by the Adopted-Basis Newton Raphson and Newton-Raphson algorithms. The default parameters were used.

### **Other instrumentation**

Mass spectra (high- and low-resolution) were measured on a Finnegan Mat 8200 mass spectrometer by electron impact ionization (EI) or fast atom bombardment (FAB) as indicated. All mass spectra were obtained with the help from either Ms. Li Li or Mr. Ed Takach at the MIT Chemistry Spectrometry Laboratory. Gas Chromatography Mass spectra were obtained by using Hewlett Packard 5890 Series II Gas Chromatograph and Hewlett Packard 5971 Series Mass Selective Detector.

Melting points were taken on a Thomas-Hoover Unimelt capillary melting point apparatus and are uncorrected.

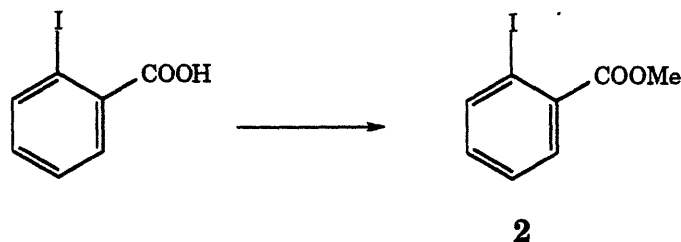


**2-Ethynylaniline (1).** A mixture of 2-iodoaniline (7.4 g, 33.9 mmol), ethynyltrimethylsilane (5.0 g, 50.9 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (194 mg, 0.17 mmol) and cuprous iodide (67 mg, 0.35 mmol) in 25 mL of diethylamine and 4 mL of DMF was stirred under nitrogen at ambient temperature overnight. The reaction mixture was then treated with 50 mL of diethyl ether and 50 mL of 0.1 N aqueous HCl solution. After the organic layer was separated, it was washed with saturated NaHCO<sub>3</sub>, then with water, dried with anhydrous MgSO<sub>4</sub>. After filtration under suction the ether solution was concentrated under aspirator pressure yielding 7.0 g of crude 2-(trimethylsilyl)ethynylaniline. The crude product was then treated with KF·2H<sub>2</sub>O (9.8 g, 0.1 mol) in 150 mL of methanol. After 8 hr at room temperature with stirring, it was concentrated under aspirator pressure, followed by the addition of 100 mL of distilled water. It was then extracted with ether (100 mL, 50 mL). The combined ether layer was washed with brine and concentrated under aspirator pressure. Flash column chromatography (silica) eluting with 8% EtOAc in hexane gave 3.1 g (80%) of the desired 2-ethynylaniline. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 3.38 (s, 1H), 4.24 (br, s, 2H), 6.67 (m, 2H), 7.13 (td, J=1.6 and 7.6 Hz, 1H), 7.31 (dd, J=1.6 and 8.2 Hz, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 80.560, 82.401, 106.490, 114.186, 117.652, 129.981, 132.454, 148.335.

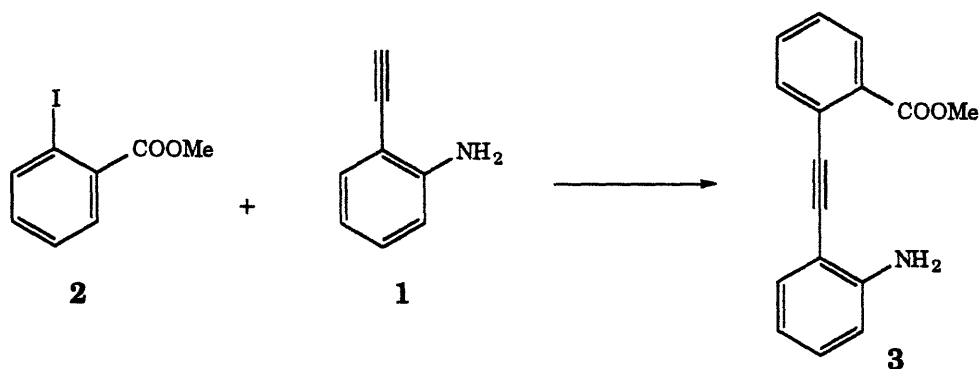


Note: 2-Ethynylaniline is quite volatile -- prolonged evaporation of the residual solvent *in vacuo* will result in some loss of the compound. An efficient way to separate the residual solvent (mostly EtOAc) is to chill the compound in a refrigerator (~ 4 °C) in a flask capped with a plastic septum. The residual solvent will eventually condense near the top of the flask.

Ref.: Arcadi *et al.*, 1989.



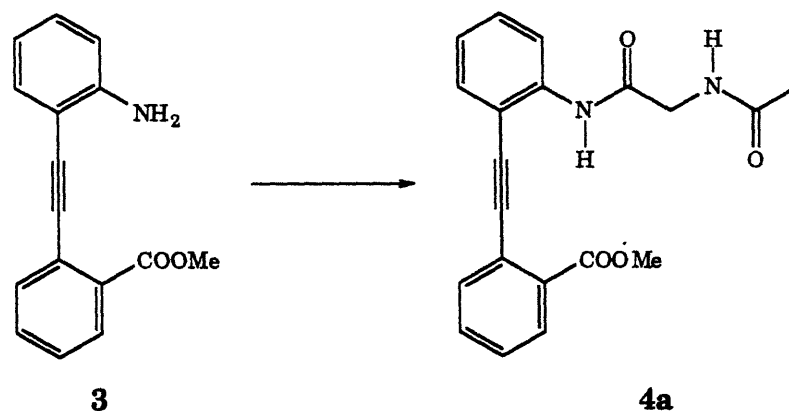
**Methyl 2-iodobenzoate (2).** In a 500 mL round bottom flask equipped with a cool water condenser, a nitrogen inlet, and a stir bar, was added 2-iodobenzoic acid (10 g, 40 mmol), anhydrous  $K_2CO_3$  (28 g, 0.2 mol) and 150 mL of acetone. The mixture was then heated to reflux. After 0.5 hr at reflux, iodomethane (14 mL, 0.2 mol) was added through the condenser. Reflux was continued overnight. After it was cooled to room temperature, the mixture was filtered and the filtrate was concentrated under aspirator pressure. EtOAc was then added, filtered, and followed by washing with hexane. Concentration of the filtrate under reduced pressure gave the desired methyl 2-iodobenzoate (10.1 g, 96%) as a clear oil.  $^1H$  NMR (300 MHz,  $CDCl_3$ ):  $\delta$  3.93 (s, 3H), 7.15 (td,  $J=1.4, 7.9$  Hz; 1H), 7.40 (td,  $J=1.4, 7.6$  Hz; 1H), 7.80 (dd,  $J=1.4, 7.6$  Hz; 1H), 8.00 (dd,  $J=1.4, 7.9$  Hz; 1H);  $^{13}C$  NMR (300 MHz,  $CDCl_3$ ):  $\delta$  52.401, 94.000, 127.843, 130.883, 132.557, 135.137, 141.248, 166.892; MS  $m/e$  (rel intensity) 262 ( $M^+$ , 67), 231 (100), 203 (35), 104 (9), 92 (10), 77 (29), 76 (86), 75 (19), 94 (19), 51 (14), 50 (88).



**2-Amino-2'-carbomethoxytolane (3).** A mixture of 2-ethynylaniline (1.60 g, 13.67 mmol), methyl 2-iodobenzoate (3.35 g, 12.78 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (0.15 g, 0.12 mmol) and cuprous iodide (0.06g, 0.27 mmol) in 25 mL of diethylamine and 6 mL of DMF was stirred under nitrogen at ambient temperature overnight. The reaction mixture was then concentrated, first under aspirator pressure, then in *vacuo*. The residue obtained was treated with 150 mL of ether and 100 mL of 0.1 N aqueous HCl solution. The organic phase was separated after filtration by gravitation, and washed with saturated NaHCO<sub>3</sub>, water, then saturated NaCl. After drying with anhydrous MgSO<sub>4</sub>, the ether filtrate was concentrated under aspirator pressure. Flash column chromatography (silica) eluting with 15% EtOAc in hexane gave the desired template (3) (yield 2.68 g, 84%). TLC R<sub>f</sub> (30% EtOAc in hexane): 0.48; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 3.92 (s, 3H), 4.90 (br, s, 2H), 6.67 (m, 1H), 6.74 (m, 1H), 7.14 (m, 1H), 7.31-7.38 (m, 2H), 7.50 (m, 1H), 7.64 (m, 1H), 8.01 (m, 1H); <sup>13</sup>C (300 MHz, CDCl<sub>3</sub>): δ 52.296, 92.423, 93.679, 107.459, 114.340, 117.372, 124.610, 127.379, 130.167, 130.218, 130.549, 132.021, 133.695, 149.241, 166.286; MS m/e (rel intensity) 251 (M<sup>+</sup>, 68), 220 (32), 219 (100), 208 (30), 190 (53), 165 (30), 110 (20), 95 (38), 82 (33); HRMS calculated for C<sub>16</sub>H<sub>13</sub>NO<sub>2</sub> 251.0946294, found 251.09480.

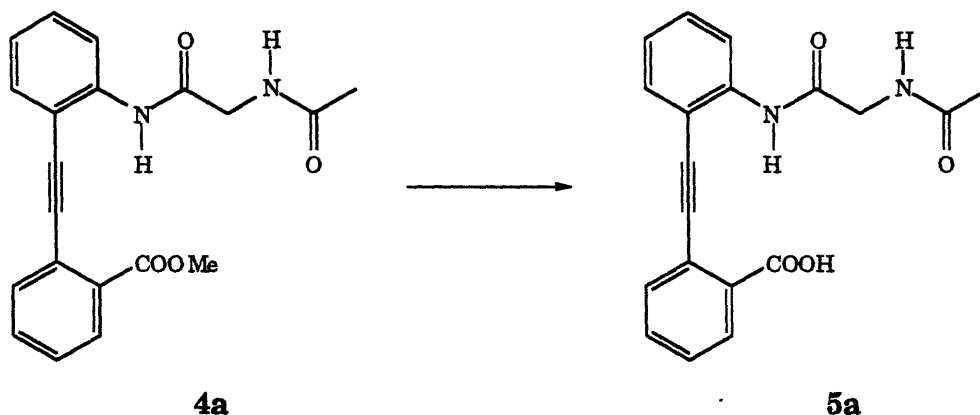
Note: This Sonogashira's procedure of Pd<sup>0</sup>-Cu<sup>I</sup> coupling reaction (Stephens & Castro, 1963; Sonogashira *et al.*, 1975; Arcadi *et al.*, 1989) was later improved in terms of both yields and shortened reaction time for more complex aryl acetylenes and aryl iodides by employing a sealed reaction tube

with controlled heating under an inert atmosphere and by replacing DEA with TEA.



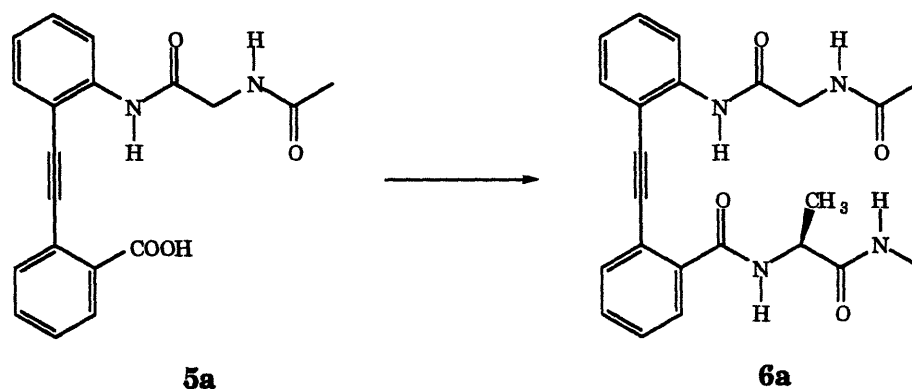
**Ac-Gly-Taa-OMe (4a).** The tolanoic amino acid ester **3** (0.69 g, 2.75 mmol) and N-acetylglycine (0.64 g, 5.50 mmol) were dissolved in 12 mL of DMF at 0 °C in an ice/water bath. Dicyclohexylcarbodiimide (0.57 g, 2.75 mmol) was quickly added to the solution. The ice bath was allowed to slowly warm up to room temperature. After a week at room temperature, the reaction was concentrated in *vacuo*. EtOAc was added and filtered. The filtrate was concentrated under reduced pressure. Preparative thin layer chromatography (TLC) eluting with 10 EtOAc/1 MeOH gave Ac-Gly-Taa-OMe (350 mg, 36%) as a white powder.  $^1\text{H}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  1.84 (s, 3H), 3.92 (s, 3H), 4.18 (d,  $J=5.6$  Hz, 2H), 7.16 (t,  $J=7.6$  Hz, 1H), 7.43 (m, 1H), 7.6 (m, 2H), 7.71 (m, 1H), 7.81 (d,  $J=7.6$  Hz, 1H), 8.05 (d,  $J=7.9$  Hz, 1H), 8.2-8.4 (m, 2H), 9.34 (s, 1H).

Note: Peptide coupling reactions of the free aryl amino group of the template with symmetric anhydride (Chen *et al*, 1978; Yamashiro, 1987) of the reacting Boc-, instead of acetyl-, protected amino acid derivatives, either preformed prior to the coupling or formed *in situ*, proved to be much superior in yields and were adopted in almost all of the subsequent coupling reactions.



**Ac-Gly-Taa-OH (5a).** To a solution of Ac-Gly-Taa-OMe (**4a**) (350 mg, 1.0 mmol) in 9 mL of methanol was added 3 mL of 1.75 M aqueous LiOH. After stirring at room temperature overnight, it was concentrated under reduced pressure. It was then acidified with a pH 3.5 citric acid buffer. Centrifugation gave the crude Ac-Gly-Taa-OH, which was used without further purification.

Note: This hydrolysis procedure is adopted from Corey *et al.* (1977).



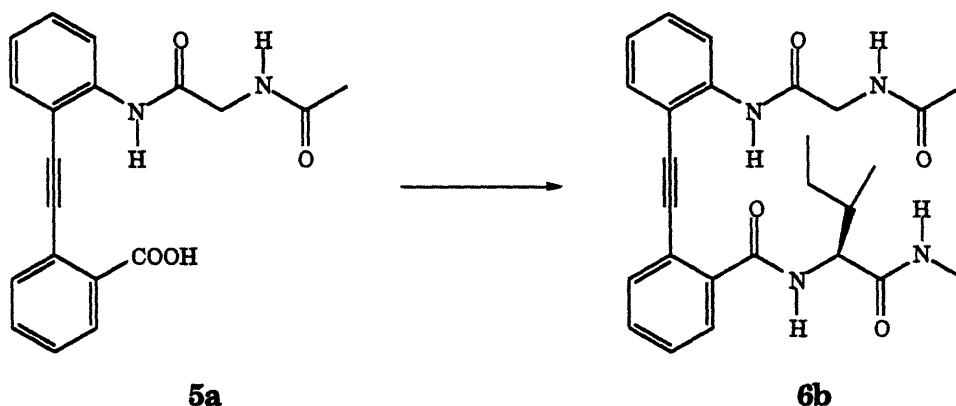
**Ac-Gly-Taa-Ala-NHMe (6a).** To a solution of the crude Ac-Gly-Taa-OH (40 mg, 0.12 mmol), H-Ala-NHMe·HCl (66 mg, 0.48 mmol), 1-hydroxybenzotriazole (32 mg, 0.24 mmol) in 8 mL of dry DMF, was added diisopropylethylamine (0.082 mL, 0.48 mmol). The solution was then cooled to 0 °C with an ice/water bath. Diisopropylcarbodiimide (0.038 mL, 0.24 mmol) was then added and the reaction was allowed to proceed at room temperature for 2 days. It was then concentrated under reduced pressure.

Preparative TLC eluting with 10 EtOAc/1 MeOH gave the desired tripeptide (14 mg, 28%) as a white powder. Further purification by reversed phase HPLC gave 11.1 mg of the desired product. HPLC (from 20 to 35% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 17.7min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  1.296 (d,  $J=7.15$  Hz, 3H), 1.872 (s, 3H), 2.520 (d,  $J=4.63$  Hz, 3H), 4.158 (dd,  $J=17.15, 5.79$  Hz; 1H), 4.231 (dd,  $J=17.10, 5.95$  Hz; 1H), 4.459 (quint,  $J=7.16$  Hz, 1H), 7.126 (td,  $J=7.57, 0.98$  Hz; 1H), 7.405 (m, 1H), 7.53 (m, 2H), 7.589 (m, 1H), 7.710 (dd,  $J=7.56, 1.22$  Hz; 1H), 7.783 (dd,  $J=7.69, 1.1$  Hz; 1H), 7.856 (quart,  $J=4.64$  Hz, 1H), 8.292 (t,  $J=5.86$  Hz, 1H), 8.324 (d,  $J=8.32$  Hz, 1H), 8.783 (d,  $J=7.15$  Hz, 1H), 9.328 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.096, 23.199, 26.329, 44.311, 48.934, 89.803, 95.214, 111.876, 119.051, 122.494, 123.258, 127.122, 128.234, 129.836, 131.086, 131.922, 133.853, 133.933, 140.401, 165.896, 167.600, 170.548, 173.456; MS  $m/e$  (rel intensity) 420 ( $\text{M}^+$ , 2), 321(77), 289 (22), 263 (42), 246 (27), 221 (17), 220 (93), 219 (100), 190 (25), 165 (28), 132 (28), 44 (23), 43 (21); HRMS calculated for  $\text{C}_{23}\text{H}_{24}\text{N}_4\text{O}_4$  420.17976, found 420.1801.

**N-BocIleNHMe.** A heterogeneous mixture of N-BocIleOpNP (3.4 g, 9.6 mmol),  $\text{K}_2\text{CO}_3$  (4.2 g, 28.8 mmol), triethylamine (7 mL, 50.2 mmol),  $\text{H}_2\text{NMe}\cdot\text{HCl}$  (3.3 g, 48.9 mmol) and 60 mL of acetonitrile was stirred vigorously at 5  $^\circ\text{C}$  (internal temperature). The ice bath was allowed to slowly warm up to room temperature overnight. After the reaction mixture was concentrated under reduced pressure, 100 mL of methylene chloride was added to the residue. The organic phase was then washed with water (3 x 60 mL), saturated  $\text{NaHCO}_3$  (2 x 70 mL), water (60 mL), aqueous HCl solution (1N, 60 mL), water (60 mL), saturated NaCl (60 mL) and dried with anhydrous  $\text{MgSO}_4$ . The filtrate was then concentrated under reduced pressure and dried further *in vacuo* yielding 2.1 g (90%) of N-BocIleNHMe as a white powder.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.80-0.86 (m, 6H), 1.04 (m, 1H), 1.37 (3, 9H), 1.50 (br, 1H), 1.75 (br, 1H), 2.73 (d,  $J=4.4$  Hz, 3H), 3.91 (dd,  $J=7.1, 9.4$  Hz; 1H), 5.28 (d,  $J=9.1$

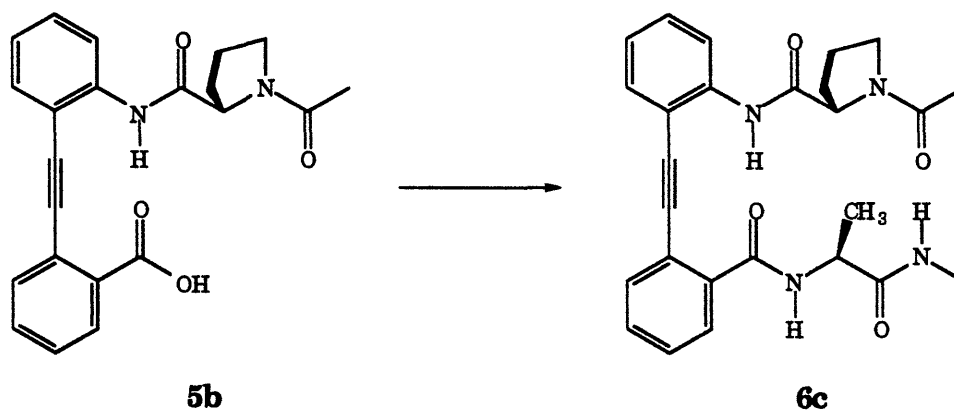
Hz; 1H), 6.72 (br, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  11.247, 15.393, 24.760, 25.897, 28.246, 37.290, 59.218, 79.580, 155.897, 172.497.

**H-Ile-NHMe·HCl.** 100 mL of 6 N HCl in THF was added to a 500 mL round bottom flask charged with N-BocIleNHMe (2.1 g, 8.6 mmol). The solution was then stirred under nitrogen overnight. It was then concentrated under reduced pressure and triturated with diethyl ether. It was further dried *in vacuo* to give a white powder (1.7 g, 100%).  $^1\text{H}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.837 (t,  $J=7.3$  Hz, 3H), 0.858 (d,  $J=6.8$  Hz, 3H), 1.08 (m, 1H), 1.50 (m, 1H), 1.81 (m, 1H), 2.622 (d,  $J=4.57$  Hz, 3H), 3.579 (d,  $J=6.08$  Hz, 1H), 8.295 (br,s, 3H), 8.601 (quart,  $J=4.56$  Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  11.218, 14.510, 24.362, 25.433, 35.965, 56.416, 167.824.



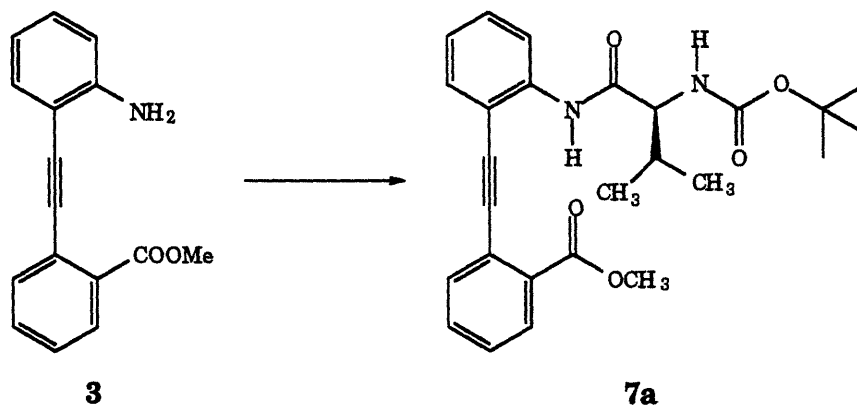
**Ac-Gly-Taa-Ile-NHMe (6b).** To a solution of the crude Ac-Gly-Taa-OH (45 mg, 0.13 mmol), H-Ile-NHMe·HCl (108 mg, 0.54 mmol), 1-hydroxybenzotriazole (37 mg, 0.27 mmol) in 7 mL of dry DMF, was added diisopropylethylamine (0.095 mL, 0.54 mmol). The solution was then cooled to 0 °C with an ice/water bath. Diisopropylcarbodiimide (0.042 mL, 0.27 mmol) was then added and the reaction was allowed to proceed at room temperature overnight. The reaction mixture was then poured into 30 mL of methylene chloride and 40 mL of 0.3 N HCl aqueous solution. The organic

layer was separated, washed with 20 mL each of water, saturated  $\text{NaHCO}_3$ , saturated  $\text{NaCl}$  and dried with anhydrous  $\text{MgSO}_4$ . The filtrate was then concentrated under reduced pressure and dried further *in vacuo* yielding 41.7 mg of crude tripeptide. Purification by reversed phase HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 30 min) gave the desired product (19.8 mg, 32%). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 12.94 min;  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.744 (m, 3H), 0.822 (d,  $J=6.81$  Hz, 3H), 1.115 (m, 1H), 1.496 (m, 1H), 1.829 (m, 1H), 1.884 (s, 3H), 2.520 (d,  $J=4.57$  Hz, 3H), 4.178 (dd,  $J=5.83, 17.11$  Hz; 1H), 4.323 (dd,  $J=8.6, 10.66$  Hz; 1H), 4.347 (dd,  $J=6, 17.11$  Hz; 1H), 7.122 (td,  $J=1.16, 7.53$  Hz; 1H), 7.407 (m, 1H), 7.504-7.537 (m, 2H), 7.565 (m, 1H), 7.684-7.711 (m, 2H), 7.992 (br, quart,  $J=4.6$  Hz, 1H), 8.342 (d,  $J=8.42$  Hz, 1H), 8.392 (t,  $J=5.83$  Hz, 1H), 8.689 (d,  $J=8.51$  Hz, 1H), 9.272 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  11.186, 15.531, 23.163, 24.929, 25.973, 38.826, 44.262, 57.560, 89.853, 95.135, 111.949, 119.151, 122.591, 123.298, 127.182, 128.349, 129.900, 131.082, 132.004, 134.015, 134.522, 140.542, 166.631, 167.967, 170.439, 172.143; MS  $m/e$  (rel intensity) 462 ( $\text{M}^+$ , 8); HRMS calculated for  $\text{C}_{26}\text{H}_{30}\text{N}_4\text{O}_4$  462.226707, found 462.22661.



**AcPro-Taa-AlaNHMe (6c)** was synthesized from AcPro-Taa-OH (**5b**) by the same procedure. Flash column chromatography (silica, 6% MeOH in EtOAc) gave a white powder. Reversed phase HPLC afforded the pure product (44 mg, 73%). TLC (9 EtOAc: 1 MeOH):  $R_f$  0.29; HPLC (20 to 100% of

MeCN in 0.1% aqueous TFA in 10 min): RT 12.51 min;  $^1\text{H}$  NMR (500 MHz,  $\text{CD}_2\text{Cl}_2$ ):  $\delta$  1.410 (d,  $J=6.83$  Hz, 3H), 2.01 (m, 1H), 2.060-2.16 (m, 1H), 2.126 (s, 3H), 2.171-2.254 (m, 2H), 2.539 (d,  $J=4.88$  Hz, 3H), 3.564 (m, 1H), 3.736 (m, 1H), 4.801 (m, 1H), 5.526 (m, 1H), 7.074 (td,  $J=1.22, 7.57$ ; 1H), 7.346 (td,  $J=1.71, 7.92$ ; 1H), 7.443 (td,  $J=1.22, 7.57$ ; 1H), 7.489-7.520 (m, 2H), 7.601 (m, 1H), 7.643 (m, 2H), 7.993 (br, 1H), 8.552 (d,  $J=8.3$  Hz, 1H), 9.298 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.565, 22.257, 25.042, 26.743, 30.643, 49.296, 50.119, 61.123, 89.525, 94.246, 112.079, 119.900, 121.848, 123.836, 127.376, 129.179, 130.527, 131.160, 132.561, 133.090, 137.908, 141.417, 169.340, 171.785, 171.911, 174.677; HRMS calculated for  $\text{C}_{26}\text{H}_{28}\text{N}_4\text{O}_4$  460.211057, found 460.21008.

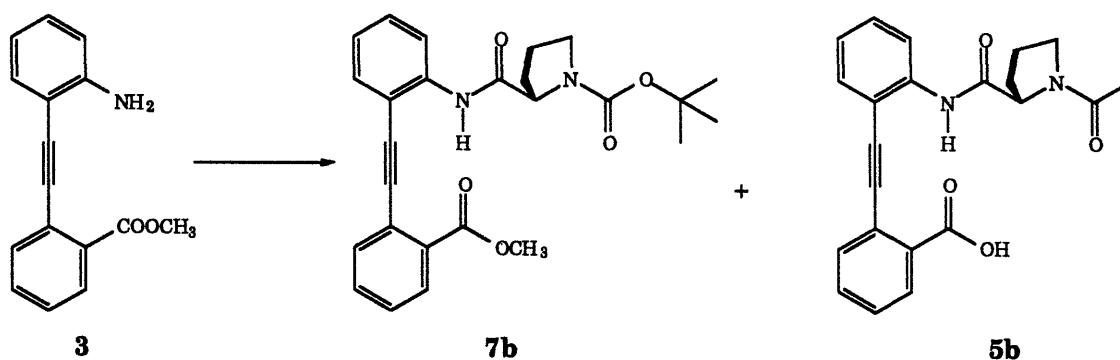


**N-BocVal-Taa-OMe (7a).** To a solution of N-BocValOH (0.868 g, 4 mmol) in 12 mL of methylene chloride at 0 °C was added dicyclohexylcarbodiimide (0.412 g, 2 mmol). After 0.5 hr at 0 °C with stirring under nitrogen, it was cooled at -20 °C for 1.5 hr. After it was filtered cold, the filtrate was concentrated under reduced pressure and further dried *in vacuo* for 0.5 hr to obtain N-BocVal) $_2$ O as a foam. The crude anhydride was then dissolved in 10 mL of chloroform at 0 °C followed by the addition of H-Taa-OMe (3) (150 mg, 0.60 mmol). The ice bath was allowed to slowly warm up to room temperature and stirring was continued overnight under nitrogen. The solvent was removed by evaporation under reduced pressure and the residue was taken up in 20 mL of EtOAc, washed with 20 mL each of saturated



NaHCO<sub>3</sub>, water, saturated NaCl. After it was dried with anhydrous MgSO<sub>4</sub>, the filtrate was then concentrated under reduced pressure. Preparative TLC eluting with 30% EtOAc in hexane yielded 223 mg (83%) N-BocVal-Taa-OMe. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 0.97 (d, J=6.7 Hz, 3H), 1.03 (d, J=6.7 Hz, 3H), 1.38 (s, 9H), 2.25 (m, 1H), 4.07 (s, 3H), 4.79 (br, m, 1H), 5.61 (br, d, J=8.6 Hz), 7.07 (t, J=7.57 Hz, 1H), 7.34-7.44 (m, 2H), 7.51-7.58 (m, 2H), 7.69 (d, 7.74 Hz, 1H), 8.09 (d, J=7.97 Hz, 1H), 8.57 (d, J=8.34 Hz, 1H), 9.36 (br, s, 1H).

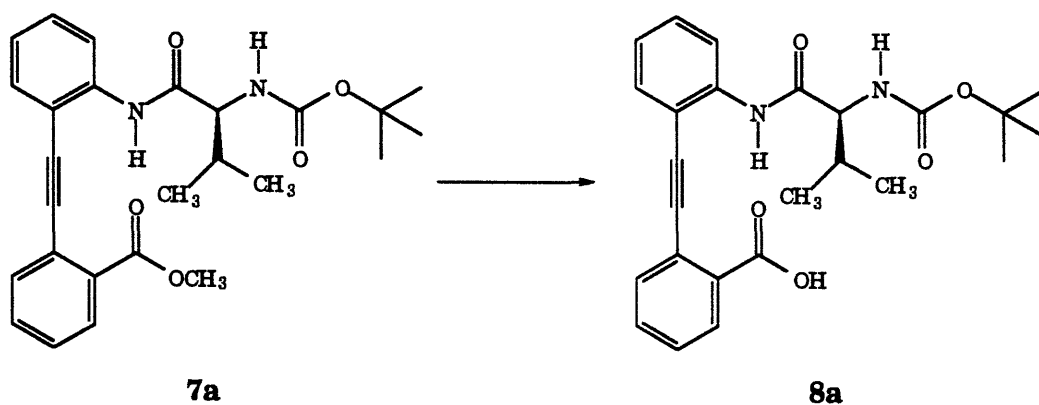
(For advantages of using symmetrical anhydrides see Rebek & Feitler, 1973; 1974; Chen *et al.*, 1978; Yamashiro, 1987.)



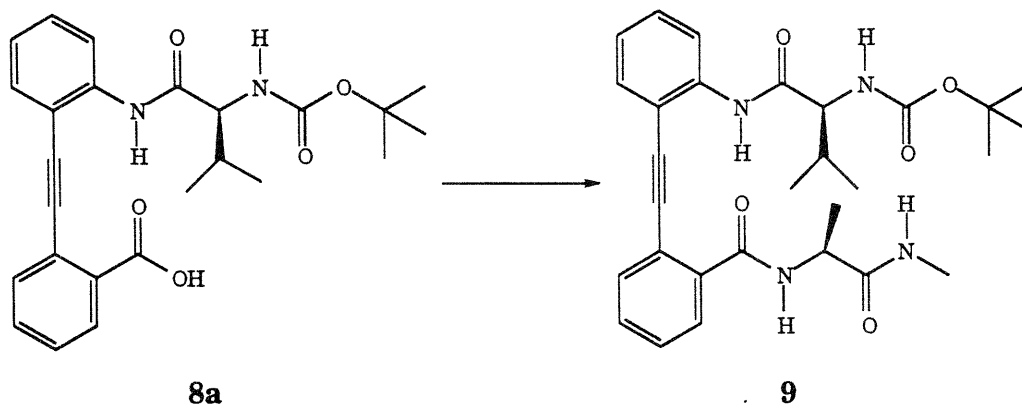
**N-BocPro-Taa-OMe (7b) & N-AcPro-Taa-OH (5b).** To a solution of H-Taa-OMe (0.49 g, 1.96 mmol) and N-BocProOH (6.9 g, 32 mmol) in 40 mL of dry THF cooled to 0 °C in an ice/water bath was added EDC (3.75 g, 19.6 mmol). The reaction was allowed go at room temperature overnight. The solvent was then removed under reduced pressure. EtOAc (70 mL) was added and the solution was washed with dilute HCl twice, water, saturated NaHCO<sub>3</sub> twice, water and saturated NaCl. It was then dried with anhydrous MgSO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude product mixture was then treated with NaOH (1 N, 20 mL) and MeOH (20ml) for 2.5 hr. The solvent volume was reduced by evaporation under reduced pressure and the aqueous solution was extracted with CH<sub>2</sub>Cl<sub>2</sub> three times (75 mL total). The organic phase was washed with dilute HCl solution, saturated NaCl and dried with MgSO<sub>4</sub> to give the crude **7b** as a yellow powder (0.57 g, 65.1%).

The aqueous phase was concentrated. More water was added and the solution was acidified to pH ~3 before it was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The solvent was then removed under reduced pressure and the crude mixture was further dried *in vacuo*. CH<sub>2</sub>Cl<sub>2</sub> (6 mL) was added at 0 °C followed by TFA (6 mL). Stirring was continued for 1 hr before it was concentrated and dried. Water (100 mL) was added and the solution was basified with 1 N NaOH to pH ~13 with vigorous stirring in an ice bath. Acetic anhydride (4ml) was added dropwise. Stirring was continued for 1 hr before it was acidified with 4 N HCl. Filtration under suction yielded the crude acid (**5b**) as a white powder (0.20 g, 27.4%).

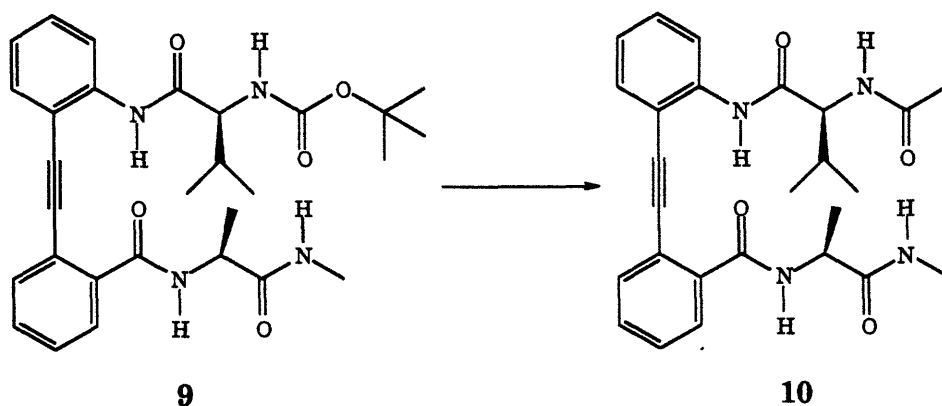
Total combined yield: 92.5%.



**N-BocVal-Taa-OH (8a).** To a solution of N-BocVal-Taa-OMe (220 mg, 0.49 mmol) in 6 mL of methanol was added 2 mL of 1.75 M aqueous LiOH. After stirring at room temperature overnight, it was concentrated under reduced pressure. It was then acidified with 5 N HCl at 0 °C. Cold filtration under suction gave a light orange powder (182 mg), which was used without further purification.

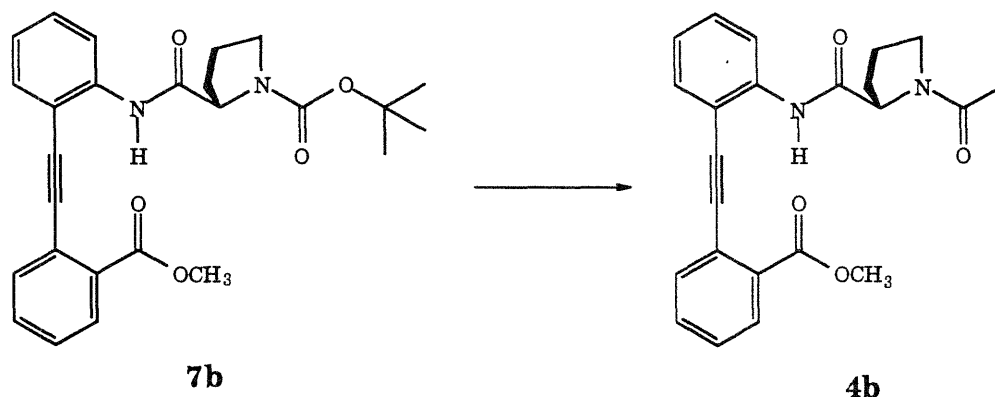


**N-BocVal-Taa-Ala-NHMe (9).** To a solution of the crude N-BocVal-Taa-OH (179 mg, 0.41 mmol), H-Ala-NHMe·HCl (227 mg, 1.64 mmol), 1-hydroxybenzotriazole (111 mg, 0.82 mmol) in 30 mL of dry DMF, was added diisopropylethylamine (0.28 mL, 1.64 mmol). The solution was then cooled to 0 °C with an ice/water bath. Diisopropylcarbodiimide (0.13 mL, 0.82 mmol) was then added and the reaction was allowed to proceed at room temperature for 2 days. It was then concentrated under reduced pressure. Preparative TLC eluting with EtOAc/hexane (5:18) gave the desired tripeptide (91 mg, 43%) as a white powder. TLC (4 hexane/1 EtOAc)  $R_f$  0.46;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.93 (d,  $J=7.0$  Hz, 3H), 1.07 (d,  $J=6.7$  Hz, 3H), 1.47 (s, 9H), 1.48 (d,  $J=6.4$  Hz, 3H), 2.32 (m, 1H), 2.77 (d,  $J=4.6$  Hz, 3H), 5.20 (dd,  $J=4.86, 9.73$  Hz; 1H), 5.30 (m, 1H), 5.56 (d,  $J=9.73$  Hz, 1H), 7.03-7.08 (m, 2H), 7.32-7.41 (m, 2H), 7.45-7.52 (m, 2H), 7.63-7.67 (m, 3H), 8.67 (d,  $J=8.5$  Hz, 1H), 9.28 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  17.030, 19.287, 20.650, 26.029, 28.484, 32.752, 48.566, 58.764, 79.533, 89.636, 94.687, 111.791, 119.268, 122.335, 123.079, 127.086, 128.104, 129.751, 130.767, 132.187, 133.703, 135.220, 140.459, 156.284, 166.448, 171.549, 173.262; MS  $m/e$  (rel intensity) 520 ( $\text{M}^+$ , 1), 420 (10), 349 (17), 322 (16), 321 (71), 318 (24), 317 (21), 291 (13), 290 (50), 289 (33), 364 (13), 263 (58), 248 (18), 247 (18), 246 (24), 221 (14), 220 (73), 219 (100), 190 (20), 165 (24), 158 (12), 132 (15), 116 (18), 72 (95), 69 (16), 57 (91), 55 (30); HRMS calculated for  $\text{C}_{29}\text{H}_{36}\text{N}_4\text{O}_5$  520.2685722, found 520.2682.



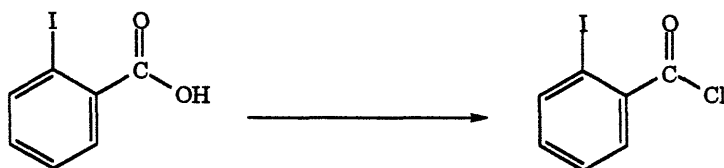
**Ac-Val-Taa-Ala-NHMe (10).** Trifluoroacetic acid (3 mL) was slowly added to a solution of N-BocVal-Taa-Ala-NHMe (78 mg, 0.15 mmol) cooled to 0 °C in an ice/water bath. After 20 min at 0 °C with stirring under nitrogen, the solvents were removed by evaporation under reduced pressure and dried *in vacuo*. Dry DMF (4 mL) was added, followed by the addition of diisopropylethylamine (0.037 mL, 0.21 mmol) and acetic anhydride (0.057 mL, 0.6 mmol). After stirring overnight under nitrogen, the reaction mixture was concentrated *in vacuo*. Water (6 mL) and 1 mL of 1 N NaOH were added to the residue. White precipitate was collected by filtration and washed copiously with water. After drying *in vacuo* it gave a white powder (56 mg, 81%). Further purification by reversed phase HPLC gave 43 mg (63%) of the desired product. HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 13.68 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.877 (d,  $J=6.84$  Hz, 3H), 0.895 (d,  $J=6.84$  Hz, 3H), 1.297 (d,  $J=7.05$  Hz, 3H), 1.913 (s, 3H), 2.109 (m, 1H), 2.492 (d,  $J=4.63$  Hz, 3H), 4.585 (m, 1H), 4.874 (dd,  $J=6.6, 8.7$  Hz; 1H), 7.147 (td,  $J=1.15, 7.58$  Hz; 1H), 7.408 (td,  $J=1.57, 7.89$  Hz; 1H), 7.516 (td,  $J=1.37, 7.58$  Hz; 1H), 7.534 (dd,  $J=1.26, 7.79$  Hz; 1H), 7.571 (td,  $J=1.47, 7.57$  Hz; 1H), 7.698 (dd,  $J=1.1, 7.7$  Hz; 1H), 7.792 (dd,  $J=1.1, 7.7$  Hz; 1H), 7.903 (quart,  $J=4.7$  Hz, 1H), 8.212 (d,  $J=8.63$  Hz, 1H), 8.230 (dd,  $J=0.7, 6.6$  Hz; 1H), 8.771 (d,  $J=7.57$  Hz, 1H), 9.461 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  17.297, 19.217, 20.491, 23.670, 25.973, 32.791, 48.515, 57.790, 89.607, 94.920, 112.164, 119.489, 122.407, 123.405,

127.290, 128.257, 129.793, 130.914, 132.342, 133.908, 135.275, 140.342, 166.523, 170.500, 171.345, 173.479; MS  $m/e$  (rel intensity) 462 ( $M^+$ , 4), 348 (17), 322 (24), 321 (98), 317 (26), 290 (30), 289 (27), 264 (15), 263 (59), 248 (15), 246 (33), 221 (15), 220 (77), 219 (100), 190 (24), 165 (22), 114 (25), 72 (72); HRMS calculated for  $C_{26}H_{30}N_4O_4$  462.226707, found 462.22606.

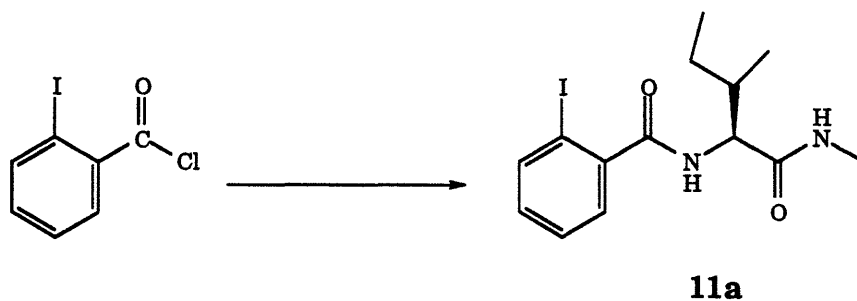


**AcPro-Taa-OMe (4b)** was synthesized from **7b** under the similar conditions shown above. After the reaction was complete, it was concentrated.  $CH_2Cl_2$  (20 mL) was added and washed with dilute NaOH, dilute HCl and water. Solvent was removed under reduced pressure and the residue was subject to flash column chromatography (EtOAc) to yield a white foam (19 mg, 100%). TLC (100% EtOAc):  $R_f$  0.15;  $^1H$  NMR (500 MHz,  $Me_2SO-d_6$ ): two conformers with a ratio of 3.6:1;  $\delta$  1.785-2.440 (m, 4H), 1.965 and 1.892 (s, 3H), 3.419-3.607 (m, 2H), 3.900 and 3.925 (s, 3H), 4.800 (m,  $J=3.35, 3.91, 8.00, 8.56$ Hz) and 4.873 (m,  $J=3.16, 3.35, 8.36, 8.55$  Hz; 1H), 7.167 and 7.221 (m,  $J=1.12, 7.44, 7.62$ Hz; 1H), 7.409-7.474 (m, 1H), 7.562-7.620 (m, 2H), 7.721 (td,  $J=1.30, 7.62$  Hz; 1H), 7.78 and 7.857 (m, 1H), 8.037-8.068 (m 1H), 8.161 and 8.288 (d,  $J=8.18$  and 7.81 Hz, 1H), 9.422 and 9.555 (s, 1H);  $^{13}C$  NMR (300 MHz,  $Me_2SO-d_6$ ): two conformers with a ratio of 3.6 : 1;  $\delta$  22.081, 22.527 and 24.477, 29.498 and 32.139, 46.328 and 47.633, 52.455, 59.841 and 60.854, 89.922 and 89.999, 94.160 and 94.298, 112.234, 120.111, 121.401, 122.568, 122.629, 123.612, 124.380, 128.910, 129.954,

130.276, 130.399, 130.461, 132.181, 132.749, 133.762, 133.839, 139.751, 165.947, 168.450, 168.711, 170.969, 171.230.

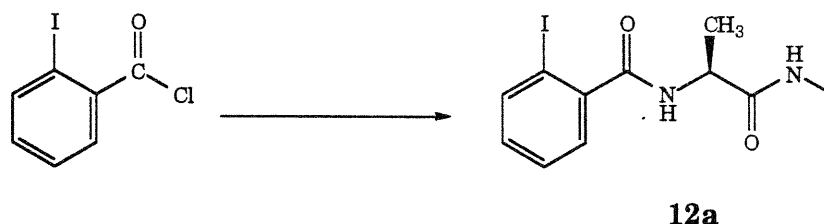


**2-Iodobenzoyl chloride.** A solution of 2-iodobenzoic acid (5 g, 20 mmol) and thionyl chloride (7 mL, 91.6 mmol) in a 50 mL round bottom flask equipped with a cool water condenser and a CaCl<sub>2</sub> drying tube was heated on a steam bath for 3 hr. The reaction mixture was then cooled to room temperature and concentrated under reduced pressure. Kugelrohr distillation *in vacuo* gave 5.36 g (99.8 %) of a clear oil.



**2-Iodobenzoyl-IleNHMe (11a).** To a solution of H-Ile-NHMe·HCl (0.53 g, 2.9 mmol) in 4 mL of DMF and 30 mL of THF was added K<sub>2</sub>CO<sub>3</sub> (2 g, 14.5 mmol). The above mixture was then vigorously stirred at 0 °C in an ice bath. 2-Iodobenzoyl chloride (0.97 g, 3.6 mmol) was added. Stirring was continued and the ice bath was allowed to warm up to room temperature overnight. Filtration followed by concentration of the filtrate under reduced pressure yielded 1.2 g of the crude product. Washing with diethyl ether afforded the title compound as a white powder (0.9 g, 83%). <sup>1</sup>H NMR (300 MHz, DCON(CD<sub>3</sub>)<sub>2</sub>): δ 0.881 (t, J=7.5 Hz, 3H), 0.975 (d, J=6.9 Hz, 3H), 1.307 (m, 1H), 1.621 (m, 1H), 1.973 (m, 1H), 2.728 (d, J=4.62 Hz, 3H), 4.407 (dd, J=7.3, 8.6 Hz;

1H), 7.181 (dd,  $J=2.1, 7.84$  Hz; 1H), 7.377-7.477 (m, 2H), 7.837 (br, quart,  $J=4.6$  Hz, 1H), 7.898 (dd,  $J=1.0, 7.8$  Hz, 1H), 8.211 (d,  $J=8.5$  Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{DCON}(\text{CD}_3)_2$ ):  $\delta$  11.703, 16.384, 25.753, 26.121, 37.825, 58.984, 93.529, 128.604, 128.986, 131.315, 140.052, 144.078, 169.687, 171.958.



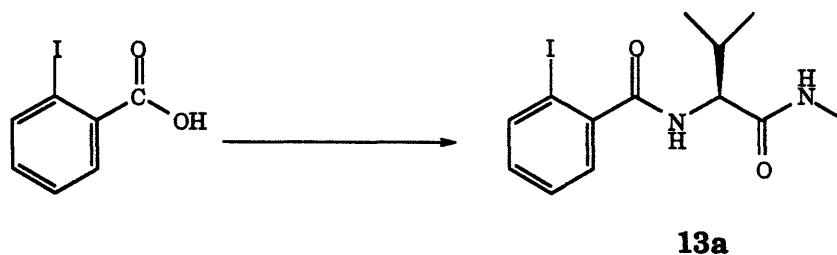
**2-Iodobenzoyl-AlaNHMe (12a).** Same procedure. 81.3% yield.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.493 (d,  $J=7.60$  Hz, 3H), 2.832 (d,  $J=4.88$  Hz, 3H), 4.668 (m, 1H), 6.36 (br s, 1H), 6.44 (br d,  $J=6.8$  Hz, 1H), 7.09 (m, 1H), 7.36 (m, 2H), 7.847 (d,  $J=7.99$  Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.198, 26.350, 49.345, 92.401, 128.210, 128.254, 131.388, 139.950, 141.248, 169.133, 172.220.

**Z-ValNHMe.** THF (70 mL) and 1 mL of water were added to a 200 mL round bottom flask charged with Z-ValOpNP (2.7 g, 7.3 mmol),  $\text{NH}_2\text{Me}\cdot\text{HCl}$  (2.5 g, 67.5 mmol) and  $\text{K}_2\text{CO}_3$  (5.1 g, 36.9 mmol) at 0 °C in an ice bath with vigorous stirring. The ice was allowed to slowly warm up to room temperature overnight. It was then filtered and the filtrate was concentrated under reduced pressure. 100 mL of methylene chloride was added and it was washed with water (100 mL, 60 mL), saturated  $\text{NaHCO}_3$  (4 x 60 mL), water (2 x 60 mL), saturated  $\text{NaCl}$  (60 mL), dried with anhydrous  $\text{MgSO}_4$ . The filtrate was concentrated and the residue was washed with diethyl ether until it was free of *p*-nitrophenol (yellow color under basic condition). White powder (1.8 g, 94%) was obtained.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.918, (d,  $J=6.9$  Hz, 3H), 0.951 (d,  $J=6.8$  Hz, 3H), 2.101 (m, 1H), 2.803 (d,  $J=4.8$  Hz, 3H), 3.935 (dd,  $J=6.4, 8.8$  Hz; 1H), 5.091 (s, 2H), 5.379 (br, 1H), 6.020 (br, 1H), 7.343 (m, 5H);  $^{13}\text{C}$  NMR (300

MHz, CDCl<sub>3</sub>):  $\delta$  17.998, 19.364, 26.253, 30.971, 60.680, 67.082, 127.924, 128.112, 128.441, 136.039, 171.608.

**HOBt·HValNHMe.** To a solution of Z-Val-NHMe (1.8 g, 6.8 mmol) in 80 mL of 95% aqueous ethanol was added a drop of acetic acid and 1-hydroxybenzotriazole (1.3 g, 7.8 mmol). After the hydrogenation bottle was purged with nitrogen, a catalytic amount of 10% Pd/C was added. The mixture was then agitated under a 45 psi pressure of hydrogen for 5 hr. The reaction mixture was then filtered through a thin cake of Celite. Evaporation of the filtrate under reduced pressure gave 2.24 g of the title compound which was used as a crude powder in the next step without further purification.

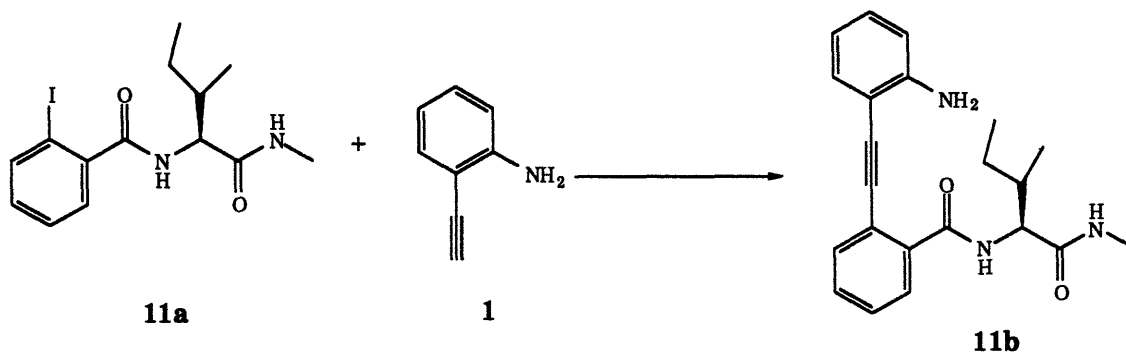
Note: The advantage of using HOBt as a weak acid in the hydrogenolysis of Cbz (Z) protected amino acids and peptides has been discussed by Bodanszky *et al.* (1982).



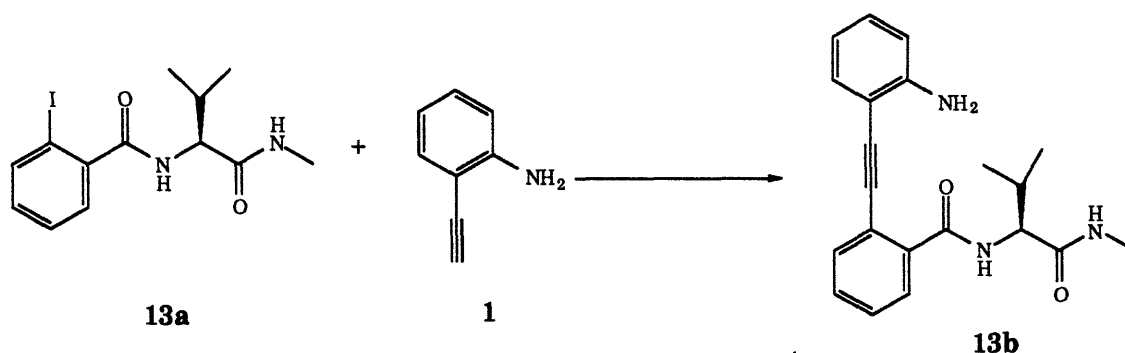
**2-Iodobenzoyl-ValNHMe (13a).** To a solution of 2-iodobenzoic acid (3.48 g, 14 mmol) in 30 mL of dry THF was added dicyclohexylcarbodiimide (1.44 g, 7.0 mmol) at 0 °C in an ice bath. White precipitate formed immediately and stirring was continued at 0 °C for 50 min. The crude anhydride (2-IC<sub>6</sub>H<sub>6</sub>CO)<sub>2</sub>O was then filtered directly into a 200 mL round bottom flask charged with 2.24 g of the crude HOBt·HValNHMe and cooled in an ice bath. The white precipitate (urea) was washed with 20 mL of dry THF. An additional 40 mL of dry THF was added. After diisopropylethylamine (1.2 mL, 7.0 mmol) was added, stirring was continued



at 0 °C for 20 min, then at room temperature overnight. The reaction mixture was then filtered and washed with THF to obtain 1.7 g of white powder. The filtrate was concentrated under reduced pressure and the residue was washed with diethylether and a small amount of THF to obtain 0.4 g of white powder. Total yield: 2.1 g, 86%.  $^1\text{H}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_5$ ):  $\delta$  0.91 (d,  $J=5.6$  Hz, 3H), 0.93 (d,  $J=6.5$  Hz, 3H), 2.02 (m, 1H), 2.62 (d,  $J=4.6$  Hz, 3H), 4.19 (t,  $J=8.5$  Hz, 1H), 7.15 (m, 1H), 7.30 (dd,  $J=1.6, 7.6$  Hz; 1H), 7.43 (m, 1H), 7.84-7.87 (m, 2H), 8.36 (d,  $J=8.6$  Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  18.715, 19.462, 25.464, 30.148, 58.691, 93.313, 127.614, 128.064, 130.449, 138.755, 142.639, 168.486, 170.746.



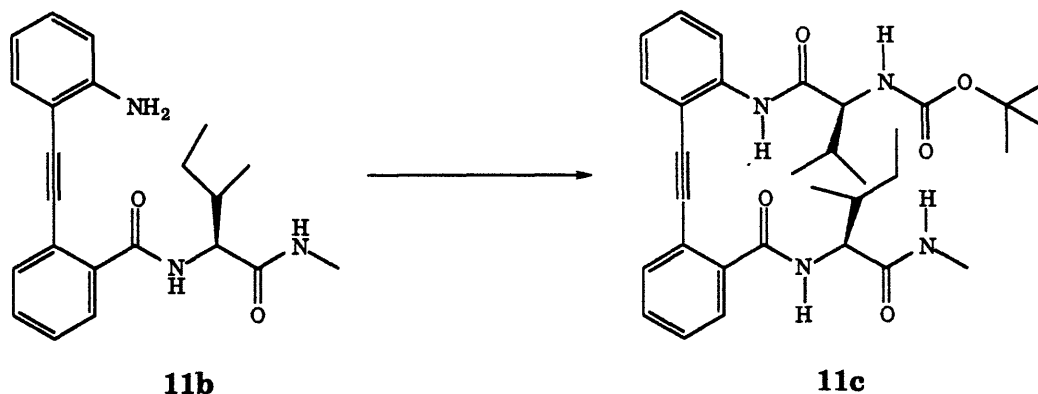
**H-Taa-Ile-NHMe(11b).** A mixture of 2-ethynylaniline (1) (50 mg, 0.43 mmol), 2-iodobenzoyl-IleNHMe (11a) (99 mg, 0.26 mmol),  $\text{Pd}(\text{PPh}_3)_4$  (30 mg, 0.026 mmol) and cuprous iodide (20 mg, 0.09 mmol) in 2 mL of diethylamine and 1 mL of DMF was stirred under nitrogen at ambient temperature for 5 days. The reaction mixture was then concentrated, first under aspirator pressure, then in *vacuo*. The residue was treated with 20 mL of EtOAc and 20 mL of 0.1 N aqueous HCl solution. The organic phase was separated and the aqueous layer was extracted twice with EtOAc. The combined organic layer was washed with saturated  $\text{NaHCO}_3$ , saturated NaCl, dried with anhydrous  $\text{MgSO}_4$ . The filtrate was then concentrated under aspirator pressure. Preparative TLC (silica) eluting with 2 EtOAc/1 hexane afforded 108 mg of the crude product which was used without further purification.



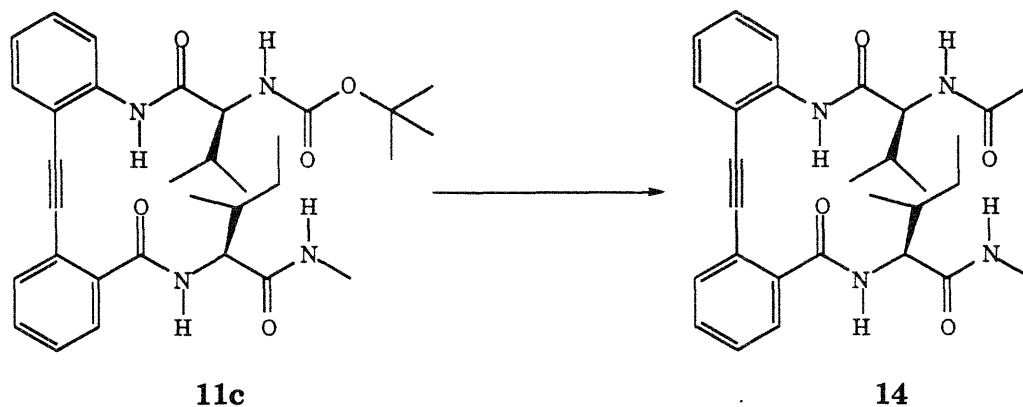
**H-Taa-Val-NHMe(13b).** A mixture of 2-ethynylaniline (1) (0.7 g, 5.98 mmol), 2-iodobenzoyl-ValNHMe (13a) (1.7 g, 4.72 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (60 mg, 0.052 mmol) and cuprous iodide (40 mg, 0.18 mmol) in 10 mL of triethylamine and 4 mL of DMF in a reaction tube (See diagram below) charged with a stir bar, was degassed, purged with argon before the tube was sealed. The mixture was then heated with a oil bath to a gentle reflux for 6 hr, then at 50-60 °C (oil bath) for 24 hr. After it was cooled to room temperature, the reaction mixture was transferred with chloroform to a 200 mL round bottom flask and was concentrated under reduced pressure. Chloroform (40 mL) was then added to the residue and the solution was washed with 40 mL of 0.1 N HCl. The aqueous layer was then back-extracted with chloroform (3 x 30 mL). The combined organic layer was washed with saturated NaHCO<sub>3</sub> (2 x30 mL), saturated NaCl (40 mL), dried with anhydrous MgSO<sub>4</sub>. The filtrate was then concentrated under aspirator pressure and dried *in vacuo* to give 1.8 g of the crude product, which was used without further purification. TLC (4 EtOAc/1 hexane) R<sub>f</sub> 0.40; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 0.972 (d, J= 6.76 Hz, 3H), 0.980 (d, J=6.76 Hz, 3H), 2.235 (m, 1H), 2.676 (d, J=4.8 Hz, 3H), 4.463 (dd, J=7.3, 8.7 Hz; 1H), 4.638 (br, s, 2H), 6.564 (br, quart, J=4.5 Hz, 1H), 6.648-6.699 (m, 2H), 7.132 (m, 1H), 7.325-7.461 (m, 4H), 7.601 (dd, J=1.29, 7.6 Hz; 1H), 7.739 (dd, J=1.5, 7.6 Hz; 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 18.511, 19.526, 26.200, 30.809, 59.574, 92.294, 106.646, 114.193, 117.444, 121.028, 128.115, 128.150, 130.249, 130.438, 132.008, 133.203, 135.776, 148.816, 167.470, 171.324.



CDCl<sub>3</sub>):  $\delta$  17.910, 26.310, 49.527, 92.370, 92.462, 106.650, 114.482, 117.706, 128.332, 128.501, 128.609, 130.497, 130.743, 132.079, 133.292, 135.503, 148.909, 167.474, 172.418.

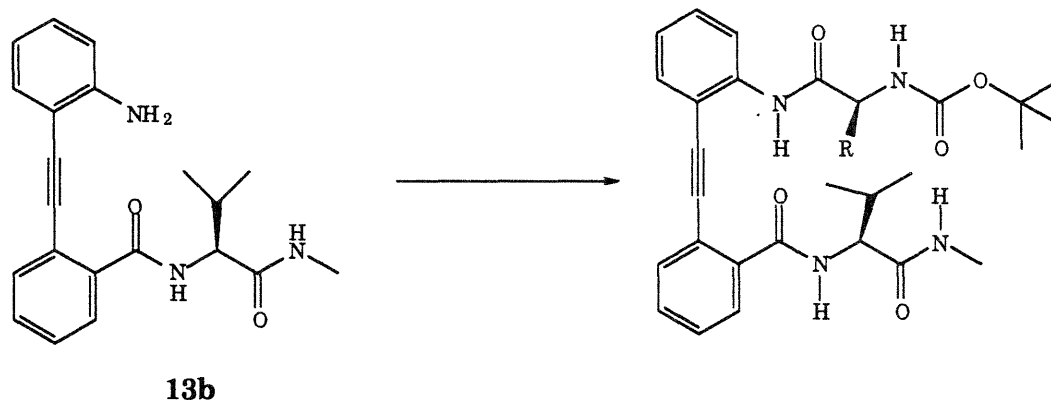


**N-BocVal-Taa-Ile-NHMe (11c).** To a solution of N-BocValOH (0.434 g, 2 mmol) in 10 mL of methylene chloride at 0 °C was added dicyclohexylcarbodiimide (0.206 g, 1 mmol). After 1 hr at 0 °C with stirring under nitrogen, it was cooled at -20 °C for 1.5 hr. After it was filtered cold, the filtrate was concentrated under reduced pressure and further dried *in vacuo* for 0.5 hr to obtain N-BocVal)<sub>2</sub>O as a foam. The crude anhydride was then dissolved in 10 mL of dry THF at 0 °C followed by the addition of H-Taa-Ile-NHMe (**11b**) (108 mg, crude, 0.26 mmol). The ice bath was allowed to slowly warm up to room temperature and stirring was continued overnight under nitrogen. Preparative TLC eluting with 3 EtOAc : 1 hexane yielded N-BocVal-Taa-Ile-NHMe (96 mg, 65% in 2 steps from 2-iodobenzoyl-IleNHMe). <sup>1</sup>H NMR (300 MHz, CD<sub>2</sub>Cl<sub>2</sub>):  $\delta$  0.87-1.03 (m, 6H), 1.23 (m, 1H), 1.41-1.48 (m, 15H), 1.73 (m, 1H), 1.82 (m, 1H), 2.20 (m, 1H), 2.68 (d, J=5.0 Hz, 3H), 4.86 (t, J=8.5 Hz, 1H), 5.24 (dd, J=5.21, 10.07 Hz; 1H), 5.52 (d, J=10.11 Hz, 1H), 6.92 (br, 1H), 7.07 (m, 1H), 7.32-7.53 (m, 4H), 7.66-7.72 (m, 3H), 8.66 (d, J=8.51 Hz, 1H), 9.22 (s, 1H).



**Ac-Val-Taa-Ile-NHMe (14).** Trifluoroacetic acid (4 mL) was slowly added to a solution of N-BocVal-Taa-Ile-NHMe (94 mg, 0.17 mmol) cooled to 0 °C in an ice/water bath. After 30 min at 0 °C with stirring under nitrogen, the solvents were removed by evaporation under reduced pressure and dried *in vacuo*. Dry THF (10 mL) was added, followed by the addition of diisopropylethylamine (0.07 mL, 0.4 mmol), K<sub>2</sub>CO<sub>3</sub> (60 mg, 0.4 mmol) and acetic anhydride (0.11 mL, 1.2 mmol). After stirring overnight under nitrogen, the reaction mixture was filtered and washed with chloroform. After the filtrate was concentrated under reduced pressure, 4 mL of water and 1 mL of 1 N NaOH were added. The white precipitate was collected by filtration and washed with copious water. After drying *in vacuo* it gave a white powder (73 mg, 86%). A fraction of it was further purified by reversed phase HPLC to give 29 mg of the desired product. TLC (4 EtOAc/1 Hex) R<sub>f</sub> 0.35; HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 15.58 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.801 (t, J=7.37 Hz, 3H), 0.839 (d, J=6.82 Hz, 3H), 0.878 (d, J=6.76 Hz, 3H), 0.891 (d, J=6.76 Hz, 3H), 1.157 (m, 1H), 1.554 (m, 1H), 1.868 (m, 1H), 1.941 (s, 3H), 2.099 (m, 1H), 2.487 (d, J=4.58 Hz, 3H), 4.396 (t, J=8.82 Hz, 1H), 4.987 (dd, J=6.81, 8.49 Hz; 1H), 7.135 (m, 1H), 7.402 (m, 1H), 7.508 (td, J=1.23, 7.5 Hz; 1H), 7.524 (dd, J=1.34, 7.71 Hz; 1H), 7.564 (m, 1H), 7.687 (dd, J=1.0, 7.6 Hz; 1H), 7.750 (d, J=7.59 Hz; 1H), 8.072 (quart, J=4.64 Hz, 1H), 8.264 (d, J=8.54 Hz, 1H), 8.296 (d, J=8.32 Hz; 1H), 8.698 (d, J=8.65 Hz, 1H), 9.406 (s, 1H); <sup>13</sup>C NMR (300 MHz, CD<sub>2</sub>Cl<sub>2</sub>): δ 11.804, 15.497, 17.742, 19.407,

23.735, 25.646, 25.953, 32.821, 39.267, 57.594, 58.179, 89.524, 94.996, 112.151, 119.519, 122.321, 123.459, 127.676, 128.574, 129.967, 131.090, 132.354, 133.911, 136.112, 140.795, 167.230, 170.757, 171.651, 172.090; MS  $m/e$  (rel intensity) 504 ( $M^+$ , 17); HRMS calculated for  $C_{29}H_{36}N_4O_4$  504.273657, found 504.27364.



Method A: N-protected amino acid N-BocAAOH (8-10 equiv) was dissolved in 4-5 mL of methylene chloride and cooled to 0 °C in an ice bath. Dicyclohexylcarbodiimide (4-5 equiv) was added and stirring was continued at 0 °C for 0.5 hr. It was then filtered and concentrated under reduced pressure. This crude anhydride was then added to a solution of H-Taa-Val-NHMe (**13b**) (1 equiv) in 8-10 mL of methylene chloride. The solution was then stirred at 0 °C for 2 hr, then at room temperature overnight. The solvent was removed by evaporation under reduced pressure. Flash column chromatography (silica) afforded crude N-Boc-AA-Taa-Val-NHMe, which was used in the following step without further purification.

**N-Boc-Ala-Taa-Val-NHMe**. TLC (2 EtOAc/1 Hex)  $R_f$  0.46.

**N-Boc-Gly-Taa-Val-NHMe**. TLC (2 EtOAc/1 Hex)  $R_f$  0.36.

**N-Boc-Val-Taa-Val-NHMe**. TLC (2 EtOAc/1 Hex)  $R_f$  0.53.

**N-Boc-Ile-Taa-Val-NHMe**. TLC (2 EtOAc/1 Hex)  $R_f$  0.49.

**N-Boc-Phe-Taa-Val-NHMe**. TLC (2 EtOAc/1 Hex)  $R_f$  0.50.

Method B: A mixture of H-Taa-Val-NHMe (**13b**) (1 equiv) and N-protected amino acid N-BocAAOH (8-10 equiv) was dissolved in 5-10 mL of dry THF and cooled to 0 °C in an ice bath. Water soluble carbodiimide EtN=C=N(CH<sub>2</sub>)<sub>3</sub>NMe<sub>2</sub>·HCl (4.3-5.4 equiv) was added and stirring was continued at 0 °C for 2-3 hr, then at room temperature overnight. The solvent was removed by evaporation under reduced pressure. EtOAc (20 mL) was then added and the solution was washed with dilute HCl solution, water, saturated NaHCO<sub>3</sub>, saturated NaCl, dried with anhydrous MgSO<sub>4</sub>. The filtrate was then concentrated under aspirator pressure and dried *in vacuo* to give the crude N-Boc-AA-Taa-Val-NHMe which was used in the next step without further purification.

**N-Boc-Leu-Taa-Val-NHMe.** TLC (1 EtOAc/1 Hex) R<sub>f</sub> 0.43.

**N-Boc-Pro-Taa-Val-NHMe.** TLC (1 EtOAc/1 Hex) R<sub>f</sub> 0.21.

**N-Boc-Met-Taa-Val-NHMe.** TLC (1 EtOAc/1 Hex) R<sub>f</sub> 0.32.

**N-Boc-Gln-Taa-Val-NHMe.** TLC (4 EtOAc/1 Hex) R<sub>f</sub> 0.20.

**N-Boc-Asn-Taa-Val-NHMe.**

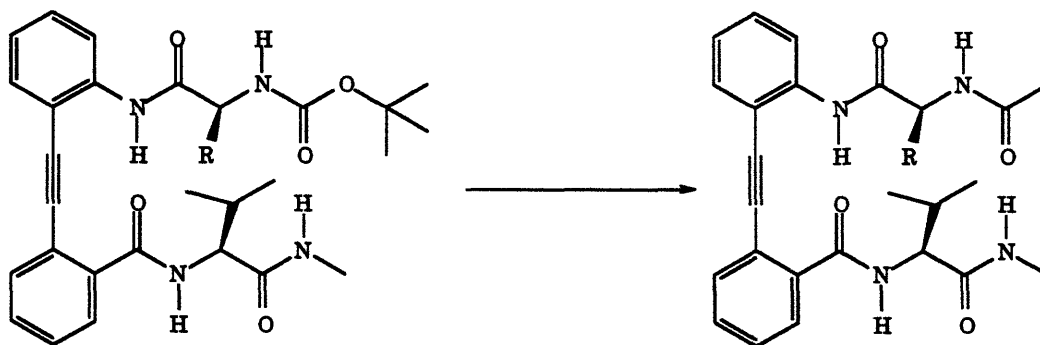
**N-Boc-Glu(OtBu)-Taa-Val-NHMe.**

**N-Boc-D-Val-Taa-Val-NHMe.**

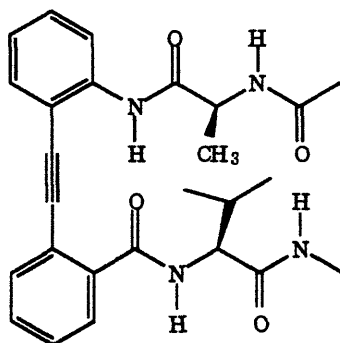
**N-Boc-D-Ala-Taa-Val-NHMe.**

**N-Boc-D-Pro-Taa-Val-NHMe.**

**N-Boc-Me<sup>N</sup>-Ala-Taa-Val-NHMe.**



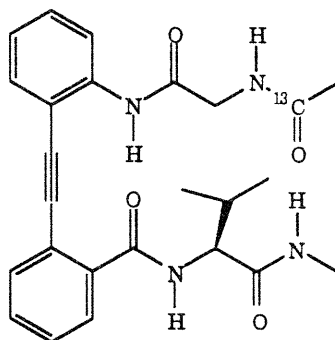
Crude N-Boc-AA-Taa-Val-NHMe (1 equiv) was dissolved in 2 mL of methylene chloride and cooled under nitrogen or argon to 0 °C in an ice bath. Trifluoroacetic acid (2 mL) was then added and stirring was continued at 0 °C for 0.5 hr. The solvents were then removed by evaporation under reduced pressure followed by trituration with ether and drying *in vacuo*. To this residue was then added K<sub>2</sub>CO<sub>3</sub> (2-3 equiv), 4 mL of dry THF and diisopropylethylamine (2-3 equiv). After the above reaction mixture was cooled to 0 °C in an ice bath, acetic anhydride (8-10 equiv) or acetyl chloride (<sup>13</sup>C=O) (5-6 equiv) was added and the ice was allowed to warm up to room temperature. Stirring was continued at room temperature overnight before it was concentrated under reduced pressure. Water (5 mL) was then added to the residue, sonicated, then followed by the addition of 1.5 mL of 1 N NaOH before it was cooled in a ice bath. It was then filtered and the solid obtained was washed with cold water to give a white powder. In the cases of Ac-Pro-Taa-Val-NHMe and Ac-Me<sup>N</sup>-Ala-Taa-Val-NHMe, the aqueous filtrate was extracted with methylene chloride and concentrated under reduced pressure to combine with the solid. It was then further purified by preparative reversed phase HPLC. Percentage yield was calculated from 2-iodobenzoylValNHMe (13a) in 3 steps.

**16a**

**Ac-Ala-Taa-Val-NHMe (16a).** Yield 56.5 mg, 85% after HPLC (in 3 steps). TLC (4 EtOAc/1 Hex) R<sub>f</sub> 0.20; HPLC (from 20 to 100% MeCN in 0.1%



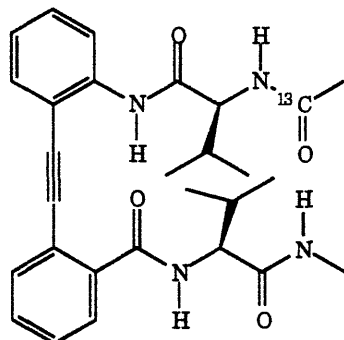
aqueous trifluoroacetic acid in 10 min): RT 13.32 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.863 (d,  $J=6.72$  Hz, 3H), 0.914 (d,  $J=6.71$  Hz, 3H), 1.342 (d,  $J=7.08$  Hz, 3H), 1.866 (s, 3H), 2.046 (m, 1H), 2.483 (d,  $J=4.65$  Hz, 3H), 4.278 (t,  $J=8.51$  Hz, 1H), 5.035 (m, 1H), 7.129 (m, 1H), 7.403 (m, 1H), 7.498-7.532 (m, 2H), 7.565 (m, 1H), 7.709-7.745 (m, 2H), 8.041 (br, quart,  $J=4.66$  Hz, 1H), 8.316 (dd,  $J=0.72$ , 8.42 Hz; 1H), 8.441 (d,  $J=7.34$  Hz; 1H), 8.684 (d,  $J=8.69$  Hz, 1H), 9.357 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.618, 19.425, 20.667, 23.589, 25.996, 32.817, 49.606, 57.993, 89.835, 95.141, 112.115, 119.358, 122.429, 123.294, 127.161, 128.186, 129.693, 130.936, 132.079, 134.053, 134.529, 140.332, 166.593, 169.563, 171.868, 172.255 MS  $m/e$  (rel intensity) 462 ( $\text{M}^+$ , 16); HRMS calculated for  $\text{C}_{26}\text{H}_{30}\text{N}_4\text{O}_4$  462.226707, found 462.22661.



**16b**

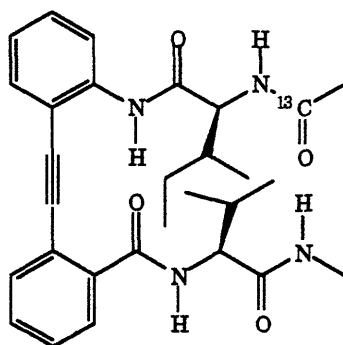
**Ac( $^{13}\text{C}=\text{O}$ )-Gly-Taa-Val-NHMe (16b).** Yield 28.6 mg, 86% after HPLC (in 3 steps). HPLC (from 20 to 100%  $\text{MeCN}$  in 0.1% aqueous trifluoroacetic acid in 10 min): RT 13.17 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.865 (d,  $J=6.76$  Hz, 3H), 0.890 (d,  $J=6.76$  Hz, 3H), 1.889 (d,  $J=6.1$  Hz, 3H), 2.037 (m, 1H), 2.522 (d,  $J=4.6$  Hz, 3H), 4.178 (ddd,  $J=4.32$ , 5.64, 17.1 Hz; 1H), 4.277 (dd,  $J=8.27$ , 8.35 Hz; 1H), 4.353 (ddd,  $J=3.85$ , 5.91, 17.1 Hz; 1H), 7.122 (m, 1H), 7.406 (m, 1H), 7.506-7.539 (m, 2H), 7.569 (td,  $J=1.41$ , 7.51 Hz; 1H), 7.696-7.723 (m, 2H), 7.983 (br, quart,  $J=4.6$  Hz, 1H), 8.337 (d,  $J=7.98$  Hz; 1H), 8.382 (m, 1H), 8.671 (d,  $J=8.54$  Hz, 1H), 9.278 (s,

1H); MS m/e (rel intensity) 449 (M<sup>+</sup>, 11); HRMS calculated for C<sub>24</sub>C\*H<sub>28</sub>N<sub>4</sub>O<sub>4</sub> 449.214411, found 449.21424.



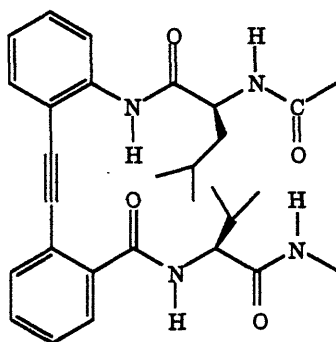
**16c**

Ac(<sup>13</sup>C=O)-Val-Taa-Val-NHMe (16c). Yield 38 mg, 80% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.64 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.888 (br, m, 9H), 0.936 (d, J=6.67 Hz, 3H), 1.949 (d, J=6.01 Hz, 3H), 2.082 (m, 2H), 2.5 (br, 3H), 4.347 (dd, J=8.55, 8.73 Hz; 1H), 5.018 (m, 1H), 7.139 (m, 1H), 7.407 (t, J=7.6 Hz, 1H), 7.497-7.538 (m, 2H), 7.572 (t, J=7.93 Hz; 1H), 7.696 (d, J=7.42 Hz, 1H), 7.786 (d, J=7.46 Hz, 1H), 8.086 (br, quart, J=4.7 Hz, 1H), 8.304 (br, d, J=7.93 Hz; 2H), 8.721 (d, J=8.64 Hz, 1H), 9.426 (s, 1H); MS m/e (rel intensity) 491 (M<sup>+</sup>, 17); HRMS calculated for C<sub>27</sub>C\*H<sub>34</sub>N<sub>4</sub>O<sub>4</sub> 491.263606, found 491.26293.



**16d**

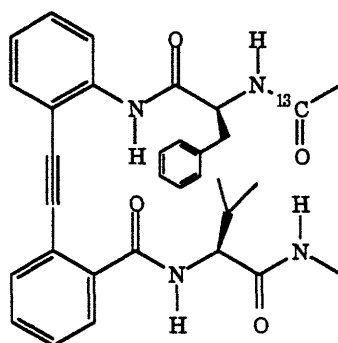
**Ac(<sup>13</sup>C=O)-Ile-Taa-Val-NHMe (16d).** Yield 42 mg, 86% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.75 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.764 (t, J=7.4 Hz, 3H), 0.869 (d, J=6.73 Hz, 3H), 0.874 (d, J=6.73 Hz, 3H), 0.928 (d, J=6.73 Hz, 3H), 1.150 (m, 1H), 1.464 (m, 1H), 1.830 (m, 1H), 1.927 (d, J=6.06 Hz, 3H), 2.059 (m, 1H), 2.493 (d, J=4.81 Hz, 3H), 4.349 (dd, J=8.74,8.65 Hz; 1H), 4.985 (m, 1H), 7.141 (m, 1H), 7.404 (m, 1H), 7.493-7.536 (m, 2H), 7.566 (td, J=1.44, 7.59 Hz; 1H), 7.700 (m, 1H), 7.775 (m, 1H), 8.096 (br, quart, J=4.71 Hz, 1H), 8.261 (d, J=8.36 Hz; 1H), 8.298 (m, 1H), 8.703 (d, J=8.65 Hz, 1H), 9.444 (s, 1H); MS m/e (rel intensity) 505 (M<sup>+</sup>, 29); HRMS calculated for C<sub>28</sub>H<sub>36</sub>N<sub>4</sub>O<sub>4</sub> 505.277011, found 505.27694.



16e

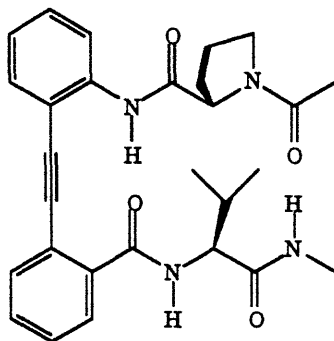
**Ac-Leu-Taa-Val-NHMe (16e).** Yield 40 mg, 83% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.38 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.770 (d, J=6.57 Hz, 3H), 0.813 (d, J=6.58 Hz, 3H), 0.868 (d, J=6.67 Hz, 3H), 0.914 (d, J=6.66 Hz, 3H), 1.561 (m, 2H), 1.65 (m, 1H), 1.861 (s, 3H), 2.047 (m, 1H), 2.479 (d, J=4.69 Hz, 3H), 4.306 (dd, J=8.55,8.64 Hz; 1H), 4.968 (m, 1H), 7.140 (td, J=1.1, 7.5 Hz; 1H), 7.401 (m, 1H), 7.493-7.529 (m, 2H), 7.558 (m, 1H), 7.709-7.739 (m, 2H), 8.103 (br, quart, J=4.7 Hz, 1H), 8.228 (d, J=8.36 Hz; 1H), 8.357 (d, J=7.89 Hz, 1H), 8.672 (d, J=8.64 Hz, 1H), 9.502 (s, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 18.695, 19.171, 22.749, 23.317, 23.470, 24.960, 25.850, 32.254, 42.388, 52.677, 58.082, 89.699, 94.674, 112.272,

119.735, 122.238, 123.436, 127.259, 128.318, 129.777, 130.837, 132.219, 133.816, 135.751, 140.342, 166.969, 170.116, 172.097, 172.497; MS m/e (rel intensity) 504 (M<sup>+</sup>, 22), 376 (39), 349 (66), 345 (39), 318 (34), 317 (50), 291 (47), 263 (25), 248 (24), 236 (22), 220 (100), 219 (95), 165 (21), 128 (22), 86 (70), 72 (25); HRMS calculated for C<sub>29</sub>H<sub>36</sub>N<sub>4</sub>O<sub>4</sub> 504.273657, found 504.27250.

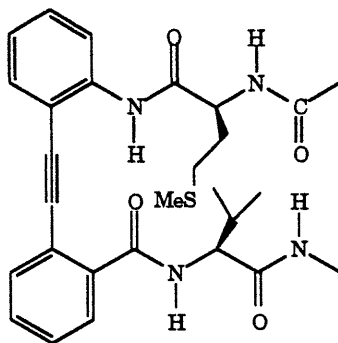


**16f**

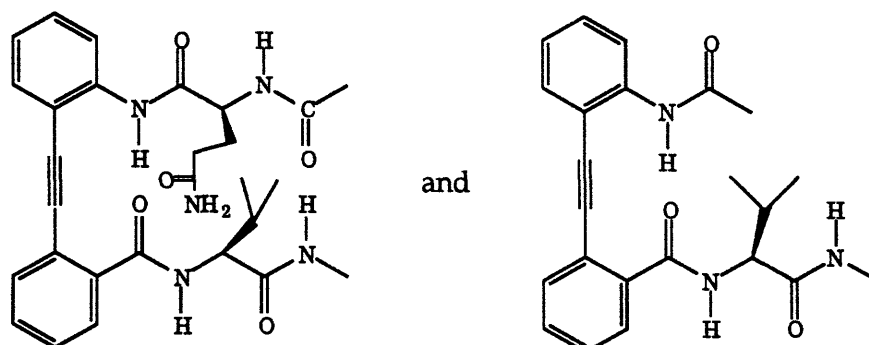
**Ac(<sup>13</sup>C=O)-Phe-Taa-Val-NHMe (16f).** Yield 48.7 mg, 94% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.78 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.837 (d, J=6.76 Hz, 3H), 0.908 (d, J=6.67 Hz, 3H), 1.790 (d, J=6.1 Hz, 3H), 2.037 (m, 1H), 2.461 (d, J=4.6 Hz, 3H), 2.867 (dd, J=10.1, 13.7 Hz; 1H), 3.128 (dd, J=4.6, 13.7 Hz; 1H), 4.282 (t, J=8.5 Hz, 1H), 5.284 (m, 1H), 7.126-7.159 (m, 4H), 7.212-7.231 (m, 2H), 7.411 (m, 1H), 7.505-7.574 (m, 3H), 7.715-7.759 (m, 2H), 8.054 (br, quart, J=4.51 Hz, 1H), 8.254 (d, J=8.36 Hz; 1H), 8.491 (dd, J=3.7, 8.2 Hz; 1H), 8.756 (d, J=8.64 Hz, 1H), 9.557 (s, 1H); MS m/e (rel intensity) 539 (M<sup>+</sup>, 25); HRMS calculated for C<sub>31</sub>C<sup>\*</sup>H<sub>34</sub>N<sub>4</sub>O<sub>4</sub> 539.261361, found 539.26118.

**16g**

**Ac-Pro-Taa-Val-NHMe (16g).** Yield 40 mg, 97% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 13.54 min.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.910 (d,  $J=6.69$  Hz, 3H), 0.975 (d,  $J=6.69$  Hz, 3H), 1.95-2.3 (m, 5H), 2.130 (s, 3H), 2.633 (d,  $J=5.0$  Hz, 3H), 3.579 (m, 1H), 3.774 (m, 1H), 4.654 (t,  $J=8.7$  Hz; 1H), 5.743 (dd,  $J=2.9, 7.7$  Hz; 1H), 6.801 (d,  $J=9.19$  Hz, 1H), 7.007 (td,  $J=1.35, 7.67$  Hz; 1H), 7.288 (m, 1H), 7.358 (m, 1H), 7.409-7.474 (m, 2H), 7.533 (dd,  $J=1.3, 7.7$  Hz; 1H), 7.613 (dd,  $J=1.1, 7.8$  Hz; 1H), 8.069 (br, quart,  $J=4.3$  Hz, 1H), 8.674 (d,  $J=8.0$  Hz, 1H), 9.126 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.921, 19.366, 22.606, 24.786, 25.892, 30.376, 32.311, 48.588, 58.646, 60.197, 89.526, 93.825, 111.423, 119.223, 121.819, 122.955, 126.379, 128.268, 129.911, 130.464, 132.214, 132.782, 137.558, 141.044, 168.868, 169.774, 171.832, 172.139; MS  $m/e$  (rel intensity) 488 ( $\text{M}^+$ , 13), 376 (22), 349 (15), 345 (18), 317 (24), 220 (14), 219 (13), 140 (25), 112 (47), 70 (100); HRMS calculated for  $\text{C}_{28}\text{H}_{32}\text{N}_4\text{O}_4$  488.242357, found 488.24301.

**16h**

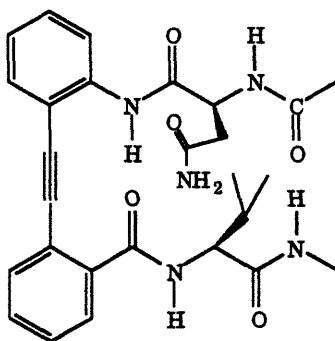
**Ac-Met-Taa-Val-NHMe (16h).** Yield 43 mg, 77% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.08 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.876 (d,  $J=6.73$  Hz, 3H), 0.928 (d,  $J=6.73$  Hz, 3H), 1.843-1.919 (m, 1H), 1.888 (s, 3H), 1.894 (s, 3H), 2.022-2.091 (m, 2H), 2.398-2.550 (m, 2H), 2.490 (d,  $J=4.76$  Hz, 3H), 4.242 (dd,  $J=8.31, 8.71$  Hz; 1H), 5.054 (m, 1H), 7.144 (td,  $J=1.19, 7.52$  Hz; 1H), 7.407 (td,  $J=1.58, 7.92$  Hz; 1H), 7.496-7.577 (m, 3H), 7.732 (m, 2H), 8.041 (quart,  $J=4.75$  Hz, 1H), 8.240 (d,  $J=7.52$  Hz, 1H), 8.446 (d,  $J=7.92$  Hz, 1H), 8.700 (d,  $J=8.71$  Hz, 1H), 9.465 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CD}_2\text{Cl}_2$ ):  $\delta$  11.204, 15.719, 19.020, 19.619, 23.811, 26.237, 30.245, 33.132, 34.591, 58.607, 89.917, 95.552, 112.674, 119.953, 122.747, 123.991, 128.075, 129.012, 130.348, 131.592, 132.682, 134.478, 135.999, 141.081, 167.723, 171.040, 171.577, 172.468; MS  $m/e$  (rel intensity) ( $\text{M}^+$ , 15); HRMS calculated for  $\text{C}_{28}\text{H}_{34}\text{N}_4\text{O}_4\text{S}$  522.230077, found 522.23016.

**16i**

**Ac-Gln-Taa-Val-NHMe (16i) and Ac-Taa-ValNHMe.** Yield 12 mg, 22.4% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 12.00 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.866 (d,  $J=6.73$  Hz, 3H), 0.906 (d,  $J=6.73$  Hz, 3H), 1.864 (s, 3H), 1.894 (m, 1H), 2.042 (m, 2H), 2.169 (m, 2H), 2.502 (d,  $J=3.96$  Hz, 3H), 4.301 (dd,  $J=8.32, 8.71$  Hz; 1H), 4.901 (m, 1H), 6.803 (s, 1H), 7.138 (td,  $J=1.19, 7.52$  Hz; 1H), 7.312 (s, 1H), 7.404 (td,  $J=1.58, 7.82$  Hz; 1H), 7.488-7.568 (m, 3H), 7.694 (dd,  $J=1.19, 7.52$  Hz; 1H), 7.725 (dd,  $J=1.59, 7.92$  Hz; 1H), 8.039 (quart,  $J=4.75$  Hz, 1H), 8.276 (d,  $J=7.52$  Hz, 1H), 8.400 (d,  $J=7.91$  Hz, 1H), 8.669 (d,  $J=8.71$  Hz, 1H), 9.350 (s, 1H); MS  $m/e$  (rel intensity) ( $\text{M}^+$ , 9); HRMS calculated for  $\text{C}_{28}\text{H}_{33}\text{N}_5\text{O}_5$  519.248169, found 519.24815.

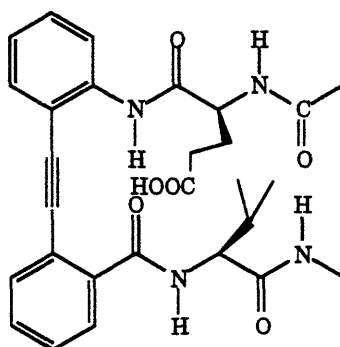
Ac-Taa-ValNHMe (6.9 mg, 17.1%) was isolated as a side product.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.880 (d,  $J=6.83$  Hz; 3H), 0.884 (d,  $J=6.59$  Hz; 3H), 2.067 (m, 1H), 2.307 (s, 3H), 2.557 (d,  $J=4.64$  Hz; 3H), 4.263 (dd,  $J=7.32, 8.3$  Hz; 1H), 7.101 (m, 1H), 7.382 (m, 1H), 7.500-7.540 (m, 2H), 7.566 (m, 1H), 7.683 (m, 1H), 7.744 (m, 1H), 7.932 (q,  $J=4.6$  Hz; 1H), 8.304 (d,  $J=8.3$  Hz; 1H), 8.512 (d,  $J=8.79$  Hz; 1H), 9.240 (s, 1H).

Note:  $\text{D}_2\text{O}$  exchange was done on Ac-Taa-ValNHMe. Rough exchange rate: MeNH, 1 hr; aryl NH, 1 day; ValNH, 3 days.



**16j**

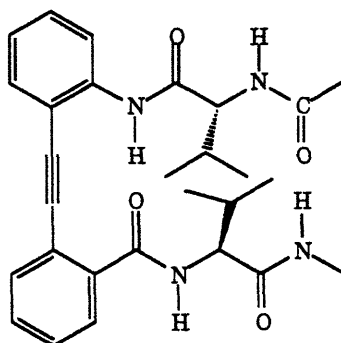
**Ac-Asn-Taa-Val-NHMe (16j).** Yield 27.1 mg, 55.4% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 11.66 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.868 (d,  $J=6.73$  Hz, 3H), 0.892 (d,  $J=6.73$  Hz, 3H), 1.784 (s, 3H), 2.018 (m, 1H), 2.514 (d,  $J=4.75$  Hz, 3H), 2.532 (dd,  $J=7.52, 15.64$  Hz; 1H), 2.696 (dd,  $J=61.4, 15.64$  Hz; 1H), 4.324 (dd,  $J=8.31, 8.71$  Hz; 1H), 4.989 (m, 1H), 6.975 (s, 1H), 7.125 (td,  $J=0.79, 7.52$  Hz; 1H), 7.3714-7.409 (m, 2H), 7.483-7.557 (m, 3H), 7.638 (dd,  $J=1.19, 7.53$  Hz; 1H), 7.824 (dd,  $J=0.99, 7.72$  Hz; 1H), 8.024 (quart,  $J=4.75$  Hz, 1H), 8.237 (d,  $J=8.31$  Hz, 1H), 8.428 (d,  $J=7.52$  Hz, 1H), 8.565 (d,  $J=8.71$  Hz, 1H), 9.352 (s, 1H); MS  $m/e$  (rel intensity) ( $M^+$ , 3); HRMS calculated for  $\text{C}_{27}\text{H}_{31}\text{N}_5\text{O}_5$  505.232519, found 505.23269. Ac-Taa-ValNHMe is also a side product here.



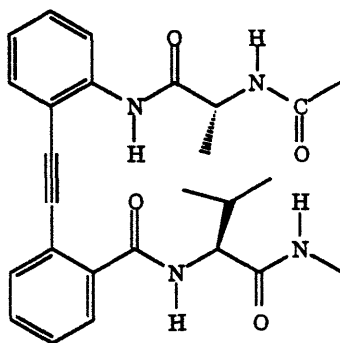
**16k**

**Ac-Glu-Taa-Val-NHMe (16k).** Yield 13.6 mg, 24.5% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 12.28 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.873 (d,  $J=6.73$  Hz, 3H), 0.918 (d,  $J=6.73$  Hz, 3H), 1.879 (s, 3H), 1.902 (m, 1H), 2.035 (m, 2H), 2.297 (m, 2H), 2.498 (d,  $J=4.73$  Hz, 3H), 4.300 (dd,  $J=8.32, 8.71$  Hz; 1H), 4.977 (m, 1H), 7.144 (td,  $J=1.19, 7.52$  Hz; 1H), 7.407 (td,  $J=1.58, 7.92$  Hz; 1H), 7.491-7.568 (m, 3H), 7.724 (dd,  $J=1.18, 7.52$  Hz; 2H), 8.010 (quart,  $J=4.75$  Hz, 1H), 8.250 (d,  $J=8.32$  Hz, 1H), 8.419 (d,  $J=7.91$  Hz, 1H), 8.662 (d,  $J=8.71$  Hz, 1H), 9.383 (s, 1H); MS  $m/e$  (rel intensity) ( $M^+$ , 6); HRMS calculated for  $\text{C}_{28}\text{H}_{32}\text{N}_4\text{O}_6$  520.232185, found 520.23221.

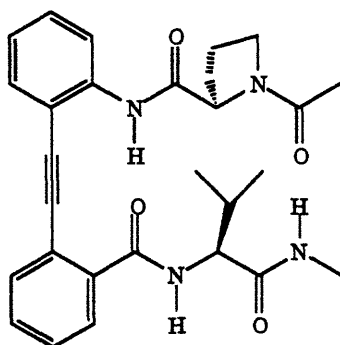


**16l**

**Ac-D-Val-Taa-Val-NHMe (16l).** Yield 37.4 mg, 71% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min, RCM column): RT 11.08 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.809 (d,  $J=6.84$  Hz, 3H), 0.839 (d,  $J=6.84$  Hz, 3H), 0.873 (d,  $J=6.84$  Hz, 3H), 0.887 (d,  $J=6.35$  Hz, 3H), 1.930 (s, 3H), 1.995 (m, 1H), 2.120 (m, 1H), 2.604 (d,  $J=4.39$  Hz, 3H), 4.475 (dd,  $J=8.30, 8.79$  Hz; 1H), 5.008 (dd,  $J=6.35, 8.79$  Hz; 1H), 7.143 (t,  $J=7.57$  Hz, 1H), 7.409 (td,  $J=1.47, 8.06$  Hz; 1H), 7.503-7.573 (m, 3H), 7.648 (dd,  $J=1.47, 7.32$  Hz; 1H), 7.702 (dd,  $J=0.98, 7.32$  Hz; 1H), 7.957 (q,  $J=4.4$  Hz, 1H), 8.262 (d,  $J=8.3$  Hz, 2H), 8.653 (d,  $J=8.78$  Hz, 1H), 9.377 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  17.136, 18.492, 19.257, 23.735, 25.865, 32.230, 33.361, 57.311, 57.940, 89.485, 94.264, 111.969, 119.284, 122.338, 123.231, 126.737, 128.383, 129.828, 130.716, 132.016, 133.179, 136.136, 140.883, 166.772, 170.018, 171.508, 172.433.

**16m**

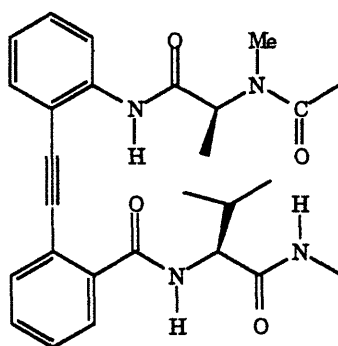
**Ac-D-Ala-Taa-Val-NHMe (16m).** Yield 43.2 mg, 77.1% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min, RCM column): RT 9.81 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  0.831 (d,  $J=7.32$  Hz, 3H), 0.846 (d,  $J=7.81$  Hz, 3H), 1.344 (d,  $J=7.81$  Hz, 3H), 1.829 (s, 3H), 2.016 (m, 1H), 2.595 (d,  $J=4.89$  Hz, 3H), 4.392 (t,  $J=8.3$  Hz, 1H), 4.920 (quint,  $J=7.32$  Hz, 1H), 7.137 (td,  $J=0.98, 7.57$  Hz; 1H), 7.406 (td,  $J=1.46, 8.06$  Hz; 1H), 7.505-7.535 (m, 2H), 7.561 (td,  $J=1.47, 7.57$  Hz; 1H), 7.674 (dd,  $J=1.46, 7.33$  Hz; 1H), 7.735 (dd,  $J=1.22, 7.57$  Hz; 1H), 7.920 (br q,  $J=4.88$  Hz, 1H), 8.270 (d,  $J=8.3$  Hz, 1H), 8.369 (d,  $J=7.32$  Hz, 1H), 8.580 (d,  $J=8.79$  Hz, 1H), 9.334 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.526, 19.281, 20.846, 23.608, 25.976, 32.393, 49.497, 58.042, 89.583, 94.845, 112.123, 119.333, 122.455, 123.359, 126.929, 128.370, 129.833, 130.878, 132.025, 133.509, 135.395, 140.769, 166.818, 169.533, 172.343, 172.498.



**16n**

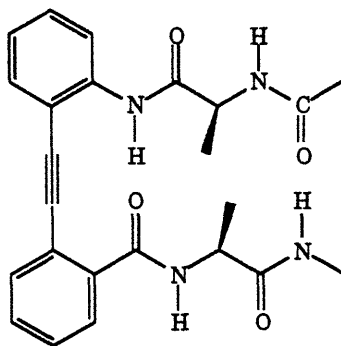
**Ac-D-Pro-Taa-Val-NHMe (16n).** Yield 43.6 mg, 74% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min, RCM column): RT 9.58 min.  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ): two conformers with a ratio of 3.0 : 1;  $\delta$  0.85 and 1.174 (t,  $J=6.84$  Hz and 7.32 Hz, 6H), 1.894 and 1.942 (s, 3H), 1.772-2.185 (m, 5H), 3.101 and 3.378-3.603 (m, 2H), 4.270 and 4.314 (dd,  $J=7.32, 8.3$  Hz; 1H), 4.864 (dd,  $J=2.93, 3.91, 7.81, 8.79$ ) and 4.897 (dd,  $J=2.93, 3.41, 8.31, 8.79$ ; 1H), 7.124 and 7.185 (td,  $J=0.98, 7.33$  Hz; 1H), 7.369-7.438 (m, 1H), 7.507-7.590 (m, 3H), 7.671 (dd,  $J=1.46, 7.32$ ) and 7.700 (dd,  $J=0.98, 1.46,$

7.33, 7.81 Hz; 1H), 7.738-7.764 (m, 1H), 7.834 and 7.955 (q, J=4.4 Hz, 1H), 8.059 and 8.215 (d, J=8.3 Hz, 1H), 8.477 and 8.487 (d, J=7.81 and 8.79 Hz, 1H), 9.566 and 9.383 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.060, 19.147, 22.205, 24.394, 25.762, 29.367, 31.094, 48.232, 59.156, 60.085, 89.643, 93.370, 111.550, 119.642, 120.711, 123.041, 126.962, 128.562, 129.828, 129.970, 131.770, 132.304, 138.258, 140.514, 168.090, 170.410, 171.161, 172.014.



**16o**

**Ac-Me<sup>N</sup>-Ala-Taa-Val-NHMe (16o).** Yield 31 mg, 69% (in 3 steps) after flash column chromatography purification (silica, twice, first with 5% MeOH in EtOAc, then with 2% MeOH in EtOAc).  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.931 (d, J=6.35 Hz, 3H), 0.977 (d, J=6.84 Hz, 3H), 1.565 (d, J=7.32 Hz, 3H), 2.03 (m, 1H), 2.131 (s, 3H), 2.678 (d, J=4.88 Hz, 3H), 3.179 (s, 3H), 4.803 (dd, J=8.3, 8.8 Hz; 1H), 6.136 (q, J=7.33 Hz, 1H), 6.867 (br d, J=8.79 Hz, 1H), 7.009 (td, J=0.98, 7.57 Hz; 1H), 7.289 (td, J=1.46, 8.06 Hz; 1H), 7.363 (td, J=1.46, 7.57 Hz; 1H), 7.426-7.472 (m, 2H), 7.599 (dd, J=0.98, 7.81 Hz; 1H), 7.623 (dd, J=0.98, 7.58 Hz; 1H), 8.028 (br m, J=3.91 Hz, 1H), 8.558 (dd, J=0.97, 8.3 Hz; 1H), 9.147 (s, 1H) 1.942.

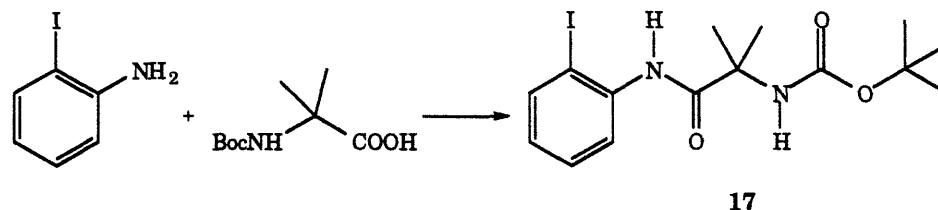


15

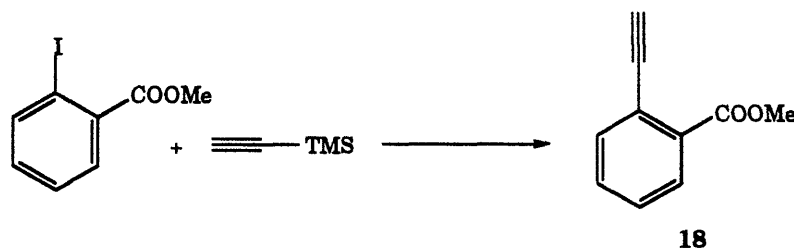
**Ac-Ala-Taa-Ala-NHMe (15)** was synthesized by the same procedure. Yield 52 mg, 100% after HPLC (in 3 steps). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.22 min.  $^1\text{H}$  NMR (500 MHz, DMSO- $d_6$ ):  $\delta$  1.317 (d,  $J=7.08$  Hz, 3H), 1.365 (d,  $J=7.08$  Hz, 3H), 1.855 (s, 3H), 2.523 (d,  $J=4.4$  Hz, 3H), 4.526 (m, 1H), 4.905 (quint,  $J=7.08$  Hz, 1H), 7.142 (td,  $J=1.22, 7.57$  Hz; 1H), 7.414 (m, 1H), 7.51-7.59 (m, 3H), 7.736 (dd,  $J=1.22, 7.57$  Hz; 1H), 7.802 (m, 1H), 7.927 (br m, 1H), 8.316 (d,  $J=7.82$  Hz, 1H), 8.371 (d,  $J=7.32$  Hz, 1H), 8.791 (d,  $J=7.32$  Hz, 1H), 9.430 (s, 1H).

**N-BocAibOH.**  $\alpha$ -Aminoisobutyric acid (10.3 g, 0.1 mol) was added to a solution of 1 N NaOH (100 mL) in 50 mL of water and 150 mL of 1,4-dioxane at 0 °C in an ice bath. Di-*t*-butyl dicarbonate (27.6 g, 0.12 mol) was added and it was vigorously stirred at 0 °C for 3 hr and at room temperature for 18 hr. The solvents were then removed by evaporation under reduced pressure and the residue was taken up in 200 mL of EtOAc, washed with 200 mL of 1N KHSO<sub>4</sub>. The aqueous layer was then extracted with EtOAc (2 x 100 mL). The combined organic layer was washed with water, saturated NaCl, dried with MgSO<sub>4</sub> and filtered. The filtrate was concentrated under reduced pressure and dried *in vacuo* to afford a white powder (12.9 g, 64%). A fraction of the product was recrystallized from EtOH/water. Mp 124-124.5 °C;  $^1\text{H}$  NMR (300 MHz, CDCl<sub>3</sub>):

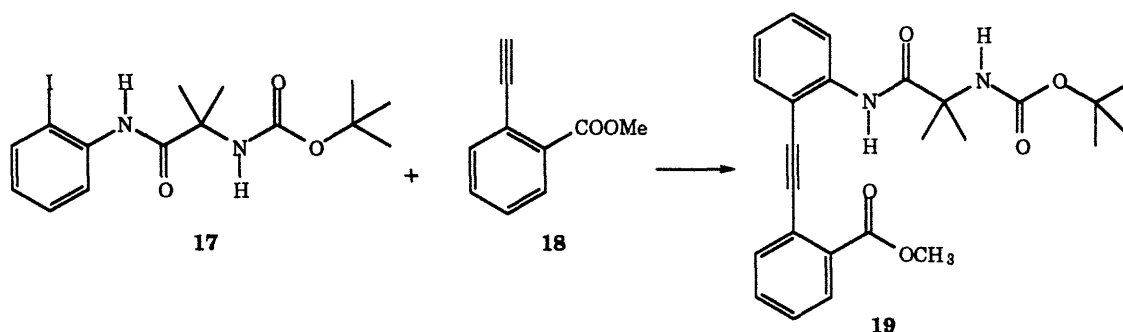
$\delta$  1.420 (s, 9H), 1.510 (s, 6H), 5.08 (br s, 1H), 11.10 (br s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  25.398, 28.361, 56.184, 80.489, 155.228, 179.425.



**N-BocAib-(2-iodo)anilide (17).** Dicyclohexylcarbodiimide (0.51 g, 2.47 mmol) was added to a solution of Boc-AibOH (0.52 g, 2.56 mmol) and 2-iodoaniline (1.00 g, 4.57 mmol) in 25 mL of dry THF at 0 °C. The solution was stirred at 0 °C for 4 hr and at room temperature for 20 hr under nitrogen. It was concentrated under reduced pressure and the residue was taken up in 50 mL of EtOAc, washed with pH 3.5 citric buffer, saturated  $\text{NaHCO}_3$ , saturated  $\text{NaCl}$ , dried with  $\text{MgSO}_4$ , and filtered. The filtrate was concentrated under reduced pressure. Flash column chromatography (15% EtOAc in hexane) afforded a slightly beige solid (0.29 g, 30%). TLC (15% EtOAc in hexane)  $R_f$  0.13;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.417 (s, 9H), 1.571 (s, 6H), 4.92 (br s, 1H), 6.803 (td,  $J=1.5, 7.8$  Hz; 1H), 7.319 (td,  $J=1.4, 7.8$  Hz; 1H), 7.742 (dd,  $J=1.4, 7.9$  Hz; 1H), 8.311 (dd,  $J=1.5, 8.2$  Hz; 1H), 8.541 (br s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  25.843, 28.460, 57.854, 80.796, 89.416, 121.079, 125.457, 129.130, 138.263, 138.551, 154.297, 172.697.



**Methyl 2-ethynylbenzoate (18).** A mixture of methyl 2-iodobenzoate (11.7 g, 44.6 mmol), ethynyltrimethylsilane (5.0 g, 50.9 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (200 mg, 0.17 mmol) and cuprous iodide (70 mg, 0.37 mmol) in 25 mL of diethylamine and 5 mL of DMF was stirred under nitrogen at ambient temperature overnight. The reaction mixture was concentrated under reduced pressure, then treated with 50 mL of diethyl ether and 50 mL of 0.1 N aqueous HCl solution. After the organic layer was separated, the aqueous layer was extracted twice with ether. The combined ether layer was washed with saturated NaHCO<sub>3</sub>, then with saturated NaCl, dried with anhydrous MgSO<sub>4</sub>. After filtration under suction the ether solution was concentrated under aspirator pressure. The crude product was then treated with KF·2H<sub>2</sub>O (13 g, 0.14 mol) in 200 mL of methanol. After 14 hr at room temperature with stirring, it was concentrated under aspirator pressure, followed by the addition of 50 mL of distilled water. It was then extracted with ether (3 x 50 mL). The combined ether layer was washed with brine, dried with anhydrous MgSO<sub>4</sub>, filtered and concentrated under aspirator pressure. Flash chromatography (silica) eluting with 1 EtOAc/9 hexane gave methyl 2-ethynylbenzoate (5.0 g, 70%). In addition, 3.37 g of methyl 2-iodobenzoate was recovered (recovered yield: 98%). TLC (1 EtOAc/9 hexane) R<sub>f</sub> 0.29; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 3.368 (s, 1H), 3.889 (s, 3H), 7.358 (td, J=1.5, 7.6 Hz; 1H), 7.434 (td, J=1.5, 7.6Hz; 1H), 7.584 (dd, J=1.5, 7.6Hz; 1H), 7.896 (dd, J=1.5, 7.6Hz; 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 52.093, 81.929, 82.175, 122.575, 128.411, 130.207, 131.635, 132.464, 134.860, 166.339.



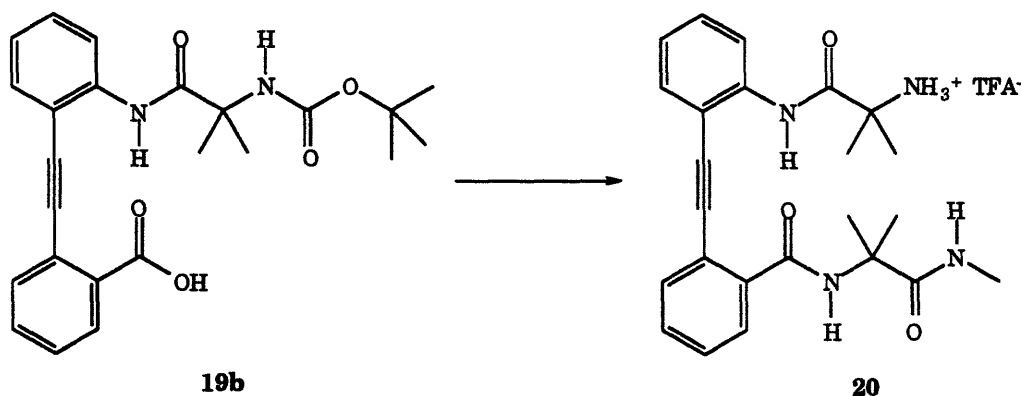
**N-BocAib-Taa-OMe (19).** A mixture of N-BocAib-(2-iodo)anilide (0.24 g, 0.61 mmol), methyl 2-ethynylbenzoate (0.12 g, 0.75 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (60 mg, 0.05 mmol) and cuprous iodide (40 mg, 0.21 mmol) in 6 mL of triethylamine and 2 mL of DMF was stirred under nitrogen in a sealed tube. It was heated with an oil bath at 80 °C (bath temp) for 2.5 hr, then at 60 °C for 21 hr. The reaction mixture was then concentrated under reduced pressure, and treated with 50 mL of EtOAc and 50 mL of pH 3.5 citric buffer solution. After the organic layer was separated, the aqueous layer was extracted twice with EtOAc. The combined organic layer was washed with saturated NaHCO<sub>3</sub> and saturated NaCl, then dried with anhydrous MgSO<sub>4</sub>. After filtration under suction the EtOAc solution was concentrated under aspirator pressure. Flash chromatography (silica) (1 EtOAc/3 hexane) gave the desired tolanoic peptide **19** (0.23 g, 85%). TLC (1 EtOAc/4 hexane) R<sub>f</sub> 0.20; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 1.172 (s, 9H), 1.625 (s, 6H), 3.956 (s, 3H), 5.914 (br s, 1H), 7.021 (t, J=7.4 Hz, 1H), 7.31-7.42 (m, 2H), 7.47-7.56 (m, 2H), 7.691 (d, J=7.6Hz, 1H), 8.038 (d, J=7.9 Hz, 1H), 8.569 (br d, J=7.1 Hz, 1H), 9.175 (br s, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 25.467, 27.970, 52.447, 57.453, 79.473, 90.329, 94.659, 111.719, 119.198, 122.868, 124.081, 128.119, 130.085, 130.162, 130.591, 132.250, 132.327, 134.461, 139.974, 154.316, 166.539, 174.094.





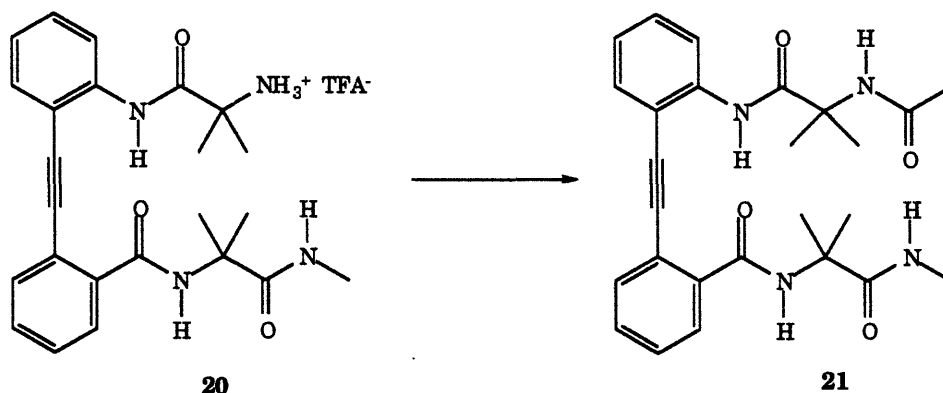
filtrate was then concentrated under reduced pressure and was added to 40 mL of water and 4 mL of 5 N NaOH. The mixture was then sonicated to give a fine suspension and filtered. The filtrate was then extracted with chloroform twice. Evaporation of the solvent under reduced pressure gave pure N-BocAibNHMe as a white powder (0.47 g, 82%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.40 (s, 9H), 1.45 (s, 6H), 2.78 (d,  $J=4.9$  Hz, 3H), 4.93 (br s, 1H), 6.44 (br, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  25.774, 26.434, 28.307, 56.761, 80.163, 154.838, 175.245.

**HCl·HAibNHMe.** N-BocAibNHMe (0.47 g, 2.16 mmol) in 6 mL of dry 1,4-dioxane was treated with 2 mL of 6.5 M HCl in 1,4-dioxane for 0.5 hr. Then the solvent was removed by evaporation under reduced pressure; the residue was triturated with ether and dried *in vacuo* (100%).  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  1.425 (s, 6H), 2.638 (d,  $J=4.1$  Hz, 3H), 8.188 (br s, 3H), 8.269 (br quart,  $J=4.1$  Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{DMF-d}_7$ ):  $\delta$  24.337, 26.471, 58.020, 172.649.



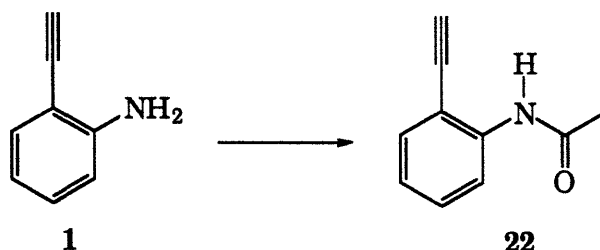
**TFA·HAib-Taa-AibNHMe (20).** A solution of N-BocAib-Taa-OH (0.13 g, 0.31 mmol), HCl·HAibNHMe (0.20 g, 1.28 mmol) and 1-hydroxybenzotriazole (0.17 g, 1.28 mmol) in 8 mL of DMF was treated with diisopropylethylamine (0.22 mL, 1.26 mmol) and cooled to 0 °C in an ice bath. Diisopropylcarbodiimide (0.20 mL, 1.28 mmol) was added. After the solution

was stirred at room temperature for 9 days, it was concentrated *in vacuo*. Preparative TLC (3 EtOAc/1 hexane) gave 28 mg of the crude product, which was then treated with 2 mL of trifluoroacetic acid in 2 mL of methylene chloride at 0 °C for 0.5 hr. After the solvent was removed by evaporation under reduced pressure, the crude salt was purified by reversed phase HPLC to give the title compound (24 mg, 14%). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 3.62 min.

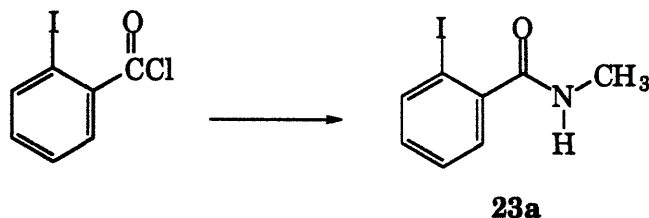


**Ac-Aib-Taa-Aib-NHMe (21).** **20** (24 mg, 0.04 mmol) was treated with 2 mL of 0.05 N NaOH aqueous solution; white precipitate was immediately observed. THF (2 mL) was then added to form a clear solution, which was cooled to 0 °C in an ice bath. Acetic anhydride (0.035 mL, 0.37 mmol) was added via a syringe. The ice was allowed to warm up to room temperature and stirring was continued overnight. The solvents were then removed under reduced pressure. Water (1 mL) and 0.5 mL of 1 N NaOH were added to the residue. Extraction with chloroform and subsequent evaporation of the solvent yielded 20 mg of the crude product. Preparative HPLC afforded the pure product (18 mg, 89 %). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 11.20 min;  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ):  $\delta$  1.414 (s, 6H), 1.443 (s, 6H), 1.553 (s, 3H), 2.514 (d,  $J=4.64$  Hz, 3H), 7.119 (td,  $J=1.10, 7.57$  Hz; 1H), 7.396 (m, 1H), 7.5-7.6 (m, 4H), 7.70-7.73 (m, 2H), 8.103 (s, 1H), 8.187 (d,  $J=7.44$  Hz; 1H), 8.480 (s, 1H), 9.021 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$

22.749, 25.159, 26.772, 57.821, 58.635, 89.546, 92.755, 111.750, 120.026, 120.257, 123.773, 127.766, 129.317, 130.453, 130.714, 131.712, 133.324, 137.440, 139.467, 168.074, 172.113, 173.003, 175.046; MS  $m/e$  (rel intensity) 462 ( $M^+$ , 14), 404 (12), 386 (24), 362 (30), 344 (47), 335 (40), 332 (57), 331 (26), 329 (20), 305 (22), 304 (69), 303 (19), 278 (28), 277 (100), 263 (28), 260 (27), 248 (38), 237 (22), 231 (18), 221 (25), 220 (100), 219 (99), 217 (18), 201 (18); HRMS calculated for  $C_{26}H_{30}N_4O_4$  462.226707, found 462.22652.

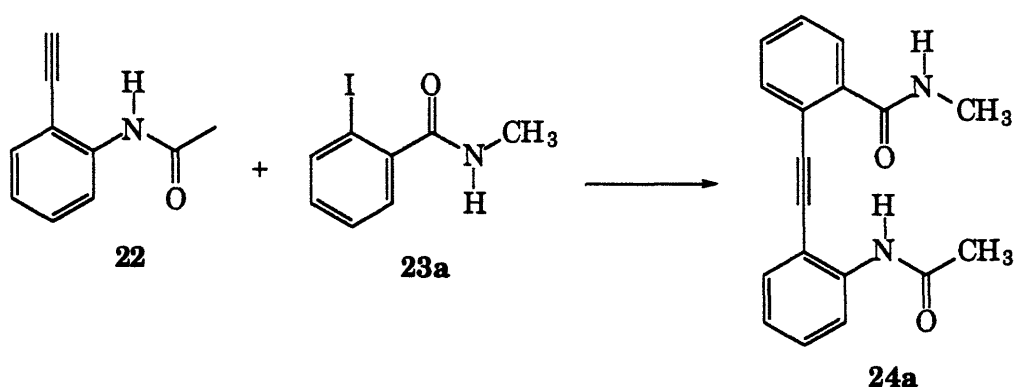


**Acetyl 2-ethynylanilide (22).** Acetic anhydride (6 mL, 63.6 mmol) was added to a solution of 2-ethynylaniline (16.45 mmol) in 50 mL of chloroform at 0 °C. Stirring was continued under nitrogen at room temperature overnight. The solution was then concentrated under reduced pressure and dried over NaOH *in vacuo* to give a slightly yellow foam. (3.2 g, 95%).  $^1H$  NMR (300 MHz,  $CDCl_3$ ):  $\delta$  2.200 (s, 3H), 3.486 (s, 1H), 7.010 (m, 1H), 7.332 (m, 1H), 7.424 (d,  $J=7.6$  Hz, 1H), 7.887 (br s, 1H), 8.371 (d,  $J=8.3$  Hz, 1H);  $^{13}C$  NMR (300 MHz,  $CDCl_3$ ):  $\delta$  24.942, 79.218, 84.281, 110.386, 119.199, 123.150, 130.083, 132.031, 139.477, 168.045.



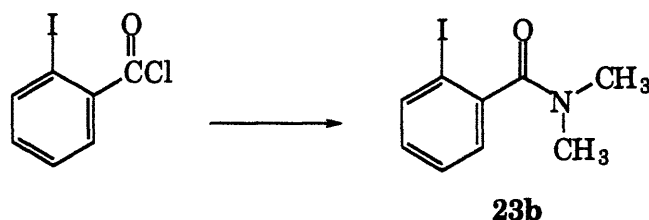
**N-Methyl 2-iodobenzamide (23a).** A mixture of  $MeNH_2 \cdot HCl$  (2 g, 29.6 mmol) and  $K_2CO_3$  (5 g, 36.2 mmol) in 40 mL of dry THF was stirred

vigorously at 5 °C. 2-Iodobenzoyl chloride (0.95 g, 3.6 mmol) was added, the reaction flask was sealed with a rubber septum under nitrogen. Vigorous stirring was maintained overnight. After it was filtered, the filtrate was concentrated under reduced pressure to give a white powder (0.78 g, 84%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 2.98 (d, J=4.8 Hz, 3H), 5.81 (br, 1H), 7.06 (m, 1H), 7.31-7.97 (m, 2H), 7.83 (d, J=7.9 Hz, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 26.741, 92.386, 128.103, 128.226, 131.021, 139.789, 142.354, 170.009.

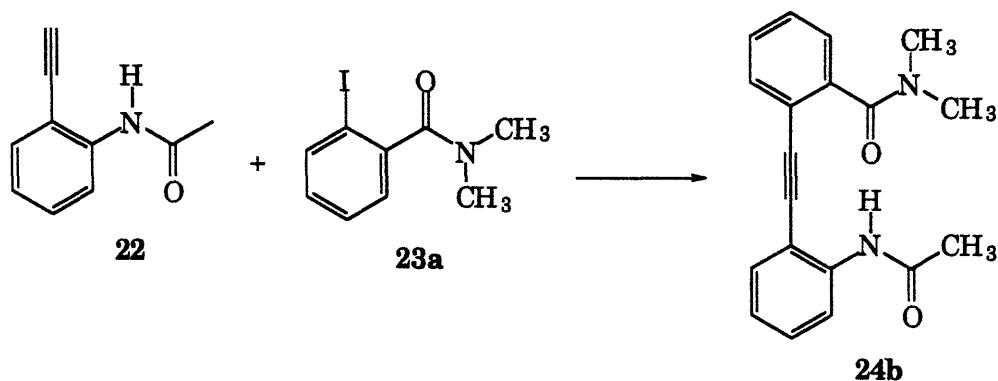


**Ac-Taa-NHMe (24a).** A mixture of acetyl 2-ethynylanilide (0.53 g, 3.29 mmol), N-methyl 2-iodobenzamide (0.77 g, 2.95 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (50 mg, 0.04 mmol) and cuprous iodide (30 mg, 0.14 mmol) in 20 mL of diethylamine and 5 mL of DMF was stirred under nitrogen at ambient temperature for 6 days. The reaction mixture was concentrated, first under aspirator pressure, then in *vacuo* before it was treated with 100 mL of chloroform and filtered. The filtrate was washed with 0.1 N aqueous HCl solution, saturated NaHCO<sub>3</sub> and water. After drying with anhydrous MgSO<sub>4</sub>, the filtrate was concentrated under aspirator pressure. Extraction with 20 mL of 4 EtOAc/1 hexane gave a yellowish powder (0.65 g, 76%). A fraction of it was subject to flash chromatography (silica) eluting with 4 EtOAc/1 hexane followed by recrystallization from chloroform (slow evaporation) to yield white fine needles. HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 12.53 min; <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 2.385 (s, 3H), 2.826 (d,

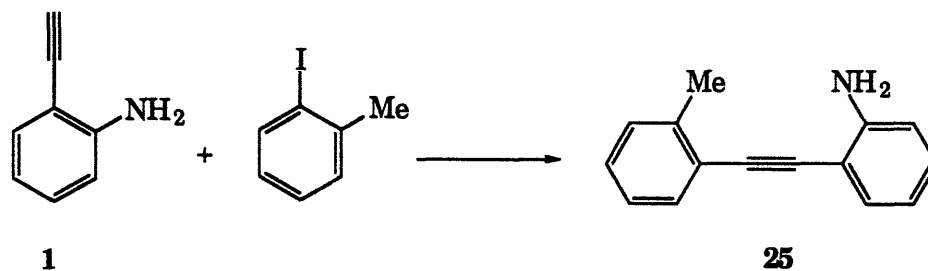
$J=4.59$  Hz, 3H), 7.103 (td,  $J=1.15, 7.52$  Hz; 1H), 7.390 (m, 1H), 7.500 (m, 2H), 7.569 (td,  $J=1.58, 7.44$  Hz; 1H), 7.684 (m, 1H), 7.712 (m, 1H), 8.400 (d,  $J=8.52$  Hz, 1H), 8.661 (quart,  $J=4.49$  Hz, 1H), 9.436 (s, 1H);  $^{13}\text{C}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  24.591, 27.064, 90.006, 93.876, 111.443, 119.796, 121.946, 122.775, 126.676, 128.272, 129.977, 130.576, 131.697, 133.324, 136.779, 141.079, 168.489, 170.424; MS  $m/e$  (rel intensity) 292 ( $\text{M}^+$ , 9), 250 (19), 232 (22), 220 (24), 219 (100), 190 (24), 165 (15); HRMS calculated for  $\text{C}_{18}\text{H}_{16}\text{N}_2\text{O}_2$  292.12118, found 292.1210.



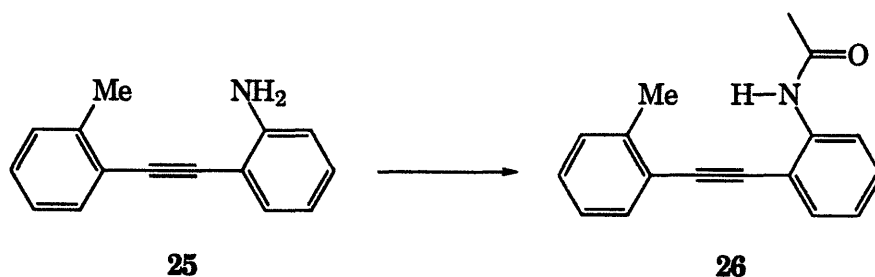
***N,N*-Dimethyl 2-iodobenzamide (23b).** A mixture of  $\text{Me}_2\text{NH}\cdot\text{HCl}$  (1 g, 12 mmol) and  $\text{K}_2\text{CO}_3$  (2 g, 14.5 mmol) in 30 mL of dry THF was stirred vigorously at 5 °C. 2-Iodobenzoyl chloride (1.0 g, 3.75 mmol) was added, then the reaction flask was sealed with a rubber septum under nitrogen. Vigorous stirring was maintained overnight. After it was filtered, the filtrate was concentrated under reduced pressure to give a clear oil (0.95 g, 92%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  2.831 (s, 3H), 3.122 (s, 3H), 7.047 (m, 1H), 7.195 (dd,  $J=1.8, 7.6$  Hz; 1H), 7.371 (td,  $J=1.0, 7.5$  Hz; 1H), 7.800 (m, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  34.636, 38.339, 92.212, 126.761, 128.154, 129.824, 138.761, 142.498, 170.383.



**Ac-Taa-N(Me)<sub>2</sub> (24b).** Same as for Ac-Taa-NHMe. Acetyl 2-ethynylanilide (0.61 g, 3.84 mmol), N,N-dimethyl 2-iodobenzamide (0.95 g, 3.44 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (50 mg, 0.04 mmol) and cuprous iodide (30 mg, 0.14 mmol) in 20 mL of diethylamine and 5 mL of DMF was stirred under nitrogen at ambient temperature for 4 days. Standard workup and flash chromatography (silica) eluting with 4 EtOAc/1 hexane gave the title compound (0.77 g, 74%). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 12.98 min; <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 2.255 (s, 3H), 2.852 (s, 3H), 3.042 (s, 3H), 7.146 (td, J=1.1, 7.54 Hz; 1H), 7.366-7.520 (m, 5H), 7.667-7.697 (m, 1H), 8.065 (d, J=8.12 Hz, 1H), 9.078 (s, 1H); <sup>13</sup>C (300 MHz, CDCl<sub>3</sub>): δ 24.561, 35.064, 38.918, 88.624, 93.062, 111.043, 119.674, 120.579, 122.744, 125.969, 128.303, 129.056, 129.808, 131.512, 131.973, 138.776, 140.511, 170.086, 170.454; MS m/e (rel intensity) 306 (M<sup>+</sup>, 24), 291 (25), 263 (17), 220 (42), 219 (100), 191 (17), 190 (36), 165 (45), 43 (34); HRMS calcd for C<sub>19</sub>H<sub>18</sub>N<sub>2</sub>O<sub>2</sub> 306.1368287, found 306.11310.

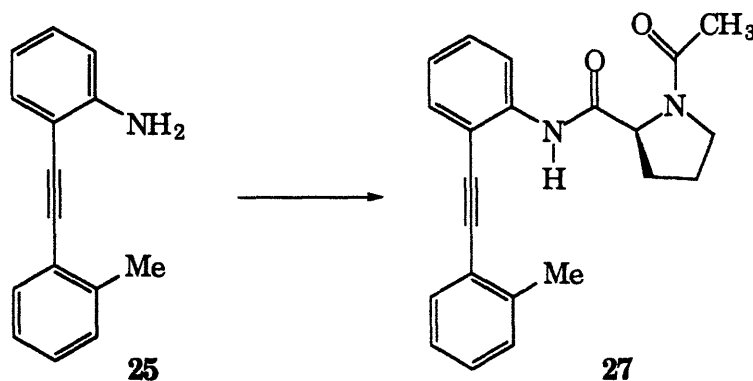


**2-Amino-2'-methyltolane(25).** A mixture of 2-ethynylaniline (0.41 g, 3.5 mmol), 2-iodotoluene (0.73 g, 3.3 mmol), Pd(PPh<sub>3</sub>)<sub>4</sub> (60 mg, 0.052 mmol) and cuprous iodide (40 mg, 0.18 mmol) in 7 mL of diethylamine and 2 mL of DMF was stirred under nitrogen at ambient temperature for 2 days. The reaction mixture was concentrated, first under aspirator pressure, then in *vacuo*, before it was treated with 50 mL of ether and 50 mL of 0.1 N aqueous HCl solution. The organic phase was separated and washed with saturated NaHCO<sub>3</sub>, water, then saturated NaCl. After drying with anhydrous MgSO<sub>4</sub>, the ether filtrate was concentrated under aspirator pressure. Flash chromatography (silica) eluting with 1 EtOAc/9 hexane gave the title compound (0.53 g, 76%) as a slightly yellow solid. TLC R<sub>f</sub> (15% EtOAc in hexane): 0.36; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 2.51 (s, 3H), 4.27 (br s, 2H), 6.70-6.75 (m, 2H), 7.11-7.24 (m, 4H), 7.37 (m, 1H), 7.49 (d, J=7.0 Hz, 1H); <sup>13</sup>C (300 MHz, CDCl<sub>3</sub>): δ 21.076, 89.852, 93.585, 108.227, 114.259, 117.906, 123.062, 125.547, 128.135, 129.389, 129.532, 131.644, 132.009, 139.503, 147.468.



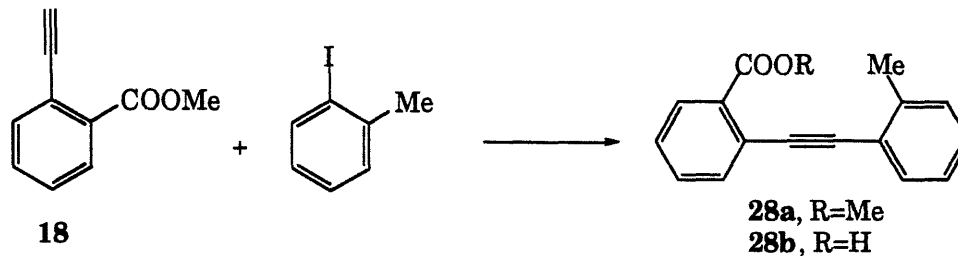
**2-Acetamido-2'-methyltolane [AcNH-Tln-Me] (26).** A solution of 2-amino-2'-methyl-tolane (0.1 g, 0.48 mmol) and acetic anhydride (0.2 mL, 2.12 mmol) in 4 mL of chloroform was stirred under nitrogen at room temperature for 20 hr. Evaporation of the solvents under reduced pressure followed by drying *in vacuo* gave a white powder (0.12 g, 100%). HPLC (from 20 to 100% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 14.84 min; <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 2.111 (s, 3H), 2.497 (s, 3H), 7.192 (t, J=7.47 Hz, 1H), 7.253 (m, 1H), 7.302-7.342 (m, 2H), 7.383 (td, J=1.46, 7.82 Hz; 1H),

7.562 (dd,  $J=1.1, 7.66$  Hz; 2H), 7.74 (d,  $J=7.84$  Hz, 1H), 9.455 (s, 1H);  $^{13}\text{C}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  21.099, 25.004, 88.247, 95.301, 112.021, 119.237, 122.152, 123.313, 125.812, 128.863, 129.554, 129.585, 131.535, 131.781, 138.628, 139.474, 167.893; MS  $m/e$  (rel intensity) 249 ( $\text{M}^+$ , 28), 234 (18), 207 (42), 206 (100), 204 (24), 179 (15), 178 (22), 43 (81); HRMS calculated for  $\text{C}_{17}\text{H}_{15}\text{NO}$  249.115364, found 249.11496.



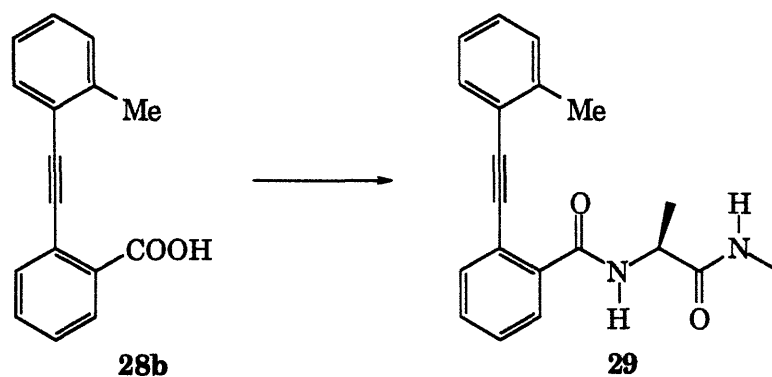
**Ac-Pro-NH-Tln-Me (27).** The title compound was obtained as a white foam from amino-methyltolane and BocProline in two steps: the standard water soluble carbodiimide EDC assisted amide bond formation followed by deprotection in TFA/ $\text{CH}_2\text{Cl}_2$  and acetylation with acetic anhydride. Overall yield 77.3 mg, 94.4% after flash column chromatography (silica, 10% EtOAc in hexane). TLC  $R_f$  (EtOAc): 0.48;  $^1\text{H}$  NMR (500 MHz,  $\text{Me}_2\text{SO}-d_6$ ): Two conformers with an average ratio of 3.2.  $\delta$  1.773-2.45 (m, 4H), 1.894 and 1.925 (s, 3H), 2.495 and 2.484 (s, 3H), 3.38-3.573 (m, 2H), 4.573 and 4.660 (m, 1H), 7.163-7.437 (m, 5H), 7.530-7.609 (m, 2H), 7.718 and 7.909 (d,  $J=8.0$  and 8.19 Hz respectively; 1H), 9.545 and 9.763 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{Me}_2\text{SO}-d_6$ ): Two conformers with an average ratio of 3.2.  $\delta$  20.4269, 22.0836, 22.49999, 24.4159, 28.8381, 31.8975, 46.3124, 47.6278, 59.8711, 60.7309, 89.0851, 89.5201, 93.6763, 114.868, 116.762, 122.086, 122.422, 124.356, 125.286, 125.828, 125.907, 128.908, 129.050, 129.200, 129.548, 131.668, 131.895, 132.245, 132.424, 137.978, 138.600, 139.614, 139.706, 168.570, 169.406, 170.398, 171.014.



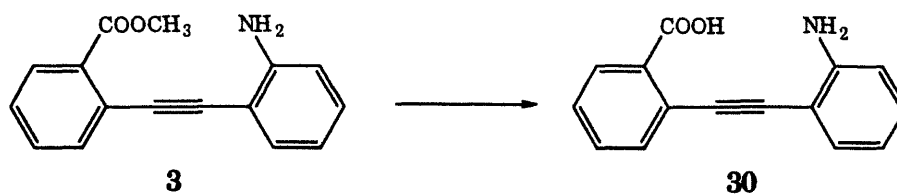


**2-Carbomethoxy-2'-methyltolane (28a).** Same as for 2-amino-2'-methyl-tolane (**25**). Yield 0.62 g (79%) as an oil.  $R_f$  (15% EtOAc in hexane): 0.39;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  2.57 (s, 3H), 3.96 (s, 3H), 7.15-7.27 (m, 3H), 7.38 (m, 1H), 7.50 (td,  $J=1.5, 7.6$  Hz; 1H), 7.55 (d,  $J=7.3$  Hz, 1H), 7.66 (m, 1H), 7.97 (dd,  $J=1.4, 7.8$  Hz, 1H);  $^{13}\text{C}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.795, 52.240, 91.893, 93.370, 122.981, 123.812, 125.455, 127.666, 128.456, 129.370, 130.307, 131.494, 131.586, 132.096, 133.991, 140.293, 166.621.

**2-Carboxy-2'-methyltolane(28b).** 2-Carbomethoxy-2'-methyltolane (0.60 g, 2.4 mmol) was treated with 21 mL of MeOH and 7 mL of 1.75 M aqueous LiOH at room temperature overnight. The solvents were removed by evaporation under reduced pressure and the residue was treated with 20 mL of water, cooled in an ice bath. Precipitate formed after 14 mL of 5N HCl was added, filtered and washed copiously with cold water to obtain a powder (0.52 g, 92%).  $^1\text{H}$  NMR (300 MHz,  $(\text{CD}_3)_2\text{CO}$ ):  $\delta$  2.53 (s, 3H), 7.20-7.27 (m, 3H), 7.46-7.53 (m, 2H), 7.61 (t,  $J=7.6$  Hz, 1H), 7.71 (d,  $J=7.6$  Hz, 1H), 8.02 (d,  $J=7.9$  Hz, 1H);  $^{13}\text{C}$  (300 MHz,  $(\text{CD}_3)_2\text{CO}$ ):  $\delta$  20.977, 93.031, 93.937, 124.031, 124.598, 126.518, 128.958, 129.538, 130.399, 131.322, 132.625, 132.857, 133.132, 134.789, 141.198, 167.175.



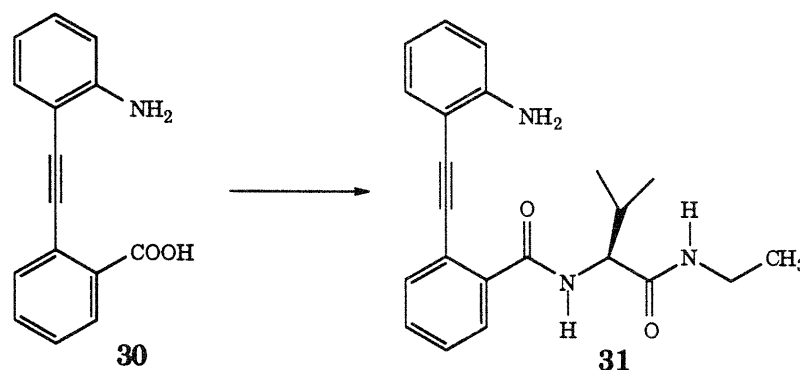
**Me-Tln-(C=O)-AlaNHMe (29).** Me-Tln-COOH (109 mg, 0.46 mmol), and H-Ala-NHMe·HCl (256 mg, 1.85 mmol) were dissolved in 8 mL of dry DMF. 1-Hydroxybenzotriazole (126 mg, 0.92 mmol) was added to the above solution, followed by diisopropylethylamine (0.32 mL, 1.85 mmol). The solution was then cooled to 0 °C with an ice/water bath. Water soluble carbodiimide EtN=C=N(CH<sub>2</sub>)<sub>3</sub>NMe<sub>2</sub>·HCl (177 mg, 0.92 mmol) was then added and the reaction was allowed to proceed at room temperature overnight. The solution was concentrated *in vacuo* and the residue was suspended in water and 1 mL of 1 N NaOH. A white powder (96 mg) was collected by filtration. Reversed phase HPLC gave the title compound (94 mg, 64%). HPLC (from 20 to 35% MeCN in 0.1% aqueous trifluoroacetic acid in 10 min): RT 13.37 min. <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-*d*<sub>6</sub>): δ 1.269 (d, J=7.05 Hz, 3H), 2.437 (s, 3H), 2.568 (d, J=4.61 Hz, 3H), 4.442 (m, 1H), 7.232 (m, 1H), 7.308-7.321 (m, 2H), 7.453-7.523 (m, 3H), 7.613 (dd, J=1.41, 5.08 Hz; 1H), 7.628 (dd, J=1.41, 5.27 Hz; 1H), 7.845 (br quart, J=4.52 Hz, 1H), 8.478 (d, J=7.52 Hz); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>): δ 17.804, 20.768, 26.219, 49.544, 90.928, 95.058, 120.303, 121.685, 125.862, 128.748, 129.240, 129.654, 129.731, 130.821, 131.850, 133.693, 134.706, 140.465, 166.631, 172.358; MS *m/e* (rel intensity) 320 (M<sup>+</sup>, 3), 263 (22), 262 (100), 220 (18), 219 (91), 218 (20), 189 (54), 165 (23), 44 (41); HRMS calculated for C<sub>20</sub>H<sub>20</sub>N<sub>2</sub>O<sub>2</sub> 320.1524788, found 320.15279.



**2-Amino-2'-carboxytolane [H-Taa-OH] (30).** 2-Amino-2'-carbo-methoxytolane (0.98 g, 3.9 mmol) was stirred in 30 mL of MeOH and 11 mL of 1.75 M aqueous LiOH overnight. After MeOH was removed under reduced pressure, the aqueous solution was acidified to *ca.* pH 2 with 5 N HCl. A yellow precipitate was collected by filtration and the filtrate was extracted with chloroform (8 x 10 mL), washed with saturated NaCl, dried with MgSO<sub>4</sub>, filtered and concentrated under reduced pressure. Total yield 0.91 g, 98%. <sup>1</sup>H NMR (300 MHz, (CD<sub>3</sub>)<sub>2</sub>CO): δ 6.57 (m, 1H), 6.75 (dd, J=1.2, 8.2 Hz; 1H), 7.10 (m, 1H), *ca.* 7.2 (br s, *ca.* 2H), 7.30 (m, 1H), 7.46 (dd, J=1.6, 7.6 Hz; 1H), 7.61 (m, 1H), 7.69 (m, 1H), 8.11 (m, 1H); <sup>13</sup>C (300 MHz, (CD<sub>3</sub>)<sub>2</sub>CO): δ 93.725, 94.232, 106.992, 114.516, 116.651, 125.511, 128.336, 131.039, 131.162, 131.699, 132.574, 133.066, 134.279, 151.769, 167.309.

**N-BocValNH<sub>2</sub>Et.** A mixture of N-BocValOpNP (1.66 g, 4.9 mmol) and K<sub>2</sub>CO<sub>3</sub> (2 g, 14.5 mmol) in 20 mL of THF was cooled to 0 °C. Ethylamine (1 mL, 70% aq., 15 mmol) was added and stirring was continued for 3.5 hr. The reaction mixture was then filtered and the filtrate was concentrated under reduced pressure. The residue was dissolved in 50 mL methylene chloride, washed with saturated NaHCO<sub>3</sub> (4 x 20 mL), aqueous HCl solution (1 N, 2 x 20 mL), saturated NaCl (2 x 20 mL) and dried with anhydrous MgSO<sub>4</sub>. The filtrate was then concentrated under reduced pressure and dried further *in vacuo* yielding 1.16 g (97%) of N-BocValNH<sub>2</sub>Et as a white powder. TLC R<sub>f</sub> (10 EtOAc : 1 MeOH): 0.66; <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 0.911 (d, J=6.62 Hz, 3H), 0.948 (d, J=6.62 Hz, 3H), 1.134 (t, J=7.30 Hz, 3H), 1.436 (s, 9H), 2.111 (m, 1H), 3.297 (m, 2H), 3.831 (dd, J=6.38, 8.82 Hz; 1H), 5.060 (br, 1H), 5.984 (br, 1H).

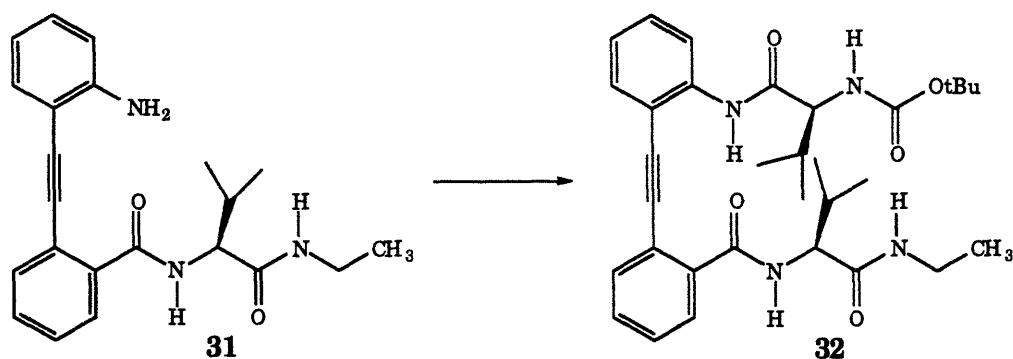
**H-Val-NHEt·HCl.** 3 mL of 6.5 N HCl in 6 mL of THF was added to a 100 mL round bottom flask charged with N-BocValNHEt (1.2 g, 4.8 mmol). The solution was then stirred under nitrogen overnight. It was then concentrated under reduced pressure and triturated with diethyl ether. It was further dried *in vacuo* to give a white powder (0.9 g, 100%). Used as a crude.



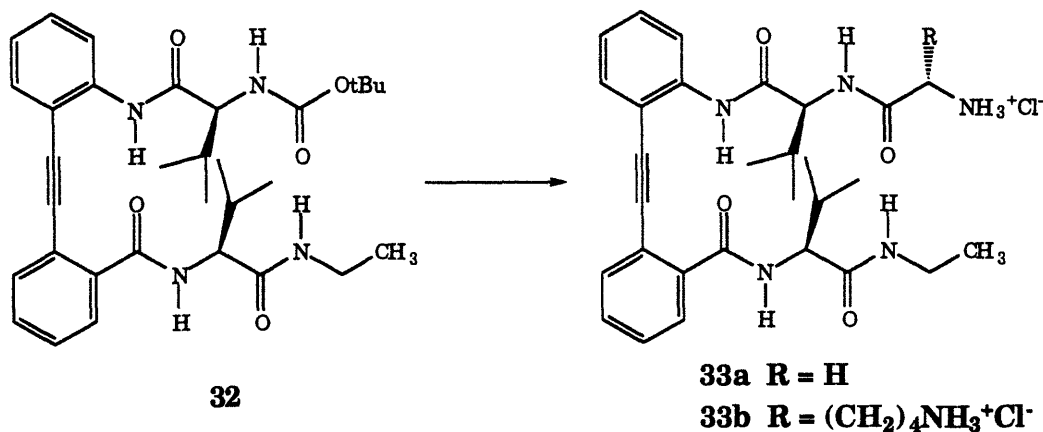
**H-Taa-Val-NHEt (31).** H-ValNHEt·HCl (0.64 g, 3.5 mmol) was suspended in 30 mL of CH<sub>2</sub>Cl<sub>2</sub>, dissolved after the addition of diisopropylethylamine (0.73 mL, 4.2 mmol). To this solution was then added 6 mL of DMF, 1-hydroxybenzotriazole (0.24 g, 1.5 mmol) and H-TaaOH (0.30 g, 1.26 mmol). The solution was then cooled to 0 °C with an ice/water bath. Dicyclohexylcarbodiimide (0.29 g, 1.4 mmol) was then added and the reaction was allowed to proceed at room temperature overnight. After it was filtered and washed with CH<sub>2</sub>Cl<sub>2</sub> (30 mL), the filtrate was washed with 1 N aqueous HCl solution (2 x 25 mL), saturated NaHCO<sub>3</sub> (2 x 25 mL), saturated NaCl (2 x 25 mL), dried with anhydrous MgSO<sub>4</sub> and filtered. The filtrate was then concentrated under reduced pressure. Preparative TLC (2 EtOAc : 1 hexane) afforded the amine (0.25 g, 55% crude) which was used without further purification. TLC (1 EtOAc : 1 hexane) R<sub>f</sub> 0.22.

**Note:** The low reactivity of the template aryl amino group is evidently shown. Peptide coupling between a free alkyl amine (amino acid here) and

the free acid of the template was little affected, if not at all, in the presence of the free aryl amino group.



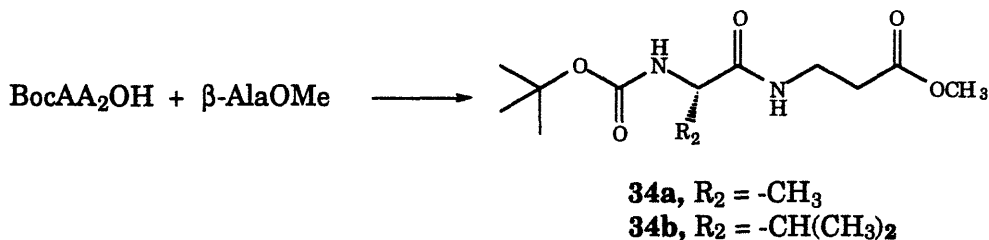
**N-BocVal-Taa-Val-NHEt (32).** N-BocValOH (0.13 g, 0.61 mmol) was dissolved in 6 mL of methylene chloride and cooled to 0 °C in an ice bath. Dicyclohexylcarbodiimide (63 mg, 0.305 mmol) was added and stirring was continued at 0 °C for 0.5 hr. It was then filtered and concentrated under reduced pressure. This crude symmetrical anhydride was then added to a solution of H-Taa-Val-NHEt (46 mg, 0.13 mmol) in 6 mL of chloroform. The solution was stirred at 0 °C for 2 hr, then at room temperature overnight. The solvent was removed by evaporation under reduced pressure. Flash chromatography (40 g of silica, 3 EtOAc/7 hexane) afforded N-Boc-Val-Taa-Val-NHEt (49 mg, 68%). TLC (1 EtOAc : 1 Hex)  $R_f$  0.47;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.891 (d,  $J=6.59$  Hz, 3H), 1.013-1.079 (m, 12H), 1.454 (s, 9H), 2.081 (m, 1H), 2.259 (m, 1H), 3.112 (m, 1H), 3.278 (m, 1H), 4.799 (m, 1H), 5.249 (dd,  $J=4.86$ , 10.44 Hz; 1H), 5.448 (d,  $J=10.41$  Hz, 1H), 6.838 (d,  $J=9.43$  Hz, 1H), 7.029 (td,  $J=1.32$ , 7.67 Hz; 1H), 7.291-7.376 (m, 2H), 7.403-7.491 (m, 2H), 7.610-7.653 (m, 2H), 7.688 (br t,  $J\sim 5.6$  Hz, 1H), 8.657 (d,  $J=8.21$  Hz, 1H), 9.251 (s, 1H);  $^{13}\text{C}$  (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  14.763, 16.944, 18.863, 19.047, 19.247, 32.161, 32.391, 33.958, 58.174, 58.619, 79.472, 89.453, 94.612, 111.786, 119.212, 122.360, 123.036, 127.090, 128.118, 129.777, 130.713, 132.049, 133.662, 135.980, 140.664, 156.419, 167.183, 171.160, 171.989.



**HCl·HGly-Val-Taa-Val-NHEt (33a).** Same procedure as for 33b shown below except the following: N-BocGly)<sub>2</sub> was used instead to give N-BocGly-Val-Taa-Val-NHEt (60 mg, 53%). TFAH·HGly-Val-Taa-Val-NHEt (28 mg, HPLC: RT 12.31 min). HCl·HGly-Val-Taa-Val-NHEt (23 mg, 42%). <sup>1</sup>H NMR (500 MHz, Me<sub>2</sub>SO-d<sub>6</sub>): δ 0.889 (d, J=6.85 Hz, 3H), 0.908-0.929 (m, 9H), 0.937 (t, J=7.21 Hz, 3H), 2.090 (m, 1H), 2.160 (m, 1H), 3.045 (m, 2H), 3.688 (m, 2H), 4.346 (dd, J=8.11, ~8.37 Hz; 1H), 5.00 (dd, J=6.12, 8.74 Hz; 1H), 7.165 (td, J=1.17, 7.57 Hz; 1H), 7.413 (td, J=1.5, 7.56 Hz; 1H), 7.527 (td, J=1.44, 7.57 Hz; 1H), 7.543 (dd, J=1.35, 7.74 Hz; 1H), 7.581 (td, J=1.41, 7.54 Hz; 1H), 7.703 (dd, J=1.3, 7.57 Hz; 1H), 7.744 (dd, J=1.17, 7.57 Hz; 1H), 7.893 (t, J=5.59 Hz, 1H), 8.047 (br, 3H), 8.105 (d, J=7.97 Hz, 1H), 8.603 (d, J=8.74 Hz, 1H), 8.672 (d, J=8.82 Hz, 1H), 9.629 (s, 1H).

**(2HCl)·HLys-Val-Taa-Val-NHEt (33b).** N-BocVal-Taa-Val-NHEt (49 mg, 0.087 mmol) in 2 mL of CH<sub>2</sub>Cl<sub>2</sub> at 0 °C was treated with 2 mL of trifluoroacetic acid for 1 hr. The solvents were then removed under reduced pressure; the residue was triturated with hexane and dried *in vacuo*. Freshly prepared active ester N-BocLys(N-Boc)OpNP (130 mg, 0.28 mmol) and diisopropylethylamine (0.03 mL, 0.17 mmol) in 5 mL of THF and 1 mL of DMF were added. After the solution was stirred under nitrogen at room temperature overnight, it was concentrated under reduced pressure and dried *in vacuo*. Water was added and the precipitate was collected by filtration.

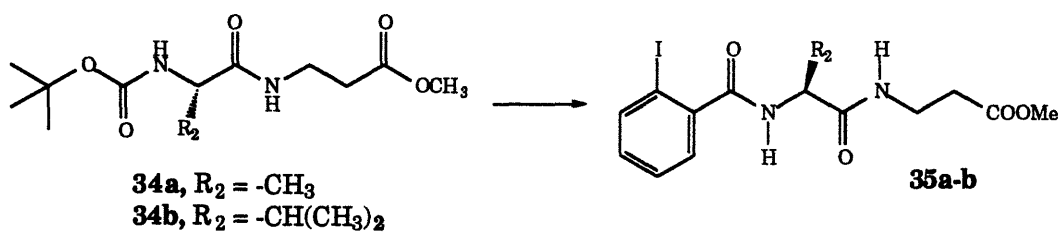
Preparative TLC (2 EtOAc/1 hexane,  $R_f$  0.54) gave N-BocLys(N-Boc)-Val-Taa-Val-NHEt (26 mg, 38%). This tetrapeptide was then deprotected with 1 mL of TFA in 1 mL of  $\text{CH}_2\text{Cl}_2$  at 0 °C for 45 min. After the solvents were removed, it was purified by HPLC to afford 28 mg of the TFA salt. Reversed phase HPLC (from 20 to 100% MeCN in 0.1% aq. TFA in 20 min): RT 11.6 min. Aqueous 1 N HCl was added and removed under reduced pressure to exchange the anion.  $(2\text{HCl})\cdot\text{HLys-Val-Taa-Val-NHEt}$  was obtained as the final product (19 mg, 89%).  $^1\text{H NMR}$  (500 MHz,  $\text{D}_2\text{O}$ ):  $\delta$  0.945-1.029 (m, 15H), 1.399 (m, 2H), 1.610 (m, 2H), 1.864 (m, 2H), 2.084 (m, 1H), 2.185 (m, 1H), 2.889 (m, 2H), 3.068 (m, 2H), 4.100 (t,  $J=6.7$  Hz, 1H), 4.266 (d,  $J=7.8$  Hz, 1H), 4.481 (d,  $J=7.6$  Hz, 1H), 7.374 (td,  $J=1.14, 7.6$  Hz; 1H), 7.509 (td,  $J=1.52, 7.61$  Hz; 1H), 7.543-7.640 (m, 5H), 7.698 (ddd,  $J=0.57, 1.33, 7.6$  Hz; 1H).



**N-BocAla- $\beta$ -AlaOMe (34a).** N-BocAlaOH (1 g, 5.29 mmol),  $\beta$ -AlaOMe $\cdot\text{HCl}$  (0.74 g, 5.30 mmol), N-hydroxysuccinimide (0.6 g, 5.21 mmol) and diisopropylethylamine (0.92 mL, 5.28 mmol) were dissolved in 20 mL of dry THF and 4 mL of DMF. The above solution was cooled to 0 °C in an ice bath. Water soluble carbodiimide  $\text{EtN}=\text{C}=\text{N}(\text{CH}_2)_3\text{NMe}_2\cdot\text{HCl}$  (1.0 g, 5.22 mmol) was added and stirring was continued at 0 °C for 2-3 hr, then at room temperature overnight. The solvent was removed by evaporation under reduced pressure. EtOAc (100 mL) was added and the solution was washed with dilute HCl solution, water, saturated  $\text{NaHCO}_3$ , saturated NaCl, and dried with anhydrous  $\text{MgSO}_4$ . The filtrate was concentrated under aspirator pressure and dried *in vacuo* to give N-BocAla- $\beta$ -AlaOMe (0.39 g, 27.2%).  $^1\text{H}$

NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  1.285 (d,  $J=7.54$  Hz, 3H), 1.385 (s, 9H), 2.494 (m, 2H), 3.471 (m, 2H), 3.635 (s, 3H), 4.078 (m, 1H), 5.128 (d,  $J=7.33$  Hz, 1H), 6.747 (br, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  18.501, 28.267, 33.734, 34.854, 50.118, 51.684, 79.939, 155.365, 172.671, 172.717.

**N-BocVal- $\beta$ -AlaOMe (34b)**. N-BocValOpNP (2 g, 5.91 mmol) was added to a solution of  $\beta$ -AlaOMe-HCl (0.9 g, 6.45 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.9 g, 6.51 mmol) in 15 mL of water and 100 mL of THF. The solution was rigorously stirred under nitrogen overnight before it was concentrated under reduced pressure. The residue was dissolved in 100 mL of EtOAc, washed with water, saturated NaCl, saturated NaHCO<sub>3</sub>, dilute NaOH, saturated NaCl, and dried (MgSO<sub>4</sub>) to give 1.44 g of white powder (80.7%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  0.853 (d,  $J=6.84$  Hz, 3H), 0.921 (d,  $J=6.89$  Hz, 3H), 1.419 (s, 9H), 2.099 (m, 1H), 2.528 (m, 2H), 3.518 (m, 2H), 3.676 (s, 3H), 3.838 (m, 1H), 4.990 (br, 1H), 6.395 (br, 1H); <sup>13</sup>C NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  17.616, 19.187, 28.274, 30.857, 33.725, 34.788, 51.802, 59.925, 79.861, 155.785, 171.491, 172.837.

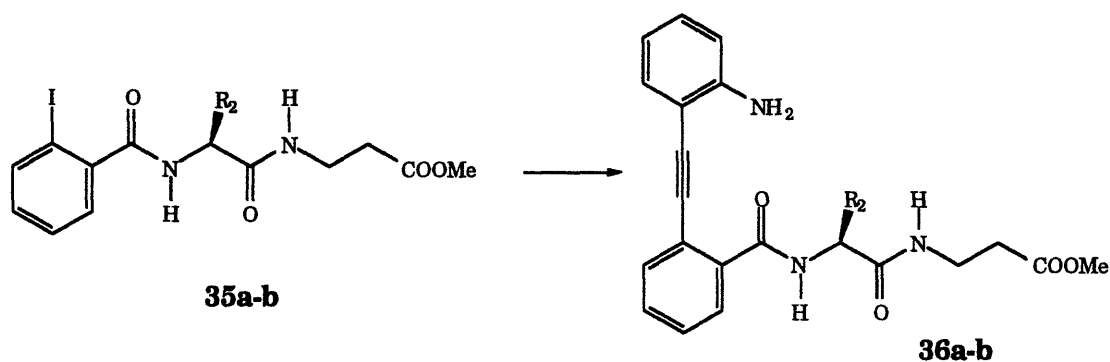


**2-Iodobenzoyl-Val- $\beta$ -AlaOMe (35b)**. Trifluoroacetic acid (4 mL) was added to a solution of N-BocVal- $\beta$ -AlaOMe (0.8 g, 2.6 mmol) in 4 mL of methylene chloride at 0 °C. Stirring was continued under nitrogen for 40 min before it was concentrated under reduced pressure and dried in *vacuo*. The residual oil was dissolved in dry THF and cooled to 0 °C with an ice bath. K<sub>2</sub>CO<sub>3</sub> (1.6 g, 11.6 mmol) was added to the above solution followed by 2-



iodobenzoyl chloride (1.0 g, 3.7 mmol). The mixture was vigorously stirred overnight and the ice bath was allowed to warm up slowly to room temperature. Standard workup gave a white powder (yield 0.64 g, 57%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.009 (d,  $J$ = 6.84 Hz, 3H), 1.024 (d,  $J$ = 5.92 Hz, 3H), 2.178 (m, 1H), 2.554 (t,  $J$ = 5.7 Hz, 2H), 3.54 (m, 2H), 3.681 (s, 3H), 4.401 (dd,  $J$ = 6.3, 8.9 Hz; 1H), 6.425 (d,  $J$ = 8.41 Hz, 1H), 6.517 (br t, 1H), 7.09 (m, 1H), 7.36 (m, 2H), 7.847 (d,  $J$ = 7.32 Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.341, 19.308, 31.347, 33.804, 35.017, 51.924, 58.987, 92.309, 128.164, 128.287, 131.297, 140.004, 141.800, 169.195, 170.454, 172.726.

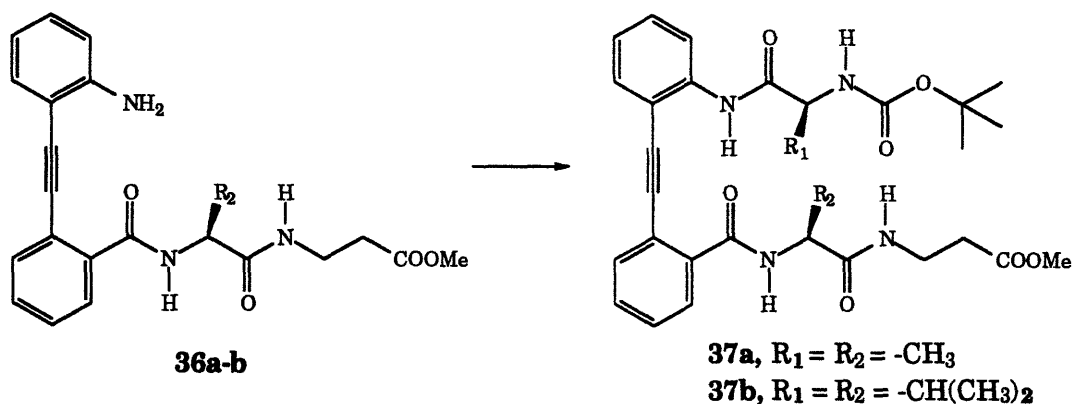
**2-Iodobenzoyl-Ala- $\beta$ -AlaOMe (35a).** Yield 0.1 g (18.2%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.479 (d,  $J$ = 6.97 Hz, 3H), 2.554 (t,  $J$ = 6.1 Hz, 2H), 3.547 (quart,  $J$ = 6.1 Hz, 2H), 3.680 (s, 3H), 4.619 (m, 1H), 6.407 (d,  $J$ = 7.9 Hz, 1H), 6.653 (br t, 1H), 7.091 (m, 1H), 7.33-7.40 (m, 2H), 7.843 (d,  $J$ = 7.68 Hz, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.382, 33.784, 35.089, 49.416, 51.904, 92.381, 128.190, 128.328, 131.338, 139.891, 141.381, 168.975, 171.631, 172.675.



**H-Taa-Ala- $\beta$ -AlaOMe (36a).** Followed the standard sealed-tube  $\text{Pd}^0\text{-Cu}^{\text{I}}$  coupling method. Yield 0.17 g, 95.5%. TLC (3 EtOAc/1 hexane)  $R_f$  0.17;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.428 (d,  $J$ = 7.01 Hz, 3H), 2.457 (m, 2H), 3.421 (m, 2H), 3.601 (s, 3H), 4.647 (m, 1H), 4.727 (s, 2H), 6.637-6.697 (m, 2H), 7.064 (br t,

1H), 7.123 (m, 1H), 7.310-7.365 (m, 2H), 7.417 (td,  $J = 1.5, 7.6$  Hz; 1H), 7.479 (d,  $J = 7.38$  Hz, 1H), 7.584 (m, 1H), 7.723 (m, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.112, 33.559, 34.980, 49.529, 51.636, 92.266, 106.555, 114.283, 117.310, 121.183, 128.044, 128.286, 130.285, 130.516, 131.913, 133.082, 135.564, 149.061, 167.408, 172.071, 172.553.

**H-Taa-Val- $\beta$ -AlaOMe (36b).** Followed the standard sealed-tube Pd-Cu coupling method. Yield 0.52 g, 86.8%. Yellow powder. TLC (2 EtOAc/1 hexane)  $R_f$  0.31;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.955 (d,  $J = 6.72$  Hz, 6H), 2.176 (m, 1H), 2.455 (m, 2H), 3.354 (m, 1H), 3.481 (m, 1H), 3.590 (s, 3H), 4.475 (dd,  $J = 7.2, 8.6$  Hz; 1H), 4.757 (s, 2H), 6.623-6.683 (m, 2H), 7.080-7.137 (m, 2H), 7.30-7.43 (m, 3H), 7.496 (d,  $J = 8.47$  Hz, 1H), 7.582 (m, 1H), 7.709 (dd,  $J = 1.2, 7.7$  Hz; 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  18.325, 19.185, 30.947, 33.543, 34.940, 51.585, 59.310, 92.201, 92.278, 106.620, 114.221, 117.169, 121.239, 127.980, 128.133, 130.160, 130.360, 131.972, 133.170, 135.872, 149.094, 167.597, 170.991, 172.480.

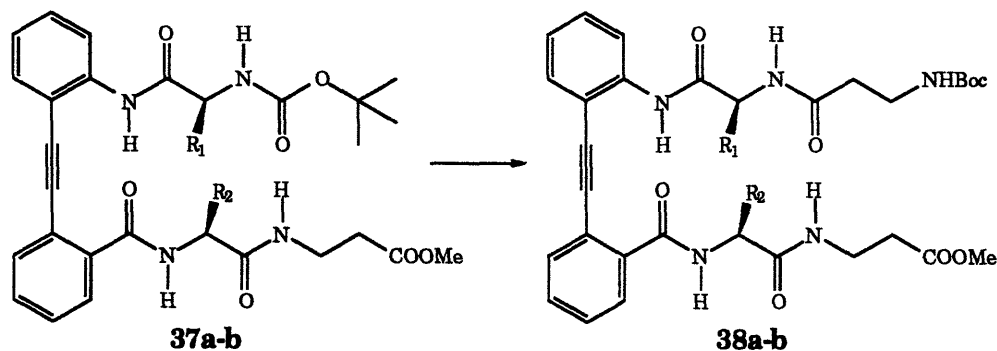


**N-BocAla-Taa-Ala- $\beta$ -AlaOMe (37a).** Water soluble carbodiimide coupling method in methylene chloride. Yield 0.20 g, 88.4%. TLC (1 EtOAc/1 hexane)  $R_f$  0.27;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.465 (s, 9H), 1.506 (d,  $J = 6.72$  Hz, 3H), 1.557 (d,  $J = 6.89$  Hz, 3H), 2.547 (m, 2H), 3.504 (m, 2H), 3.636 (s, 3H), 5.139 (m, 1H), 5.265 (m, 1H), 5.830 (d,  $J = 8.25$  Hz, 1H), 7.060 (m, 1H), 7.317-7.387 (m

2H), 7.4-7.52 (m, 3H), 7.642 (m, 1H), 7.721 (d,  $J = 7.81$  Hz, 1H), 7.771 (br t, 1H), 8.648 (d,  $J = 8.31$  Hz, 1H), 9.531 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  19.831, 20.967, 28.323, 33.881, 35.310, 48.976, 50.343, 51.525, 79.534, 89.853, 94.997, 111.03, 119.336, 122.422, 123.098, 127.137, 128.104, 129.777, 130.883, 132.019, 133.846, 134.261, 140.680, 155.498, 166.185, 171.744, 172.466, 173.080.

**N-BocVal-Taa-Val- $\beta$ -AlaOMe (37b).** Water soluble carbodiimide coupling method in methylene chloride. Yield 0.30 g, 76.2%. Flash chromatography did not separate the desired product from contaminated by-product N-BocVal-Val-Taa-Val- $\beta$ -AlaOMe at this stage. TLC (1 EtOAc/1 Hex)  $R_f$  0.26.

Note: The BocValOH starting material was an old sample contaminated with the dipeptide BocVal-ValOH, which can be easily separated by flash chromatography in the next step.

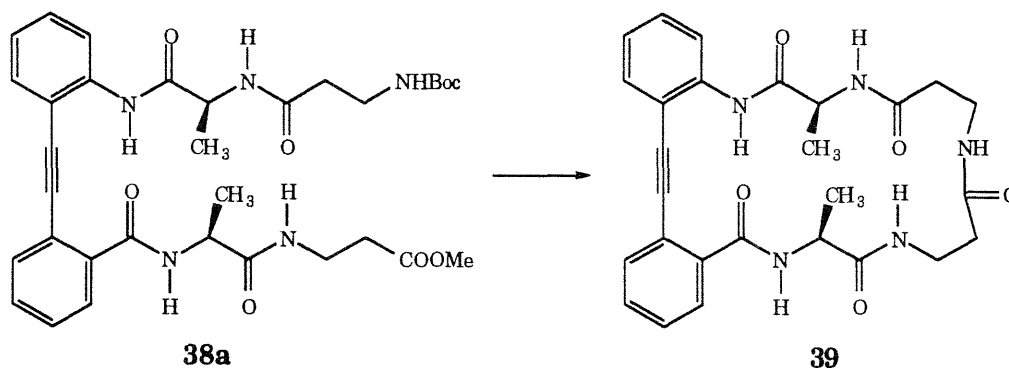


**N-Boc- $\beta$ -Ala-Val-Taa-Val- $\beta$ -AlaOMe (38b).** To a solution of N-BocVal-Taa-Val- $\beta$ -AlaOMe (**37b**) (0.30 g, 0.48 mmol) in 4 mL of methylene chloride at 0 °C was added trifluoroacetic acid (4 mL). Stirring was continued at 0 °C for 1 hr before the solution was concentrated. The residue was dissolved in THF (20 mL) before  $\text{K}_2\text{CO}_3$  (0.18 g, 1.3 mmol) and diisopropylethylamine (0.18 mL, 1.03 mmol) were added. The mixture was cooled to 0 °C with an ice bath followed by the addition of N-Boc- $\beta$ -AlaOSu (0.28 g, 0.97 mmol). Stirring

continued overnight and the ice bath was allowed to gradually warm up to room temperature. After the reaction mixture was concentrated under reduced pressure, distilled water was added to the residue. Precipitate was collected by filtration and was subjected to flash chromatography (silica, 1:1 then 2:1 EtOAc/hexane) to afford 0.19 g (58.1%) of the desired product plus 0.08 g (21.4%) of the by-product N-Boc- $\beta$ -Ala-Val-Val-Taa-Val- $\beta$ -AlaOMe (TLC, 2.5 EtOAc/1 hexane:  $R_f$  0.18). TLC (2.5 EtOAc/1 Hex)  $R_f$  0.34;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  0.910 (d,  $J$ = 6.84 Hz, 3H), 0.99-1.04 (m, 9H), 1.408 (s, 9H), 2.079 (m, 1H), 2.268 (m, 1H), 2.51 (m, 4H), 3.44 (m, 4H), 3.552 (s, 3H), 4.970 (dd,  $J$ = 6.9, 9.3 Hz; 1H), 5.46 (br, 1H), 5.544 (m, 1H), 6.572 (d,  $J$ = 9.93 Hz, 1H), 6.872 (d,  $J$ = 9.17 Hz, 1H), 7.040 (m, 1H), 7.284-7.390 (m, 2H), 7.424-7.501 (m, 2H), 7.624-7.670 (m, 2H), 8.134 (br t, 1H), 8.611 (d,  $J$ = 8.79 Hz, 1H), 9.249 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  17.347, 18.790, 19.005, 19.297, 28.388, 32.380, 32.442, 33.747, 35.006, 36.480, 36.956, 51.636, 57.517, 57.655, 79.184, 89.503, 94.801, 111.953, 119.339, 122.349, 123.301, 127.278, 128.276, 129.842, 130.902, 132.207, 133.896, 135.708, 140.438, 155.977, 167.110, 171.394, 171.671, 172.024, 172.546.

Note: The by-product N-Boc- $\beta$ -Ala-Val-Val-Taa-Val- $\beta$ -AlaOMe was separated at this stage.

**N-Boc- $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -AlaOMe (38a).** Same as above without flash chromatography. Quantitative yield, 0.22 g.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  1.424 (s, 9H), 1.513 (d,  $J$ = 6.89 Hz, 3H), 1.580 (d,  $J$ = 6.89 Hz, 3H), 2.55 (m, 4H), 3.4-3.5 (m, 4H), 3.656 (s, 3H), 5.226 (m, 1H), 5.526 (m, 1H), 5.576 (br, 1H), 6.995 (d,  $J$ = 7.0 Hz, 1H), 7.704 (m, 1H), 7.318-7.536 (m, 5H), 7.667 (m, 1H), 7.745 (d,  $J$ = 7.82 Hz, 1H), 8.165 (br t, 1H), 8.614 (d,  $J$ = 8.30 Hz, 1H), 9.589 (s, 1H);  $^{13}\text{C}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  20.185, 20.692, 28.446, 33.790, 35.157, 36.385, 36.708, 49.054, 49.698, 51.818, 79.151, 89.961, 95.320, 112.180, 119.459, 122.638, 123.391, 127.291, 128.273, 129.901, 131.114, 132.235, 134.124, 140.624, 156.051, 166.094, 171.146, 172.236, 172.651, 173.219.



**Cyclic [ $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -Ala] (39).** LiOH•H<sub>2</sub>O (34 mg, 0.8 mmol) was added to a solution of N-Boc- $\beta$ -Ala-Ala-Taa-Ala- $\beta$ -AlaOMe (**38a**) (0.17 g, 0.26 mmol) in 3 mL of THF, 1 mL of MeOH and 1 mL of water. The resulting solution was stirred overnight and concentrated under reduced pressure. pH 3 citric acid buffer was added to the residue and white powder was collected. The aqueous phase was extracted with EtOAc (3 x 12 mL), washed with brine and dried with MgSO<sub>4</sub>, and the residue was combined with the precipitate to give 161 mg of the acid (99.6%). The crude acid (0.12 g, 0.19 mmol) was then dissolved in 10 mL of CH<sub>2</sub>Cl<sub>2</sub> and was cooled to 0 °C. To this cold solution was added perfluorophenol (0.14 g, 0.76 mmol) and then followed by the addition of DCC (40 mg, 0.19 mmol). The reaction was run under argon overnight (0 °C to room temperature), after which it was concentrated under reduced pressure and *in vacuo*. The residue was redissolved in 2 mL of CH<sub>2</sub>Cl<sub>2</sub>, cooled to 0 °C with an ice bath, and then 2 mL of TFA was added to the chilled solution under argon. Stirring was continued for 45 min at 0 °C, after which it was concentrated under reduced pressure and dried *in vacuo*. The residue obtained was then dissolved in 60 mL of dry THF and was injected over a period of 4 hrs by using a syringe pump (*ca.* 15 mL/hr) into a mixture of 0.4 mL of DIEA and 0.8 g of K<sub>2</sub>CO<sub>3</sub> in 800 mL of dry THF with vigorous stirring. Then the reaction mixture was concentrated and the residue was dissolved in EtOAc, filtered and the organic phase was washed

with dilute HCl solution, saturated NaHCO<sub>3</sub>, brine and dried (MgSO<sub>4</sub>). The residue was purified by HPLC (20-100% CH<sub>3</sub>CN in 0.1% aqueous TFA over 20 min) to give 56 mg (58%) of crude cyclized product, which was further purified by preparative TLC (twice, 500 micron plates, 9 CHCl<sub>3</sub>/1 MeOH). The collected silica gel was filtered with a fritted glass funnel using 9 CHCl<sub>3</sub>/1 MeOH solution, and the resultant filtrate was concentrated. The resulting residue was then dissolved in a small amount of chloroform and filtered through a kimwipe-pipet plug to give the desired product (14 mg, 14.6%). TLC (9 CHCl<sub>3</sub>: 1 MeOH): R<sub>f</sub> 0.40; HPLC (20 to 100% of MeCN in 0.1% aqueous TFA in 20 min): RT 15.68 min; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ 0.859 (m, 1H), 1.349 (d, J=7.33 Hz, 3H), 1.747 (d, J=6.84 Hz, 3H), 1.805 (dt, J=3.91, 14.16 Hz; 1H), 2.275 (m, J=3.42, 4.88, 3.91, 8.3, 8.79 Hz; 2H), 2.605 (m, 1H), 3.055 (m, 1H), 3.798 (m, 1H), 4.262 (m, 1H), 4.337 (quint, J=6.84, 7.33 Hz; 1H), 4.673 (q, J=6.84, 7.33 Hz; 1H), 5.724 (d, J=6.84 Hz, 1H), 6.079 (dd, J=3.42, 3.91, 8.3, 8.79; 1H), 6.438 (s, 1H), 6.547 (d, J=8.3 Hz, 1H), 7.203 (td, J= 0.98, 7.33 Hz; 1H), 7.239-7.294 (m, 2H), 7.399-7.462 (m, 2H), 7.840 (d, J=7.33 Hz, 1H), 8.094 (dd, J=2.44, 9.28 Hz; 1H), 8.534 (d, J=8.79 Hz, 1H), 8.900 (s, 1H); <sup>13</sup>C NMR (300 MHz, CD<sub>2</sub>Cl<sub>2</sub>): δ 15.074, 15.181, 34.176, 36.786, 37.232, 38.092, 50.315, 51.374, 105.042, 120.736, 123.207, 123.699, 124.236, 124.943, 129.902, 130.271, 130.332, 133.526, 136.413, 137.718, 139.623, 168.000, 170.364, 170.687, 172.959, 174.864; HRMS calculated for C<sub>27</sub>H<sub>29</sub>N<sub>5</sub>O<sub>5</sub> 503.216869, found 503.21695.

## Appendix A: a mathematical derivation of the molar ellipticity

There exist two principal laws governing the process of light absorption: Lambert's law and Beer's law.

According to Lambert's law, the proportion of light absorbed by a transparent medium is independent of the intensity of the incident light and each successive layer of the medium absorbs an equal fraction of the light passing through it. Let's define variable  $z$  as the path length of the light in the medium up to point  $z$ , variable  $I_z$  as the intensity of the light at that point  $z$  of the medium, and the proportionality constant  $\alpha$  of this particular medium at a specific wavelength of the light. Following Lambert's law, we will have

$$dI_z = -\alpha I_z dz, \text{ or } dI_z/I_z = -\alpha dz, \quad (\text{A.1})$$

where the negative sign corresponds to a decrease of intensity of the incident light being absorbed by the medium. Assigning  $l$  as the thickness of the medium layer,  $I_0$  as the intensity of the incident light, and  $I$  as the intensity of the transmitted light, then upon integration of the above equation, i.e.,

$$\int_{I_0}^I \frac{dI_z}{I_z} = -\int_0^l \alpha dz,$$

the following relationship is obtained:

$$\ln(I_0/I) = \alpha l, \text{ or } \log(I_0/I) = \alpha l / \ln 10 = Kl, \quad (\text{A.2})$$

where the new constant  $K$  is the so-called Bunsen-Roscoe extinction coefficient. The quantity  $\log(I_0/I)$  is called the absorbance  $A$  or the optical density, and  $I_0/I$  the transmittance  $T$ . Intensity  $I$  is proportional to the intensity of electronic wave  $E^2$ ,  $[I/I_0 \propto (E/E_0)^2]$ , in which  $E$  is the amplitude of the electronic wave and can be measured directly. Because  $e^{-\alpha l} \leq 1$ ,  $\alpha$  is, in general, a very small number, which will be derived later.

Beer's law states that the amount of light absorbed is proportional to the number of molecules of the absorbing substance in the light path. Therefore, in a transparent solvent the absorption coefficient of solution will be proportional to the concentration,  $c$ , namely,  $K = kc$ . This concentration-independent new constant  $k$  is called absorptivity. The molar absorptivity or the molar absorption (extinction) coefficient, denoted by  $\epsilon$ , is the absorptivity when  $c$  and  $l$  are expressed in units of moles per liter (M) and in centimeters respectively. The following equation represents the relationship of the absorbance  $A$  with the molar absorption coefficient:

$$\epsilon = \frac{A}{lc} \quad (\text{M}^{-1} \text{cm}^{-1}), \quad (\text{A.3})$$

in units of square centimeters per millimole or square meters per decimole, which are mis-stated occasionally as square centimeters per decimole in the literature.

The molar absorption coefficient,  $\epsilon$ , is usually below  $10^5$ . From the previous derivation, we have

$$A = \alpha l / \ln 10 = \epsilon lc, \quad (\text{A.4a})$$

$$\alpha = \epsilon \ln 10 / 100, \quad (\text{A.4b})$$



where the factor 100 reflects different unit conventions used on the two sides of the equation. Since sample concentration is usually on the order  $10^{-5}$  M, we have, in general,  $\alpha < 0.02$ . This inequality will be useful in the following section when CD is discussed.

Circular Dichroism (CD) is closely related to UV absorption spectroscopy. However, since CD measures absorption difference of left- and right-circularly polarized light, it is a very sensitive photospectroscopic technique at detecting minute conformational changes in asymmetrical molecules.

Only an optically active medium can give rise to CD, and optical activity arises from the fact that most molecules are chiral, i.e., they have handedness and are not mirror images of each other. An optically active substance could be an enantiomeric molecule, a protein secondary structure, or an achiral chromophore in a chiral environment.

Light consists of electromagnetic transverse waves: the directions of the vibrating electric and magnetic vectors are orthogonal to the direction of propagation. In a right-handed Cartesian coordinate system, imagine the light is propagating in the direction of the positive  $z$  axis, the electric vector  $\mathbf{E}$  is along the  $x$  axis with sinusoidal oscillation in both amplitudes and directions, and the magnetic vector  $\mathbf{B}$  will be oscillating sinusoidally along the  $y$  axis. The plane containing the vibrating  $\mathbf{E}$  vector and the direction of propagation is called the plane of vibration, and the light wave is termed plane-polarized.

Light from common sources consists of independent wavetrains whose planes of vibration are randomly oriented about the direction of propagation. Such light, though transverse, is unpolarized. However, when two plane-polarized light waves are of equal amplitude, are at right angles to each other,

are moving in the same direction, and differ in phase by  $90^\circ$ , the resultant vector  $\mathbf{E}$  will have a constant amplitude and will rotate either clockwise or counterclockwise depending on the sign of the  $90^\circ$  phase difference between these two plane-polarized waves. This type of light is called circularly polarized. By convention, when the vector  $\mathbf{E}$  representing the circularly polarized light is rotating clockwise, it is called right-circularly polarized; and by contrast, the counterclockwise rotating circularly polarized light is called left-circularly polarized. Circularly polarized light has chirality (handedness), and therefore will exhibit chiral discrimination against chiral molecules.

When the medium is optically active and when the light is circularly polarized, not only do the left- and the right-circularly polarized rays travel at different speeds, but also are they absorbed to different extents. Optical rotatory dispersion (ORD) is related to circular birefringence, which is the difference in refractive indices for the right- and left-circularly polarized light. On the other hand, circular dichroism (CD) is the difference in absorbance of right- and left-circularly polarized light:

$$\Delta A = A_l - A_r = (\epsilon_l - \epsilon_r) cl, \quad (\text{A.5})$$

in which  $A_l$  and  $A_r$  are the absorbances for left- and right-circularly polarized light, respectively, and  $\epsilon_l$  and  $\epsilon_r$  are the corresponding molar absorption coefficients. The molar circular dichroism is defined as the difference between  $\epsilon_l$  and  $\epsilon_r$ :

$$\Delta\epsilon = \epsilon_l - \epsilon_r, \quad (\text{A.6})$$

which is generally referred to as the circular dichroism. The CD is small compared with the molar absorption coefficients  $\epsilon_l$  or  $\epsilon_r$ . The anisotropy factor (or dissymmetry factor)  $g$ ,

$$g = \frac{\Delta\varepsilon}{\varepsilon} = \frac{2(\varepsilon_l - \varepsilon_r)}{\varepsilon_l + \varepsilon_r}, \quad (\text{A.7})$$

is usually around  $10^{-4}$  and is rarely larger than  $\sim 10^{-2}$  (Kuhn, 1930).

A different quantity, rather than molar CD, is commonly used in representing CD data. It is called molar ellipticity, which is proportional to molar CD but has different units. Understanding the difference between these two quantities is key to grasping Circular Dichroism. CD is a complex phenomenon and a rather detailed mathematical treatment of the subject is truly essential. The following is a mathematical treatment on this subject.

As shown earlier, light can be expressed as an electromagnetic plane wave. The electric field of a plane wave can be represented by a vector  $\mathbf{E}$ , whose amplitude is modulated by the equation:

$$\mathbf{E} = \mathbf{E}_o \cos 2\pi \left( \frac{z}{\lambda} - \frac{z}{\lambda'} \right), \quad (\text{A.8})$$

where  $z$  is the coordinate along which the light is traveling (from 0 to  $l$ , the thickness of the medium),  $\lambda$  and  $\lambda'$  are the wavelengths of light in vacuum (and approximately, in air) and in the medium, respectively. According to Huygens' principle and the law of refraction,  $\lambda' = \lambda/\eta$ , in which  $\eta$  is the refractive index for the medium and we obtain

$$\mathbf{E} = \mathbf{E}_o \cos \frac{2\pi z}{\lambda} (\eta - 1). \quad (\text{A.9})$$

In a right-handed Cartesian coordinate system  $z$  is assigned as the axis of propagation. The plane of the light wave may then be arbitrarily chosen as

parallel to one of the two other perpendicular axes, say the  $x$  axis. In this case, if  $\mathbf{i}$  and  $\mathbf{j}$  are unit vectors in the direction of the  $x$  axis and  $y$  axis, respectively, then the amplitude of the wave is given by

$$\mathbf{E} = E_o \mathbf{i} \cos \theta, \quad (\text{A.10})$$

in which  $\theta = 2\pi z(\eta-1)/\lambda$ , and  $E_o$  is the maximum amplitude of the wave.

By analogy, circularly polarized light waves propagating in the  $z$  direction can be expressed simply as a vector sum of two plane waves:

$$\begin{aligned} \mathbf{E}_r &= E_o (\mathbf{i} \cos \theta_r - \mathbf{j} \sin \theta_r), \\ \mathbf{E}_l &= E_o (\mathbf{i} \cos \theta_l + \mathbf{j} \sin \theta_l), \end{aligned} \quad (\text{A.11})$$

where the subscripts  $r$  and  $l$  stand for right- and left-circularly polarized light, respectively. It is noteworthy that a plane wave can also be expressed as a vector sum of left- and right-circularly polarized waves, given that  $\theta_l = \theta_r$ . However, when  $\theta_l \neq \theta_r$ , and by definition of  $\theta$ , it results in a phase lag  $\Delta\theta$  due to the difference in refractive indexes,  $\eta_l$  and  $\eta_r$ , for left- and right-circularly polarized light waves:

$$\Delta\theta = \theta_l - \theta_r = 2\pi z(\eta_l - \eta_r)/\lambda, \quad (\text{A.12})$$

and the resultant vector sum becomes

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_r + \mathbf{E}_l = E_o \mathbf{i} (\cos \theta_r + \cos \theta_l) - E_o \mathbf{j} (\sin \theta_r - \sin \theta_l) \\ &= 2E_o \cos \frac{(\theta_r + \theta_l)}{2} \left[ \mathbf{i} \cos \frac{(\theta_r - \theta_l)}{2} - \mathbf{j} \sin \frac{(\theta_r - \theta_l)}{2} \right], \end{aligned} \quad (\text{A.13})$$

after the trigonometric identities:

$$\begin{aligned}\cos \alpha + \cos \beta &= 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}, \text{ and} \\ \sin \alpha - \sin \beta &= 2 \cos \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}.\end{aligned}$$

were applied. It is easy to see that this resultant light wave is again a plane wave, but rotated by an angle  $= (\theta_r - \theta_i)/2$ . This rotation is the basis for optical rotatory dispersion and is commonly converted into degrees per decimeter as  $\phi$ :

$$\begin{aligned}\delta &= \frac{\theta_i - \theta_r}{2} = \frac{\pi z}{\lambda} (\eta_l - \eta_r), \text{ in radians, and} \\ \phi &= \frac{180\delta}{0.1\pi z} = \frac{1800}{\lambda} (\eta_l - \eta_r), \text{ in degrees per decimeter.}\end{aligned}\tag{A.14}$$

The above equations are true only if the medium is optically active and transparent, i.e., the absorption constant  $\alpha$  (or absorptivity  $k$ ) is small. However, in the case when the medium is no longer transparent and the absorbance of light is significant, the amplitude  $E_o$  itself in the above vector presentations of circularly polarized light waves must also be modulated.

Since the intensity  $I$  of light is proportional to the square of its amplitude for its wave function, i.e.,

$$\frac{I}{I_o} = \left( \frac{E}{E_o} \right)^2, \text{ and from Eqn A.2 we then have}$$

$$E = E_o e^{-\alpha z/2}.\tag{A.15}$$

Therefore for the two circularly polarized light waves in a non-transparent medium, we now have

$$\begin{aligned}\mathbf{E}_r &= E_o e^{-\alpha_r z/2} (\mathbf{i} \cos \theta_r - \mathbf{j} \sin \theta_r), \\ \mathbf{E}_l &= E_o e^{-\alpha_l z/2} (\mathbf{i} \cos \theta_l + \mathbf{j} \sin \theta_l),\end{aligned}\tag{A.16}$$

where  $\alpha_l$  and  $\alpha_r$  are the absorption constants for the left- and right-circularly polarized light, respectively. The new light wave then is the vector sum of the above circularly polarized light:

$$\mathbf{E} = \mathbf{E}_r + \mathbf{E}_l = E_o \mathbf{i} (e^{-\alpha_r z/2} \cos \theta_r + e^{-\alpha_l z/2} \cos \theta_l) - E_o \mathbf{j} (e^{-\alpha_r z/2} \sin \theta_r - e^{-\alpha_l z/2} \sin \theta_l).$$

The amplitude of the new sum vector becomes

$$\begin{aligned}|\mathbf{E}| &= E_o \sqrt{(e^{-\alpha_r z/2} \cos \theta_r + e^{-\alpha_l z/2} \cos \theta_l)^2 + (e^{-\alpha_r z/2} \sin \theta_r - e^{-\alpha_l z/2} \sin \theta_l)^2}, \\ &= E_o \sqrt{(e^{-\alpha_r z/2})^2 + (e^{-\alpha_l z/2})^2 + 2e^{-\alpha_r z/2} e^{-\alpha_l z/2} \cos(\theta_r + \theta_l)}\end{aligned},\tag{A.17}$$

where the trigonometric identity,

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta,$$

was applied.

From the expression for the amplitude  $|\mathbf{E}|$  of the sum vector, we can see that its amplitude now has a maximum and a minimum, and the sum vector  $\mathbf{E}$  traces an ellipse that has a major and a minor axis. The major axis is defined by the major amplitude,

$$E_o(e^{-\alpha_r z/2} + e^{-\alpha_i z/2}) = |\mathbf{E}_r| + |\mathbf{E}_i|, \text{ when } \cos(\theta_r + \theta_i) = 1, \quad (\text{A.18})$$

and the minor axis defined by the minor amplitude,

$$E_o(e^{-\alpha_r z/2} - e^{-\alpha_i z/2}) = |\mathbf{E}_r| - |\mathbf{E}_i|, \text{ when } \cos(\theta_r + \theta_i) = -1, \quad (\text{A.19})$$

It is clear that the wave form is no longer circularly polarized but rather “elliptically” polarized resulting from differential absorption of circularly polarized light waves. The ellipticity  $\psi$ , which measures the eccentricity of the ellipse, is an angle whose tangent is defined as equal to the ratio of the minor to the major amplitudes of the ellipse. Therefore, the following expressions hold:

$$\tan \psi = \frac{E_o(e^{-\alpha_r z/2} - e^{-\alpha_i z/2})}{E_o(e^{-\alpha_r z/2} + e^{-\alpha_i z/2})} = \frac{1 - e^{z(\alpha_r - \alpha_i)/2}}{1 + e^{z(\alpha_r - \alpha_i)/2}}. \quad (\text{A.20})$$

From Section 4.1.1 (Eqn A.4a & b),  $\alpha\lambda(\text{m})/\ln 10 = \epsilon l(\text{cm})c$ , we then have  $\alpha = \epsilon c \ln 10 / 100$ . Since the molar absorption coefficient  $\epsilon$  is rarely above  $10^5$  and concentration  $c$  is usually on a micromolar scale, we know that  $\alpha$  is a small number and is usually not larger than 0.02. The thickness of the medium is usually 0.01 m, therefore the product  $z\alpha$  is an even smaller number (on the order of  $10^{-4}$ ). The quantity  $z(\alpha_r - \alpha_i)/2$  is, therefore, very small because it is the difference between two very small numbers. As a result, we can approximate both sides of Eqn A.20 with their respective Taylor series, and we obtain the following expression for the ellipticity,

$$\psi \approx \frac{1 - [1 + z(\alpha_r - \alpha_l)/2]}{1 + 1 + 0} = \frac{z}{4}(\alpha_l - \alpha_r). \quad (\text{A.21})$$

Replacing  $\alpha_l$  and  $\alpha_r$  with their respective absorptivities  $k_l$  and  $k_r$  according to the relationship  $\alpha = kc \ln 10$  obtained in Section 4.1.1, the ellipticity  $\psi$  can be expressed as follows:

$$\psi \approx \frac{cz \ln 10}{4}(k_l - k_r), \text{ in radians} \quad (\text{A.22})$$

or more commonly presented as molar ellipticity  $[\theta]$  in units of degrees-square centimeters-per decimole (notice the different units and the factor 100 as compared to those of square centimeters-per millimole for molar absorption coefficient  $\epsilon$ ):

$$\begin{aligned} [\theta] &= \left(\frac{\psi}{cz}\right) \times \left(\frac{180}{\pi}\right) \times 100 \\ &= \frac{18,000 \ln 10}{4\pi}(\epsilon_l - \epsilon_r) \\ &\approx 3298.2102 \Delta \epsilon \quad . \end{aligned} \quad (\text{A.23})$$

We have proved that molar ellipticity is proportional to molar circular dichroism by a constant number of 3298. In this thesis, molar ellipticity,  $[\theta]$ , is used throughout and occasionally, molar CD,  $\Delta \epsilon$ , is used to illustrate certain points.



**Appendix B: Tabulation for Figure 4.10 and 4.15**

Tabulation for Figure 4.10

Tolan derivatives	A-band		B-band	
	$\lambda_{\max}$	$\epsilon$	$\lambda_{\max}$	$\epsilon$
AcAla-Taa-ValNHMe	317	14400	255	33600
	292	18900	248	34100
AcLeu-Taa-ValNHMe	315	12400	254	30500
	292	16500	247	31300
AcMet-Taa-ValNHMe	316	14300	255	33900
	292	19000	248	34800
AcVal-Taa-ValNHMe	317	13900	255	32000
	293	18000	248	32500
AcIle-Taa-ValNHMe	317	12800	255	29700
	293	16700	248	30300
AcPhe-Taa-ValNHMe	316	12300	255	29900
	293	16000	248	30700
AcGln-Taa-ValNHMe	315	12700	255	29900
	292	17000	247	30700
AcAsn-Taa-ValNHMe	314	11400	(255)	(27500)
	292	15400	247	28500
AcGlu-Taa-ValNHMe	315	12800	255	30200
	292	17100	247	31200

Tabulation for Figure 4.15

Tolan derivative	A-band		B-band				
	$\lambda_{\max}$	- $[\theta]$	$\lambda_{\max}$	$[\theta]_1$	$\lambda_{\max}$	- $[\theta]_2$	$-[\theta]_1/[\theta]_2$
AcVal-Taa-AlaNHMe	317	16800	253	106900	228	90200	1.19
AcGly-Taa-AlaNHMe	321	1400	249	7900	225	11700	0.68
AcVal-Taa-ValNHMe	318	25200	253	171900	228	124500	1.38
AcLeu-Taa-ValNHMe	318	29400	252	175400	229	126800	1.38
AcGly-Taa-ValNHMe	321	5700	250	19500	226	29200	0.67
AcIle-Taa-ValNHMe	318	26000	253	159000	229	114000	1.39
AcPro-Taa-ValNHMe	320	16400	252	112000	228	112500	1.00
AcPhe-Taa-ValNHMe	319	32900	253	200500	229	108900	1.84
AcAla-Taa-ValNHMe	319	18200	252	133000	229	113700	1.17
AcVal-Taa-IleNHMe	318	24100	253	157100	228	114900	1.37
AcGly-Taa-IleNHMe	321	3600	248	11100	224	20700	0.54
AcPro-Taa-AlaNHMe	316	5400	252	49800	228	67000	0.74
AcMet-Taa-ValNHMe	318	35700	253	192000	229	118800	1.62
AcGln-Taa-ValNHMe	318	18700	252	124600	229	83100	1.50
AcAsn-Taa-ValNHMe	316	8900	252	80600	229	67100	1.20
AcGlu-Taa-ValNHMe	318	18000	252	108000	229	76100	1.42
Gly-Val-Taa-ValNHMe	319	24000	253	163200	228	122300	1.33
Ac-Taa-ValNHMe	317	4900	253	10200	225	10900	0.94

## Bibliography

- Abramenkov, A.V.; Almenningen, A.; Cyvin, B.N.; Cyvin, S.J.; Jonvik, T.; Khaikin, L.S.; Romming, C.; Vilkov, L.V. *Acta Chem. Scand.* **1988**, A42, 674.
- Allerhand, A.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1963**, 85, 866.
- Anfinsen, C.B. *Science* **1973**, 181, 223, and references cited therein.
- Anfinsen, C.B.; Scheraga, H.A. *Adv. Protein Chem.* **1975**, 29, 205, and references cited therein.
- Arcadi, A; Cacchi, S; Marinelli, F. *Tetrahedron Lett.* **1989**, 30, 2581.
- Arico-Muendel, C.C. Ph.D. thesis **1992**, Massachusetts Institute of Technology.
- Ashida, T.; Tamaka, O.; Yamane, T. *Int. J. Peptide Protein Res.* **1981**, 17, 322.
- Aue, W.P.; Bartholdi, E.; Ernest, R.R. *J. Chem. Phys.* **1976**, 64, 2229.
- Baker, E.N.; Hubbard, R.E. *Prog. Biophys. Molec. Biol.* **1984**, 44, 97.
- Ballardin, A.; Fischman, A.J.; Gibbons, W.A.; Roy, J.; Schwartz, I.J.; Smith, C.W.; Walter, R.; Wyssbrod, H.R. *Biochemistry* **1978**, 17, 4443.
- Bara, Y.A.; Friedrich, A.; Kessler, H.; Molter, M. *Chem. Ber.* **1978**, 111, 1045.
- Baranovic, G.; Colombo, L.; Skare, D. *J. Mol. Struct.*, **1986**, 147, 275.
- Bax, A. *Two-Dimensional Nuclear Magnetic Resonance in Liquids* **1982**, Reidel: Boston.
- Bax, A.; Davis, D.G. *J. Magn. Reson.* **1985**, 63, 207.
- Beale, R.N.; Roe, E.M.F. *J. Chem. Soc.*, **1953**, 2755.
- Belanger, P.C.; Dufresne, C.; Scheigetz, J.; Young, R.N.; Springer, J.P.; Dmitrienko, G.I. *Can. J. Chem.* **1982**, 60, 1019.
- Bianchi, E.; Venturini, S.; Pessi, A.; Tramontano, A.; Sollazzo, M. *J. Mol. Biol.* **1994**, 236, 649, and references cited therein.

- Bierzynski, A.; Kim, P.S.; Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2470.
- Binsch, G.; Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 411.
- Blanchard, D.E. Ph.D. thesis **1992**, Massachusetts Institute of Technology.
- Bodanszky, M.; Bednarek, M.A.; Bodanszky, A. *Int. J. Peptide Protein Res.* **1982**, *20*, 387.
- Bodanszky, M. *Principles of Peptide Synthesis*, Springer-Verlag: Berlin, 1984.
- Bodanszky, M.; Martinez, J. *Synthesis* **1981**, 333.
- Bothner-By, A.A.; Stephens, R.L.; Lee, J.; Warren, C.D.; Jeanloz, R.W. *J. Am. Chem. Soc.* **1984**, *106*, 811.
- Bowen, B.R. Ph.D. thesis **1988**, Massachusetts Institute of Technology.
- Bracken, C.; Gulyas, J.; Taylor, J.W.; Baum, J. *J. Am. Chem. Soc.* **1994**, *116*, 6431, and references cited therein.
- Bradley, E.K.; Thomason, J.F.; Cohen, F.E.; Kosen, P.A.; Kuntz, I.D. *J. Mol. Biol.* **1990**, *215*, 607.
- Brahms, S.; Brahms, J.; Spach, G.; Brack, A. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 3208.
- Brandts, J.F.; Halvorson, H.R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953.
- Bryson, J.W.; Beta, S.F.; Lu, H.S.; Suich, D.J.; Zhou, H.X.; O'Neil, K.T.; DeGrado, W.F. *Science* **1995**, *270*, 935, and references cited therein.
- Bystrov, V.F. *Progress in NMR Spectroscopy* **1976**, *10*, 41.
- Cassar, L. *J. Organometal. Chem.* **1975**, *93*, 253.
- Castro, C.E.; Stephens, R.D. *J. Org. Chem.* **1963**, *28*, 2163.
- Castro, C.E.; Havlin, R.; Honwad, V.K.; Malte, A.; Moje, S. *J. Am. Chem. Soc.* **1969**, *91*, 6464.
- Chen, F.M.F.; Kuroda, K.; Benoiton, N.L. *Synthesis* **1978**, 928.
- Chothia, C.; Janin, J. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4146.
- Chothia, C.; Janin, J. *Biochemistry* **1982**, *21*, 3955.

- Chou, P.Y.; Fasman, G.D. *Biochemistry* **1974**, *13*, 222.
- Conover, W.W. *Topics in Carbon-13 NMR Spectroscopy*, Wiley: New York, 1984.
- Corey, E.J.; Cheng, X.-M. *The Logic of Chemical Synthesis* **1989**, Wiley & Sons: New York.
- Corey, E.J.; Szekely, I.; Shiner, C.S. *Tetrahedron Lett.* **1977**, *18*, 3529.
- Coste, J.; Le-Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, *31*, 205.
- Creighton, T.E. *Biochem. J.* **1990**, *270*, 1.
- Currie, B.L.; Krstenansky, J.L.; Lin, Z.-l.; Ungwitayatorn, J.; Lee, Y.-H.; del Rosario-Chow, M.; Sheu, W.-S.; Johnson, M.E. *Tetrahedron* **1993**, *49*, 3489.
- Dale, J. *Acta Chem. Scand.* **1957**, *11*, 971.
- Dale, J.; Titlestad, K. *Chem. Commun.* **1969**, 656.
- Dale, J.; Titlestad, K. *Chem. Commun.* **1970**, 1403.
- Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B.H. *Science* **1994**, *266*, 776.
- Deber, C.M.; Madison, V.; Blout, E.R. *Acc. Chem. Res.* **1976**, *9*, 106.
- Diaz, H.; Kelly, J.W. *Tetrahedron Lett.* **1991**, *32*, 5725.
- Diaz, H.; Espina, J.R.; Kelly, J.W. *J. Am. Chem. Soc.* **1992**, *114*, 8316.
- Diaz, H.; Tsang, K.Y.; Choo, D.; Kelly, J.W. *Tetrahedron* **1993**, *49*, 3533.
- Dickerson, R.E.; Geis, I. *The Structure and Action of Proteins* **1969**, Harper & Row Pub.: New York.
- Dieck, H.A.; Heck, F.R. *J. Organometal. Chem.* **1975**, *93*, 259.
- Dormann, D.E.; & Bovey, F.A. *J. Org. Chem.* **1973**, *38*, 1719, 2379.
- Dyson, H.J.; Rance, M.; Houghten, R.A.; Lerner, R.A.; Wright, P.E. *J. Mol. Biol.* **1988**, *201*, 161.
- Espiritu, A.A.; White, J.G. *Z. Kristallographie* **1978**, *147*, 177.
- Feigel, M. *J. Am. Chem. Soc.* **1986**, *108*, 181.
- Fersht, A.R.; Matouschek, A.; Serrano, L. *J. Mol. Biol.* **1992**, *244*, 771.
- Ferstandig, L.L. *J. Am. Chem. Soc.* **1962**, *84*, 3553.

- Fischer, S.; Dunbrack, R.L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931.
- Freidinger, R.M.; Veber, D.F.; Hirschmann, R.; Paegle, L.M. *Int. J. Pept. Protein Res.* **1980**, *16*, 464.
- Freidinger, R.M.; Veber, D.F.; Perlow, D.S.; Brooks, J.R. Saperstein, R. *Science* **1980**, *210*, 656.
- Freidinger, R.M.; Perlow, D.S.; Veber, D.F. *J. Org. Chem.* **1982**, *47*, 104.
- Frerot, E.; Coste, J.; Pantaloni, A.; Dufour, M.-N.; Jouin, P. *Tetrahedron* **1991**, *47*, 259.
- Gellman, S.H. *Biochemistry* **1991**, *30*, 6633.
- Genin, M.J.; Johnson, R.L. *J. Am. Chem. Soc.* **1992**, *114*, 8778.
- Genin, M.J.; Mishra, R.K.; Johnson, R.L. *J. Med. Chem.* **1993**, *36*, 3481.
- Gierasch, L.M.; King, J.; Ed. *Protein Folding: Deciphering the Second Half of the Genetic Code*, **1990**, American Association for the Advancement of Science: Washington, D.C..
- Goldenberg, D.P.; Frieden, R.W.; Haack, J.A.; Morrison, T.B. *Nature* **1989**, *338*, 127.
- Grathwohl, C.; Wüthrich, K. *Biopolymers* **1976**, *15*, 2025.
- Groebke, K.; Renold, P.; Tsang, K.Y.; Allen, T.J.; McClure, K.F.; Kemp, D.S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4025.
- Huyghues-Despointes, B.M.P.; Scholtz, J.M.; Baldwin, R.L. *Protein Sci.* **1993**, *2*, 80, and references cited therein.
- IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.* **1970**, *52*, 1; *J. Biol. Chem.* **1970**, *245*, 6489; *Biochemistry*, **1970**, *9*, 3471.
- Iwamoto, T.; Grove, A.; Montal, M.O.; Montal, M.; Tomich, J.M. *Int. J. Pept. Protein Res.* **1994**, *43*, 597, and references cited therein.
- Jeener, J.; Meier, B.H.; Bachmann, P.; Ernest, R.R. *J. Chem. Phys.* **1979**, *71*, 4546.
- Kahn, M. *Synlett* **1993**, 821.
- Kahn, M.; Wilke, S.; Chen B.; Fujita, K. *J. Am. Chem. Soc.* **1988**, *110*, 1638.
- Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11.

- Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870.
- Kauzmann, W. *Advan. Protein Chem.* **1959**, *14*, 1.
- Kemp, D.S.; Allen, T.J.; Oslick, S.L. *J. Am. Chem. Soc.* **1995**, *117*, 6641.
- Kemp, D.S.; Allen, T.J.; Oslick, S.L. *J. Am. Chem. Soc.* **1996**, *118*, 4249.
- Kemp, D.S.; Blanchard, D.E.; Muendel, C.C. in *Peptides, Chemistry, Structure, Biology*; **1992**, Smith, J.; Rivier, J. Eds.; ESCOM: Leiden.
- Kemp, D.S.; Bowen, B.R. *Tetrahedron Lett.* **1988**, *29*, 5077.
- Kemp, D.S.; Bowen, B.R.; Muendel, C.C. *J. Org. Chem.* **1990**, *55*, 4650.
- Kemp, D.S.; Boyd, J.G.; Muendel, C.C. *Nature* **1991**, *352*, 451.
- Kemp, D.S.; Carey, R.I. *J. Org. Chem.* **1993**, *58*, 2216.
- Kemp, D.S.; Curran, T.P.; Boyd, J.G.; Allen, T. *J. Org. Chem.* **1991**, *56*, 6683.
- Kemp, D.S.; McNamara, P. *Tetrahedron Lett.* **1982**, *23*, 3761.
- Kemp, D.S.; McNamara, P. *J. Org. Chem.* **1984**, *49*, 2286.
- Kemp, D.S.; Stites, W.E. *Tetrahedron Lett.* **1988**, *29*, 5057.
- Kemp, D.S.; Sun, E.T. *J. Org. Chem.* **1984**, *49*, 2286.
- Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1970**, *9*, 219.
- Kessler, H.; Kondor, P. *Chem. Ber.* **1979**, *112*, 3541.
- Kessler, H.; Hölzemann, G. *Liebigs Ann. Chem.* **1981**, 2028.
- Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1982**, *21*, 512.
- Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernest, R.R. *J. Am. Chem. Soc.* **1987**, *109*, 607.
- Kim, C.A.; Berg, J.M. *Nature* **1993**, *362*, 267.
- King, J. *BioTechnology* **1986**, *4*, 297, and references cited therein.
- Kopple, K.D.; Ohnishi, M.; Go, A. *J. Am. Chem. Soc.* **1969**, *91*, 4264.
- Kopple, K.D.; Go, A.; Pilipauskas, D.R. *J. Am. Chem. Soc.* **1975**, *97*, 6830
- Kortüm, G.; Dreesen, G. *Chem. Ber.* **1951**, *84*, 182.
- Krstenansky, J.L.; Baranowski, R.L.; Currie, B.L. *Biochem. Biophys. Res. Commun.* **1982**, *109*, 1368.
- Kuhn, W. *Trans. Faraday Soc.* **1930**, *26*, 293.

- Kumar, A.; Wagner, G.; Ernest, R.R.; Wüthrich, K. *J. Am. Chem. Soc.* **1981**, *103*, 3654.
- Kuroda, Y.; Nakai, T.; Ohkubo, T. *J. Mol. Biol.* **1994**, *236*, 862, and references cited therein..
- Lecomte, J.T.J.; Matthews, C.R. *Protein Engng* **1993**, *6*, 1, and references cited therein.
- Levinthal, C. *J. Chim. Phys.* **1968**, *65*, 44.
- Liang, G.B.; Rito, C.J.; Gellman, S.H. *Biopolymers* **1992**, *32*, 293.
- Liu, C.-F.; Tam, J.P. *J. Am. Chem. Soc.* **1994**, *116*, 4149.
- Llinas, M.; Klein, M.P. *J. Am. Chem. Soc.* **1975**, *97*, 4731, and references cited therein.
- Lombart, H.-G.; Lubell, W.D. *J. Org. Chem.* **1994**, *59*, 6147, and and references cited therein.
- Luthman, K.; Hacksell, U. *Acta. Chem. Scand.* **1993**, *47*, 461.
- Lyu, P.C.; Liff, M.I.; Marky, L.A.; Kallenbach, N.R. *Science* **1990**, *250*, 669.
- MacArthur, M.W.; Thornton, J.M. *J. Mol. Biol* **1991**, *218*, 397.
- Macura, S.; Huang, Y.; Suter, D.; Ernst, R.R. *J. Magn. Reson.* **1981**, *43*, 259.
- Macura, S.; Wüthrich, K.; Ernst, R.R. *J. Magn. Reson.* **1982**, *46*, 269.
- Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochimica Et Biophysica Acta* **1955**, *16*, 1.
- Marsh, R.E.; Donohue, J. *Advan. Protein Chem.* **1967**, *22*, 235.
- Matthews, C.R. *Methods Enzymol.* **1987**, *154*, 498, and references cited therein.
- Mattice, W.L. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 93.
- Mavridis, A.; Moustakali-Mavridis, I. *Acta Cryst.* **1977**, *B33*, 3612.
- Merkel, E.; Wiegand, C. *Naturwissenschaften*, **1947**, *34*, 122.
- Merkel, E.; Wiegand, C. *Z. Naturforsch.* **1948**, *3b*, 93.
- Miick, S.M.; Todd, A.P.; Millhauser, G.L. *Biochemistry* **1991**, *30*, 9498.
- Milner-White, E.J.; Bell, L.H.; Maccallum, P.H., *J.Mol.Biol.* **1992**, *228*,725



- Minor, D.L., Jr.; Kim, P.S. *Nature* **1994**, *367*, 660; *371*, 264.
- Monera, O.D.; Kay, C.M.; Hodges, R.S. *Biochemistry* **1994**, *33*, 3862, and references cited therein.
- Munson, M.; O'Brian, R.; Sturtevant, J.M.; Regan, L. *Protein Sci.* **1994**, *3*, 2015, and references cited therein.
- Myszka, D.G.; Chaiken, I.M. *Biochemistry* **1994**, *33*, 2363, and references cited therein.
- Nagai, U.; Sato, K. *Tetrahedron Lett.* **1985**, *26*, 647.
- Nagai, U.; Sato, K.; Nakamura, R.; Kato, R. *Tetrahedron* **1993**, *49*, 3577.
- Nutt, R.F.; Veber, D.F.; Saperstein, R. *J. Am. Chem. Soc.* **1980**, *102*, 6539.
- Nagayama, K.; Kumar, A.; Wüthrich, K.; Ernest, R.R. *J. Magn. Reson.* **1980**; *40*, 321.
- Noggle, J.H.; Schirmer, R.E. *The Nuclear Overhauser Effect* **1971**, Academic Press: New York.
- Nowick, J.S.; Mahrus, S.; Smith, E.M.; Ziller, J.W. *J. Am. Chem. Soc.* **1996**, *118*, 1066.
- Nowick, J.S.; Holms, D.L.; Mackin, G.; Noronha, G.; Shaka, A.J.; Smith, E.M. *J. Am. Chem. Soc.* **1996**, *118*, 2764.
- Nyanguile, O.; Mutter, M.; Tuchscherer, G. *Lett. Pept. Sci.* **1994**, *1*, 9, and references cited therein.
- Oki, M.; Iwamura, H. *Bull. Chem. Soc. Jpn.* **1959**, *32*, 1135, and references cited therein.
- Okuyama, K.; Hasegawa, T.; Ito, M.; Mikami, N. *J. Phys. Chem.* **1984**, *88*, 1711.
- Olofsson, S.; Johansson, G.; Baltzer, L. *J. Chem. Soc. Perkin Trans.* **1995**, *2*, 1.
- Otzen, D.E.; Fersht, A.R. *Biochemistry* **1995**, *34*, 5718.
- Ovchinnikov, Y.A.; Ivanov, V.T. *Tetrahedron* **1974**, *30*, 1871.
- Ovchinnikov, Y.A.; Ivanov, V.T. *Tetrahedron* **1975**, *31*, 2177.
- Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741.

- Pauling, L.; Corey, R.B.; Branson, H.R. *Proc. Natl. Acad. Sci. U.S. A.* **1951**, *37*, 205.
- Pavia, D.L.; Lampman, G.M.; Kriz, G.S. Jr. *Introduction to Spectroscopy* **1979**, Saunders: Philadelphia.
- Pease L.G.; Niu, C.H.; Zimmermann, G. *J. Am. Chem. Soc.* **1979**, *101*, 184.
- Prey, V.; Berbalk, H. *Monatsh.* **1951**, *82*, 990.
- Quinn, T.P.; Tweedy, N.B.; Williams, R.W.; Richardson, J.S.; Richardson, D.C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747.
- Ramachandran, G.N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
- Ramachandran, G.N.; Venkatachalam, C.M.; Krimm, S. *Biophysic. J.* **1966**, *6*, 849.
- Ramachandran, G.N.; Sasisekharan, V. *Adv. Protein. Chem.* **1968**, *23*, 283.
- Ramakrishnan, C.; Ramachandran, G.N. *Biophysic. J.* **1965**, *5*, 909.
- Rebek, J.; Feitler, D. *J. Am. Chem. Soc.* **1973**, *95*, 4052.
- Rebek, J.; Feitler, D. *J. Am. Chem. Soc.* **1974**, *96*, 1606.
- Rich, D.H.; Bhatnagar, P.K. *J. Am. Chem. Soc.* **1978**, *100*, 2212.
- Ripka, W.C.; De Lucca, G.V.; Bach, A.C., II; Pottorf, R.S.; Blaney, J.M. *Tetrahedron* **1993**, *49*, 3593.
- Rohl, C.A.; Scholtz, J.M.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *Biochemistry* **1992**, *31*, 1263, and references cited therein.
- Robertson, J.M.; Woodward, I. *Proc. R. Soc. London* **1938**, *Ser A*, *164*, 436.
- Saebo, S.; Almlöf, J.; Boggs, J.E.; Stark, J.G. *J. Mol. Struct. (Theochem)* **1989**, *200*, 361.
- Sarkar, P.K.; Doty, P. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 981.
- Sato, K.; Nagai, U. *J. Chem. Soc. Perkin Trans. I*, **1986**, 1231.
- Schimmel, P.R.; Flory, P.J. *J. Mol. Biol.* **1968**, *34*, 105.
- Schleyer, P. v. R.; Trifan, D.S.; Backskai, R. *J. Am. Chem. Soc.* **1958**, *80*, 6691.
- Schlubach, V. H. H.; Franzen, V. *Ann.* **1951**, *573*, 110.

- Scholtz, J.M.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *J. Am. Chem. Soc.* **1991**, *113*, 5102, and references cited therein.
- Schultz, G.D.; Schirmer, R.H. *Principles of Protein Structure* **1978**, Springer-Verlag: New York.
- Sheehan, J.C.; Cruickshank, P.A.; Boshart, G.L. *J. Org. Chem.* **1961**, *26*, 2525.
- Shoemaker, K.R.; Kim, P.S.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *Nature* **1987**, *326*, 563, and references cited therein.
- Shortle, D.; Meeker, A.K. *Biochemistry* **1989**, *28*, 936.
- Siemion, I.Z. *Liebigs Ann. Chem.* **1971**, *748*, 88.
- Smith, C.K.; Withka, J.M.; Regan, L. *Biochemistry* **1994**, *33*, 5510.
- Smith, C.K.; Regan, L. *Science* **1995**, *270*, 980.
- Smith, E.L., in *Structure and Function of Cytochromes*, Okunuki, K., Kamen, M.D., Sekuzu, I., Eds., **1968**, Univ. Tokyo Press: Tokyo.
- Sonogashira, K.; Toda, Y.; Hagihara, N. *Tetrahedron Lett.* **1975**, *16*, 4467.
- Spencer, J.R.; Antonenko, V.V.; Delaet, N.G.J.; Goodman, M. *Int. J. Peptide Protein Res.* **1992**, *40*, 282.
- States, D.J.; Haberkorn, R.A.; Ruben, D.J. *J. Magn. Reson.* **1982**, *48*, 286.
- Stephens, R.D.; Castro, C.E. *J. Org. Chem.* **1963**, *28*, 3313.
- Still, W.C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923.
- Suzuki, H. *Bull. Chem. Soc. Jpn.* **1960**, *33*, 379; 389; 396; 406; 944.
- Suzuki, H.; Koyano, K.; Shida, T.; Kira, A. *Bull. Chem. Soc. Jpn.* **1982**, *55*, 3690.
- Tanaka, T.; Hayashi, M.; Kimura, H.; Oobatake, M.; Nakamura, H. *Biophys. Chem.* **1994**, *50*, 47, and references cited therein.
- Titlestad, K.; Groth, P.; Dale, J.; Ali, M.I. *J. Chem. Soc. Chem. Commun.* **1973**, 346.
- Weisman, J.S.; Kim, P.S. *Science* **1991**, *253*, 1386.
- West, R. *J. Am. Chem. Soc.* **1959**, *81*, 1614.
- Wiegand, C.; Merkel, E. *Ann.* **1947**, *557*, 242.
- Wouters, M.A.; Curmi, P.M.G. *Proteins Struct. Funct. Genet.* **1995**, *22*, 119.
- Wüthrich, K.; Grathwohl, C. *FEBS Lett.* **1974**, *43*, 337.

Yamashiro, D. *Int. J. Peptide Protein Res.* **1987**, *30*, 9.

Yan, Y.; Erickson, B.W. *Protein Sci.* **1994**, *3*, 1069, and references cited therein.

Yapa, K; Weaver, D.L.; Karplus, M. *Proteins Struct. Funct. Gen.* **1992**, *12*, 237.

Yoshida, Z.-i.; Osawa, E. *J. Am. Chem. Soc.* **1965**, *87*, 1467.

### Biographical Note

Zhong Qi Li comes from Qiqihar, a remote city in the northeastern China, where he attended elementary schools and high schools. During the high schools, he spent the summers on the city's model plane team, in which he built model airplanes and competed against other city teams. In 1985, he claimed the number one place in the National College Entrance Examinations in the Province (State) of Heilongjiang and as a result, was selected by the Harbin Institute of Technology as the first of the top three students sent to study in the United States. After a one year stay at the institute, he came to her sister school Phillips Academy at Andover, MA, where he completed one year of post graduate studies. In 1987, he was accepted to Harvard University, where he spent four memorable years and in 1991 received his B.A., *magna cum laude*, with highest honors, and M.A. in chemistry. While in college, he worked in Professor J.J. Grabowski's group on Liquid Phase Pulsed Photoacoustic Calorimetry. For two summers (1989 and 1990), he also interned in pharmaceutical chemistry with Dr. T. Kline at Biogen Inc. in Cambridge, MA. In 1991 he began his graduate studies at the Massachusetts Institute of Technology in Cambridge, MA; in January, 1992, he joined the research group of Professor D.S. Kemp. While at the graduate school, he also served as a translator and consultant for China-related businesses.

In the spring of 1995, Mr. Li received postdoctoral offers from Professor Y. Kishi and Professor D. A. Evans, and a half offer from Professor E. J. Corey, all of Harvard University. Shortly after, another postdoctoral offer came from Professor P. A. Wender of Stanford University.

Mr. Li has the following publications:

"2-Amino-2'-Carboxydiphenylacetylenes as  $\beta$ -Turn Mimetics. Synthesis and Conformational Properties." D. S. Kemp and Z. Q. Li, *Tetrahedron Letters*, **1995**, *36*, 4175.

"A Short  $\beta$ -Sheet Containing Proline Nucleated by a 2,2'-Substituted Tolane  $\beta$ -Turn Mimetic." D. S. Kemp and Z. Q. Li, *Tetrahedron Letters*, **1995**, *36*, 4179.