Pergamon

Bioorganic & Medicinal Chemistry, Vol. 4, No. 6, pp 739–766, 1996

0968-0896/96 $15.00+0.00

S0968-0896(96)00051-X

Review Article

Progress Towards Understanding β-Sheet Structure

Carey L. Nesloney and Jeffery W. Kelly*
Department of Chemistry, Texas A&M University, College Station, TX 77843-3255, U.S.A.

Abstract—This review is focused on recent advances in our understanding of β-sheet structure. It is intended to supplement previous surveys describing the early characterization and study of β-sheet structure. The first two sections of this review provide a brief introduction to β-sheet structure referencing the prior comprehensive reviews in this area as well as integrating new concepts. The next part outlines the typical problems encountered in solution studies on β-sheet structures. The most useful spectroscopic and biophysical techniques used to characterize β-sheet structures are described in the fourth section. Current hypotheses regarding the folding of predominantly β-sheet proteins are discussed in some detail in the fifth segment. The efforts of a number of laboratories to utilize peptides or peptidomimetics to serve as small β-sheet model systems are reviewed in the penultimate section. Finally, the efforts of a number of research groups focusing on the de novo design of β-sheet-based proteins are outlined.

Contents

1. Introduction

Extended β-sheet structures were first observed in keratin fibers in 1933.[1] X-ray analysis of these fibers suggested that the structures were composed of straight, extended peptide strands which were hydrogen bonded to each other in an antiparallel fashion. It was noted that the side chains of consecutive amino acid residues in a strand alternately project above and below the plane of the β-sheet. In 1951, Pauling and Corey proposed detailed structures for both parallel and antiparallel sheets in which the backbones of the β-strands adopted a pleated conformation (Fig. 1).[2] The correct hydrogen bonding patterns for both types of sheets were also proposed at this time. Since then, analysis of the crystal structures of β-sheet regions of proteins and computational studies of β-sheet structures have greatly enhanced our understanding of the

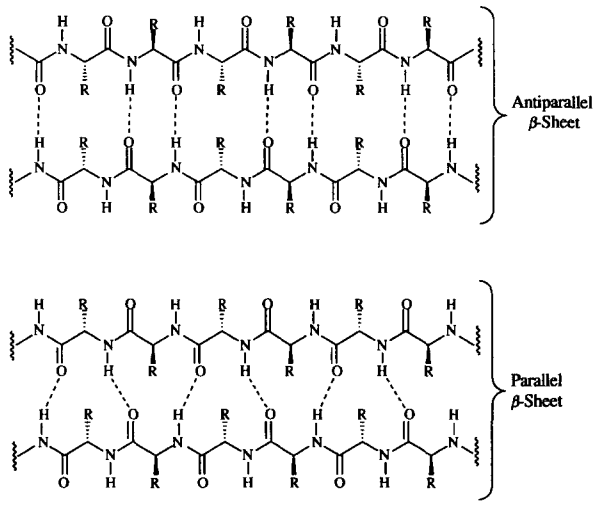


Figure 1. Schematic representation of parallel and antiparallel β-sheets.

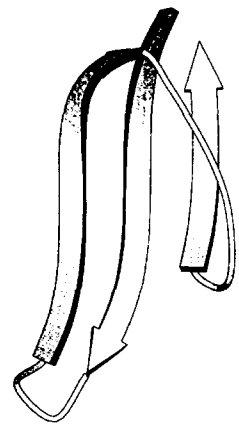


Figure 2. Ribbon diagram of a three-stranded β-sheet showing a β-turn and a right-handed crossover connection.[230] The diagram was created using coordinates (PDB entry 10XY, 1995) for hemocyanin subunit II.[233]

structural characteristics of β-sheets. However, the mechanisms of β-sheet folding are still not fully understood.

The aim of this review is to provide a current survey of β-sheet structures. The work is not intended to be comprehensive as several good surveys are available which describe the early characterization and study of β-sheet structure.[3–9] Instead, this review will be biased by recent results which contribute to our understanding of β-sheet structures with an emphasis on attempts to understand the mechanism(s) by which β-sheet formation occurs.

2. Background

2.1 β-Sheet structure

The β-sheet structure commonly observed in proteins is composed of nearly fully extended polypeptide chains (β-strands) which interact with each other through hydrophobic and electrostatic forces including van der Waals forces and hydrogen bonding involving the backbone carbonyl and amide proton functionalities on neighboring strands (Fig. 1). The dipoles of the amide bonds alternate along the chain, providing additional favorable electrostatic interactions. The peptide dipole moments of the alternating strands within an antiparallel sheet cancel each other. Parallel β-sheets have an overall dipole moment in which the positive end of the dipole is in the direction of the N-terminus of the β-strands.[10]

Antiparallel β-sheets are composed of strands which alternate in orientation and have an interstrand distance of ~5 Å. The strands are almost fully extended with a repeat period of 7.0 Å/residue pair (where the repeat period refers to the average distance along the chain axis of each dipeptide unit). The interstrand hydrogen bonds form alternating 10- and 14-membered rings. Antiparallel β-sheets can be composed of as few as two β-strands and generally have several strands with approximately six residues per strand. The resulting sheets frequently have one face which is solvent exposed.[3,11] The simplest antiparallel β-sheet folding motif has β-turns intervening between neighboring β-strands. More complex topologies place non-contiguous portions of the sequence in neighboring strands, typically by a right-handed crossover connection as shown in Figure 2.[11]

Parallel β-sheets generally exhibit more pronounced pleating due to the fact that the best hydrogen bonds are formed when the peptide chains are slightly less extended. This results in the sheets having a repeat period of 6.5 Å/residue pair with interstrand hydrogen bonds leading to uniform 12-membered rings. Parallel β-sheets are proposed to be less stable than antiparallel sheets because they are generally composed of five or more strands and are usually buried within the hydrophobic interior of proteins.[3,11] However, the validity of this assumption is not clear. Computational studies of homopolymers and copolymers suggest that the nature of the amino acids may determine the relative stabilities of the two types of sheets (*vide infra*).[12] That parallel sheets are less common than antiparallel sheets may also be due, in part, to the necessity of crossover connections in parallel β-sheet structures. Whereas a β-turn may serve to reverse the direction of a peptide chain and allow formation of an antiparallel β-sheet, the formation of a parallel sheet requires two peptide segments which are not close in sequence to be brought in proximity of each other. This may be accomplished by looping the intervening polypeptide sequence connecting the strands above or below the plane of the sheet. The intervening peptide segment connecting the strands may be a loop or it may exhibit secondary structure (most commonly an α-helix). The observation of approximately half as many parallel β-strand connectivities as antiparallel connections in a study of 144 proteins supported the idea that the longer loops required to connect parallel β-strands are both kinetically and thermodynamically unfavorable.[9]

The occurrence of mixed sheets is rare due to the different geometrical requirements for interstrand hydrogen bonding.[11] Once one side of a β-strand has become hydrogen bonded to an adjacent strand in either a parallel or antiparallel fashion, the other side of the strand has its hydrogen bond donors and acceptors in a conformation best suited for the same type of hydrogen bonding. Thus, optimal hydrogen bonding occurs in a sheet that is either all parallel or antiparallel. When mixed sheets are observed, the interface strand adopts an intermediate conformation and the sheets usually have the general appearance characteristics of the predominant hydrogen bonding type.[3] The hydrophobic ordering of strands, preferences of neighboring strands (i.e. the pairing of hydrophobic groups, unlike charges, or branched and unbranched residues), as well as evidence for specific residue pairing have been thoroughly discussed by Richardson.[3] Briefly, the central strands of β-sheets are usually the most hydrophobic with the remaining strands decreasing in hydrophobicity towards the surface of the protein.[13] The specific pairing of β-strands due to direct (non-random) interactions between their side chains, rather than random interactions or simple hydrophobic contact, has been suggested by statistical evaluations of residue pairs observed in proteins.[14–17] In a recent study of a series of 32 pairwise mutations at the solvent exposed face of the B1 domain of streptococcal protein G, Smith and Regan observed a correlation between statistical and experimental data. Substitutions were made at positions 44 and 53 which form a hydrogen bonded residue pair.[18] Interactions between residues which have a high propensity to form β-sheet structure (Trp, Phe, Tyr, Thr, Ile, Val) and between charged residues were examined (Arg, Lys, Glu). The range of free energies associated with side chain interactions was determined to be ~1.8 kcal/mol, with pairs containing exclusively aliphatic and aromatic residues having the higher $\Delta(\Delta\Delta G)$ values. The study suggests that the overall stability is dependent on both side chain interaction energy and the intrinsic propensity of each residue to adopt a β-sheet structure.

Most β-sheets observed in proteins are not planar as proposed in Pauling's original models. Instead, they exhibit a macromolecular twist.[19] The ϕ and ψ values for twisted β-sheets are somewhat more positive than those of planar sheets (Table 1), leading to a right-handed twist of 0–30° defined by the orientation of the backbone of the nearest β-strand relative to an adjacent strand (Fig. 3).[3,20] Twisted β-sheets are energetically more favorable than planar β-sheets.

Table 1. Typical dihedral angles observed in β-sheets

β-Sheet	ϕ°	ψ°
Parallel (ideal)	−119	113
Parallel (non twisted)	−132 to −100	132 to 103
Parallel (heavily twisted)	−94 to −65	158 to 143
Antiparallel (non twisted)	−160 to −133	161 to 118
Antiparallel (heavily twisted)	−108 to −84	176 to 140

Figure 3. Ribbon diagram depicting a typical right-handed twist in a parellel β-sheet. The diagram was created using coordinates (PDB entry 2FX2, 1991) for flavodoxin.[234]

Early explanations for the enhanced stability of twisted sheets have been surveyed elsewhere.[3] Proposals to explain the twist include: the observation that the accessible conformations of amino acid residues located in the β-area of the Ramachandran plot lie to the right of the $n = 2$ line; the tetrahedral distortion at the amide nitrogen combined with optimal hydrogen bonding geometry results in a twisted conformation; and considerations of hydrogen bonding and non-bonded interactions result in twisted sheets having a lower calculated free energy. Recently, it has been suggested that the main reason for twisting is the tendency of hydrophobic groups to tightly pack on the surface of a β-sheet.[21] Computational studies on the relative stabilities of twisted and non-twisted β-sheets involving two and three-stranded sheets of polyalanine ($n = 4$ and 8) suggest that twisting was favored by both intra- and inter-chain nonbonded interactions between the amino acid side chains which are optimized in the twisted structures.[12,22,23] Twisting of the β-sheets is opposed by intrachain electrostatic interactions, as the most favorable orientation of peptide dipoles occurs in the non-twisted structure. The sheet twist is also destabilized by non-optimal hydrogen-bonding due to the distortion of the hydrogen bonds within the twisted structure.[22] This study and analogous studies of polyvaline and polyisoleucine peptides[23] demonstrate that the amount of twist within a sheet decreases with increasing length and number of strands within a sheet so as to permit hydrogen bonding at the ends of the strands, consistent with previous observations.[24] Evaluation of additional homopolypeptides and polydipeptides suggests that twisted sheets are more stable for antiparallel sheets when the sheets are composed of small, unbranched residues (Gly, Ala, Abu, Leu) while twisted parallel sheets are more stable when the amino acids are β-branched (Val, Ile, Thr), polar (Ser, Thr), aromatic (Phe, Tyr), or contain a long side chain (Lys).[12] Thus, the relative stabilities of the twisted

sheets are dependent on the packing efficiency of the strands of the β-sheet (both intra- and inter-strand interactions) and hence, are determined by the amino acid composition of the strands and the frequency of various intra- and inter-strand neighboring pairs.

β-Sheet structures commonly exhibit irregular features such as β-bulges and β-bends. A β-bulge occurs when two residues on one strand are positioned opposite a single residue on the neighboring strand.[25] β-Bulges are observed almost exclusively in antiparallel β-sheets and rely on near optimal hydrogen bonding. A strand containing a two-residue β-bulge cannot propagate a sheet. Therefore, β-bulges are always at either an edge strand or the end of a sheet. Bulges disrupt the normal alternation of side chains by positioning the side chains of the two bulge residues on the same side of the sheet and increase the normal twist of the β-strands by imposing a slight bend in the sheet. The different types of β-bulges were originally classified by Richardson.[25] Recently, Chan et al. have extended this classification system.[26] Their study of 170 proteins indicated that most proteins contain at least two bulges with less than 10% of bulges occurring between parallel strands. It was also noted that bulges are often conserved within protein families, suggesting that they may contribute to the function of proteins by affecting the direction of the β-strands and the orientation of side chains. A set of 247 protein structures was examined by Daffner et al. in their evaluation of bent β-strands (strands which exhibit a change in direction but not a reverse turn).[27] Bent strands were found in 85% of the 946 strands studied. Large β-bends (those having a bend angle of greater than 25°) occurred most commonly in antiparallel strands of more than five residues which were within a sheet of at least three strands and did not contain a β-bulge. A preference for charged and polar residues and for Gly and Pro in the bend region was observed, implying that these residues are important in the stabilization of β-bends.

2.2 Common topologies

β-Sheet structures occur within proteins in a relatively small number of topologies, the most common being β-sandwiches (including immunoglobin domains), β-barrels, and α/β arrangements (Fig. 4a–c). Recently, the β-clam, β-propeller, β-helix, and β-roll structures have been characterized as well (Fig. 5a–d). As thorough reviews are available for most of these structures, only brief attention will be given to the discussion of the β-sheet topologies. Cohen and Chothia have evaluated the characteristics of observed β-sandwiches in naturally occurring proteins. The removal of the hydrophobic side chains from contact with aqueous solvent is generally considered to be the stabilizing force in the association of two sheets into a sandwich structure. In a survey of 10 β-sandwiches, Cohen observed the two sheets to be spaced an average of 8.3–10.3 Å apart and to be oriented such that the strands of the 'bottom sheet' are rotated counterclockwise by 20–50° relative to the strands of the 'top sheet' (Fig. 4a).[4] Sheets which were composed of medium sized residues were found to have smooth interacting surfaces, where as the surfaces of sheets composed of both large and small residues were found to interdigitate. In their survey of 12 β-sandwiches, Chothia and Janin observed a similar twist of 30–65° between the orientation of the strands in the two sheets.[28] Additionally, the rows of side chains from the top and bottom strands were found to be aligned even though the main chains of the β-strands are not. The edge strands of the sheets generally packed like the internal strands except that one side of the strand was

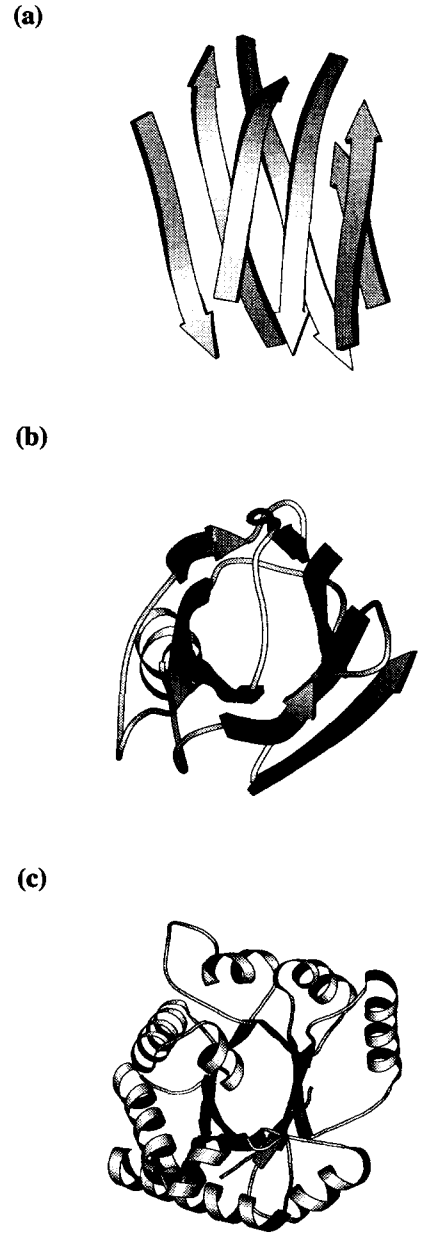


Figure 4. (a) Ribbon diagram of a β-sandwich. The diagram was created using coordinates (PDB entry 1MCS, 1993) for immunoglobin λ light chain dimer.[235] (b) Representation of an eight-stranded antiparallel β-barrel. The diagram was created using coordinates (PDB entry 2PAB, 1977) for prealbumin.[231] (c) Ribbon diagram of an α/β barrel protein, triose phosphate isomerase[236] (PDB entry 1TIM, 1976).

not involved in hydrogen bonding, allowing somewhat greater conformational flexibility. Some edge strands were observed to hydrogen bond to the top sheet on one end of the strand and to the bottom sheet at the other end. In fewer examples, edge strands were observed to be hydrogen bonded to each other (generally considered to be a β-barrel). A twisted edge strand has been observed in several β-sandwich proteins and probably serves to prevent unwanted sheet mediated self-assembly.

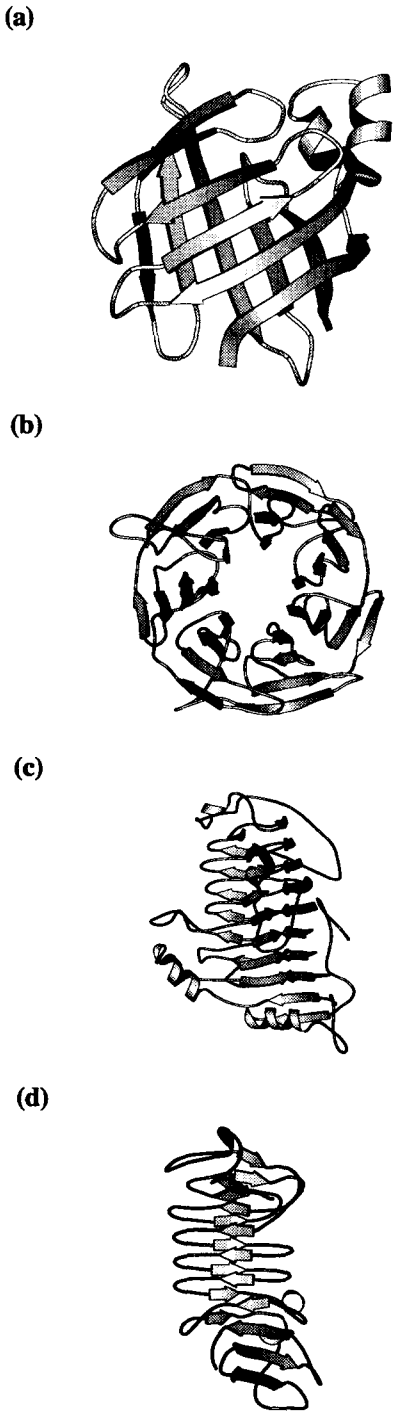


Figure 5. (a) Ribbon diagram of a β-clam protein, rat intestinal fatty acid binding protein[237] (PDB entry 1IFC, 1991). (b) Ribbon diagram of a seven-bladed β-propeller. The diagram was created using coordinates (PDB entry 1GOG, 1993) for galactose oxidase.[38] (c) Ribbon diagram of pectate lyase C showing the right-handed β-helix[43] (PDB entry 2PEC, 1994). (d) Ribbon diagram of the β-sandwich domain of alkaline protease containing the β-roll structure[54] (PDB entry 1KAP, 1995).

Due to differences in the size of the two interacting β-sheets and/or the interruption of sheet to sheet facial side chain interactions by side chains from other secondary structures which pack between the sheets, less than 50% of the amino acid residues interact in the interior of a β-sandwich. Interestingly, 61% of the total residues involved in the face to face interactions are Leu, Val, Ile, and Phe, whereas a significantly lower frequency of occurrence is observed in an average β-sheet (31%) and in protein interiors (38%). The authors suggest that the high percentage of these residues allows for smooth, well-packed surfaces. Finkelstein and Nakamura have noted the importance of filling voids in the surfaces of β-sheets with hydrophobic side chains.[29] Aromatic groups, which have strong van der Waals interactions and whose rigidity allows them to be fixed with only a small entropic cost, are considered the best candidates for obtaining uniform packing within β-sheets. Conformational minimum energy calculations on homopolypeptides have suggested two low energy classes of β-sandwich packing.[30] The lowest energy class consists of sheets in which the strands are nearly parallel or antiparallel to each other with the sheets at a negative angle to each other. As seen in natural systems, this packing allows the best complementarity of the sheet interfaces. The second packing arrangement consists of two twisted sheets where the strands in each sheet are oriented nearly perpendicular to each other allowing contacts near two opposite corners with fewer contacts at the center of the sheets. Although slightly higher in energy, these types of structures are observed in proteins, usually when a sheet folds back upon itself or when the two sheets are covalently linked.

An early survey of 11 immunoglobin domains by Lesk and Chothia has shown this domain to consist of either of two types of β-sandwiches, the first having sheets of four and five strands and the second having sheets of four and three strands.[8] In all of the domains surveyed, the two sheets were linked by a disulfide bridge with a nearby tryptophan residue packing against the cysteines. Changes in the size of the interior residue side chains required extensive structural adjustments to maintain the overall packing and, in some cases, insertion of side chains from the interstrand loops provided a means of compensating for small interior residues. In 1994, Bork et al. compared 23 immunoglobin domains and further categorized these structures into four sub-types which contain a common structural core but which are distinguished by the number of strands within each sheet and the location of the conformationally flexible edge-strands.[5] The authors suggest that the lack of a conserved loop region between families which could serve to initiate folding disfavors the turn nucleation theory for the folding of these domains. Also, while the disulfide bridges may stabilize the fold,

they do not appear to be essential. The one conserved feature among the sub-types was the hydrophobicity of the central strand of one of the sheets, consistent with previous observations of hydrophobicity patterns within β-sheets.[13] Aromatic residues near the termini of the core β-strands were also conserved within sub-types (although not between sub-types, particularly for those which lacked disulfide bridges). Although the extreme sequence diversity among immunoglobin domains implies that there is no single interaction or localized set of interactions which appears to be responsible for the immunoglobin fold, the hydrophobicity of the central strand and the terminal aromatic residues suggest that hydrophobic interactions may be involved in the nucleation or stabilization of the folding of the β-sheet sandwich motif.

β-Barrels are a structure in which the β-strands are wrapped around each other to form a cylindrical or nearly cylindrical shape (Fig. 4b).[3] Barrels are usually composed of 5–13 strands, which exhibit a right-handed twist relative to the axis of the barrel and are characterized by hydrophobic interiors. Those composed of parallel strands are the most symmetrical, while the more common antiparallel barrels are usually less cylindrical and may exhibit a disruption in the regular hydrogen bonding pattern. Antiparallel barrels containing an odd number of strands have one parallel interaction. However, this is the only example of mixed sheet interactions which occurs in β-barrel structures. The connectivities between the antiparallel strands can be simple up and down connectivities but are more commonly Greek key or jelly roll topologies (e.g. a +3, −1, −1, +3 connectivity).[11] The sheets can also be joined by segments of loops or helices as in the α/β proteins described below. Lasters et al. have determined the average sheet separations within six parallel β-barrels to be 11.6–13 Å compared to 8.5–11.5 Å which was calculated for orthogonally packed antiparallel sheets.[31] Additionally, the average angle of the β-strand axes relative to the barrel axis was 36° and the average twist angle of the strands was −26°. Optimal packing of side chains within the interior of the parallel β-barrel is obtained in barrels composed of five to eight β-strands.

The α/β barrel proteins are most commonly characterized by a parallel β-sheet barrel surrounded by α-helices which are formed by the peptide segments which intervene between the β-strands (Fig. 4c). The α-helices are usually packed such that their axes run parallel to the strand direction.[32] The first 14 examples of this fold contained β-barrels which were composed of eight strands and had cross-sections which varied from being circular (with a diameter of 14.5 Å) to being elliptical (with dimensions of 11.5 × 16.5 Å). Based on an analysis of the sheet geometry and the residue packing inside the β-barrels of three of the known α/β barrel proteins, Lesk and co-workers determined that the normal right-handed twist (−26°) of the β-sheets results in ideal packing of the hydrophobic side chains within the cavity of the barrel with optimal packing occurring when the barrel contains eight β-strands.[7]

β-Clam structures (Fig. 5a) were first identified by Sacchettini and co-workers in their structure elucidation of the rat intestinal fatty acid binding protein.[33] The classification has been adopted for other acid binding proteins including the cellular retinoic acid binding protein, CRABP.[34] The β-clam is composed of two antiparallel β-sheets with nearly orthogonal strands (usually five) with two short α-helices near the N-terminus. The interior of the β-clam is lined with nonpolar amino acids and serves as the binding site for the fatty acid ligand.

Another recently identified β-sheet structure is the propeller assembly which has been observed in several proteins. The first example of the propeller fold was observed in the crystal structure of the influenza virus glycoprotein antigen, neuraminidase, which consists of six β-sheets.[35] An additional six-bladed propeller structure, bacterial sialidase,[36] and two seven-bladed structures, methylamine dehydrogenase[37] and galactose oxidase,[38] have subsequently been determined. The β-propeller fold consists of six to eight twisted, four-stranded, antiparallel β-sheets that are arranged in a circular fashion resembling the blades of a propeller (Fig. 5b). The innermost strands of each sheet are arranged parallel to each other with a hole centered around the axis of the assembly. Good packing of each pair of adjacent sheets is obtained by the sheets being oriented as in a normal β-sandwich. Calculations suggest that the twist of the β-sheets in the propeller is slightly greater than in a typical β-sandwich.[39] This twist decreases as the number of sheets in the propeller increases while the radius of the central hole increases as the number of sheets increases. These two effects appear to be optimally balanced when the propeller is composed of seven sheets although both six- and sevenfold propellers have been observed. Recently, two eightfold propellers were observed in the crystal structures of methanol dehydrogenase from two closely related bacteria,[40] validating the prediction that an eightfold propeller may be possible. Two four-bladed propeller structures have now been observed as well.[41,42] The four-bladed propeller is similar to the larger ones but contains a disulfide bridge between the first and fourth blades. The first three strands are well-structured while the outermost strand of each blade is quite irregular. Additionally, there is short α-helix between each pair of blades. The small number of blades appears to be compensated for by the disulfide bridge and the incorporation of some of the α-helix side chains at the blade interfaces as well as by the presence of larger residues (especially aromatics) in the outer strands.

The crystal structures of three pectate lyases have revealed a new fold, the right-handed parallel β-helix.[43–46] This unusual fold consists of parallel strands which are wound into right handed coils with a rise of ~4.86 Å per turn (Fig. 5c).[47–50] Each of the seven complete coils contains three parallel β-strands

with an average of 22 residues per turn. The coils are not circular. Instead, two of the parallel β-sheets form an antiparallel β-sandwich while the third sheet is nearly perpendicular to one end of the sandwich such that the cross-section of the helix somewhat resembles an 'L' shape. The distance between the C_α's of the two sheets is 9.1 Å (which is within the range normally observed for antiparallel β-sheet sandwiches) and the side chains of the sheet residues are interdigitated across the sandwich interface. With typical twists of −20° to 40°, the β-strands are less twisted than in a normal parallel β-sheet. The parallel β-helix is stabilized by hydrogen bonding between the parallel β-strands as well as from hydrogen bonding between asparagines at the two reside elbow turns which connect the second and third sheets. Each residue in the center of these asparagine stacks forms five hydrogen bonds (three to adjacent asparagines and two to main chain atoms). Serine stacks are also observed. Both appear to add stability to the sharp turns. Additionally, aromatic stacks in which the rings are nearly planar to each other and aliphatic stacks (most often of valine and isoleucine and less frequently of alanine and leucine) are observed both in the interior and on the exterior of the β-helix. The P22 tailspike protein consists of three identical subunits each of which is largely composed of a right-handed parallel β-helix (13 complete turns).[51,52] The β-helix of the P22 tailspike protein tapers near the C-terminus and does not contain Asn ladders. It is believed that the loss of stability is recovered through packing between each of the β-helices in the trimeric structure. Additionally, the β-sheets at the C-terminus of each monomer interdigitate to create a single structural unit. It is noteworthy that the parallel β-helices are quite stable despite being composed almost entirely of parallel β-sheets and having largely solvent exposed exteriors.

A left-handed parallel β-helix was recently reported by Raetz and Roderick.[53] The crystal structure of UPD-*N*-acetylglucosamine acyltransferase revealed a trimer consisting of identical subunits each of which are composed of two domains. The C-terminal domain (residues 187–262) consists of four α-helices and the N-terminal domain (residues 1–186) consists of a left-handed parallel β-helix. The helix resembles an equilateral prism having dimensions of 17 × 39 Å with a long narrow channel (2–3 Å) along the core. The uniform structure of the domain is attributed to a hexapeptide repeat which occurs 28 times (accounting for 168 of the 186 residues). The repeat contains an aliphatic residue (usually Ile, Val, or Leu) at every sixth position (position *i*) and a small residue (Ala, Ser, Cys, Val, Thr, or Asn) at the $i-2$ position. Gly commonly occupies the $i+1$ position. The helix consists of 10 coils, two of which have external loops at the corners and two of which are abbreviated. The remaining coils contain 18 residues each (three hexapeptide units). Six small, hydrophobic residues of each coil are directed inward while 12 of the side chains are directed outward. As with the right-handed β-helices, residues in equivalent positions of adjacent coils stack directly on top of each other in regular ladders and polar groups form both main chain and side chain hydrogen bonds. Hydrogen bonding along both the edges and corners of the strands contributes to the stability of the structure. The most unusual feature of this structure is the large number of left-handed crossover connections. The authors suggest that the extremely flat, non-twisted nature of the parallel β-strands and the length of the crossover connections may contribute to the tolerance of the remarkable left-handed connectivities in the left-handed β-helix motif.

A similar motif to the right-handed β-helices, the β-roll, was revealed in the crystal structure of alkaline protease.[54] This motif consists of several pairs of parallel strands which fold into a parallel β-sandwich (Fig. 5d).[47,50,54] Each coil contains 18 amino acids and has an average rise of 4.8 Å. Several repeats of the coil result in a right-handed coiled structure. The coil consists of a highly repetitive sequence motif, GGxGxDxUx, where x is an arbitrary residue and U is a large hydrophobic residue. Each turn, two repeats of the GGxGxDxUx sequence, forms half of a hexacoordinated Ca^{2+} binding site.

Based on structural similarities between subunits of unrelated proteins, Efimov has recently proposed some requirements for the folding of two-layer α/β and β-sandwich proteins.[6] First, α-helices and β-strands cannot be packed into the same layer of a protein as this would prevent the free amide protons and carbonyl groups of the β-strand from either hydrogen bonding to another strand or accessing the solvent. Second, crossing of the right-handed connection regions is prohibited. Finally, βαβ units should be folded into right-handed superhelices while three consecutive β-strands should form a βββ-superhelix as long as there is at least one additional strand between the first and third strands. The author suggests that these guidelines are not a method for predicting structure but rather provide a means to search for possible structures by stepwise addition of secondary structural elements. The frequency of the two layer subunits was suggested to be a result of both the stability of the structures gained from hydrophobic interactions and hydrogen bonding and the kinetic accessibility of tertiary structure formation as rapid addition of adjacent segments of secondary structure may favor the formation of the tertiary structure.

3. Challenges in Studying and Designing β-Sheets

3.1 The topology problem

β-Sheet structures are topologically more complicated than α-helices in that they are dependent on long range interactions between residues on neighboring strands as opposed to the local i and $i+4$ intrachain interactions which stabilize an α-helix. Unlike α-helices, β-sheet structures can be formed through intermolecular hydrogen bond mediated association of two β-strands or sheets affording a β-sheet quaternary structure. The stability of a β-sheet is dependent in

large part on both intra- and inter-strand residue-residue interactions (primarily hydrogen bonding and hydrophobic interactions), as well as on intersheet associations of the hydrophobic sheet surfaces. Hence, it is perhaps most appropriate to classify β-strands as secondary structures and β-sheets as tertiary or quaternary structures.

3.2 The self-assembly problem

Unlike α-helices, β-strands are not stable as isolated secondary structures. In addition to needing neighboring strands to fulfill their hydrogen bonding requirements, β-strands also gain stability from side chain–side chain interactions of neighboring residues. A sheet composed of more than two strands typically has a hydrophobic surface associated with it which can mediate face-to-face assembly. Thus, β-sheets may self-associate both laterally (through intermolecular sheet formation) and in a face to face manner to afford β-sandwich-like structures. The tendencies of relatively small polypeptides to undergo competitive intra- and inter-molecular folding leading to heterogeneous aggregated β-sheet structures has limited our ability to understand β-sheet structure using the conformational propensities of amino acid homopolymers and sequential copolymers.[55–64] Even though intramolecular folding to a metastable β-sheet structure most likely occurs, the sheet is in equilibrium with an extended conformation both of which can associate to afford a stable associated β-sheet structure (Fig. 6). Similarly, small hydrophobic or amphiphilic peptides have proven to be poor model systems as they also form self-associated β-sheet structures.[65–75]

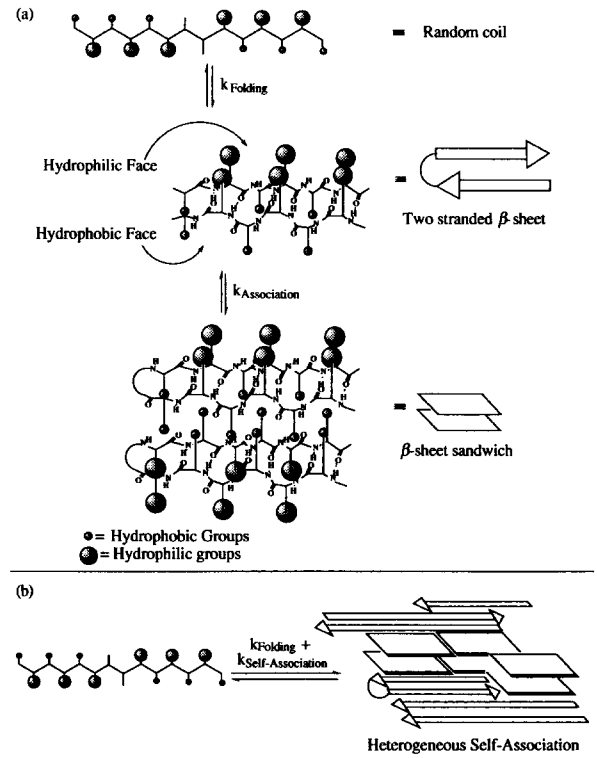


Figure 6. Schematic representation of β-sheet folding and self-association pathways. (a) Formation of homogeneous β-sheet structure occurs under ideal conditions (e.g., $-k_{\text{folding}} > k_{\text{self-association}}$) and (b) small peptides generally form heterogeneous β-sheets due to folding and self-association having comparable rates.

4. Characterizing β-Sheet Structure

4.1 Analytical equilibrium ultracentrifugation

Analytical equilibrium ultracentrifugation experiments are the most rigorous and accurate approach with which to determine the solution molecular weight of molecules or molecular assemblies.[76–83] Even though concentration dependent far-UV circular dichroism (CD) studies have historically been used for this purpose, they should be avoided because the CD signals afforded by a monomer and higher order oligomers can be similar enough to misinterpret the concentration dependent CD data. Both low- and high-speed ultracentrifugation analyses should be carried out for all polypeptides of interest. Lower rotor speeds (3000 rpm for 30 min) are used to determine the initial UV absorbance in the sample cell and to screen for high mol wt species (which will create a moving boundary) before the appropriate higher rotor speeds (up to 60,000 rpm, depending on the expected molecular weight of the molecule being examined) are used to generate an exponential distribution of the polypeptide concentration over the length of the cell. The concentration of a peptide at equilibrium under the influence of a centrifugal field as a function of radial distance in the centrifuge cell can be fit to a single exponential of the type: $c(r) = c_h e^{A N (r^2 - r_h^2)}$, where r is a given radial position, c_h is the concentration in g/L at a given radial reference position (r_h) in the cell, N is the apparent molecular weight of the solute, and A is a constant defined as $A = 1 - \bar{\upsilon}\rho\omega^2/(2RT)$, where $\bar{\upsilon}$ is the partial specific volume, ρ is the solution density, ω is the angular velocity in radians/s, R is the gas constant and T is absolute temperature. Should the β-sheet polypeptide form a well-defined oligomeric solution structure, sedimentation equilibrium experiments will provide information regarding molecular weight of the quaternary structure formed and its dissociation constant without perturbing the chemical equilibrium.[84] It is imperative that the quaternary structure, or lack thereof, be known to interpret other spectroscopic data correctly and to evaluate the suitability of 2D NMR experiments which work best for polypeptides having a monomeric or well-defined quaternary structure with a mol wt of less than 12,000.

4.2 Circular dichroism spectroscopy

Unfortunately, standard far-UV circular dichroism spectra for β-sheet structures are still not agreed upon, unlike the case for α-helical structure where standard spectra are known. The CD spectra that most textbooks cite as representative of a β-sheet structure originate from poly-L-Lys which adopts an aggregated β-sheet structure exhibiting a characteristic maxima at 195 nm and a single minima around 218 nm.[85] An

unordered peptide produces a CD spectrum that is roughly the mirror image of the aggregated β-sheet structure (i.e. spectral maxima at ~217 nm and minima at ~195 nm). The CD spectrum of a monomeric or low molecular weight β-sheet may not exhibit a maximum at ~195 nm as discerned from deconvolution of protein CD spectra and from CD studies on β-sheet model systems developed by our laboratory.[68,86,87] Evolving data suggest that the β-sheet minimum for monomeric or low molecular weight β-sheets can vary in wavelength by ±5 nm around 215 nm. Structurally well-defined β-sheets exhibit a mean residue ellipticity of approximately $-10{,}000 \pm 2000$. Mean residue ellipticities considerably higher than this appear to result from large intermolecularly associated β-sheet structures. This can be verified by analytical equilibrium ultracentrifugation. Numerous proteins and a few peptide model systems are beginning to shed some light on the expected far-UV CD spectra for β-sheet structure. In spite of our relatively poor current understanding of sheet CD spectroscopy, the technique has proven to be very useful as a structural screening method and for discerning structural changes in predominantly β-sheet proteins upon denaturation. A distinct advantage of CD over NMR is that the timescale of the CD experiment is faster than the conformational interchanges occurring in the polypeptide backbone, hence unique CD signals can be seen for random coil and β-sheet structure even though they may be interconverting.

4.3 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance spectroscopy (NMR) is very useful for characterizing the structure of a peptide or protein on a residue specific basis.[88–92] NMR facilitates identification of amide protons participating in intramolecular hydrogen bonding and through nuclear overhauser effect (NOE) analysis indentifies those residues in tertiary contact with other residues, both close and far away in sequence. A limitation associated with NMR analysis of peptide structures is that the timescale of the NMR experiment is slow (milliseconds to seconds) relative to the timescale of peptide conformational interconversions. This means that any structural data determined from an NMR spectra may represent contributions from interconverting conformations.[91,92] Furthermore, since the side chains and backbone atoms are more densely packed and sterically constrained in proteins, significantly fewer NOEs are observed, on a per residue basis, in peptides than in proteins.[92] Parallel and antiparallel β-sheets in proteins have characteristic NOEs that allow their structure to be distinguished through NMR (Fig. 7).[93–95] In order to utilize the observed NOEs, it is first necessary to make sequence specific resonance assignments. This is achieved by assigning the non-exchangeable spin systems (with respect to residue type) using COSY, Relayed-COSY, and TOCSY NMR data from a 1–10 mM solution of the peptide in D_2O. DQF-COSY and relayed-COSY spectra in water facilitate the assignment of amide protons to a specific amino acid spin system. TOCSY data allows the verification of difficult assignments. NOESY spectra are used to connect the spin systems of the aromatic residues with the C_α and C_β protons.[96] Sequence assignments are then accomplished by taking advantage of the NOSEY spectra recorded in water. Sequence connectivities are afforded by $d_{\alpha N}$ NOEs. Successive $d_{\alpha N}$ NOEs connecting the C_α-H of residue i to the N–H of residue $i+1$ (Fig. 7, $d_{\alpha N}(i, i+1)$) are used in combination with the COSY data that connects the N–H of residue $i+1$ to the C_α-H of residue $i+1$. This information is then used to identify the next sequential residue assignment through an NOE from C_α-H of residue $i+1$ to the N–H of residue $i+2$. This sequence is repeated until all possible sequence specific assignments are made. With the sequence specific resonance assignments in hand, the COSY spectra are then re-examined for $^3J_{NH_\alpha}$ coupling constants that are greater than 7.5 Hz, which are consistent with residues in an extended β-sheet conformation.[97] The distinction between peptides and proteins is quite clear with regard to $^3J_{NH_\alpha}$ coupling constants. Due to conformational averaging in peptides, it is unrealistic to expect coupling constants approaching 10 Hz, which are commonly observed in protein β-sheets. Another useful parameter for characterizing β-sheet structure is the chemical shift index, which refers to the downfield shift of $C_\alpha H$ in β-sheet structures relative to the $C_\alpha H$ shift in random coils.[98] Interstrand NOEs also prove very useful in characterizing β-sheet structure. NOESY spectra recorded with different mixing times (to eliminate spin diffusion) are employed to identify protons that are close in space but not sequence. These long range NOEs are used to identify neighboring amino acids on different strands of a β-sheet and serve as the restraints in molecuar dynamics/mechanics approaches to structure determination. The long range NOEs $d_{\alpha N}(i, j)$, $d_{NN}(i, j)$, and

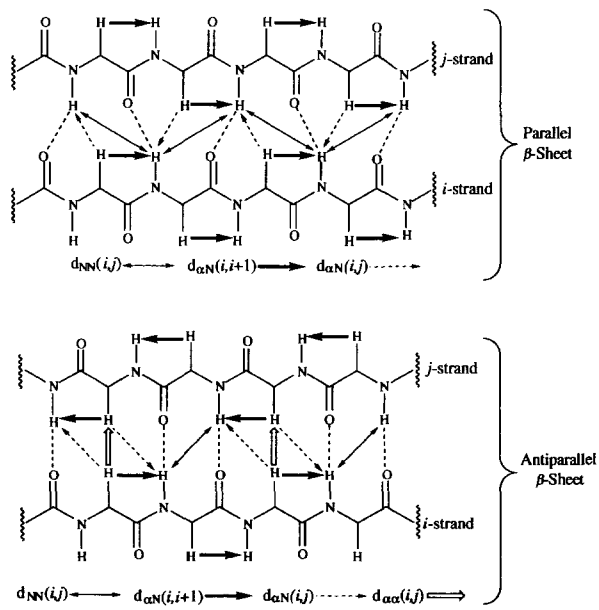


Figure 7. Schematic representation of parallel and antiparallel β-sheets showing the intra- and inter-strand NOEs. The amino acid side chains have been removed for clarity.

$d_{\alpha\alpha}(i, j)$ are strongly supportive of antiparallel β-sheet structure. Large, short range $d_{\alpha N}(i, i+1)$ NOEs are also indicators of β-sheet structure. However, these NOEs are also seen in disordered peptides. Therefore, the $d_{\alpha N}$ NOEs must be supported by other data to be viewed as indicative of β-sheet structure. In the case of peptides, particularly for β-hairpins which are conformationally much more mobile than a protein β-sheet, the observation of $d_{\alpha N}(i, j)$, $d_{NN}(i, j)$, and $d_{\alpha\alpha}(i, j)$ NOEs may be difficult even though a two-stranded antiparallel β-sheet is present. The detection of NOEs from distant protons depends on how many other protons are available for competitive relaxation and at what distance and on how small of an NOE can be detected.[99] The tertiary and quaternary structure is obtained in the case of proteins by entering the long range NOEs into a distance geometry program, which will identify the structures that best fit the NOE distance constraints. These structural conformations can then be energy minimized to obtain a consensus structure.[100] Most peptide structures are underdetermined, meaning there are insufficient numbers of NOEs to identify a unique structure. Should any of the expected NOEs associated for a given peptide be too small or non-existent, ROESY spectroscopy may be employed.[101] It is not uncommon in peptides for the product of the correlation time and the spectrometer frequency to equal one ($\omega_0\tau_c \approx 1$). This occurs in aqueous buffers when $(MW)(\omega_0) \approx 10^{12}$, resulting in the NOE being zero and necessitating the use of ROESY in lieu of NOESY spectroscopy. It is useful (especially for peptides) to do classical NOE difference spectroscopy to be sure that weak NOEs are not overlooked.

After the sequence specific resonance assignments have been made, the hydrogen-bonding behavior of the amide protons can be evaluated. Temperature coefficients are useful for evaluating the extent to which backbone amide protons participate in intramolecular hydrogen bonding.[92,102] This will provide an indication of the type of hydrogen bonding environment that is experienced by the amide proton through the changes in chemical shift as a function of sample temperature. In aqueous solutions, a temperature dependence of the chemical shift of $< |-4|$ ppb/K is indicative of significant solvent shielding. Temperature coefficients between $|-4|$ and $|-6|$ ppb/K suggest a hydrogen bond that is moderately shielded from solvent. Finally, temperature coefficients $> |-6|$ ppb/K suggest that the amide is exposed to bulk aqueous solvent (i.e. not intramolecularly hydrogen bonded).[103]

Amide protons that participate in a structural hydrogen bond within a secondary or tertiary structural element, such as a β-sheet, exchange much slower with the solvent than do the amide protons exposed to bulk solvent. Amide proton–deuterium exchange (N–H↔N–D) is dominated by base catalysis at pH greater than 3. Therefore, it is possible to control the timescale of the proton–deuterium exchange from slow (several minutes at pH ~3) to very rapid (milliseconds at pH ~9).[104] Peptide amide exchange studies facilitate the identification of amide protons which are involved in intramolecular hydrogen bonding, providing further evidence to support the antiparallel β-sheet structure. A simple, frequently used exchange experiment is to dissolve a peptide in buffered D_2O solution at pH 4. The exchange rate for the assigned protons is measured by 1- and 2-D NMR integration methods.[96] H–D exchange data must be carefully interpreted, considering competing hydrogen bonding interactions (e.g. side chain to backbone) which could lead to an incorrect interpretation of the intramolecular hydrogen bonding network. These experiments allow a nearly complete structural determination of a defined β-sheet polypeptide when used in combination with the other methods discussed here.

4.4 FT-IR spectroscopy

The examination of the secondary structure content of proteins and peptides can be accomplished by deconvolution of the amide I region (primarily C=O stretch) between 1620 and 1690 cm^{-1} of a protein or peptide in D_2O.[105–112] Deconvolution is necessary because each secondary structural component contributes a band whose width is usually greater than the separation between the maxima of adjacent bands. Several accepted methods are now available for enhancing the resolution of the FT-IR spectra of a given peptide or protein.[113–115] Polypeptide concentrations of 5% (w/v) using a cell with a 0.075 mm path length are used to record several thousand co-added interferograms which are averaged to produce an FT-IR spectrum. An antiparallel β-sheet structure is characterized by a strong band around 1632 cm^{-1} and a weak band at approximately 1680 cm^{-1}. In general, the areas of the bands around 1637 cm^{-1} associated with β-structure are more easily integrated than the bands at 1650 or 1646 cm^{-1}, which are associated with helical and random structure, respectively.

5. Folding Studies

5.1 Biophysical studies to understand the folding of predominantly β-sheet proteins

The collapse of a randomly oriented polypeptide chain into a compactly folded structure requires a considerable loss of local and non-local entropy. This opposing force is nearly equal to the free energy gained from the hydrophobic effect (which is a complex temperature dependent phenomena involving both entropic and enthalpic contributions) and from the contribution of hydrogen bonding.[116] The net stability of the folded state results from a small difference (~5–15 kcal/mol) in free energy between the folded and unfolded states. The folding of a protein into a unique, stable, well-defined structure is due to non-covalent interactions such as hydrogen bonding, hydrophobic interactions, dispersive factors, ion pairing, and the intrinsic propensities of amino acid residues and peptide fragments to adopt specific conformations.

The folding of a polypeptide into an α-helical structure is the best understood element of the protein folding process. The formation of an α-helix is known to occur through an initial nucleation step in which a hydrogen bond is formed between an i and $i+4$ residue pair. This energetically unfavorable nucleation step requires the entropically unfavorable restriction of the dihedral angles in the three residues intervening between the hydrogen bonded i and $i+4$ residues. Once the first hydrogen bond has formed, the effective concentration of hydrogen bond donors and acceptors proximal to the nucleus is increased. Also, subsequent addition of adjacent amino acids in the sequence requires the restriction of the dihedral angles in only one amino acid residue for each hydrogen bond added to the helix. Thus, once the helix nucleus is formed, propagation is thermodynamically favored and occurs rapidly.

Although β-sheet structure is commonly observed in proteins, relatively few biophysical studies regarding the folding mechanisms of all β-proteins have been carried out thus far. The fact that the mechanisms of β-sheet folding have not been well characterized is due, in part, to the difficulties inherent in creating a well-behaved peptide model system to study β-sheet formation in aqueous solution.[20,91,117,118] Those studies which have been done (typically employing small proteins) indicate that the folding of a predominantly β-sheet protein is slow when compared to predominantly α-helical proteins, implying that the folding mechanisms may be different.[34,119–121] A study of hen lysozyme by Radford and co-workers has shown that the α-domain which consists of four α-helices folds independent of, and more rapidly than, the β-domain which contains a double stranded β-sheet, a triple stranded β-sheet, a 3_{10} helix, and a long loop.[122] The tertiary structure of the α-domain is not fully organized until cooperative interactions between the two domains are established. A subsequent study indicated that although the denatured states are dependent on the denaturant (DMSO or GdmCl), the folding pathways are virtually indistinguishable.[123] Again, formation of the native state appeared to be limited by formation of the β-domain.

Additional insight into the mechanisms of β-sheet folding has been provided by studying peptide fragments of predominantly β-sheet proteins. Unlike fragments of the mostly α-helical protein, myohemerythrin, which adopt helical secondary structure under conditions where the protein is folded,[120] analogous studies on fragments of the predominantly β-sheet protein, plastocyanin, indicate that there is no tendency for the isolated peptides of β-sheet proteins to form β-sheet structure.[119] This implies that tertiary interactions are critically important in the formation of β-sheet structure. Similarly, evaluations of peptides composed of the first 21 and the first 35 residues of ubiquitin (corresponding to the β-hairpin and the β-hairpin/α-helix domains in the folded protein, respectively) indicate that the β-hairpin structure is not formed in aqueous solution in the absence of the remaining 42 C-terminal residues.[124] However, in the presence of 30–60% methanol, both peptides adopted native-like β-sheet structure. The authors suggest that addition of methanol as a cosolvent serves to facilitate the formation of a native-like β-hairpin by enhancing the stability of the hydrogen bonded structure as would normally be accomplished by tertiary interactions with the fully folded protein. Hence, tertiary interactions or some form of protection from aqueous solvent again appears to be necessary for the formation of a stable β-hairpin structure.

Both equilibrium and kinetic studies of the folding of predominantly β-sheet proteins have been carried out in efforts to determine the folding mechanisms of β-sheet proteins. In their study of cellular retinoic acid binding protein (CRABP), Gierasch and co-workers found that the β-clam protein unfolds via a two-state reversible process when denatured by either acid (pH 2.8–3.5) or urea ($\geq$4 M) although the unfolded states were different for the two denaturants.[34] The acid denatured CRABP was found to contain a higher content of α-helical structure than the native protein. This residual structure was further unfolded by addition of urea (2–4 M). The authors suggest that the formation of helical structure in the denatured state may provide initiation sites for hydrophobic collapse which could facilitate the formation of β-sheet structure in other regions of the protein. A study of the SH3 domain of spectrin, a 62 amino acid β-barrel formed by two antiparallel β-sheets, indicated a two-state mechanism in which folding was reversible at pH $<$ 4 (and at pH $>$4 when the concentration was less than 1 mg/mL).[125] The free energy of folding was calculated to be $\sim$13 kJ/mol (at pH 3.5, 298 K). Kinetics studies indicated a slow and a fast phase and the authors suggested that the slow phase may be due to isomerization (from *cis* to *trans*) of either of the two prolines in a folded non-native structure. At low urea concentrations, no intermediates were observed. Rudolph et al. studied the folding of γII-crystallin from calf eye lens, a protein which consists of two homologous domains each comprised of a β-sandwich formed by two four-stranded antiparallel Greek key sheets.[126] The protein was found to denature in 7 M urea at elevated temperatures or at low pH. Unfolding occurred via a reversible two-step process in which the kinetic and equilibrium intermediates were identical and the folding/unfolding of the two domains was independent. The authors suggest that the protein undergoes sequential folding/unfolding of the two domains with the intermediate corresponding to the partially unfolded protein in which one domain is unfolded and one domain is in its native state. Tetrameric R67 dihydrofolate reductase (DHFR) is a protein where each monomer adopts a β-barrel fold having a missing fifth strand. Association of two monomers results in a third (six-stranded) β-barrel being formed at the interface. Studies of DHFR indicate that both chemical and thermal denaturation is reversible with the simplest possible mechanism being T$\leftrightarrow$2D$\leftrightarrow$4M.[127] A stable dimeric intermediate is observed. However, a folded monomeric intermediate is not observed during unfolding/folding, suggesting

that the folded monomer is stabilized by the interactions of β-strands B-D at the dimer interface. Ropson et al. have studied the β-clam protein, rat intestinal fatty acid binding protein (IFABP).[128] Chemical denaturation (GdnHCl, urea, KSCN, and GdnSCN) indicated reversible, two-state unfolding with little or no intermediate present under equilibrium conditions. Kinetic studies, however, indicated that IFABP unfolds through a two-step mechanism in which the first phase is slow followed by a fast step. Folding of IFABP appears to proceed through at least two different pathways. Replacing Trp-82 and Trp-6 with 6-fluorotryptophan and studying the folding of these variants by ^{19}F NMR spectroscopy, CD, and fluorescence demonstrated that one or more intermediates were present during folding/unfolding.[129] The intermediate or intermediates resemble the unfolded state rather than the native protein and contain little or no regular structure. The authors suggest that the intermediate may contain a hydrophobic cluster consisting of Trp-82, four phenylalanines (each from a different β-strand) and two leucines. This cluster could then serve as a template for further folding by effectively mapping the locations and lengths of five of the β-strands. Alternatively, formation of the turn nearest in sequence to Trp-82 (composed of Glu-85–Lys-88) would position Trp-82, a phenylalanine, and a leucine within close proximity of each other and formation of the turn closest to Trp-82 in the native protein (Glu-63–Asp-67) would position Trp-82 near two phenylalanines and a leucine. Either turn would be stabilized by hydrophobic interactions and could serve as an initiation site for further hydrophobic collapse. Kinetic studies of the β-sheet protein, interleukin-1β (Il-1β) suggest that folding occurs through the rapid formation of β-like structure around a nonpolar core with at least two intermediates (with half lives of 0.7–1.5 and 15–25 s, respectively) being formed. Subsequent formation of the native β-sheet structure occurred slowly with a half-life of ~20 min.[130] In summary, a number of pathways appear to exist for the formation of β-sheet structure.

In addition to demonstrating multiple mechanisms of β-sheet folding, the studies described above also provide support to the hypothesis that hydrophobic clustering or other similar tertiary interactions (e.g., aromatic–aromatic interactions) are involved as early intermediates in the protein folding process. Hydrophobic clusters appear to be partially responsible for the nucleation of protein folding and may be especially important in β-sheet folding. The potential role of hydrophobic residues in the folding of β-sheet proteins has also been pointed out by Hazes and Hol.[131] In their study of Greek key β-barrels from unrelated protein families, the only highly conserved feature was the hydrophobic character of the barrel core. A short loop connecting the B and C strands was also conserved, suggesting that it might be involved in a nucleation mechanism common to all seven barrels. The authors proposed a 'β-zipper' folding model in which two β-strands joined by a short loop interact with one another through their hydrophobic side chains to form a β-arch conformation which is stabilized by the hydrophobic interactions. The two strands of the arch end up as the central strands of two β-sheets which similarly interact through their side chains. The β-arch structure is proposed to serve as a template to which additional flanking β-strands are added, affording the β-barrel.

Several researchers have also demonstrated the existence of hydrophobic clusters composed of aliphatic and aromatic side chains within model peptides or protein fragments under denaturing conditions.[116,132–136] Wüthrich's study of the 63 residue peptide corresponding to the amino-terminal portion of the 434-repressor protein revealed a well defined hydrophobic cluster in which the isopropyl groups of Val-54, Val-56, and Leu-59 are in contact with the indole ring of Trp-58 in 7 M urea.[134] A slight rearrangement of this local structure yields the helical structure found in the folded protein, implying that this cluster is important for 434 folding. In their study of peptide fragments of plastocyanin, Dyson and co-workers provided NMR evidence for the formation of a hydrophobic cluster between the aromatic ring of Phe-29 and the γ-methyl group of Ile-27 in two peptides (KIVFKNNA and KIVFKNNAGFPH) which are otherwise unstructured but which correspond to a region within the folded protein which exhibits β-sheet structure where the Phe-29 and Ile-27 residues are in contact with each other.[119] Lumb and Kim studied eight residue peptide models corresponding to residues 17–24 of bovine pancreatic trypsin inhibitor (BPTI) in order to investigate the possible interactions between Ile-19 and Tyr-21 which are implicated to occur in the unfolded protein.[137] The authors found that Ile-19 and Tyr-21 exist in a hydrophobic cluster under 'unfolded' conditions in an orientation similar to that found in native BPTI. The Matthews group has employed fluorescence spectroscopy to demonstrate the existence of a hydrophobic cluster in which the Trp-74 residue of dihydrofolate reductase is buried in an early folding intermediate.[136] The active forms of some small peptides, namely somatostatin and cyclosporin, also appear to be stabilized by hydrophobic clusters.[138,139] Analogous hydrophobic clusters in small, dibenzofuran-based amino acid containing peptides have also been shown to be critical for β-sheet formation (*vide infra*).[87] These and other studies are providing increasing support for the idea that hydrophobic clusters serve to direct the folding of proteins into low-energy structures which serve as a nucleus for further hydrophobic collapse and/or subsequent folding of the protein.[116,132–136]

5.2 β-Sheet propensities

The importance of side chain–side chain tertiary interactions on the stabilities of β-sheet structures is evident from recent studies evaluating the β-sheet forming propensities of the 20 naturally occurring amino acids by the Regan, Kim, Berg, and Fersht laboratories.[18,140–144] Kim and Berg employed a zinc-finger peptide (PYXCPECGKSFSQKSDLVKHQRTHTG) to

evaluate β-sheet propensities.[143] Position three was chosen as the guest site as it is within a well-defined segment of the antiparallel β-sheet which is stabilized upon binding of the peptide to cobalt(II) or zinc(II). The side chain of the guest amino acid is solvent-exposed and substitution at this position occurs without detectable changes in solubility or structural properties. The peptides are completely unfolded in the absence of metal. Hence, metal-ion affinity was used as a measure of folding free energy. The metal-ion-binding free energies were determined through a two-peptide competitive metal binding assay. The authors report a good correlation with the statistically determined propensities of Chou and Fasman (correlation coefficient of 0.83, excluding Pro and Gly) in their solvent exposed, antiparallel β-sheet peptide but point out that stability differences of amino acids within parallel and antiparallel strands and within internal and edge strands should be investigated. Their propensities are listed in Table 2 along with the statistically determined Chou–Fasman probabilities (P_β),[145,146] as well as the propensities determined by Regan and co-workers,[140] and Minor and Kim.[141,142]

Regan and co-workers have employed the IgG F_c-binding domain B1 of staphylococcal protein G to determine the β-sheet propensities of amino acids within a central strand of a β-sheet. The 56 residue protein is composed of a four-stranded β-sheet that is diagonally crossed by an α-helix. The protein lacks both proline residues and disulfide bonds, is highly soluble, and exhibits a reversible, two-state thermal denaturation transition. The β-sheet forming propensities were determined by substitution of each amino acid into position 53 of the protein. This position was deemed the optimal guest site because it is located in the middle of a central β-strand, is solvent exposed, and extends from the face of the sheet which is opposite to the α-helix (and thus, avoids packing and helix-guest interactions). Additionally, the amino acid at position 53 will be fully hydrogen bonded as opposed to an amino acid on a solvent exposed edge-strand. Two mutations of the host protein (I6A, T44A) were made to minimize host–guest residue side chain interactions, maximize solvent accessibility of the guest site, and to maintain a reasonably stable protein. The contribution of each guest residue to the stability of the resulting β-sheet was evaluated by determining the thermal denaturation curve of each single site variant which yields the difference in free energy, $\Delta\Delta G$, relative to IgGB1 containing Ala at the guest site. The structural integrity of several of the variants was confirmed by NMR studies. The β-sheet propensities determined by Regan and co-workers exhibit a 2.5 kcal/mol range in $\Delta\Delta G$ values between the best and worst amino acids. This range is larger than that obtained with substitutions in the zinc-finger peptide (0.94 kcal/mol), and exhibits a general correlation with the statistically determined propensities (Table 2).

The effect of context with regard to an internal versus an edge strand on the β-sheet forming propensities of amino acids has been directly addressed by Minor and Kim in studies employing the immunoglobin-binding domain B1 from protein G.[141,142] In the first study, amino acids were substituted into a guest site of a centrally located β-strand (residue 53) of the I6A, T44A, T51S, T55S variant.[141] With the exception of the unfolded proline variant, all of the mutants were found to exhibit structural integrity as discerned from their F_c

Table 2. Experimentally determined β-sheet propensities

Guest residue	P_β (Chou and Fasman)	$\Delta\Delta G$ (kcal/mol) (Kim and Berg edge strand)	$\Delta\Delta G$ (kcal/mol) (Smith and Regan central strand)	$\Delta\Delta G$ (kcal/mol) (Minor and Kim central strand)	$\Delta\Delta G$ (kcal/mol) (Minor and Kim edge strand)
Val	1.64	−0.53	−0.94	0.82	0.17
Ile	1.57	−0.56	−1.25	1.0	0.02
Thr	1.33	−0.48	−1.36	1.1	0.83
Tyr	1.31	−0.50	−1.63	0.96	0.11
Trp	1.24	−0.48	−1.04	0.54	−0.17
Phe	1.23	−0.55	−1.08	0.86	0.16
Leu	1.17	−0.48	−0.45	0.51	−0.24
Cys	1.07	−0.47	−0.78	0.52	0.08
Met	1.01	−0.46	−0.90	0.72	−0.02
Gln	1.00	−0.40	−0.38	0.23	0.04
Ser	0.94	−0.39	−0.87	0.70	0.63
Arg	0.94	−0.44	−0.40	0.45	−0.43
Gly	0.87	0	1.21	−1.2	−0.85
His	0.83	−0.46	−0.37	−0.02	−0.01
Ala	0.79	−0.35	0	0	0
Lys	0.73	−0.41	−0.35	0.27	−0.40
Asp	0.66	−0.41	0.85	−0.94	−0.10
Asn	0.66	−0.38	−0.52	−0.08	−0.24
Pro	0.62	0.23	ND	< -3	< -4
Glu	0.51	−0.41	−0.23	0.01	0.31

Note: Due to differences in the methods in which the $\Delta\Delta G$s were determined in the different studies, the reported $\Delta\Delta G$ values for the Minor and Kim studies are opposite in sign to the Kim and Berg and Smith and Regan studies. In the latter studies, negative $\Delta\Delta G$s indicate an increase in stability while in the Kim and Minor study, positive $\Delta\Delta G$s indicate an increase in stability.

binding affinity and by NMR characterization. The stability of the mutants was measured by thermal unfolding with the assumption that any changes in global stability resulted entirely from changes in the ability of the guest residues to adopt a β-sheet structure. The range of the determined $\Delta\Delta G$ values (2.05 kcal/mol) indicated a significant difference in the β-sheet forming propensities of the naturally occurring amino acids (Table 2). The authors found a modest correlation of their data with the statistically determined P_β values and noted that β-branching favors sheet formation. In the second study, position 44 which is located in an edge strand was chosen as the guest site and the nearest neighbor residues (42, 46, and 53) were replaced with alanine.[142] The guest variants were found to have an affinity for F_c which was sevenfold less than that of the wild-type protein. With the exception of the unfolded proline mutant, the structural integrity of each protein was confirmed by evaluation of the chemical shifts of Trp-43 and the observed NOEs. The $\Delta\Delta G$ values were obtained as before and the overall magnitude of the derived β-sheet propensity scale ($\sim$2 kcal/mol) for the edge strand was similar to that determined for centrally located amino acids. However, no correlation with the P_β values was observed. Unlike the central position, over half of the $\Delta\Delta G$ values measured at the edge position were within a range of $\sim$0.4 kcal/mol. The authors suggest that β-sheet propensity is dependent on two components. The first, an intrinsic ability to form a local extended β-strand structure, has a minor role in determining β-sheet propensity based on the small range of β-sheet propensities measured at the edge position and the relative lack of preference for particular side chain rotomers in β-strand residues.[147] The second factor, the ability to interact with the surrounding tertiary β-sheet structure, plays a dominant role in determining β-sheet propensity as the difference in β-sheet propensity between the edge and central sites correlates with both the free energy of transfer for the amino acids from octanol to water[148] and with the non-polar accessible surface area of the side chains.[149] The authors point out that the amount of buried surface area differs significantly at the edge and center positions of a β-sheet, substantiating the importance of the observed correlations. Additionally, the variations found in the β-sheet propensities of amino acids at position 53 of two different mutants of the immunoglobin-binding domain B1 from protein G (I6A T44A T51S T55S in the Minor and Kim study versus I6A T44A T51 T55 in the Smith and Regan study) points to the importance of interactions with neighboring residues.

The experimental β-sheet propensities were found to correlate with those predicted from the statistical studies of Chou and Fasman in three out of four of the determinations. An evaluation of the propensities of amino acids from 279 structures which was based on only the dihedral angles of the amino acids also agreed well with the experimentally determined β-sheet propensity scales.[150] This study suggested that a particular amino acids β-sheet propensity may be different for different parts of a β-sheet depending on the local dihedral angles. The context dependence of β-sheet propensities was further demonstrated by Otzen and Fersht in their evaluation of seventeen different mutations which were made at six different positions in the β-sheet region of chymotrypsin inhibitor 2 (CI2).[144] The mutations were designed to allow the determination of the effects of reducing the hydrophobic surface area of the native residue and included changes of Ile-49 to Val, Ala, and Gly residues, Thr-22 to Val, Ala, and Gly residues, Val-53, 79, and 82 to Ala and Gly residues, Val-53, 79, and 82 to Thr, Ile-49 to Thr, and Ala-77 to Gly. The differences in stability towards GdnHCl denaturation of the mutants versus wild type were determined. With the exception of I49V, all of the mutations resulted in decreased stability. The order of stability was Val > Ala > Gly for all the mutations but the $\Delta\Delta G$ values between each position varied greatly. Hence, the propensities of the three residues was highly dependent on their location within the β-sheet. The observed context dependence of the experimentally determined β-propensity scales supports the idea that the stability of β-sheets is dependent on both local and non-local residue–residue interactions.

The physical basis for β-sheet propensities are not clearly established, but it appears that several factors are important including side chain–side chain hydrophobic interactions. It has recently been suggested that main chain electrostatics are important in determining the conformational preference of a specific amino acid residue.[151] Unlike α-helices, the peptide dipoles of residues in a β-conformation are antiparallel to each other resulting in a weak field and hence, weaker interactions with nearby groups. Bulky and β-branched side chains are more capable of restricting access of water and protein groups to the peptide backbone. Thus, they effectively prevent screening of local and non-local electrostatic interactions by water molecules and other protein groups which act to stabilize α-helical conformations. In the absence of non-local stabilizing forces, the more favorable local electrostatics of the β-conformation are preferred. Hence, bulky and β-branched residues favor a β-conformation. Bai and Englander have also suggested that the ability of a residue to block the peptide from solvent interaction favors β-sheet formation.[152] The authors suggest that side chain-dependent destabilization of hydrogen bonding of the peptide group to water in the random coil favors β-conformation in that the steric blocking effect contributes an enthalpic stabilization to intraprotein hydrogen bonds in the range of 0–0.9 kJ/mol (relative to alanine). Although other contributions such as conformational entropy, steric factors, and hydrophobic interactions are also likely to be involved, the influence of side chains on electrostatic interactions and hydrogen bonding appear to be important factors in the context dependence of β-sheet propensities.

An interesting study was recently reported in which the conformational propensity of an amino acid residue was compared to the hydrophobic periodicity of the sequence to determine which interaction would

dominate structure formation in self-assembling peptides.[153] Two peptides were composed of amino acids with high α-helix forming propensities (peptides 1A and B) and two were composed of amino acids with high β-sheet forming propensities (peptides 2A and B). Peptides 1A and 2A were designed with a periodicity of polar and nonpolar residues corresponding to that expected for an α-helical peptide while peptides 1B and 2B had a periodicity corresponding to that of a β-sheet. Evaluation of the peptides by circular dichroism spectroscopy indicated that the observed structure was that predicted by the periodicity of the peptide regardless of the intrinsic propensities of the amino acid residues comprising the peptide.

6. β-Hairpin Model Systems

6.1 Peptides and proteins

There are very few peptides which adopt a well defined β-sheet conformation in aqueous solution in the absence of additional stabilizing factors. As previously stated, the majority of polypeptides which adopt a β-sheet structure self-associate to form large aggregates which are not amenable to detailed study. There are examples of polypeptides and small proteins which adopt a stable β-sheet structure when additional tertiary interactions are present. For instance, several small β-sheets have been observed in the presence of an α-helix.[154–158] Also, peptides based on the zinc-finger motif adopt β-sheet structure in the presence of metals.[143,157,158] We will consider here only those peptides which adopt a well defined, intramolecular β-sheet structure in aqueous solution.

A relatively small, cyclic peptide, ramoplanose, has been found to adopt a stable β-sheet structure in aqueous solution.[159] The peptide consists of 16 amino acids of which 12 are unusual (Fig. 8). The solution conformation of the peptide was determined using a combination of NMR-derived distance constraints and molecular dynamics calculations. The slow amide proton exchange, low $\Delta\delta/\Delta T$ values, and large (>8.0 Hz) $^3J_{\alpha NH}$ values of several of the residues as well as interstrand NOEs confirmed the presence of β-sheet structure. The narrow line width and wide chemical shift dispersion exhibited in the NMR spectrum was thought to be indicative of a well-defined, homogeneous structure. The peptide forms an antiparallel β-sheet with residues Thr-8–Phe-9 and Gly-14–Ala-15 forming the necessary β-turns and a β-bulge formed by Orn-4 and Thr-5 (which are both hydrogen bonded to Gly-14). The unusual composition of the peptide appears to be responsible for the observed stability of the β-sheet. The presence of both D and L amino acids prevents helix formation and serves to orient the hydrophobic side chains of residues 5–7 and 10–13 on one side of the sheet allowing efficient packing of the hydrophobic side chains and protection of the backbone–backbone hydrogen bonds from solvent. Additionally, the hydrophilic termini of the peptide provide water solubility.

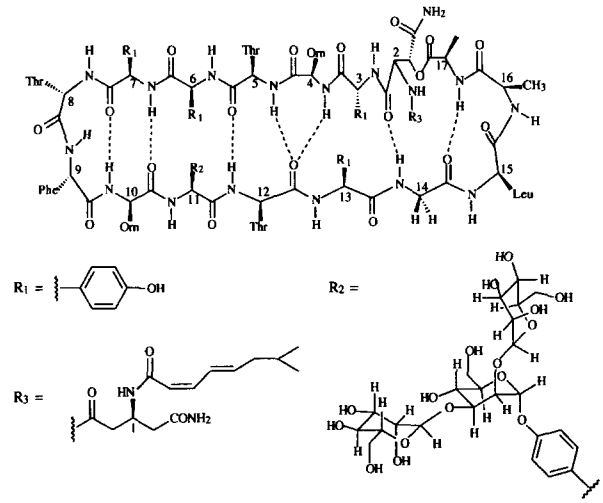


Figure 8. Solution structure of ramoplanose as determined by NMR.

An opiate receptor mimetic peptide was designed by Kullmann to adopt a conformation in which two antiparallel β-sheets (each composed of two β-strands) would form a twisted rectangular trough such that an appropriate opioid peptide could bind to the central cavity.[160] Circular dichroism spectroscopy indicated that the peptide adopted a predominantly β-sheet structure. The peptide was determined to stoichiometrically bind [Leu]-enkephalin. The binding resulted in the burial of some aromatic side chains suggesting that the desired β-sheet structure had been obtained.

Blanco et al. have reported an example of a stable β-hairpin structure.[161] The 16 residue, acyclic peptide corresponds to the second β-hairpin of the B1 domain of protein G (residues 41–56) and was found to exist in equilibrium between a random coil and β-hairpin conformation. The peptide was determined to be monomeric by CD and NMR concentration studies, size exclusion chromatography, and sedimentation equilibrium ultracentrifugation experiments. NMR studies afforded $\Delta\delta/\Delta T$ values for residues 42, 46, 48, 49, 50, and 51 which were indicative of hydrogen bonding while those for residues 44, 53, and 55 were large, suggesting interconversion of conformers. Both intra- and inter-strand NOEs were observed. The chemical shifts of the strand residues were also indicative of sheet structure and were used to calculate an average population of 42% β-hairpin structure. A smaller peptide (YQNPDGSQA) based on the native β-hairpin region of tendamistat but designed to maximize the turn structure was also found to adopt a significant population of β-hairpin structure in aqueous solution.[162] NMR studies of this peptide provided NMR evidence (intra- and inter-strand NOEs and a small temperature coefficient for the amide proton of residue 5) which suggested the presence of a β-hairpin structure. However, multiple conformations were observed in this case as well.

The solution structure of PMP-D2 was reported by Mer et al.[163] The 35 residue peptide is crosslinked by three disulfide bridges (Cys-14–Cys-32, Cys-4–Cys-19,

and Cys-17–Cys-27). NMR distance constraints were combined with distance geometry and/or simulated annealing calculations to determine the solution structure of PMP-D2. These calculations combined with large $^3J_{\alpha NH}$ values, the downfield shifted α proton chemical shifts, and the slow amide exchange rates of NHs which were expected to be hydrogen bonded suggest a three-stranded antiparallel β-sheet with the strands composed of residues 8–11, 15–19, and 25–29. A short loop composed of residues 5–8 and anchored by a disulfide bridge was also observed. The authors attribute the stability of the compact globular structure to several hydrophobic clusters which form between the valine and threonine residues on the convex side of the sheet.

The solution structure of ω-conotoxin GVIA (ω-CgTx), a 27 amino acid peptide toxin, was determined by three research groups utilizing 2D NMR techniques in combination with distance geometry and restrained molecular dynamics calculations.[164–166] The sharp NMR spectra obtained for this peptide imply a monomeric system. Although there are some small discrepancies in the three structures, the peptide contains three disulfide bridges (C1–C16, C8–C19, and C15–C26) and has been shown to consist of a three-stranded β-sheet where residues 6–8, 18–21, and 24–27 comprise the strands (Fig. 9). Amide exchange studies reveal hydrogen bonds between the following pairs: CO(Ser6) HN(Cys26), CO(Ser18) HN(Tyr27), CO(Arg25) HN(Asn20), and CO(Cys26) HN(Gly5). Additional hydrogen bonds which varied between the three structures were noted by the authors. The structure contains four well defined turns (residues 3–6, 9–12, 15–18, and 21–24) and a loop region (residues 8–14). Additionally, the termini of the peptide are constrained by disulfide bridges providing further structural rigidity to the peptide. An additional structure was observed by Pallaghy et al. in which the disulfide bridges were mismatched (C1–C8, C15–C19, C16–C26).[165] This peptide was observed to be much less structured. Thus far, the regular form of ω-conotoxin appears to be one of the most structurally well defined, small β-sheets available.

Ilyina and Mayo have studied the 33-residue β-sheet domain from platelet factor 4 (residues 23–55).[167] Although the peptide exists in a random coil in the monomeric state, the peptide undergoes a 'subunit association induced folding' which results primarily from hydrophobic collapse. β-Sheet structure was found to increase with increasing concentration or increasing temperature. The amide and α-proton chemical shifts of the tetrameric peptide were indicative of a native-like β-sheet fold and NOEs supported a tetrameric β-sheet structure. The structure was, however, molten globular in that there was no fixed tertiary structure as evidenced by the presence of multiple conformational states. (One conformation accounted for approximately 50% of the structures while four additional structures were thought to each represent 5–20% of the conformations.) Although, a native like β-sheet conformation was obtained in the absence of half the amino acid sequence of platelet factor 4, the authors concluded that folding was thermodynamically linked to subunit association induced hydrophobic collapse which resulted in multiple conformational states. The above examples of peptides which adopt a stable β-sheet structure in aqueous solution have contributed to our understanding of β-sheets. However, a truly well defined acyclic, non-crosslinked system which is amenable to detailed biophysical studies remains elusive.

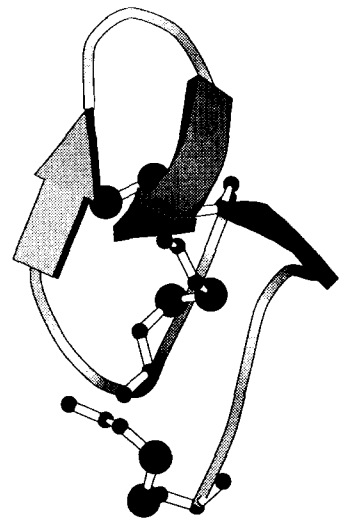


Figure 9. Ribbon diagram of ω-conotoxin. The diagram was created using coordinates (PDB entry 1CCO, 1993) for one of the 24 structures determined by NMR distance restraints and molecular dynamics calculations.[165]

6.2 Peptidomimetics

One of the greatest challenges to understanding β-sheet structure is obtaining a well behaved system which is amenable to detailed study. One approach for achieving a well defined fold in a small peptide is to design the peptide such that its conformational propensities facilitate rapid intramolecular folding. This can be accomplished by introducing a conformationally rigid amino acid or template into the peptide such that a single conformation is stabilized, facilitating formation of the desired conformation or folded structure by reducing the entropic penalty associated with constraining portions of the peptide chain. This concept was first applied to β-sheets by Hirschmann, Veber, Freidinger, Nutt, and their colleagues at Merck with their work on cyclic somatostatin mimics.[168–171] Several β-turn mimetics have since been prepared either to control peptide conformation and/or for medicinal chemistry purposes.[138,139,172–195]

The first example of a template nucleated β-sheet was presented by Kemp and co-workers in their work with peptide-functionalized diacylaminoepindolidiones.[181,183,184,199] The epindolidione nucleus was designed to act as the central strand of a parallel β-sheet by providing three appropriately spaced hydrogen bonding sites on each side of the planar

molecule (Fig. 10a). The planarity of the template molecule prohibits van der Waals interactions with the side chains of the amino acids in the adjoining peptides. Hence, the formation of β-sheet structure in these systems is dependent on effective chain reversal and hydrogen bonding interactions. The importance of sequence requirements in the region of the peptides adjacent to the epindolidione molecule has been demonstrated. In the case of the parallel nucleator, the dipeptide Pro-D-Ala was found to be sufficient for chain reversal. The parallel strand mimic was found to adopt monomeric β-sheet structure in DMSO as evidenced by NMR. Specifically, the temperature dependence of the amide protons, $^3J_{\alpha NH}$ coupling constants and observed NOEs were consistent with a β-sheet structure. Incorporation of urea functionalities resulted in a template for antiparallel β-sheet structure formation (Fig. 10b).[183] These studies suggest that the combination of the proper turn sequence and the strand mimic is sufficient for β-sheet structure formation in DMSO. Recently, the Kemp laboratory has carried out a number of interesting studies on water soluble mimetics of this type which will be published in due course.[200]

This pioneering effort has demonstrated that template assisted β-sheet formation is possible and has provided a foundation for the development of similar strategies. A tolan-based turn mimetic, 2-amino-2′-carboxydiphenylacetylene, was recently reported by the Kemp laboratory to initiate β-sheet structure when incorporated into short peptides as evidenced by CD (EtOH) and NMR data (NOE data, NH-C_αH coupling constants, and amide temperature dependence in DMSO).[201,202]

The Smith and Hirschmann groups have reported the synthesis and characterization of pyrrolinone-based β-strand peptidomimetics.[203] The conformationally restricted enamine-based scaffolding replaced the peptide backbone while appropriately chosen side chains mimicked peptidyl side chains. The trispyrrolinone shown in Figure 11 was designed to mimic the tetrapeptide H-Leu-Leu-Val-Tyr-OMe, a fragment of equine angiotensinogen which crystallizes as a parallel β-sheet. Crystal structures of the mimetic indicate that the BOC protected compound **1** forms interstrand hydrogen bonds and is oriented in an antiparallel β-sheet conformation in the solid state with both the side chain and carbonyl orientations overlapping well with those of the natural peptide. Removal of the BOC group (compound **2**) resulted in a crystal structure in which the intermolecular interactions mimicked those of a parallel β-sheet. Thus, the solid state conformations of the pyrrolinone-based β-strand mimetics suggest that the strand mimetics can adopt β-sheet-like structure in the solid state. The ability of these and similarly designed mono- and bis-pyrrolinones to act as renin and/or HIV-1 protease inhibitors has been evaluated.[204,205] It appears that the 3,5-linked pyrrolinone unit provides an effective, nonpeptidic scaffolding for proteolytic enzyme inhibitors.

The introduction of conformationally rigid templates into cyclic peptides has been addressed by Fiegel and co-workers. In one study, phenoxathiin derivatives **3** and **4** (Fig. 12a) were incorporated into cyclic peptides designed to adopt parallel β-sheet structure (Fig. 12b).[206] NMR data ($^3J_{\alpha NH}$ values and NOEs) in combination with computational studies suggest that peptides **A** and **B** exist in a conformation consistent with parallel β-sheet structure in chloroform solution. Peptides **C** and **D** which contain only one phenoxathiin spacer did not exhibit NMR behavior indicative of β-sheet structure. Compound **5** was synthesized as an antiparallel β-sheet nucleator and incorporated into peptide **E** which was determined by NMR (amide temperature dependence, coupling constants, and NOE connectivities) to exist in a conformation which

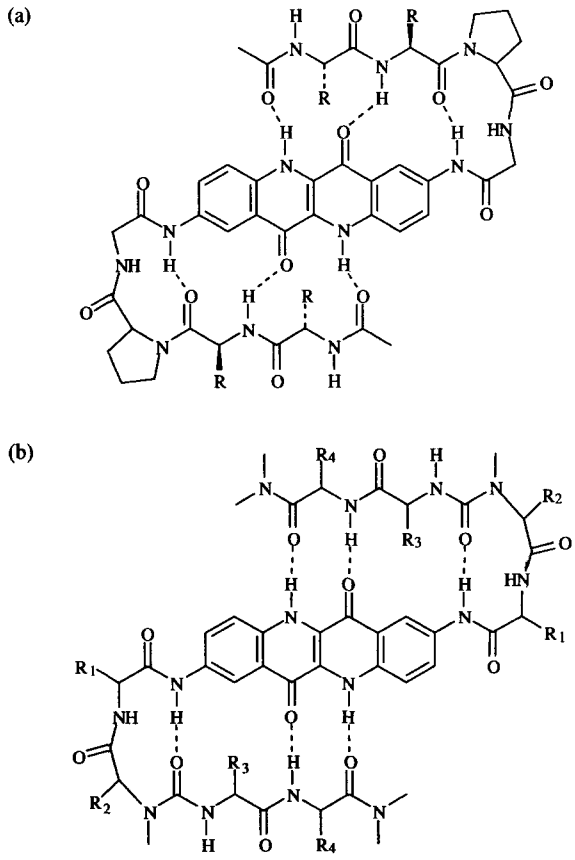
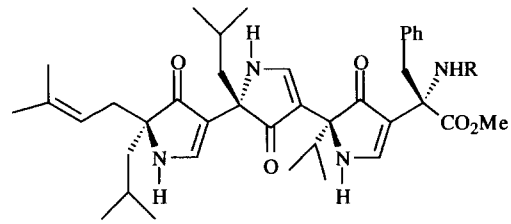


Figure 10. Epindolidione-based template for (a) parallel β-sheet formation and (b) antiparallel β-sheet formation.



1 R = $C(O)OC(CH_3)_3$
2 R = H

Figure 11. 3,5,5-Substituted pyrrolinone-based peptidomimetic designed to mimic a β-strand.

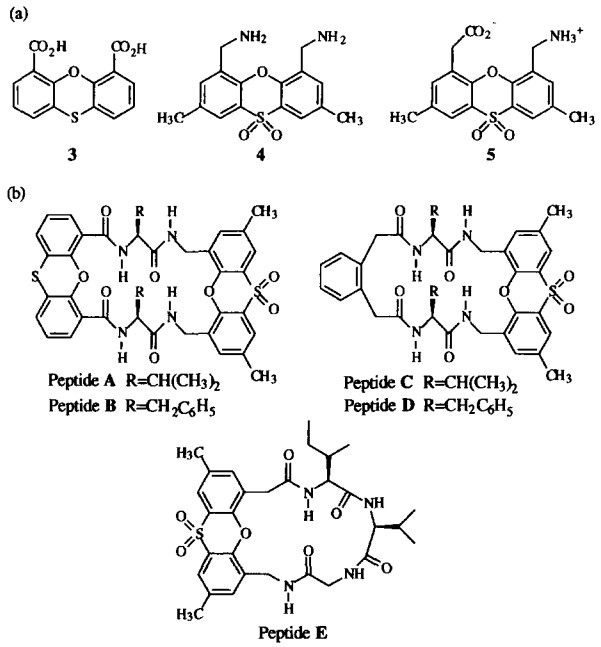


Figure 12. (a) β-Turn mimetics phenoxathin-4,6-dicarboxylic acid (**3**), 2,8-dimethyl-4,6-bis(aminoethyl)phenoxathin-1-,10-dioxide (**4**), and 2,8-dimethyl-4-(carboxymethyl)-6-(aminomethyl)phenoxathin-*S*-dioxide (**5**) and (b) peptides prepared with the phenoxathin nucleators.

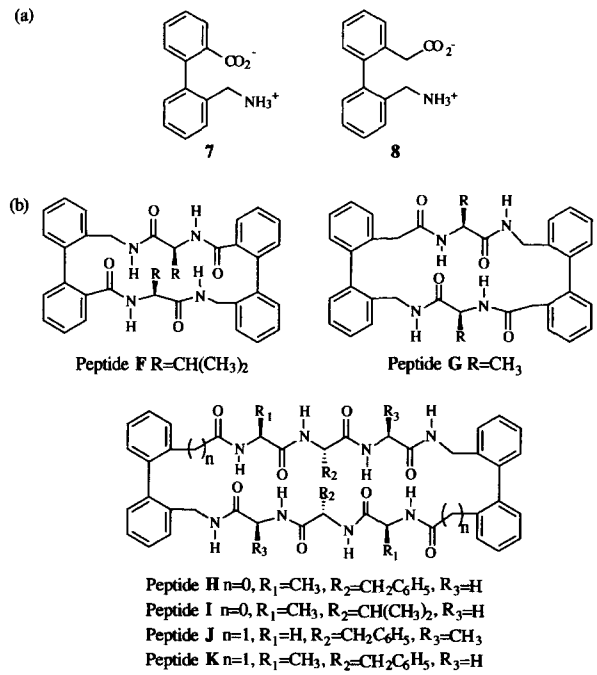


Figure 14. (a) β-Turn mimetics 2′-aminomethylbiphenyl-2-carboxylic acid (**7**) and 2-aminomethyl-2′-carboxymethylbiphenyl (**8**) and (b) peptides prepared with the bipheny-based amino acids.

contained a β-loop.[190] Recently, the hydrogen bonding abilities of a series of 2,8-dimethylphenoxathiin-4,6-dicarboxylic acid derivatives have been investigated by FT-IR and variable temperature NMR.[207] Compound **6** (Fig. 13) adopts a 10-membered hydrogen bonded ring conformation consistent with formation of a parallel β-sheet. Currently, the efficacy of this compound as a β-turn mimetic in protein-based polymers is being investigated.

Two biphenyl-based amino acids, 2′-aminomethylbiphenyl-2-carboxylic acid (**7**) and 2-aminomethyl-2′-carboxymethylbiphenyl (**8**) have also been prepared as β-turn mimetics by the Feigel group (Fig. 14a).[198,208] Compounds **7** and **8** were initially incorporated into cyclic peptides composed of two α-amino acids and two biphenyl-based amino acids (Fig. 14b, peptides **F** and **G**). NMR studies of peptide **G** in DMSO revealed three diastereomers of the peptide and indicated that the *R,S* diastereomer does not adopt a β-sheet conformation while the *R,R* and *S,S* diastereomers exist as an equilibrium between β-sheet and γ-loop containing conformations.[192] More recently, these two biphenyl-based amino acids were incorporated into larger cyclic peptides (**H**–**K**).[209] Again multiple conformations were observed by NMR in DMSO. In the case of the peptides containing **7** (peptides **H** and **I**), a major conformation was present. NMR studies of these peptides indicated that the major conformation was consistent with an extended structure (β-sheet) in which each of the two amino acids flanking the biphenyl-based amino acids were hydrogen-bonded. Additionally, an upfield shift of the methyl group of one of the alanines flanking the biphenyl moieties indicates a hydrophobic interaction between the two groups. While the incorporation of both the phenoxathiin- and biphenyl-based nucleators into cyclic peptides suggests that these compounds may be capable of nucleating β-sheet structure in DMSO, the evidence is limited to NMR data alone. Additional spectroscopic evaluation (e.g. circular dichroism and/or infrared studies) could provide further insight into the conformation of these cyclic peptides. Also, the systems are currently limited to studies in DMSO or chloroform and are complicated by the presence of multiple conformations.

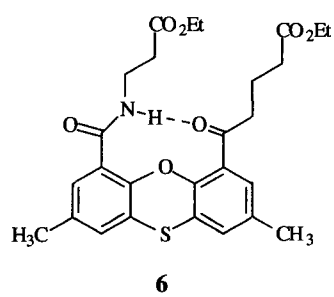
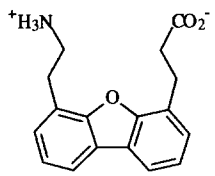


Figure 13. Hydrogen-bonding of a 2,8-dimethylphenoxathiin-4,6-dicarboxylic acid derivative.

Efforts to understand the relative importance of local conformational propensities and hydrophobic clusters in the nucleation of β-sheet folding in the Kelly laboratory have focused on acyclic peptides incorporating a single dibenzofuran-based amino acid, 4-(2′-aminoethyl)-6-dibenzofuranpropionic acid (**9**), which was designed to replace the $i+1$ and $i+2$ residues of a β-turn (Fig. 15).[87,210–212] The conformationally biased,

(a)



9

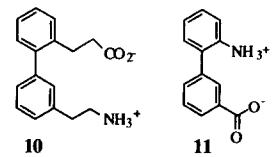
(b) Peptide **L** H_3N-V K L **9** V K L-C(O)NH_2
Peptide **M** H_3N-V K L **9** F K L-C(O)NH_2
Peptide **N** H_3N-V L K **9** K V L-C(O)NH_2
Peptide **O** H_3N-K V K V K V **9** V K VK V K-C(O)NH_2

Figure 15. (a) Dibenzofuran-based β-sheet nucleator (4-(2′-aminoethyl)-6-dibenzofuranpropionic acid (**9**) and (b) peptides prepared with the dibenzofuran-based amino acid.

dibenzofuran-based amino acid **9** was designed to reverse the chain direction of a linear peptide by facilitating hydrogen bonding between the flanking α-amino acids and by facilitating tertiary interactions (hydrophobic cluster formation) between the dibenzofuran skeleton and the flanking hydrophobic α-amino acid side chains. Initial studies of heptameric peptides incorporating **9** (i.e. peptides **L** and **M**) demonstrated that the residue satisfies these design criteria.[87,210–212] The flexibility of the ethylene spacer groups between the dibenzofuran skeleton and the carboxy and amino functionalities allows the amino acid to adopt a conformation in which the dibenzofuran skeleton is oriented perpendicular to the plane of the β-sheet, placing the flanking α-amino acid side chains in close proximity to the aromatic portion of **9**. Hydrophobic cluster conformation has been confirmed by a near-UV CD signal for the dibenzofuran chromophore, a characteristic upfield shift of the NMR signals belonging to the methyl groups of the flanking hydrophobic amino acids, and by NOEs between the aromatic protons and the methyl groups of the side chains of the flanking α-amino acids. The hydrophobic cluster conformation is also observed under conditions where β-sheet structure cannot form, implying that the cluster is poised and waiting to nucleate folding once solution conditions permit strand-strand interactions. Replacing the flanking hydrophobic α-amino acids with hydrophilic residues (as in peptide **N**) has demonstrated that the hydrophobic cluster is required for β-sheet nucleation as this substitution does not allow **9** to nucleate β-sheet formation. Evaluation of two dibenzofuran-based amino acids which lacked the flexibility to adopt a hydrophobic cluster conformation also resulted in unstructured peptides.[87] The dibenzofuran-based heptapeptide **L** is monomeric and exhibits far-UV CD spectra consistent with a mixture of β-sheet structure and random coil. NMR characterization (amide temperature dependence, $^3J_{\alpha NH}$ coupling constants, and NOEs) of peptide **L** indicates a fluctuating β-sheet which is structured proximal to **9** and unordered at the termini of the peptide. This behavior is expected in small, isolated sheets. The circular dichroism spectra of peptide **M** shows that the peptide adopts a predominantly β-sheet structure. In an effort to obtain a well-defined structure, the length of the peptide was increased to 13 residues (peptide **O**).[68] Again a hydrogen-bonded hydrophobic cluster composed of -V-**9**-V- was observed by near-UV CD and NMR at low pH. However, the high charge density of this peptide (+7 at pH 4.5) required an increase in solution pH or ionic strength to facilitate folding. It was determined that the neutralization of two charges resulted in the pH independent folding of peptide **O** by preparing numerous trifluoroacetylated lysine analogues of peptide **O**. These tridecameric peptides undergo intramolecular folding followed by intermolecular association leading to fibril formation. Although all evidence suggests that intramolecular folding precedes self-association, the two events are linked. Recently, self-assembly has been prevented by the incorporation of two N-methylated amino acids in place of selected lysine residues.[213] The N-methylated peptides have been determined to be monomeric by analytical equilibrium ultracentrifugation studies and exhibit far-UV CD spectra indicative of well defined β-sheet structure. Additionally, the $^3J_{\alpha NH}$ coupling constants and amide exchange rate data on an ^{15}N-labeled N-methylated peptide suggest a well defined β-sheet structure. The ability to nucleate β-sheet structure using residue **9** combined with the ability to control self-assembly obtained by the incorporation of N-methylated amino acids should provide a powerful method for evaluating β-sheet folding.

The generality of this β-sheet nucleation approach was demonstrated by the preparation of additional aromatic amino acids which may serve as potential β-sheet nucleators by promoting intramolecular hydrogen bonding and/or hydrophobic cluster formation. The biphenyl-based amino acids, 3′-aminoethyl-2-biphenylpropionic acid (**10**) and 2-amino-3′-biphenylcarboxylic acid (**11**), were designed to nucleate β-sheet structure in aqueous solution when incorporated into amphiphilic peptides in place of the $i+1$ and $i+2$ residues of the β-turn (Fig. 16).[214,215] NMR and IR studies on amide analogues of **10** and **11** demonstrate that **10** and, to a lesser extent, **11** have the ability to support a hydrogen bonding network capable of facilitating β-sheet folding.[214] The preferred conformation of the phenethyl portion of **10** was predicted to facilitate hydrophobic cluster formation involving the biphenyl skeleton and the side chains of the flanking amino acids. An NMR structural evaluation of peptides incorporating **10** (peptides **P** and **Q**) revealed the presence of a hydrophobic cluster involving the aromatic rings of **10** and the side chains of the flanking hydrophobic α-amino acids under conditions where these peptides appear to be unstructured by far UV-CD.[215] In addition, the slower proton/deuterium exchange rates for the amide protons of the α-amino acids flanking **10** indicate the presence of the desired hydrogen bonding scheme in aqueous solution. Incorporation of **10** into a tridecapeptide in which the two amino acids flanking **10** were hydrophobic (peptide **P**) resulted in β-sheet folding and subsequent self-assembly into high mol wt β-sheet fibrils. The incorporation of residue **11** into an identical α-amino acid sequence (peptide **S**) does not result in folding under

(a)



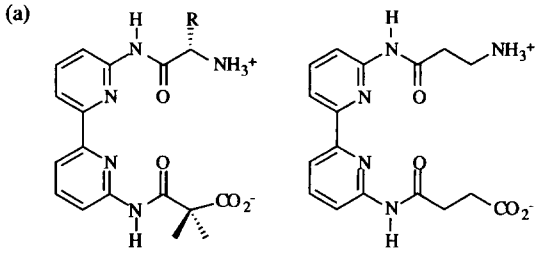
10 11

(b)

Peptide P H_3N-K V K V K V **10** V K V K V K-C(O)NH_2
Peptide Q H_3N-V O L **10** V O L-C(O)NH_2
Peptide R H_3N-V O L **10** F O L-C(O)NH_2
Peptide S H_3N-K V K V K V **11** V K V K V K-C(O)NH_2
Peptide T H_3N-K V (NMe)L V K V **10** V K V (NMe)L V K-C(O)NH_2

Figure 16. (a) Biphenyl-based amino acids 3′-(2-aminoethyl-)-2-biphenylpropionic acid (**10**) and 2-amino-3′-biphenylcarboxylic acid (**11**) and (b) peptides prepared incorporating the biphenyl-based amino acids.

(a)



12a R = H **13**
12b R = isopropyl
12c R = isobutyl

(b)

Peptide U H_3N-K V T V K **12b** K V T V K-C(O)NH_2
Peptide V H_3N-K V T V K V **13** V K V T V K-C(O)NH_2
Peptide W H_3N-K V K V K V **13** V K V K V K-C(O)NH_2
Peptide X H_3N-V O L **13** V O L-C(O)NH_2

Figure 17. 6,6′-Bis(acylamino)-2,2′-bipyridine-based amino acids **12a–c** and **13** and (b) peptides prepared with the bipytidine-based amino acids.

the same conditions, implying that the hydrogen bonded hydrophobic cluster promoted by **10**, but not **11**, is required for β-sheet folding. Unlinking the intramolecular folding equilibria from the self-association equilibria was accomplished by strategically replacing two of the exterior amide protons with methyl groups (peptide **T**). This eliminates their ability to act as hydrogen bond donors and sterically blocks the close approach of the exterior β-strands in two intramolecularly folded β-sheets, blocking self-assembly. Peptide **T** was found to be monomeric by analytical ultracentrifugation studies under conditions in which it exhibited a CD spectrum indicative of a mixture of β-sheet and random coil conformations. The peptide is most likely ordered in the region next to **10** and less ordered at the termini of the strands.

An evaluation of the sheet nucleating efficacy of 6,6′-bis(acylamino)-2,2′-bipyridine-based amino acids has recently been reported (Fig. 17).[216] These amino acids were determined to bind Cu(II) under both alkaline and acidic conditions in a 1:1 stoichiometry. Crystallographic studies of model compounds indicate that the bipyridine rings of peptides containing residues **12** and **13** should exist in the transoid conformation in the absence of Cu(II) ions or other stabilizing forces and the bipyridine rings should adopt a cisoid conformation with near perfect square planar geometry about the copper atom in the presence of Cu(II) (under both alkaline and acidic conditions). Peptide **U**, containing residue **12b** was found to adopt a β-sheet conformation when bound to one equivalent of Cu(II) in aqueous solution at pH 9.5. Residue **13** was capable of promoting β-sheet formation both in the presence and absence of Cu(II) ions (peptides **V** and **W**). This is believed to be due to the ability of **13** to adopt a hydrophobic cluster conformation in the absence of copper binding.

Incorporation of residues **10** and **13** into heptameric peptides analogous to those containing the dibenzofuran-based amino acid (peptides **L** and **M**) resulted in unstructured peptides (peptides **Q**, **R**, and **X**, respectively). This indicates that the hydrophobic cluster conformations formed with the biphenyl- and bipyridine-based amino acids contribute less stability than the hydrophobic cluster formed by the dibenzofuran-based residue which is capable of stabilizing a fluctuating β-sheet structure in heptameric peptides. Both **10** and **13** require additional stabilization from the sequence (i.e. additional length) in order to promote β-sheet structure formation.

The studies described above and studies of a dibenzofuran-based diacid designed to nucleate parallel β-sheet structure[217] have further demonstrated the importance of both local conformational preferences and hydrophobic clusters in the nucleation of β-sheet folding in aqueous solution.

A conformationally restricted dipeptide mimetic has recently been shown to promote β-hairpin formation by the Gellman group.[218] When incorporated into the sequence, AcVal-**X**-Leu-NMe_2 (peptide **Y**), compound **14** (Fig. 18a) was determined by FT-IR and NMR to adopt a compact, intramolecularly hydrogen bonded folding pattern in CH_2Cl_2 (Fig. 18b). The *trans*-5-amino-3,4-dimethylpent-3-enoate was chosen with the expectation that $A^{1,2}$ and $A^{1,3}$ strain would disfavor extended conformations. The less sterically hindered compound, **15**, does in fact show a weaker folding propensity and undergoes aggregation. Thus, it appears that the replacement of a backbone amide group with a tetrasubstituted *trans*-alkene may be useful in designing specifically folded molecules, although this has not yet been evaluated in peptide systems in an aqueous environment.

Nowick et al. have reported the synthesis and study of oligoureas which serve as a molecular scaffolding to orient different groups in a parallel fashion.[197,219] Evaluations of the 1,3-propanediamine- (**16**) and 1,2-ethanediamine-based ureas (**17**) by FT-IR and ^{1}H NMR indicate that the former adopt a well-defined geometry and exhibit substantially more intramolecular hydrogen

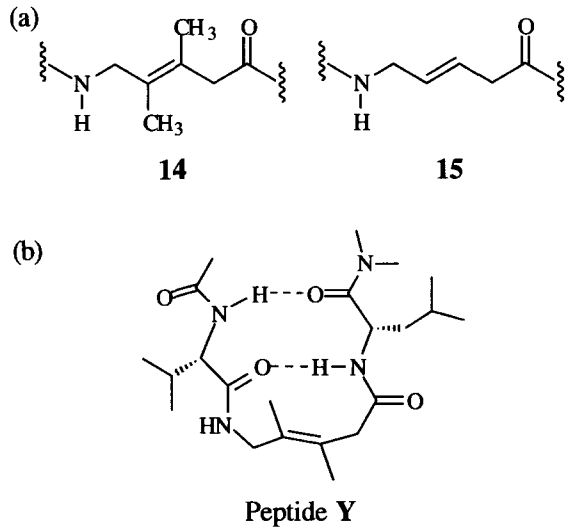


Figure 18. (a) Conformationally restricted dipeptide mimetics *trans*-5-amino-3,4-dimethylpent-3-enoate (**14**) and *trans*-5-amino-pent-3-enoate (**15**) and (b) hydrogen bonding of **14** within peptide **Y**.

bonding than the latter in chloroform solution (Fig. 19a).[219] This is presumably due to the entropic differences of forming the 9- and 10-membered hydrogen bonded rings. The directional control provided by the phenyl substituent on the diamine backbone and the intramolecular hydrogen bonding are thought to promote an ordered structure within the di- and tri-ureas. Preliminary studies are reported to suggest that the same structures are adopted in water and DMSO, suggesting that these structures may be capable of serving as scaffolds from which parallel β-sheet structures could be obtained in aqueous solution. Recent NMR data (NOEs and ^1H chemical shifts) suggest that peptide **Z** which incorporates the 1,2-ethane diamine-based urea (Fig. 19b) adopts a parallel β-sheet structure with a nine-membered hydrogen bonded U-turn in $CDCl_3$.[220]

7. De Novo Designed β-Sheet-Based Proteins

The ability to design proteins that adopt a unique tertiary structure is highly desirable in light of possible applications to biochemical and materials problems. Such a capability would also demonstrate a partial understanding of protein folding processes.[221] The de novo design of α-helical peptides has been quite successful. In comparison, the de novo design of β-sheet proteins has proven much more difficult. However, several de novo designed β-sheet proteins have recently been reported. Although a truly native-like structure has not yet been realized, these results are nonetheless impressive.

A de novo designed α/β barrel protein has been prepared and studied by Martial and co-workers.[222] The protein, octarellin, was based on eight repeating segments of 30 residues and was designed to adopt an eight-stranded parallel α/β barrel. Urea-gradient electrophoresis experiments indicated cooperative unfolding of the protein with a midpoint of ~4 M urea. Initial CD studies indicated that the protein contained an appropriate amount of α-helix structure. Subsequent studies employing FT-IR (amide I′ absorption band) and Raman spectroscopies to evaluate the secondary structure of octarellin suggested that the protein had the expected secondary structure content.[223] However, the low thermal stability of the protein combined with gel filtration experiments which indicate the most compact form of the protein to have an apparent mol. wt of 34,000 (vs 26,000, as expected) and the fact that the tyrosines which should be buried are not well protected suggest that the barrel of octarellin is loosely packed.

Ptitsyn et al. designed the de novo polypeptide albebetin to contain a four-stranded antiparallel

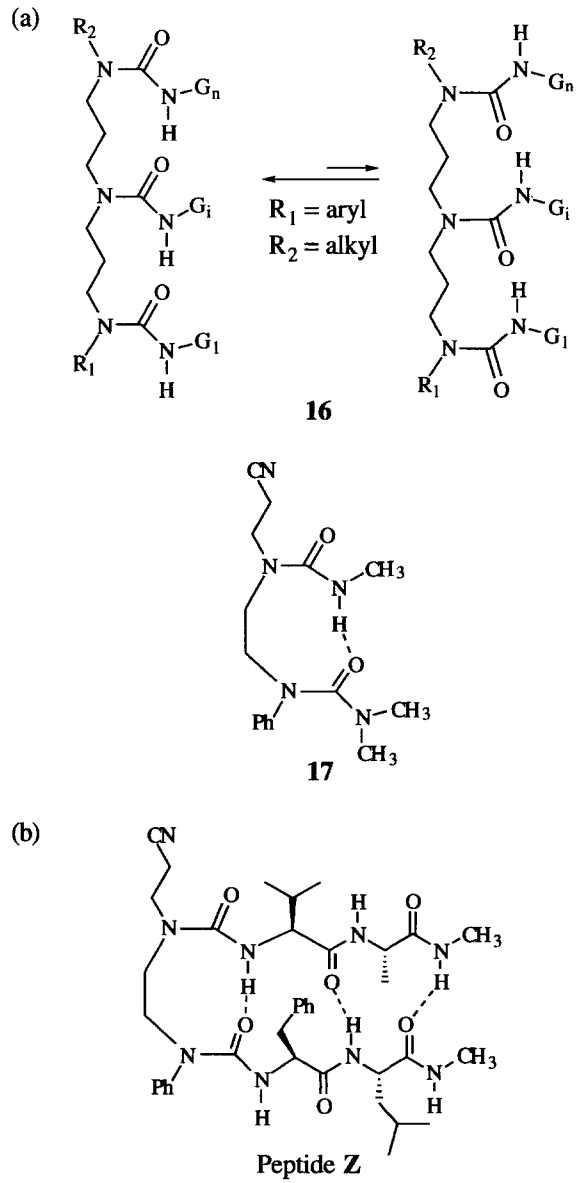


Figure 19. (a) 1,3-Propane diamine- and 1,2-ethanediamine-based ureas (**16** and **17**, respectively) designed as molecular scaffolds for parallel β-sheet formation and (b) a peptide incorporating the 1,2-ethanediamine-based urea.

β-sheet with two α-helices packed against one face of the sheet.[224] The 73 residue peptide was based on a repeating αββ unit and was designed according to several criteria. Hydrophobic packing of the two α-helices against the sheet was favored by appropriately placed hydrophobic amino acid residues in both the β-strand and the helix regions of the peptide, the shortest possible linkages (ProGly and GlyGly) were used between the α and β regions, and the outside β-strands were composed of more polar residues than the internal strands. Additionally, CPK modelling was employed to determine residues with maximal hydrophobic packing capabilities and the external surface was given a high polar character to help prevent aggregation. The polypeptide was encoded from a chemically synthesized gene. Size exclusion chromatography and trypsinolysis studies indicated that the protein was monomeric and nearly as compact as a native globular protein. Urea gradient electrophoresis experiments afforded a cooperative unfolding curve indicative of a stable, compact structure. However, additional studies are needed to determine whether the polypeptide is in a molten globular or a highly structured, native-like state.

The design, synthesis and characterization of betadoublet, a β-sandwich protein similar to betabellin 12[225,226] was recently reported by Richardson and co-workers.[227] The protein consists of two identical, four-stranded antiparallel β-sheets which are covalently linked by a disulfide bond. The internal side chains were selected based on preferences for adopting the β-sheet conformation and their ability to pack tightly within the core of the protein given the closer inter-sheet distance (7 Å) imposed by the disulfide linkage. External side chains were selected based on β-sheet preferences and solubility characteristics. The strands were connected by Type I′ turns containing Asp and Gly at the 2 and 3 positions. After expression from *E. coli* and air oxidation to obtain the disulfide linkage, betadoublet was determined by size exclusion chromatography to be a compact, monomeric structure in aqueous solution. Evaluation of the protein by circular dichroism spectroscopy indicated that the oxidized protein contained β-sheet structure while the reduced betadoublet exhibited a signal indicative of β-sheet character only at high concentrations. Analysis of the laser Raman spectrum of oxidized betadoublet indicated that the protein contained approximately 62% β-sheet and 30% turn character. Thermal denaturation indicated cooperative unfolding with a free energy of 2.5 kcal/mol at 37 °C. This value is low relative to that of native proteins but is in the range of values observed for active destabilized mutant enzymes. ANS binding assays and NMR data (chemical shift dispersion and amide exchange data) indicate that betadoublet adopts a structure which is most likely molten globular in nature. Also of importance, is the water solubility (10 mg/mL) of betadoublet. This protein successfully demonstrates the importance of hydrophobic forces and the distribution of amino acid residues as a dominant force in the folding of a peptide into a state which is close to that of a highly ordered, native protein. However, as the authors point out, the more subtle requirements for further hydrophobic collapse into a well defined, native-like structure have yet to be fully realized.

The design of betabellin 14D is based on its predecessor, betabellin 12, and is a continuation of the search for an ideal β-protein model.[228] The protein was designed to adopt a β-sandwich structure composed of two identical, four-stranded, antiparallel β-sheets. Each sheet is composed of alternating polar and non-polar residues with an overall charge of +10 to improve the water solubility (10 mg/mL) of the dimeric structure. Type I′ β-turn regions containing D-Lys-D-Ala or D-Ala-D-Lys units were incorporated. The polypeptide was synthesized using solid phase peptide synthesis techniques (FMOC chemistry) and the monomers (betabellin 14S) were covalently linked by air-oxidation of the cysteine residues to form the complete protein (betabellin 14D). Circular dichroism studies indicated betabellin 14S to exist as an unordered structure while betabellin 14D exhibited a characteristic β-sheet spectrum. The NMR data on both species confirmed the CD data as the chemical shifts of the aliphatic and aromatic side chains and the amide protons of betabellin 14S were poorly dispersed, indicative of a random coil conformation. Betabellin 14D, however, had well-dispersed aliphatic signals with several methyl resonances shifted upfield of 0.8 ppm and amide and aromatic signals shifted downfield of 8.7 ppm, consistent with a β-sheet structure in aqueous solution. Thermal denaturation of betabellin 14D revealed a reversible unfolding similar to that of a native protein structure (ΔS_m 321 ± 9 cal/mol K and ΔH_m 106 ± 5 kcal/mol). However, the authors note that while the hydrophobic interaction of the nonpolar faces is necessary to stabilize the betabellin structure, it is not sufficient to do so as is evidenced by the lack of the β-sandwich structure in the absence of the disulfide linkage (betabellin 14S).

A small, all β-sheet protein, the minibody, was designed by Sollazzo and co-workers to adopt a sandwich composed of two three-stranded sheets (representative of an immunoglobin fragment of the variable heavy domain of M_cPC603) and to bind metal upon folding.[229] The protein was found to adopt a monomeric, β-sheet structure by size exclusion chromatography and circular dichroism spectroscopy with and without metal. A sharp denaturant-induced unfolding transition was observed up to 4.0 M urea (midpoint 2.15 M). Metal binding studies indicate that metals were bound through interactions with at least one histidine residue from each loop of the protein with a selectivity order of $Cu > Zn \gg Cd > Co$. The metal binding suggests that the loops are within close proximity of each other in the folded protein. Although the limited solubility of the protein prevented more detailed analysis (i.e., NMR studies), the preliminary characterization suggests that a well defined structure has been obtained.

The attempts to create de novo designed proteins which are described above are impressive. The desired

overall structures appear to have been obtained, demonstrating that the ability to predict and design β-sheet structure is possible. Thus far, however, the finer details required to obtain a highly ordered, native-like tertiary structure remain elusive.

8. Conclusions

Significant progress has been made over the past few years regarding an emerging understanding of β-sheet structure. Several new β-folds and topologies have been discovered and there appears to be a general agreement that β-sheet structure has elements of both secondary, tertiary, and, in some cases, quaternary structure. Hydrophobic cluster formation is emerging as a common theme in the folding of predominantly β-sheet structures. Hydrophobic side chain–side chain interactions between residues far apart in sequence appear to play a critical role in the folding and stability of β-sheet structures. A variety of spectroscopic and biophysical techniques have been sufficiently developed to characterize the structures of β-sheets in natural proteins and those from de novo design efforts. Although a well-defined tertiary structural model remains elusive, several designed β-sheets and peptidomimetics offer promising results. Clearly, β-sheet structure has attracted the interest of numerous laboratories whose combined efforts should greatly increase our understanding of this common protein structure in the near future.

Acknowledgements

The authors wish to thank Drs J. F. Sinclair and L. S. Mullins for their generous help in preparing the MOLSCRIPT figures. The finanical support of the Robert A. Welch Foundation, Searle Scholars Program/The Chicago Community Trust (J.W.K.), the National Institutes of Health (Grant R01 GM51105), the Camille and Henry Dreyfus Foundation Teacher Scholars Program (J.W.K.), and the Center for Macromolecular Design is gratefully acknowledged.

References and Notes

1. Astbury, W. T. *Trans. Faraday Soc.* **1933**, *29*, 193.

2. Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729.

3. Richardson, J. S. In *Advances in Protein Chemistry*; Anfinsen, C. B.; Edsall, J. T.; Richards, F. M. Eds.; Academic: New York, 1981; Vol. 34, pp 167–339.

4. Cohen, F. E.; Sternberg, M. J. E.; Taylor, W. R. *J. Mol. Biol.* **1981**, *148*, 253.

5. Bork, P.; Holm, L.; Sander, C. *J. Mol. Biol.* **1994**, *242*, 309.

6. Efimov, A. V. *J. Mol. Biol.* **1995**, *245*, 402.

7. Lesk, A. M.; Brändén, C.-I.; Chothia, C. *Proteins: Struct., Funct., Genet.* **1989**, *5*, 139.

8. Lesk, A. M.; Chothia, C. *J. Mol. Biol.* **1982**, *160*, 325.

9. Woolfson, D. N.; Evans, P. A.; Hutchinson, E. G.; Thornton, J. M. *Protein Engng* **1993**, *6*, 461.

10. Hol, W. G. J.; Halie, L. M.; Sander, C. *Nature* **1981**, *294*, 532.

11. Richardson, J. S. *Nature* **1977**, *268*, 495.

12. Chou, K.-C.; Némethy, G.; Scheraga, H. A. *Biochemistry* **1983**, *22*, 6213.

13. Sternberg, M. J. E.; Thornton, J. M. *J. Mol. Biol.* **1977**, *115*, 1.

14. von Heijne, G.; Blomberg, C. *J. Mol. Biol.* **1977**, *117*, 821.

15. Lifson, S.; Sander, C. In *Molecular Mechanisms of Biological Recognition*; Balaban, M. Ed.; Elsevier: New York, 1979; pp 145–155.

16. Lifson, S.; Sander, C. *J. Mol. Biol.* **1980**, *139*, 627.

17. Wouters, M. A.; Curmi, P. M. G. *Proteins: Struct., Funct., Genet.* **1995**, *22*, 119.

18. Smith, C. K.; Regan, L. *Science* **1995**, *270*, 980.

19. Chothia, C. *J. Mol. Biol.* **1973**, *75*, 295.

20. Creighton, T. E. *Proteins: Structure and Molecular Properties;* 2nd ed.; W. H. Freeman: New York, 1993, p 187.

21. Vtyurin, N. *Proteins: Struct., Funct., Genet.* **1993**, *15*, 62.

22. Chou, K.-C.; Pottle, M.; Némethy, G.; Ueda, Y.; Scheraga, H. A. *J. Mol. Biol.* **1982**, *162*, 89.

23. Chou, K. C.; Neméthy, G.; Scheraga, H. A. *J. Mol. Biol.* **1983**, *168*, 389.

24. Chou, K. C.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 7047.

25. Richardson, J. S.; Getzoff, E. D.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2574.

26. Chan, A. W. E.; Hutchinson, E. G.; Harris, D.; Thornton, J. M. *Protein Sci.* **1993**, *2*, 1574.

27. Daffner, C.; Chelvanayagam, G.; Argos, P. *Protein Sci.* **1994**, *3*, 876.

28. Chothia, C.; Janin, J. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4146.

29. Finkelstein, A. V.; Nakamura, H. *Protein Engng* **1993**, *6*, 367.

30. Chou, K. C.; Némethy, G.; Rumsey, S.; Tuttle, R. W.; Scheraga, H. A. *J. Mol. Biol.* **1986**, *188*, 641.

31. Lasters, I.; Wodak, S. J.; Alard, P.; van Cutsem, E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3338.

32. Chothia, C. *Nature* **1988**, *333*, 598.

33. Sacchettini, J. C.; Gordon, J. I.; Banaszak, L. J. *J. Biol. Chem.* **1988**, *263*, 5815.

34. Liu, Z.-P.; Rizo, J.; Gierasch, L. M. *Biochemistry* **1994**, *33*, 134.

35. Varghese, J. N.; Laver, W. G.; Colman, P. M. *Nature* **1983**, *303*, 35.

36. Crennell, S. J.; Garman, E. F.; Laver, W. G.; Vimr, E. R.; Taylor, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9852.

37. Vellieux, F. M. D.; Huitema, F.; Groendijk, H.; Kalk, K. H.; Jxn, F. J.; Jongejan, J. A.; Duine, J. A.; Petratos, K.; Drenth, J.; Hol, W. G. J. *EMBO J.* **1989**, *8*, 2171.

38. Ito, N.; Phillips, S. E. V.; Stevens, C.; Ogel, Z. B.; McPhearson, M. J.; Keen, J. N.; Yadav, K. D. S.; Knowles, P. F. *Nature* **1991**, *350*, 87.

39. Murzin, A. G. *Proteins: Struct., Funct., Genet.* **1992**, *14*, 191.

40. Xia, Z.; Dai, W.; Xiong, J.; Hao, Z.; Davidson, V. L.; White, S.; Mathews, F. S. *J. Biol. Chem.* **1992**, *267*, 22289.

41. Faber, H. R.; Groom, C. R.; Baker, H. M.; Morgan, W. T.; Smith, A.; Baker, E. N. *Structure* **1995**, *3*, 551.

42. Li, J.; Brick, P.; O'Hare, M. C.; Skarzynski, T.; Lloyd, L. F.; Curry, V. A.; Clark, I. M.; Bigg, H. F.; Hazleman, B. L.; Cawston, T. E.; Blow, D. M. *Structure* **1995**, *3*, 541.

43. Yoder, M. D.; Keen, N. T.; Jurnak, F. *Science* **1993**, *260*, 1503.

44. Yoder, M. D.; Lietzke, S. E.; Jurnak, F. *Structure* **1993**, *1*, 241.

45. Pickersgill, R.; Jenkins, J.; Harris, G.; Nasser, W.; Robert-Baudouy, J. *Struc. Biol.* **1994**, *1*, 717.

46. Jenkins, J.; Nasser, W.; Scott, M.; Pickersgill, R.; Vignon, J.-C.; Robert-Baudouy, J. *J. Mol. Biol.* **1992**, *228*, 1255.

47. Yoder, M. D.; Jurnak, F. *FASEB J.* **1995**, *9*, 335.

48. Chothia, C.; Murzin, A. G. *Structure* **1993**, *1*, 217.

49. Cohen, F. E. *Science* **1993**, *260*, 1444.

50. Sprang, S. R. *Trends Biochem. Sci.* **1993**, *18*, 313.

51. Steinbacher, S.; Seckler, R.; Miller, S.; Steipe, B.; Huber, R.; Reinemer, P. *Science* **1994**, *265*, 383.

52. Goldenberg, D. P.; Creighton, T. E. *Curr. Biol.* **1994**, *4*, 1026.

53. Raetz, C. R. H.; Roderick, S. L. *Science* **1995**, *270*, 997.

54. Baumann, U.; Wu, S.; Flaherty, K. M.; McKay, D. B. *EMBO J.* **1993**, *12*, 3357.

55. Maeda, H. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 3438.

56. Maeda, H.; Ooi, K. *Biopolymers* **1981**, *20*, 1549.

57. Maeda, H.; Gatto, Y.; Ikeda, S. *Macromolecules* **1984**, *17*, 2031.

58. Brack, A.; Orgel, L. E. *Nature* **1975**, *256*, 383.

59. Brack, A.; Caille, A. *Int. J. Peptide Protein Res.* **1978**, *11*, 128.

60. Brack, A.; Spach, G. *J. Am. Chem. Soc.* **1981**, *103*, 6319.

61. Johnson, B. J. *J. Pharm. Sci.* **1974**, *63*, 313.

62. Mattice, W. L.; Lee, E.; Scheraga, H. A. *Can. J. Chem.* **1985**, *63*, 140.

63. Rippon, W. B.; Chen, H. H.; Walton, A. G. *J. Mol. Biol.* **1973**, *75*, 369.

64. Seipke, G.; Arfmann, H. A.; Wagner, K. G. *Biopolymers* **1974**, *13*, 1621.

65. Osterman, D. G.; Kaiser, E. T. *J. Cell. Biochem.* **1985**, *29*, 57.

66. DeGrado, W. F.; Lear, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7684.

67. Rajasekharan Pillai, V. N.; Mütter, M. *Acc. Chem. Res.* **1981**, *14*, 122.

68. Choo, D. W.; Schneider, J. P.; Graciani, N. R.; Kelly, J. W. *Macromolecules* **1996**, *29*, 355.

69. Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3334.

70. Zhang, S.; Lockshin, C.; Cook, R.; Rich, A. *Biopolymers* **1994**, *34*, 663.

71. Krantz, D. D.; Zidovetzki, R.; Kagan, B. L.; Zipursky, S. L. *J. Biol. Chem.* **1991**, *266*, 16801.

72. Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N. *Nature* **1993**, *366*, 324.

73. Ghadiri, M. R.; Granja, J. R.; Buehler, L. K. *Nature* **1994**, *369*, 301.

74. Granja, J. R.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1994**, *116*, 10785.

75. Sun, X.; Lorenzi, G. P. *Helv. Chim. Acta* **1994**, *77*, 1520.

76. Borchard, W. *Progress In Analytical Ultracentriugation*; Steinkopff Springer: Darmstadt, 1991.

77. Bowen, T. J.; Rowe, A. J. *An Introduction to Ultracentrifugation*; Wiley-Interscience: London, 1970.

78. McCall, J. S.; Potter, B. J. *Ultracentrifugation*; Bailliere Tindall: London, 1973.

79. Williams, J. W. *Ultracentrifugation of Macromolecules; Modern Topics*; Academic Press: New York, 1972.

80. Schachman, H. K. *Ultracentrifugation in Biochemistry*; Academic Press: New York, 1973.

81. De Felippis, M.; Alter, L. A.; Pekar, A. H.; Havel, H. A.; Brems, D. N. *Biochemistry* **1993**, *32*, 1555.

82. Harding, S. *Meth. Mol. Biol.* **1994**, *22*, 75.

83. Stankowski, S.; Pawlak, M.; Kaisheva, E.; Robert, C. H.; Schwarz, G. *Biochem. Biophys. Acta* **1991**, *1069*, 77.

84. Manavalan, P.; Momany, F. A. *Biopolymers* **1980**, *19*, 1943.

85. Woody, R. W. *Tetrahedron: Asymmetry* **1993**, *4*, 529.

86. Díaz, H.; Tsang, K. Y.; Choo, D.; Espina, J. R.; Kelly, J. W. *J. Am. Chem. Soc.* **1993**, *115*, 3790.

87. Tsang, K. Y.; Díaz, H.; Graciani, N.; Kelly, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 3988.

88. Lockhart, D. J.; Kim, P. *Science* **1993**, *260*, 198.

89. Osterhout Jr., J. J.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059.

90. Robertson, A. D.; Purisima, E. O.; Eastman, M. A.; Scheraga, H. A. *Biochemistry* **1989**, *28*, 5930.

91. Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519.

92. Williamson, M. P.; Waltho, J. P. *Chem. Soc. Rev.* **1992**, 227.

93. Kline, A. D.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1986**, *189*, 377.

94. Kline, A. D.; Wüthrich, K. *J. Mol. Biol.* **1986**, *192*, 869.

95. Kline, A. D.; Wüthrich, K. *J. Mol. Biol.* **1985**, *183*, 503.

96. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley Interscience: New York, 1986; Chaps 7–10.

97. Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741.

98. Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647.

99. Sanders, J. K. M.; Hunter, B. K. *Modern NMR Spectroscopy—A Guide for Chemists;* 2nd ed.; Oxford University Press: New York, 1993; Chap. 2.

100. Montelione, G. T.; Wüthrich, K.; Nice, E. C.; Burgess, A. W.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8594.

101. Anteunis, M. J. O.; Callens, R. E. A.; Tavernier, D. K. *Eur. J. Biochem.* **1975**, *58*, 259.

102. Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161.

103. Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. *J. Am. Chem. Soc.* **1992**, *114*, 3182.

104. Elove, G.; Roder, H. *Structure and Stability of Cytochrome c Folding Intermediates*; American Chemical Society: Washington DC, 1991; pp 51–63.

105. Haris, P. I.; Chapman, D. *Biopolymers* **1995**, *37*, 251.

106. Byler, D. M.; Susi, H. *Biopolymers* **1986**, *25*, 469.

107. Byler, D. M.; van Gusteren, W. F. *FEBS Lett.* **1993**, *323*, 215.

108. Surewicz, W. K.; Mantsch, H. H. *Biochim. Biophys. Acta* **1988**, *952*, 115.

109. Surewicz, W. K.; Mantsch, H. H.; Chapman, D. *Biochemistry* **1993**, *32*, 389.

110. Dong, A.; Caughey, W. S. *Biochemistry* **1990**, *29*, 3303.

111. Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181.

112. Takeda, N.; Kato, M.; Taniguchi, Y. *Biochemistry* **1995**, *34*, 5980.

113. Kauppinen, J. K.; Moffatt, D. J.; Mantsch, H. H. *Appl. Spectrosc.* **1981**, *35*, 271.

114. Cameron, D. G.; Moffatt, D. J. *Appl. Spectros.* **1987**, *41*, 539.

115. Savitzky, A.; Golay, J. E. *Anal. Chem.* **1964**, *36*, 1627.

116. Dill, K. A. *Biochemistry* **1990**, *29*, 7133.

117. Mattice, W. L. *Ann. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 93.

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

119. Dyson, H. J.; Sayre, J. R.; Merutka, G.; Shin, H. C.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 819.

120. Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795.

121. Finkelstein, A. V. *Proteins: Struct., Funct., Genet.* **1991**, *9*, 23.

122. Radford, S. E.; Dobson, C. M.; Evans, P. A. *Nature* **1992**, *358*, 302.

123. Kotik, M.; Radford, S. E.; Dobson, C. M. *Biochemistry* **1995**, *34*, 1714.

124. Cox, J. P. L.; Evans, P. A.; Packman, L. C.; Williams, D. H.; Woolfson, D. N. *J. Mol. Biol.* **1993**, *234*, 483.

125. Viguera, A. R.; Martínez, J. C.; Filimonov, V. V.; Mateo, P. L.; Serrano, L. *Biochemistry* **1994**, *33*, 2142.

126. Rudolph, R.; Siebendritt, R.; Nesslauer, G.; Sharma, A. K.; Jaenicke, R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 4625.

127. Zhuang, P.; Eisenstein, E.; Howell, E. E. *Biochemistry* **1994**, *33*, 4237.

128. Ropson, I. J.; Gordon, J. I.; Freiden, C. *Biochemistry* **1990**, *29*, 9591.

129. Ropson, I. J.; Frieden, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7222.

130. Varley, P.; Gronenborn, A. M.; Christensen, H.; Wingfield, P. T.; Pain, R. H.; Clore, G. M. *Science* **1993**, *260*, 1110.

131. Hazes, B.; Hol, W. G. J. *Proteins: Struct., Funct., Genet.* **1992**, *12*, 278.

132. Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.

133. Rose, G. D.; Roy, S. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4643.

134. Neri, D.; Billeter, M.; Wider, G.; Wüthrich, K. *Science* **1992**, *257*, 1559.

135. Topping, K. D.; Evans, P. A.; Dobson, C. M. *Proteins: Struct., Funct., Genet.* **1991**, *9*, 246.

136. Garvey, E. P.; Swank, J.; Matthews, C. R. *Proteins: Struct., Funct., Genet.* **1989**, *6*, 259.

137. Lumb, K. J.; Kim, P. S. *J. Mol. Biol.* **1994**, *236*, 412.

138. Hirschmann, R. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1278.

139. Wiley, R. A.; Rich, D. H. *Med. Res. Revs* **1993**, *13*, 327.

140. Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510.

141. Minor, D. L., Jr.; Kim, P. S. *Nature* **1994**, *367*, 660.

142. Minor, D. L., Jr.; Kim, P. S. *Nature* **1994**, *371*, 264.

143. Kim, C. A.; Berg, J. M. *Nature* **1993**, *362*, 267.

144. Otzen, D. E.; Fersht, A. R. *Biochemistry* **1995**, *34*, 5718.

145. Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211.

146. Chou, P. Y.; Fasman, G. D. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 45.

147. McGregor, M. J.; Islam, S. A.; Sternberg, M. J. E. *J. Mol. Biol.* **1987**, *198*, 295.

148. Fauchere, J.-L.; Pliska, V. *Eur. J. Med. Chem.-Chim. Ther.* **1983**, *18*, 369.

149. Livingstone, J. R.; Spolar, R. S.; Record, M. T. J. *Biochemistry* **1991**, *30*, 4237.

150. Munoz, V.; Serrano, L. *Proteins: Struct., Funct., Genet.* **1994**, *20*, 301.

151. Avbelj, F.; Moult, J. *Biochemistry* **1995**, *34*, 755.

152. Bai, Y.; Englander, S. W. *Proteins: Struct., Funct., Genet.* **1994**, *18*, 262.

153. Xiong, H. Y.; Buckwalter, B. L.; Shieh, H.-M.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6349.

154. Kwon, D. Y.; Kim, P. S. *Eur. J. Biochem.* **1994**, *223*, 631.

155. Mayo, K. H.; Yang, Y.; Daly, T. J.; Barry, J. K.; La Rosa, G. J. *Biochem. J.* **1994**, *304*, 371.

156. Oas, T. G.; Kim, P. S. *Nature* **1988**, *336*, 42.

157. Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635.

158. Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809.

159. Skelton, N. J.; Harding, M. M.; Mortishire-Smith, R. J.; Rahman, S. K.; Williams, D. H.; Rance, M. J.; Ruddock, J. C. *J. Am. Chem. Soc.* **1991**, *113*, 7522.

160. Kullmann, W. *J. Med. Chem.* **1984**, *27*, 106.

161. Blanco, F. J.; Rivas, G.; Serrano, L. *Structural Biol.* **1994**, *1*, 584.

162. Blanco, F. J.; Jiménez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887.

163. Mer, G.; Kellenberger, C.; Koehl, P.; Stote, R.; Sorokine, O.; Van Dorsselaer, A.; Luu, B.; Hietter, H.; Lefèvre, J.-F. *Biochemistry* **1994**, *33*, 15397.

164. Davis, J. H.; Bradley, E. K.; Miljanich, G. P.; Nadasdi, L.; Ramachandran, J.; Basus, V. J. *Biochemistry* **1993**, *32*, 7396.

165. Pallaghy, P. K.; Duggan, B. M.; Pennington, M. W.; Norton, R. S. *J. Mol. Biol.* **1993**, *234*, 405.

166. Sevilla, P.; Bruix, M.; Santoro, J.; Gago, F.; García, A. G.; Rico, M. *Biochem. & Biophys. Res. Comm.* **1993**, *192*, 1238.

167. Ilyina, E.; Mayö, K. H. *Biochem. J.* **1995**, *306*, 407.

168. Veber, D. F.; Strachan, R. G.; Bergstrand, S. J.; Holly, F. W.; Homnick, C. F.; Hirschmann, R. J. *J. Am. Chem. Soc.* **1976**, *98*, 2367.

169. Veber, D. F.; Holly, F. W.; Paleveda, W. J.; Nutt, R. F.; Bergstrand, S. J.; Torchiana, M.; Glitzer, M. S.; Saperstein, R.; Hirschmann, R. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2636.

170. Veber, D. F.; Saperstein, R.; Nutt, R. F.; Freidinger, R. M.; Brady, S. F.; Curley, P.; Perlow, D. S.; Palveda, W. J.; Colton, C. D.; Zacchei, A. G.; Toco, D. J.; Hoff, D. R.; Vandlen, R. L.; Gerich, J. E.; Hall, L.; Mandarino, L.; Cordes, E. H.; Anderson, P. S.; Hirschmann, R. *Life Sci.* **1984**, *34*, 1371.

171. Freidinger, R. M.; Verber, D. F.; Schwenk Perlow, D. *Science* **1980**, *210*, 656.

172. Hölzeman, G. *Kontalte (Darmstadt)* **1991**, *1*, 3.

173. Hölzeman, G. *Kontalte (Darmstadt)* **1991**, *2*, 55.

174. Olson, G. L.; Voss, M. E.; Hill, D. E.; Kahn, M.; Madison, V. S.; Cook, C. M. *J. Am. Chem. Soc.* **1990**, *112*, 323.

175. Olson, G. L.; Bolin, D. R.; Bonner, M. P.; Bös, M.; Cook, C. M.; Fry, D. C.; Graves, B. J.; Hatada, M.; Hill, D. E.; Kahn, M.; Madison, V. S.; Rusiecki, V. K.; Sarabu, R.; Sepinwall, J.; Vincent, G. P.; Voss, M. E. *J. Med. Chem.* **1993**, *36*, 3039.

176. Freidinger, R. M. *Trends Pharmacol. Sci.* **1989**, *10*, 270.

177. Ernest, I.; Kalvoda, J.; Rihs, G.; Mutter, M. *Tetrahedron Lett.* **1990**, *31*, 4011.

178. Kemp, D. S.; Sun, E. T, *Tetrahedron Lett.* **1982**, *23*, 3759.

179. Kemp, D. S.; McNamara, P. E. *J. Org. Chem* **1984**, *49*, 2286.

180. Kemp, D. S.; McNamara, P. E. *J. Org. Chem.* **1985**, *50*, 5834.

181. Kemp, D. S.; Bowen, B. R. *Tetrahedron Lett.* **1988**, *29*, 5077.

182. Kemp, D. S.; Stites, W. E. *Tetrahedron Lett.* **1988**, *29*, 5057.

183. Kemp, D. S.; Bowen, B. R.; Muendel, C. C. *J. Org. Chem.* **1990**, *55*, 4650.

184. Kemp, D. S. *Trends Biotechnol.* **1990**, *8*, 249.

185. Kahn, M.; Devens, B. *Tetrahedron Lett.* **1986**, *27*, 4841.

186. Kahn, M.; Chen, B. *Tetrahedron Lett.* **1987**, *28*, 1623.

187. Kahn, M.; Wilke, S.; Chen, B.; Fujita, K. *J. Am. Chem. Soc.* **1988**, *110*, 1638.

188. Kahn, M.; Bertenshaw, S. *Tetrahedron Lett.* **1989**, *30*, 2317.

189. Kahn, M. *Synlett* **1993**, 821.

190. Feigel, M. *J. Am. Chem. Soc.* **1986**, *108*, 181.

191. Feigel, M. *Liebigs Ann. Chem.* **1989**, 459.

192. Brandmeier, V.; Feigel, M. *Tetrahedron* **1989**, *45*, 1365.

193. Nagai, U.; Sato, K. *Tetrahedron Lett.* **1985**, *26*, 647.

194. Krstenansky, J. L.; Baranowski, R. L.; Currie, B. L. *Biochem. Biophys. Res. Commun.* **1982**, *109*, 1368.

195. Hinds, M. G.; Richards, N. G. J.; Robinson, J. A. *J. Chem. Soc., Chem. Commun.* **1988**, 1447.

196. Sato, K.; Nagai, U. *J. Chem. Soc., Perkin. Trans 1* **1986**, 1231.

197. Nowick, J. S.; Powell, N. A.; Martinez, E. J.; Smith, E. M.; Noronha, G. *J. Org. Chem.* **1992**, *57*, 3763.

198. Nakanishi, H.; Chrusciel, R. A.; Shen, R.; Bertenshaw, S.; Johnson, M. E.; Rydel, T. J.; Tulinsky, A.; Kahn, M. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1705.

199. Kemp, D. S.; Bowen, B. R. *Tetrahedron Lett.* **1988**, *29*, 5081.

200. Kemp, D. S. Personal communication, 1995.

201. Kemp, D. S.; Li, Z. Q. *Tetrahedron Lett.* **1995**, *36*, 4175.

202. Kemp, D. S.; Li, Z. Q. *Tetrahedron Lett.* **1995**, *36*, 4179.

203. Smith III, A. B.; Guzman, M. C.; Sprengeler, P. A.; Keenan, T. P.; Holcomb, R. C.; Wood, J. L.; Carroll, P. J.; Hirschmann, R. *J. Am. Chem. Soc.* **1994**, *116*, 9947.

204. Smith III, A. B.; Hirshmann, R.; Pasternak, A.; Akaishi, R.; Guzman, M. C.; Jones, D. R.; Keenan, T. P.; Sprengeler, P. A.; Darke, P. L.; Emini, E. A.; Holloway, M. K.; Schleif, W. A. *J. Med. Chem.* **1994**, *37*, 215.

205. Smith III, A. B.; Akaishi, R.; Keenan, T. P.; Guzman, M. C.; Holcomb, R. C.; Sprengeler, P. A.; Wood, J. L.; Hirshman, R.; Holloway, M. K. *Biopolymers* **1995**, *37*, 29.

206. Wagner, G.; Feigel, M. *Tetrahedron* **1993**, *49*, 10831.

207. Winningham, M. J.; Sogah, D. Y. *J. Am. Chem. Soc.* **1994**, *116*, 11173.

208. Brandmeier, V.; Feigel, M.; Bremer, M. *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 486.

209. Brandmeier, V.; Sauer, W. H. B.; Feigel, M. *Helv. Chim. Acta* **1994**, *77*, 70.

210. Díaz, H.; Tsang, K. Y.; Choo, D.; Kelly, J. W. *Tetrahedron* **1993**, *49*, 3533.

211. Díaz, H.; Espina, J. R.; Kelly, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 8316.

212. Díaz, H.; Kelly, J. W. *Tetrahedron Lett.* **1991**, *32*, 5725.

213. Choo, D. W.; Fiori, W. R.; Kelly, J. W. Unpublished results, 1996.

214. Nesloney, C. L.; Kelly, J. W. *J. Org. Chem.* **1996**, in press.

215. Nesloney, C. L.; Kelly, J. W. Unpublished results, 1996.

216. Schneider, J. P.; Kelly, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 2533.

217. Chitnumsub, P.; Kelly, J. W. Unpublished results, 1996.

218. Gardner, R. R.; Liang, G.-B.; Gellman, S. H. *J. Am. Chem. Soc.* **1995**, *117*, 3280.

219. Nowick, J. S.; Abdi, M.; Bellamo, K. A.; Love, J. A.; Martinez, E. J.; Noronha, G.; Smith, E. S.; Ziller, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 89.

220. Nowick, J. S.; Smith, E. M.; Noronha, G. *J. Org. Chem.* **1995**, *60*, 7386.

221. Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8729.

222. Goraj, K.; Renard, A.; Martial, J. A. *Protein Engng* **1990**, *3*, 259.

223. Beauregard, M.; Goraj, K.; Goffin, V.; Heremans, K.; Goormaghtigh, E.; Ruysschaert, J. M.; Martial, J. A. *Protein Engng* **1991**, *4*, 745.

224. Fedorov, A. N.; Dolgikh, D. A.; Chemeris, V. V.; Chernov, B. K.; Finklestein, A. V.; Schulga, A. A.; Alakhov, Y. B.; Kirpichnikov, M. P.; Ptitsyn, O. B. *J. Mol. Biol.* **1992**, *225*, 927.

225. McClain, R. D.; Yan, Y.; Williams, R. W.; Donlan, M. E.; Erickson, B. W. In *Peptides: Chemistry and Biology*; Smith, J. A.; Rivier, J. E. Eds; ESCOM: Leiden, The Netherlands, 1992; pp 364–365.

226. Richardson, J. S.; Richardson, D. C.; Tweedy, N. B.; Gernert, K. M.; Quinn, T. P.; Hecht, M. H.; Erickson, B. W.; Yan, Y.; McClain, R. D.; Donlan, M. E.; Surles, M. C. *Biophys. J.* **1992**, *63*, 1186.

227. 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.

228. Yan, Y.; Erickson, B. W. *Protein Sci.* **1994**, *3*, 1069.

229. Pessi, A.; Bianchi, E.; Crameri, A.; Ventuini, S.; Tramontano, A.; Sollazzo, M. *Nature* **1993**, *362*, 367.

230. All ribon diagrams were created using coordinates from the Protein Data Bank (PDB) at Brookhaven National Laboratory[231] and the program MOLSCRIPT.[232]

231. Blake, C. C. F.; Geisow, M. J.; Oatley, S. J.; Rerat, B.; Rerat, C. *J. Mol. Biol.* **1978**, *121*, 339.

232. Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.

233. Magnus, K. A.; Hazes, B.; Ton-That, H.; Bonaventura, C.; Bonaventura, J.; Hol, W. G. J. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 302.

234. Watt, W.; Tulinsky, A.; Swenson, R. P.; Watenpaugh, K. D. *J. Mol. Biol.* **1991**, *218*, 195.

235. Edmundson, A. B.; Harris, D. L.; Fan, Z.-C.; Guddat, L. W.; Schley, B. T.; Hanson, G.; Tribick, G.; Geysen, H. M. *Proteins: Struct., Funct., Genet.* **1993**, *16*, 246.

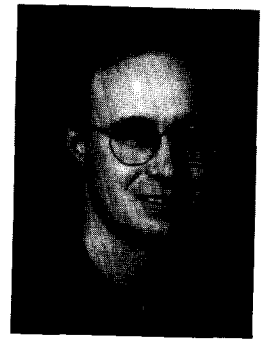
236. Banner, D. W.; Bloomer, A. C.; Petsko, G. A.; Phillips, D. C.; Wilson, I. A. *Biochem. Biophys. Res. Commun.* **1976**, *72*, 146.

237. Scapin, G.; Gordon, J. I.; Sacchettini, J. C. *J. Biol. Chem.* **1992**, *267*, 4253.

(*Received in U.S.A.* 30 *November* 1995; *accepted* 21 *February* 1996)

Biographical Sketch

Carey L. Nesloney is a postdoctoral research assistant at Texas A&M University. Carey was born in San Antonio, Texas, and received her B.S. degree from Trinity University, where she was recognized with the William Crews McGavock Award for Outstanding Research. Carey completed her Ph.D. at Texas A&M University in August 1995. She will be continuing her interest in bioorganic chemistry and organic synthesis as a postdoctoral fellow in the Center for Organic and Medicinal Chemistry at the Research Triangle Institute.



Jeffery W. Kelly is an Associate Professor of Chemistry at Texas A&M University. He was born in Medina, New York, and received a B.S. degree in 1982 from the State University of New York, College at Fredonia, and a Ph.D. from the University of North Carolina in 1986. Jeff then spent three years at the Rockefeller University, where he was a NIH postdoctoral fellow with the late E. T. Kaiser. He then moved to Texas A&M University in 1989. During his tenure at A&M, he has been recognized as a Searle Scholar and a Camille and Henry Dreyfus Scholar. His research interests are centered in bioorganic chemistry and biochemistry of peptides and proteins and include the design and synthesis of unnatural amino acids to control folding in small peptides, elucidation of the biophysical mechanism(s) of human amyloid disease, and the development of chimeric self-assembling peptide materials.