What I have learned by using chemical model systems to study biomolecular structure and interactions

James S. Nowick*

Received 23rd June 2006, Accepted 4th September 2006
First published as an Advance Article on the web 18th September 2006
DOI: 10.1039/b608953b

Chemical model systems provide valuable insights into biomolecular structure and interactions by allowing researchers to simplify, isolate, and manipulate aspects of the complex molecular machinery of living systems. This perspective describes my laboratory's design, synthesis, and study of chemical model systems that fold and self-assemble like proteins and elucidates the insights that have come from studying these systems. Many of these studies have focused on protein β-sheets, which exhibit fascinating intra- and intermolecular interactions and play important roles in protein folding, aggregation, and molecular recognition.

Introduction

Molecular structure and molecular interaction have long been my passion. The ability of complex carbon-containing molecules to adopt beautiful three-dimensional shapes and to interact specifically through non-covalent interactions is the underlying basis of much of both biology and organic chemistry. I am hardly alone in my fascination, which runs as a dominant theme in organic chemistry that has captivated greats such as Kekulé, van't Hoff, and Le Bel in the nineteenth century, and Pauling, Watson and Crick, Cram, and Lehn in the twentieth, to name only a few. In this *perspective*, I aim to share what I have learned thus far in my journey through organic, supramolecular, and bioorganic chemistry that began twenty-seven years ago and continues to thrill and delight me. I have chosen to emphasize themes that have emerged across the various research studies in my laboratory, as these have been most instructive to me and to my students. This *perspective* provides a unique opportunity to describe these themes, many of which may not be apparent from reading the individual papers that my students and I have published. These observations transcend the particular problems and topics that my students and I have studied and are equally relevant to other problems and topics in bioorganic and supramolecular chemistry.

The development of molecular scaffolds: the problems of solubility and synthesis, the importance of intramolecular hydrogen bonding, and a bit of personal history

When I began my independent faculty career at the University of California, Irvine (UCI) in 1991, I decided to explore molecular structure and interactions in the context of peptides and nucleic acids because of the inherent importance of these biomolecules. One of my initial projects focused on the development of unnatural oligomers that fold into well-defined three-dimensional structures. In this project, undergraduate students Noel Powell and Eduardo Martinez and postdoctoral student Glenn Noronha combined urea-based molecular templates with other structural elements to form organized, intramolecularly hydrogen-bonded structures. Our first paper in this area, published in 1992, describes an

Department of Chemistry, University of California, Irvine, Irvine, CA 92697-2025, USA. E-mail: jsnowick@uci.edu; Fax: +1 949-824-9920; Tel: +1 949-824-6091

James S. Nowick received his AB degree in Chemistry in 1985 from Columbia University and his PhD degree in Organic Chemistry in 1990 from MIT, where he was both an NSF Graduate Fellow and an ACS Division of Organic Chemistry Graduate Fellow. After an NSF postdoctoral fellowship in supramolecular chemistry at MIT, he began his independent career as an Assistant Professor at UCI in 1991. He was promoted to Associate Professor in 1996 and Professor in 1998. Professor Nowick has received a Camille and Henry Dreyfus Foundation New Faculty Award (1991), an American Cancer Society Junior Faculty Research Award (1992), an NSF Young Investigator Award (1992), an Arnold and Mabel Beckman Foundation Young Investigator Award (1994), a Presidential Faculty Fellow Award (1995), a Camille Dreyfus Teacher-Scholar Award (1996), an Alfred P. Sloan Research Fellowship (1997), and an American Chemical Society Arthur C. Cope Scholar Award (1998). For his contributions to research and education at UCI, he has received the Award for Outstanding Faculty Contribution to Undergraduate Research (1995), the Chancellor's Award for Excellence in Undergraduate Research (1997), and the School of Physical Sciences Award for Outstanding Contributions to Undergraduate Education (1999).

James S. Nowick

intramolecularly hydrogen-bonded oligourea *molecular scaffold* that presents a series of amino acids or other groups in an orderly fashion.[1] Fig. 1 illustrates the structure. This paper provides one of the first examples of a class of molecules that through the vision and elegant research efforts of Sam Gellman have become known as "foldamers".[2,3]

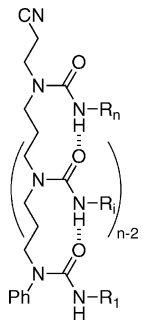
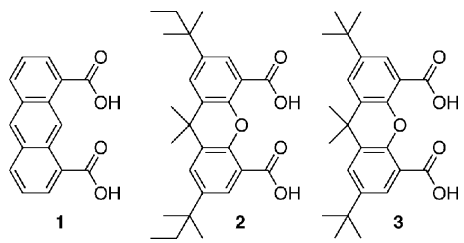


Fig. 1 An oligourea *molecular scaffold.*[1]

The idea behind this project came from unsolved problems that arose during my studies as an NSF postdoctoral fellow in the laboratories of Professor Julius Rebek, who was then at MIT. I had set out to make a U-shaped building block that pointed two carboxylic acid groups in the same direction. I initially examined and rejected anthracene-1,8-dicarboxylic acid (**1**), because the compound and its derivatives exhibited terrible solubility in organic solvents and could not be monofunctionalized selectively.[4] A solution to this problem emerged from a discussion with David Buckler, a graduate student in the laboratories of Professor Daniel Kemp, who was developing xanthene-based templates for thiol-capture based approaches to the segment-coupling of peptides.[5] On the basis of his ease of synthesizing and handling various xanthene derivatives, I decided to pursue derivatives of 9,9-dimethylxanthene-4,5-dicarboxylic acid as a substitute for anthracene-1,8-dicarboxylic acid and to introduce *tert*-alkyl groups at the 2- and 7-positions to augment solubility in organic solvents.[6] Not only did the resulting 2,7-di-*tert*-amyl- and 2,7-di-*tert*-butyl-9,9-dimethylxanthene-4,5-dicarboxylic acids (**2** and **3**) exhibit good solubility in organic solvents, but they also proved remarkably easy to selectively monofunctionalize, through either esterification of the diacid to form a monoester or treatment of a diester with anhydrous hydrogen bromide or iodide to form the corresponding monoester.



These reactions are selective, because the two carboxyl groups communicate through intramolecular hydrogen bonding. The dicarboxylic acid adopts a conformation in which one of the carboxyl groups assumes an intramolecularly hydrogen-bonded s-*cis* conformation. Selective monoesterification occurs because the OH of the second acid group is stabilized by intramolecular hydrogen bonding; selective cleavage of the diester occurs because a single proton can coordinate to the two carboxyl groups of the diester. Fig. 2 illustrates the intramolecular hydrogen bonding in this system.

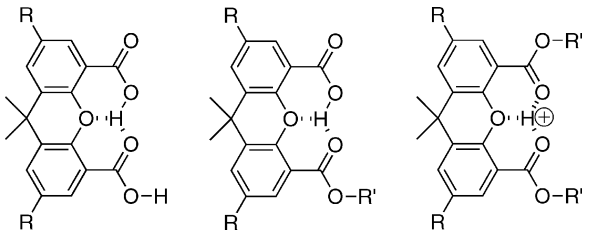
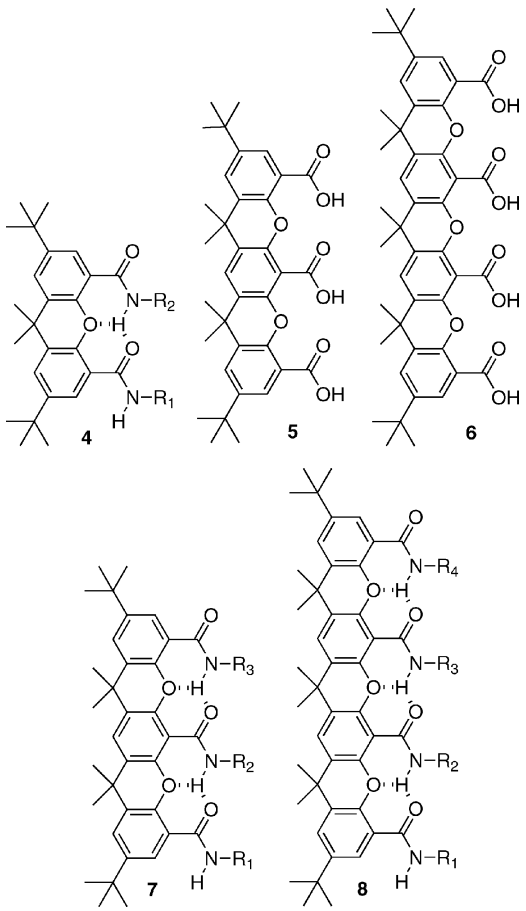


Fig. 2 Intramolecular hydrogen bonding in the 9,9-dimethylxanthene-4, 5-dicarboxylic acid system.

The effect of the interaction of the two carboxyl groups on the conformation and reactivity of the xanthenedicarboxylic acid system captivated my imagination. Mono- and diamide derivatives were also organized through intramolecular hydrogen bonding. I envisioned the diacid as a sort of scaffold that could hold two substituents in well-defined orientations, like the bases in DNA. Structure **4** illustrates this notion. I also envisioned larger homologues, such as triacid **5** and (perhaps somewhat fancifully) tetra-acid **6**, that might serve as building blocks for the creation of larger structures, such as triamides **7** and tetra-amides **8**, that also could orient groups in sequence.



Fancy quickly gave way to reality, as even triacid **5** proved too challenging to synthesize with reasonable effort. A better

approach was clearly needed. Upon beginning my studies at UCI, I envisioned retaining the organization provided by intramolecular hydrogen bonding, while replacing the polycyclic ring system with a linear chain. For structure and ease of synthesis, the carboxylic acid groups were replaced by urea groups. For stability, the oxygen atoms in the linear chain were replaced by methylene groups. A phenyl group was incorporated at the "bottom" of the molecular scaffold to control the rotational orientation of the urea groups. This feature, the alignment of the carbonyl groups to point just "upward," instead of either "upward" or "downward", could not be achieved in the original xanthene system, and was inspired by a paper by Shudo and coworkers describing the effect of *N*-alkyl groups on the conformation of benzanilides.[7] Fig. 3 illustrates my thought processes in designing triurea derivatives as surrogates for triamides **7**.

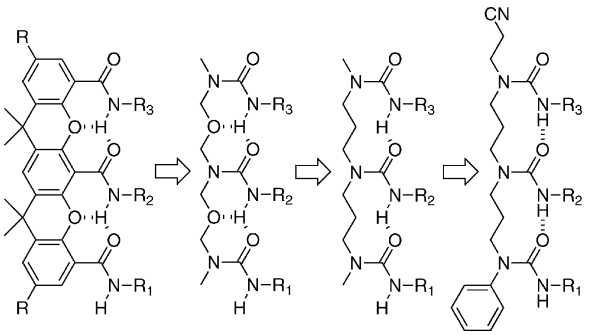
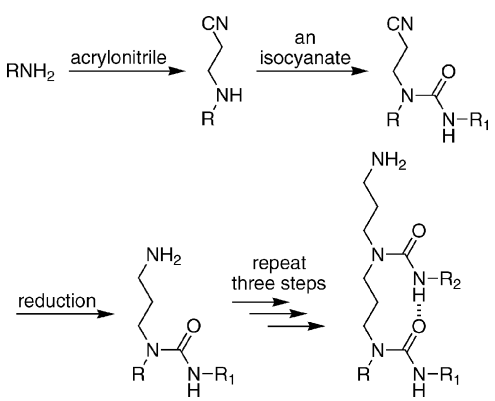


Fig. 3 Conceptual evolution of oligourea molecular scaffolds.

The last structural element of the molecular scaffold, a cyanoethyl group, was incorporated as part of an iterative synthesis that was developed to allow these unnatural oligomers to be constructed by a repeated series of steps, much akin to the iterative synthesis of peptides. The iterative synthesis involves three steps and is outlined in Scheme 1. The first step is addition of an amine to acrylonitrile; the second is reaction of the resulting aminonitrile with an isocyanate to form a urea; and the third is reduction of the nitrile group to a primary amine. Repetition allows the construction of di- and triureas. The cyano group was initially envisioned only as a synthetic element, but subsequent studies established that it further stabilizes the intramolecularly hydrogen-bonded conformation of the molecular scaffold.[8] In



Scheme 1 Iterative synthesis of oligourea molecular scaffolds.

further studies, we developed lower-homologues of the molecular scaffold, based on 1,2-diaminoethane bridges between the urea groups instead of 1,3-diaminopropane bridges, and alternative solution-phase and solid-phase routes to these compounds.[9,10]

The studies described in this section taught me the value of using intramolecular hydrogen bonding to achieve structure and control of conformation and reactivity "on the cheap". They also taught me the value of designing molecules with simple and efficient syntheses and the importance of developing stratagems to overcome problems of poor solubility and aggregation in designed molecules.

The problem of hydrogen bonding in aqueous solution, the theme of dimers, and a bit more personal history

A second of my initial projects focused upon the molecular recognition of nucleic acids in aqueous solution. Although a variety of papers from leading laboratories in the 1980s had focused upon the recognition of nucleic acid bases in non-competitive organic solvents, the recognition of individual nucleic acid bases and related compounds in aqueous solution proved elusive, because water competes with hydrogen-bonding groups of typical synthetic receptors. To address this limitation, undergraduate student Jenny Chen and postdoctoral student Glenn Noronha developed a strategy to achieve molecular recognition of nucleic acid derivatives in water by incorporating the hydrogen-bonding groups of the receptor inside micelles. The resulting self-assembling supramolecular receptors bind complementary nucleic acid bases and nucleotides in water by means of hydrogen-bonding and hydrophobic interactions (Fig. 4).[11,12]

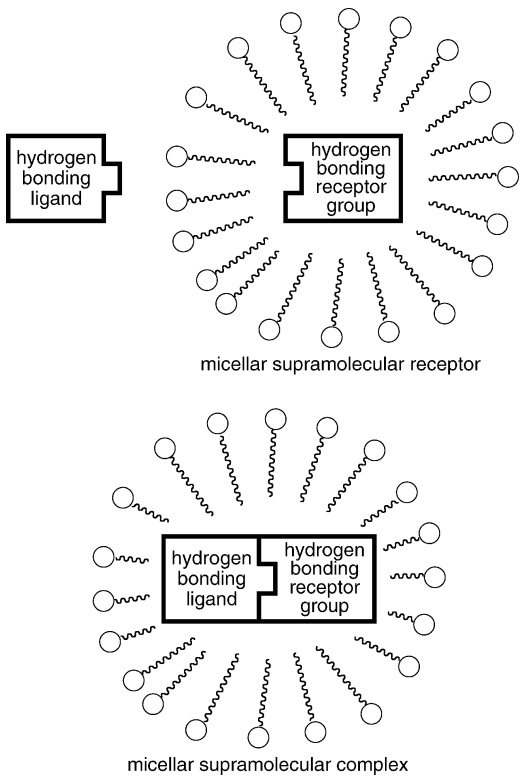


Fig. 4 Cartoon diagram of micellar supramolecular receptors. The hydrogen-bonding groups are based on adenine and thymine.[11]

An unexpected and tangential result from these studies, in the form of an X-ray crystallographic structure, portended what has become a dominant theme in my research laboratory. In studying the nucleobase ligand N^6-acetyl-9-propyladenine, Glenn Noronha discovered that the molecule forms a hydrogen-bonded dimer with four hydrogen bonds, in which the acetamide group adopts an unusual *cis* conformation (Fig. 5 and 6).[11b] This conformation endows the molecule with a self-complementary hydrogen-bonding surface and permits dimer formation. As *cis* amides are generally less stable than *trans* amides, it is likely that hydrogen-bonded dimer formation in the solid state stabilizes the *cis* geometry.

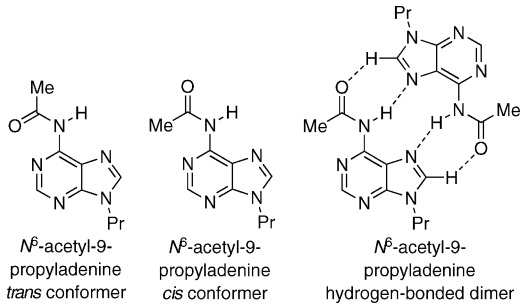


Fig. 5 Conformations and dimerization of N^6-acetyl-9-propyladenine.

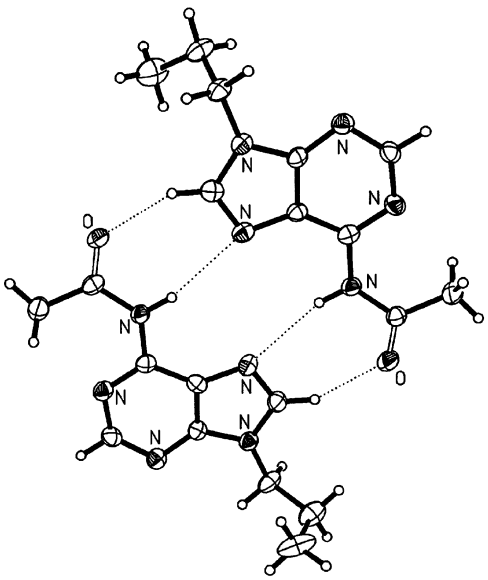
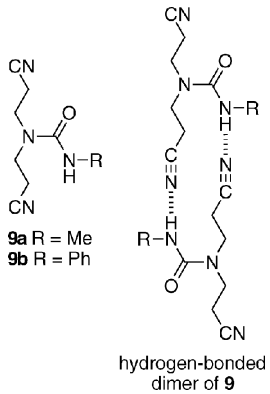


Fig. 6 X-Ray crystallographic structure of the N^6-acetyl-9-propyladenine dimer.[11b]

The symmetrical beauty of self-complementary hydrogen-bonded dimers reemerged when undergraduate student Valerie Antonovich and postdoctoral student Glenn Noronha prepared a series of control compounds associated with the molecular scaffolds described in the preceding section.[13] In studying ureas **9a** and **b**, they noticed an unusual propensity to form hydrogen-bonded dimers in solution and in the solid state. Fig. 7 shows the X-ray crystallographic structure of the dimer of **9b**.



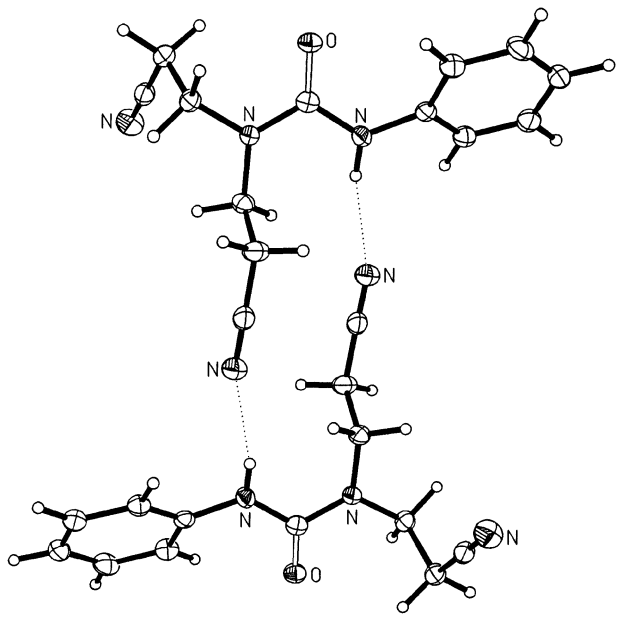


Fig. 7 X-Ray crystallographic structure of the hydrogen-bonded dimer of **9b**.[13]

When my students and I published the structures of these dimers in 1993 and 1995, I did not fully appreciate their relevance to the mimicry of biomolecular structure and interactions. After a talk I presented at UCLA in 1997, Professor J. Fraser Stoddart commented that I should use self-assembly in my designs. This comment, in conjunction with my prior experience with dimer formation, primed me to better appreciate dimerization and use it in subsequent molecular design. When, the next year, graduate student Jim Tsai discovered yet another structure that formed a dimer, I realized that dimerization provided another way of creating molecular structure "on the cheap", as well as an important way in which nature creates higher-order structure in biomacromolecules.

On solubility, aggregation, and unsatisfied hydrogen-bond donors

A third initial project met a painful demise, but in its failure taught a valuable lesson to my students and me. We had hoped to create molecular receptors that served as catalysts for the asymmetric conjugate addition of nucleophiles to Michael acceptors based on acrylic acid and acrylamide. Our basic idea was to develop a molecular cleft with hydrogen-bonding functionality complementary

to that of a carboxylic acid or amide group and chiral general-acid and general-base functionality capable of deprotonating a nucleophile and protonating the enolate anion that would result from conjugate addition. Fig. 8 illustrates the structure of one of the receptors we had designed and illustrates how it might bind to an α,β-substituted acrylic acid. Although we ultimately hoped to prepare chiral versions of the receptor bearing acidic and basic groups "R" based on histidine, we initially embarked on achiral versions bearing simple alkyl groups "R" to test whether the receptors could be prepared and whether they would bind carboxylic acids.

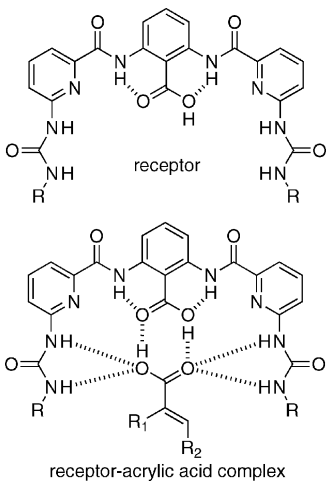


Fig. 8 Design of a receptor for acrylic acids.

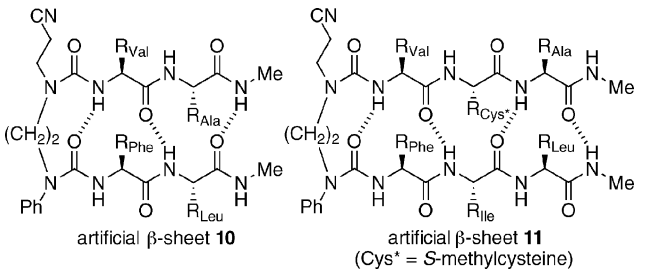
Postdoctoral student Byungwoo Yoo prepared the methyl-substituted complex (R = CH_3), but found it to be insoluble in the sort of non-competitive organic solvents that would be necessary for hydrogen-bonding complexation of the acrylic acid (*e.g.*, $CDCl_3$). He tried without success to increase the solubility by preparing the propyl (R = C_3H_7) and octyl (R = C_8H_{17}) versions of the receptor. By the time he embarked on the octadecyl (R = $C_{18}H_{37}$) version, it was clear that the design for this receptor was hopeless and that we would never be able to use this system to make asymmetric catalysts for conjugate addition.

These painful studies helped teach us that molecules with unsatisfied hydrogen-bonding valences, and particularly hydrogen-bond-donor groups, often self-associate strongly and exhibit poor solubility in organic solvents. In subsequent efforts, we avoided *N*,*N*′-disubstituted ureas, which appear to be particularly prone to self-association through interactions of the urea groups.

Among the toughest decisions that I have had to make is when to abandon a project. Often the original concept behind the project is good, but its reduction to practice just does not work with a given molecular system. The subsequent rise of the important and popular area of *organocatalysis* has taught me that the overarching goal of this project was good, even though our molecular design was not.[14]

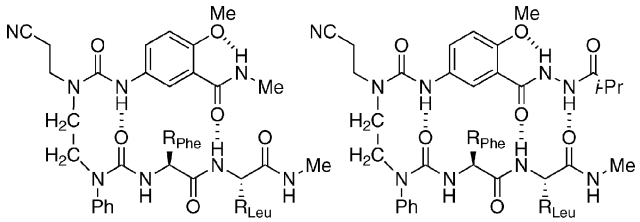
Controlling *intermolecular* hydrogen bonding through designed *intramolecular* hydrogen bonding

My students and I envisioned using our molecular scaffolds to mimic the structures of biopolymers, such as DNA and proteins. In one project, graduate student Eric Smith and postdoctoral student Glenn Noronha used the urea-based scaffold as a template for the creation of β-sheets. Our initial efforts resulted in *artificial β-sheet* **10**, which folded in chloroform solution to mimic the structure and hydrogen-bonding pattern of a parallel β-sheet.[15] ^{1}H NMR chemical shift studies suggested the population of β-sheet conformation to be about 50%. The successful folding of this structure prompted us to pursue bigger and more complex structures. Toward this end, Eric synthesized but did not extensively study larger artificial β-sheet **11**.[10] Concerns about the potential for aggregation, poor solubility, and incomplete folding of such structures prompted us to try to develop additional tactics to improve the conformational homogeneity of our artificial β-sheets.



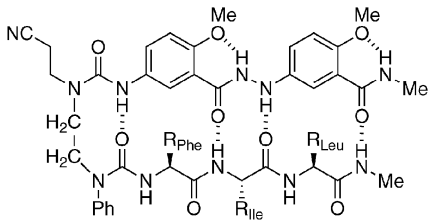
To create larger, better-folded structures, graduate students Eric Smith, Amy Sue Waldman, Jim Tsai, and Mason Pairish, and postdoctoral students In Quen Lee and Darren Holmes began to develop additional molecular templates to satisfy hydrogen-bonding valences, and particularly hydrogen-bond-donor groups, through *intramolecular* hydrogen bonding. Through synthetic, IR and NMR spectroscopic, and X-ray crystallographic studies, they developed a series of molecular templates based on 5-amino-2-methoxybenzoic acid, and related hydrazine derivatives, to mimic and complement the hydrogen-bonding functionality of peptide strands and block unfilled hydrogen-bonding valences. The 2-methoxy group played the dual roles of blocking a hydrogen-bond-donor group and providing organization through intramolecular hydrogen bonding within these *β-strand mimics.* By combining the molecular scaffold with these β-strand mimics, they created a variety of *artificial β-sheets* (*e.g.*, **12–15**) that fold into well-defined structures in chloroform and other non-competitive organic solvents.[16] My students and I were particularly pleased to find that these molecules did not aggregate significantly and had good solubility.

The relative ease with which these and similar artificial β-sheets could be prepared and the reliability with which they folded suggested that we needed to move on to bigger challenges by testing the limits of the strategy of combining oligourea molecular scaffolds and β-strand mimics to make molecules that mimic the structure and hydrogen-bonding patterns of protein β-sheets. Three-stranded artificial β-sheets **16a** and **b**, and four-stranded artificial β-sheet **17** proved worthy targets for these studies. These structures combine features of doubly-templated artificial β-sheets **12–15** with features of singly-templated artificial β-sheets **10** and **11**. Singly-templated structures **10** and **11** contain parallel β-sheets, while doubly-templated structures **12–15** mimic the structure and hydrogen-bonding pattern of an antiparallel β-sheet. The three- and four-stranded structures are mixed β-sheets, containing both parallel and antiparallel elements.

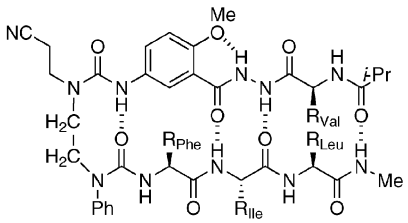


artificial β-sheet **12**

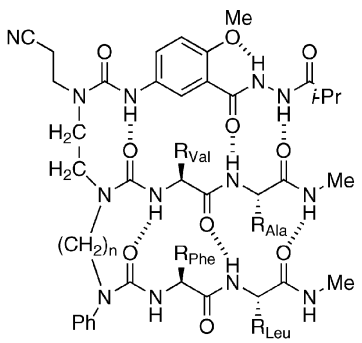
artificial β-sheet **13**



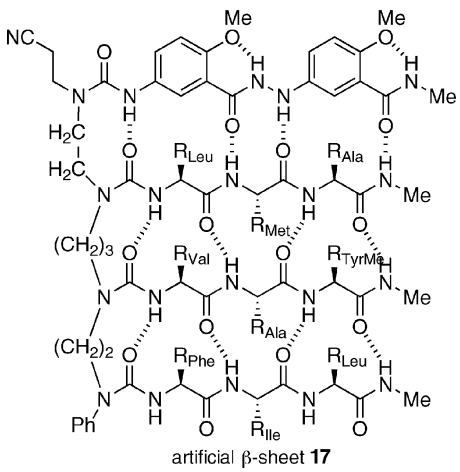
artificial β-sheet **14**



artificial β-sheet **15**



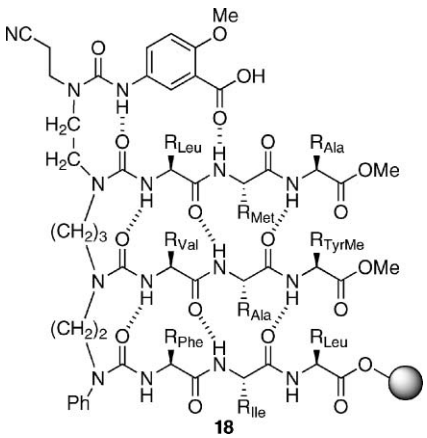
artificial β-sheet **16a** (n = 2)
artificial β-sheet **16b** (n = 3)



artificial β-sheet **17**

While structures **16a** and **b** readily succumbed to chemical synthesis and folded into well-defined β-sheet structures in chloroform solution, structure **17** proved more challenging.[10,17]

The final steps of the lengthy solid-phase synthesis of **17** involved construction of the β-strand mimic by formation of an acyl hydrazide linkage on the solid support, followed by aminolytic cleavage from the resin. Difficulties in carrying out these final steps on advanced resin-bound intermediate **18** led to failure during our first and only attempt to synthesize this four-stranded artificial β-sheet.



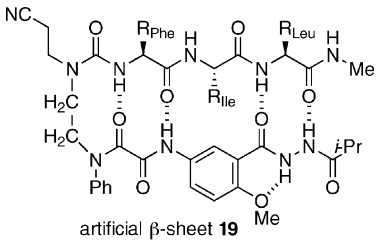
18

On discovery, dimerization, and the importance of self-assembly

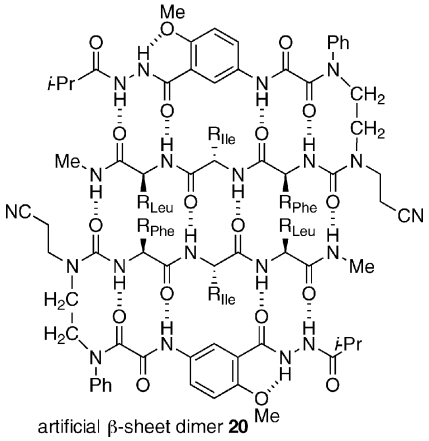
Sometimes frustration and serendipity combine in interesting ways to yield new opportunities. We never went back and repeated our attempts to synthesize artificial β-sheet **17**, although I suspect we could have been successful with just a bit more effort. Instead, a discovery by graduate student Jim Tsai provided a new and more powerful strategy for generating larger and more complex β-sheet structures.

Jim had set out to ask a relatively simple question: whether the β-strand mimic could be incorporated at the "bottom" of the artificial β-sheet, rather than at the "top."[18] To provide suitable connectivity, he incorporated an oxalamide linker unit to connect the 5-amino-2-methoxybenzoic hydrazide unit to the 1,2-diaminoethane-based molecular scaffold. These design features gave rise to artificial β-sheet **19**. Through ^{1}H NMR studies, he found that **19** not only folds to adopt a β-sheet structure, but that this structure dimerizes in chloroform solution to form a well-defined hydrogen-bonded dimer, artificial β-sheet dimer **20**.

The discovery that artificial β-sheet **19** forms well-defined hydrogen-bonded dimers was a revelation that provided my students and me with a new paradigm for easily and efficiently creating large and complex β-sheet structures. When I had embarked upon working with β-sheets, reviewers of grant proposals had cautioned that β-sheets are notorious in their propensity to aggregate. This aggregation occurs, in part, because peptide strands present two hydrogen-bonding edges and can form infinite two-dimensional hydrogen-bonded networks. By blocking one of the hydrogen-bonding edges of the peptide strand with the β-strand mimic, while leaving the other edge exposed, we had developed a structure that was primed to dimerize but not aggregate.



artificial β-sheet **19**



artificial β-sheet dimer **20**

The formation of such well-defined dimers was essentially unprecedented in peptides and arises because one hydrogen-bonding edge is preorganized while the other edge is blocked. We have subsequently incorporated this feature into many of our structures. (Similar dimerization behavior does not occur in doubly-templated artificial β-sheets **12–15**, because the phenyl group on the "bottom" edge partially blocks the hydrogen-bonding surface.) We have generally found that we can build β-sheets with one or two unsatisfied hydrogen-bond-donor groups that are well-behaved and do not undergo uncontrolled aggregation in chloroform and other non-competitive organic solvents. When aggregation and poor solubility problems arise, we can generally mitigate them by blocking exposed unsatisfied hydrogen-bond-donor groups. For this reason, 2-alkoxybenzoic acid groups have figured heavily in many of our subsequent molecular designs.

Our discovery also helped make me aware of the importance of dimerization and self-assembly in proteins.[19] Nature uses controlled interactions between the edges of β-sheets extensively to create functional quaternary structures in proteins, such as HIV-1 protease. Our dimer reminded me of the defensin HNP-3 dimer, which was characterized by my colleagues at UCI and UCLA.[20] The monomer of this protein, like artificial β-sheet **19**, presents a hydrogen-bonding edge consisting of two pairs of hydrogen-bond-donor and acceptor groups that come together to form the functional form of the protein. The other edge of the defensin HNP-3 β-sheet is blocked and is unable to participate in intermolecular hydrogen bonding (Fig. 9).

Interactions between β-sheets are also important in protein aggregation, which occurs widely in neurodegenerative diseases such as Alzheimer's disease, Huntington's disease, and the prion diseases.[19] Although the structures of protein aggregates, such as the β-amyloid associated with Alzheimer's disease, are not generally known precisely, at atomic resolution, they appear to involve uncontrolled interactions between peptide strands that present two hydrogen-bonding edges and can form infinite two-dimensional hydrogen-bonded networks.[22] In subsequent efforts

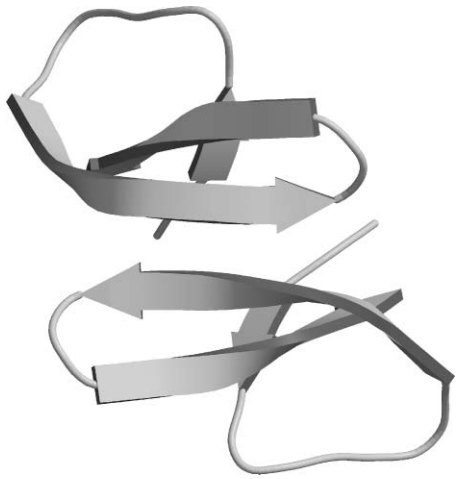


Fig. 9 The defensin HNP-3 dimer.[20,21]

to understand molecular recognition between protein β-sheets, we have surveyed the scope of β-sheet interactions between proteins and have developed a database of these interactions in collaboration with Professor Pierre Baldi and his students in UCI's School of Information and Computer Science.[23,24] By developing molecules that mimic these interactions, my students and I aim to not only understand these interactions but ultimately to control them, with the promise of developing new therapies to diseases that involve interactions between β-sheets.

On NOEs, testable hypotheses, and spectroscopic probes

Once we began preparing and studying a variety of artificial β-sheets, it became increasingly clear that each of our designs centered around the *testable hypothesis*: "If the molecule folds into a β-sheet structure or dimerizes through β-sheet formation, then groups that would otherwise be remote will be brought into proximity." Such proximity is readily elucidated through ^{1}H NMR NOE studies, with two-dimensional ROESY-type NOE experiments being particularly well suited to the intermediate-sized molecules that we were studying. (*ca.* 500–1500 molecular weight). Parallel and antiparallel β-sheets exhibit a series of short proton–proton distances involving the NH-, α-, and β-protons of adjacent peptide strands. Fig. 10 shows characteristic values of these distances (in angstroms), with Newman projections of methyl groups (alanine) representing typical side chains; Fig. 11 provides attractive molecular models of these β-sheets.[25] We have found that NOEs associated with distances of 3–3.5 angstroms or less can generally be observed in our systems if the resonances associated with these protons are well resolved. Because NOEs vary as distance to the inverse sixth power, a longer distance gives a much weaker NOE, which is more difficult to detect. An NOE from a 3.2 Å interproton distance, for example, is expected to be about fourteen percent of the strength of an NOE from a 2.3 Å interproton distance.

In contrast to proteins, the side-chains of peptides and peptide-based model systems generally rotate freely and populate multiple rotamers. For this reason, and because our artificial β-sheets lack many of the tertiary contacts typical of proteins (*e.g.*, three-dimensional packing of an α-helix against a β-sheet or face-to-face packing of β-sheets), we have not heavily pursued the

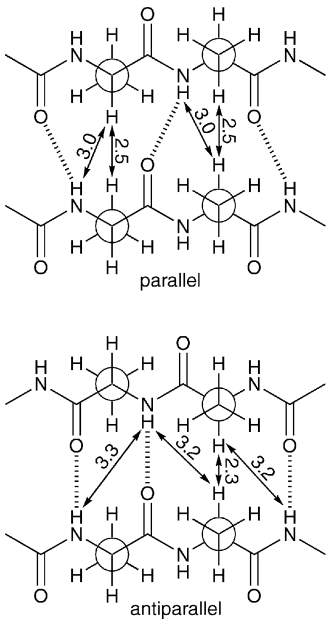


Fig. 10 Typical short interstrand proton–proton distances in parallel and antiparallel β-sheets, in angstroms, which give rise to characteristic NOEs. Distances shown are based on idealized β-sheet structures.[25]

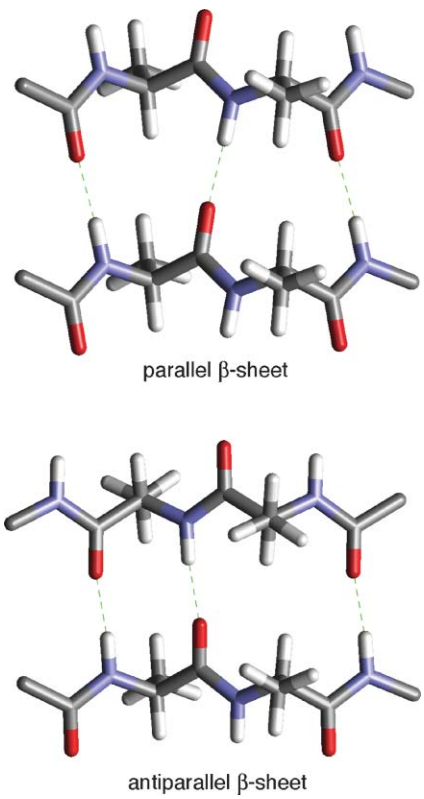


Fig. 11 Models of parallel and antiparallel β-sheets.

elucidation of three-dimensional model structures thorough NMR spectroscopy. Instead, we tend to view two- and three-dimensional molecular models as hypotheses to be tested, both through NOE studies and through other indices of structure, such as chemical shifts and coupling constants of NH and α-protons. While three-dimensional models could be made from these data, they would lack the resolution and detail of protein NMR structures and would provide little meaningful information about the shape, curvature, and side-chain orientations of the artificial β-sheets.

We had been particularly fortunate thus far in obtaining NOE data to establish folding, in part because the proton at the 6-position of the aromatic ring of each of our 5-amino-2-methoxybenzoic acid β-strand mimics appears downfield in the ^{1}H NMR spectrum and is well removed from the amino acid α-protons, which appear midfield. As a result, the β-strand mimic serves not only as a template but also as a *spectroscopic probe* that provides evidence for folding. The β-strand mimic provides the additional advantage of telling whether the NOE experiment is working properly, as the methoxy group of the β-strand mimic is very close to the 3-position of the aromatic ring and gives a very strong NOE that cannot be missed if the experiment is working properly. Fig. 12 illustrates the NOEs that are typically observed.

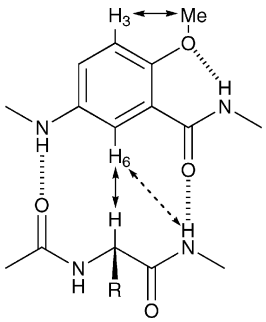


Fig. 12 Typical key NOEs associated with the 5-amino-2-methoxybenzoic acid β-strand mimic. An NOE between the proton at the 6-position of the β-strand mimic and the α-proton of the adjacent peptide is always seen, while an NOE between the proton at the 6-position of the β-strand mimic and the NH-proton of the adjacent peptide is often seen. A strong NOE between the methoxy group and the 3-position of the β-strand mimic indicates that the NOE experiment is working properly.

The observation of a strong NOE between the leucine and phenylalanine α-protons of artificial β-sheet dimer **20** taught my students and me the importance of deliberately building testable hypotheses into our designs by incorporating additional spectroscopic probes. Our choices of amino acids in this and prior artificial β-sheets had been largely arbitrary among the typical non-polar amino acids. Our observation of the NOE between the α-protons of phenylalanine and leucine drove home the point that amino acids can be deliberately picked to allow testing of the hypothesis stated above. For NOEs to be observed between protons requires that the protons not have the same chemical shift, with separations of 0.2 ppm or more being ideal. Also required to uniquely identify NOEs between two protons is that the resonances for these protons not be precisely coincident with resonances of other protons; separation of the centers of the resonances by a couple of hundredths of a ppm is sufficient for two-dimensional NOE studies.

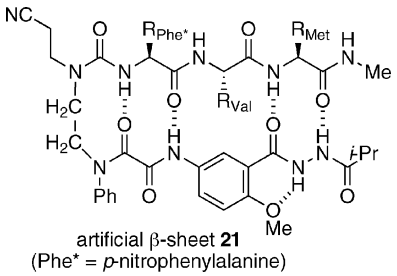
Typical chemical shifts of the various protons of amino acids in unstructured peptides are known and provide useful starting points for picking amino acids with resonances that do not overlap and are at suitable positions.[26] Table 1 lists the positions of the α-protons of the of the twenty proteinogenic amino acids in unstructured peptides (GGXA).[26] In β-sheet structures, the α-protons generally shift downfield by several tenths of a ppm.[27]

Table 1 Chemical shifts of α-protons of amino acids in unstructured peptides[26]

Amino acid	Chemical shift	Amino acid	Chemical shift	Amino acid	Chemical shift	Amino acid	Chemical shift
Ala (A)	4.35	Gly (G)	3.97	Met (M)	4.52	Ser (S)	4.50
Cys (C)	4.69	His (H)	4.63	Asn (N)	4.75	Thr (T)	4.35
Asp (D)	4.76	Ile (I)	4.23	Pro (P)	4.44	Val (V)	4.18
Glu (E)	4.29	Lys (K)	4.36	Gln (Q)	4.37	Trp (W)	4.70
Phe (F)	4.66	Leu (L)	4.38	Arg (R)	4.38	Tyr (Y)	4.60

That the α-proton resonance of phenylalanine typically occurs significantly downfield of the α-proton resonance of leucine (4.66 ppm *vs.* 4.38 ppm in the unstructured peptide) allowed us to readily observe the NOE associated with the short (*ca.* 2.3 Å) distance between these protons in artificial β-sheet dimer **20**.

Jim Tsai performed an elegant experiment that further enhanced my awareness of the importance of well-chosen chemical shifts and the power and precision of NOE data to elucidate β-sheet structure and interaction.[18] To prove that artificial β-sheet **19** was forming a dimer, and that the NOE observed between the phenylalanine and leucine α-protons was intermolecular, rather than intramolecular, Jim prepared "triple mutant" artificial β-sheet **21**, with *p*-nitrophenylalanine (Phe*) in place of phenylalanine, valine in place of isoleucine, and methionine in place of leucine. He chose methionine because he anticipated its α-proton resonance (4.52 ppm in the unstructured peptide) would fall between that of phenylalanine and leucine and would allow the observation of NOE signals associated with the heterodimer. A ROESY experiment on a mixture of artificial β-sheets **19** and **21** revealed NOEs between the phenyalanine and methionine α-protons and between the *p*-nitrophenylalanine and leucine α-protons, thus proving the formation of β-sheet dimers (Fig. 13 and 14).



artificial β-sheet **21**
(Phe* = *p*-nitrophenylalanine)

This ROESY experiment, in conjunction with ^{1}H NMR titration studies, yielded an unanticipated result that set the stage for further research. Jim found that artificial β-sheet **19** self-associates much more strongly than "triple-mutant" artificial β-sheet **21** (600 M^{-1} *vs.* 90 M^{-1} self-association constant in $CDCl_3$). From the relative sizes of the crosspeaks of the ROESY spectrum (Fig. 13), it appeared that the heterodimer forms with a self-association constant of intermediate magnitude.

These findings tickled my curiosity and posed a question that my students and I were not well equipped to answer until three years later: "Why do different β-sheets interact with different affinities and what is the effect of different amino acid side chains upon the strength of these interactions?" Although Jim Tsai prepared three additional "single mutant" artificial β-sheets and measured their self-association constants, the months of effort required for these

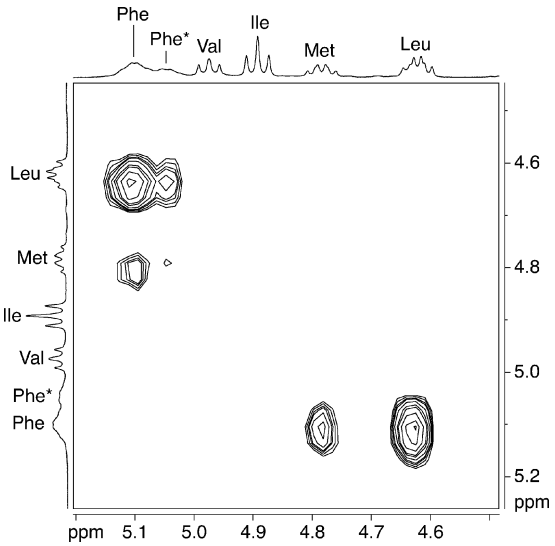


Fig. 13 ^{1}H NMR ROESY spectrum (α-proton region) of a solution of artificial β-sheets **19** and **21** in $CDCl_3$ at 303 K (15 and 10 mM, respectively). The Phe–Leu crosspeaks arise from the **19** homodimer (**20**), the Phe*–Met crosspeaks arise from the **21** homodimer, and the Phe*–Leu and Phe–Met crosspeaks arise from the **19**·**21** heterodimer.[18]

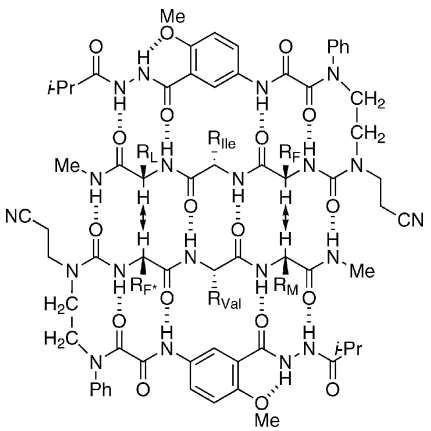
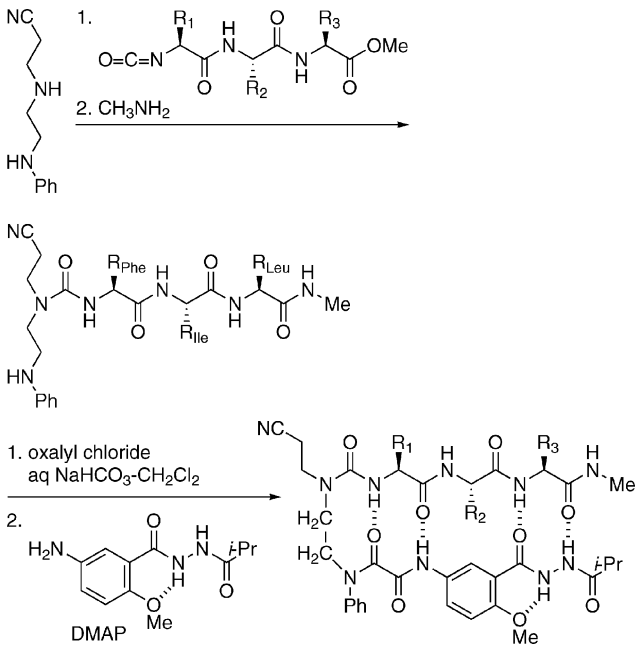


Fig. 14 Key NOEs associated with the **19**·**21** heterodimer.

studies seemed too great to justify a full-fledged attempt to answer this question.[28] Further frustrating our efforts to understand these interactions was that we could only quantify the formation of the homodimers with accuracy, and not the heterodimers, and that each measurement of a self-association constant required four hours of tedious NMR titration studies, which had to be repeated in triplicate to ensure accuracy.

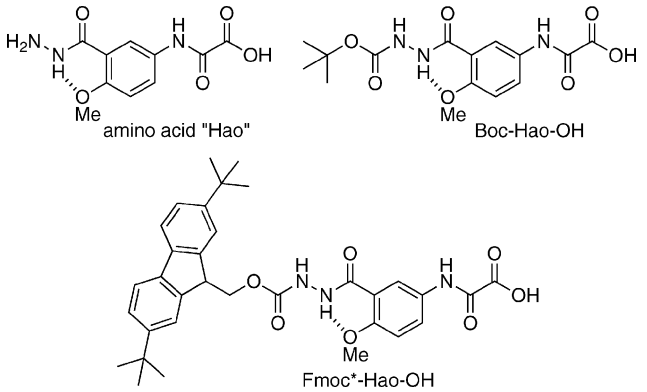
The value and limitations of chemical synthesis and the benefits of modular building blocks

Jim's syntheses of artificial β-sheet **19**, "triple-mutant" **21**, and the "single mutants" demonstrated to me both the importance and the limitations of chemical synthesis in our approach to understanding biomolecular structure and interactions (Scheme 2).[18] The synthesis of each of these artificial β-sheets required two technically demanding steps. The first involved the synthesis of a peptide isocyanate, a hitherto unreported class of compounds that postdoctoral students Darren Holmes and Glenn Noronha and graduate student Eric Smith had specifically invented to allow the synthesis of our artificial β-sheets.[29] The second involved the reaction of an amino group with oxalyl chloride to form an oxamoyl chloride adduct and its trapping with an amine to form an oxamide linkage.

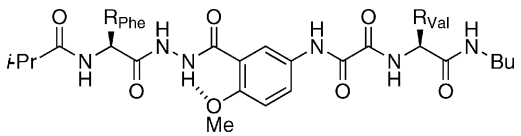


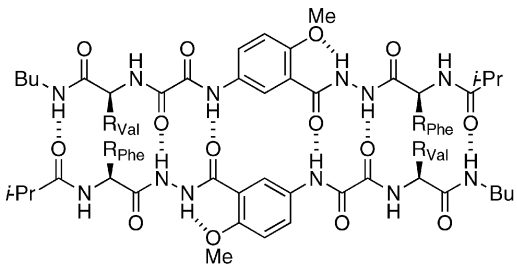
Scheme 2 The synthesis of artificial β-sheets **19** and **21**.

At our research group meeting one evening, postdoctoral student Santanu Maitra amazed me with a molecule that offered a new strategy to overcoming many of the limitations that we were encountering in chemical synthesis.[30] Santanu had recognized that the β-strand mimic of Jim's artificial β-sheet was, itself, an amino acid and developed protected versions of this amino acid, which could be used as building blocks in standard peptide synthesis techniques. We subsequently came to call this amino acid "Hao" to reflect the three components that make it up: hydrazine, 5-amino-2-methoxybenzoic acid, and oxalic acid. Santanu developed a Boc-protected version of Hao, Boc–Hao–OH, that is compatible with either solution-phase or solid-phase peptide synthesis. Using a *tert*-butylated version of the Fmoc protecting group (Fmoc*) that graduate student Matt Koutrolis and postdoctoral student Kim Stigers had developed to overcome the poor solubility of many Fmoc-protected compounds, Santanu created a protected version of this amino acid that is ideal for solid-phase peptide synthesis: Fmoc*–Hao–OH.[31]



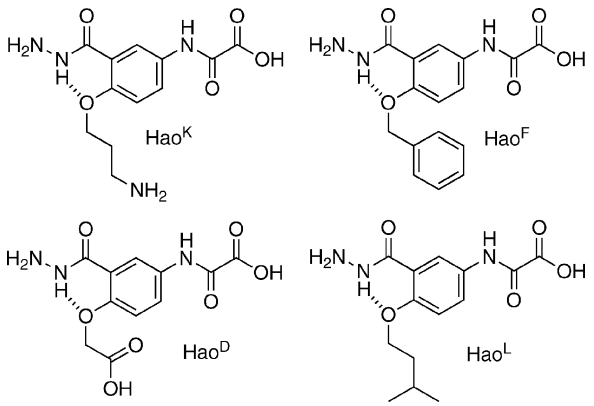
Santanu's invention of Hao also built upon and complemented efforts by graduate student Ye Sun, who had developed peptides containing this β-strand mimic using a cumbersome synthetic route that involved assembling the β-strand mimic as part of the peptide synthesis.[30] Studies of a peptide that contained the Hao amino acid that had been constructed by this route (**22**, "Ye's peptide") had already established that Hao mimics a β-strand conformation and induces β-sheet dimerization when incorporated into peptides. Santanu's introduction of the Boc- and Fmoc*-protected Hao amino acid now made Ye's peptide and related peptides readily accessible for studies.





The idea of functional modular building blocks that can reliably be coupled to form larger and more complex structures with predictable properties is so powerful that it has now thoroughly permeated all of our structures and molecular designs. Postdoctoral student Tania Khasanova went on to develop a family of four Hao analogues in which the methoxy group is replaced with other alkoxy groups to impart enhanced solubility and folding properties.[32] Because the modified alkoxy groups resemble the side chains of lysine, aspartic acid, phenylalanine, and leucine, we call these derivatives Hao^{K}, Hao^{D}, Hao^{F}, and Hao^{L}. Tania prepared these modified Hao amino acids in protected forms suitable for use in solid-phase peptide synthesis, and they have proven useful in preparing β-sheet peptides with

modified properties. For example, Hao^K-containing peptides have enhanced water solubility and diminished aggregation in aqueous solution.



Postdoctoral students Bill Kemnitzer and Santanu Maitra further reinforced the value of functional modular building blocks by combining the amino acid ornithine with the amino acid Hao to create a composite building block Orn(i-PrCO–Hao) that can be incorporated into peptides, through coupling of the Fmoc-protected building block **23**, to induce folding and dimerization in organic solvents.[33] Peptide **24** (*o*-anisoyl–Val–Orn(i-PrCO–Hao)–Phe–Ile–Leu–NHMe) and the corresponding dimer illustrate the role of this building block and the structure of the peptides and dimers that form. Building block Orn(i-PrCO–Hao) permits the construction of dimeric structures even more complicated than that of artificial β-sheet **19** in a day or two, rather than in weeks. The simplicity of this building block has even permitted us to prepare as many as eight different structures at a time using simple parallel synthesis techniques and equipment. A valuable and unanticipated reward of the Orn(i-PrCO–Hao) building block was the discovery that the δ-linked ornithine amino acid subunit provides additional spectroscopic probes through the formation of a well-defined turn structure with distinct NMR properties (NOEs, magnetic anisotropy, and *J*-coupling) that indicate its folding.

The δ-linked ornithine turn unit was inspired by our earlier work with urea-based molecular scaffolds, as well as by a discussion with Professor Kit Lam of the UC Davis Cancer Center. Kit Lam was intrigued by the Hao amino acid building block but lamented that it could not be decoded in one-peptide-one-bead combinatorial libraries by Edman sequencing. He suggested that we append the Hao amino acid to the side chain of lysine. Our experiences with urea-based molecular scaffolds indicated that lysine was not likely to fold well into a turn structure, but that ornithine, which has one carbon atom fewer in the side chain, was likely to work. Subsequent studies by postdoctoral student Justin Brower have established that this building block is comparable to the best peptide β-turns at creating a well-defined peptide β-hairpin structure (Fig. 15).[34] The δ-linked ornithine turn unit has become another mainstay in our toolbox of modular amino acid building blocks, both because of its structural integrity and because the α-amino group provides water solubility and an extra linkage point that a conventional peptide β-turn does not have.

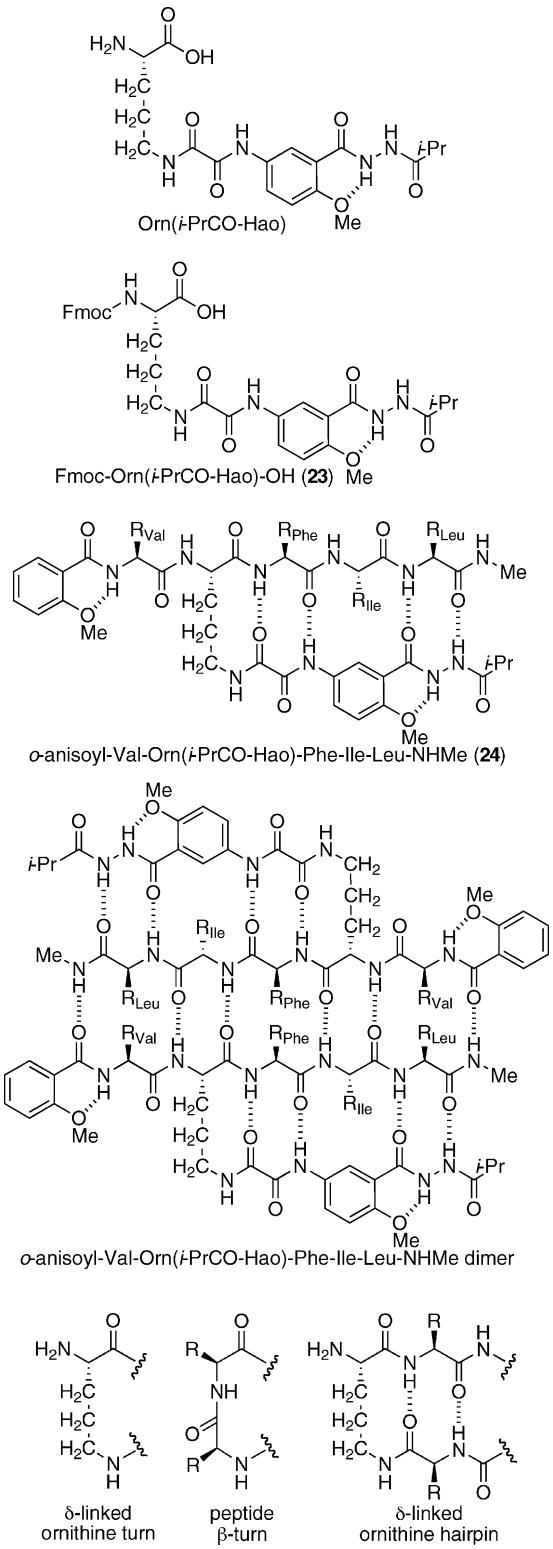
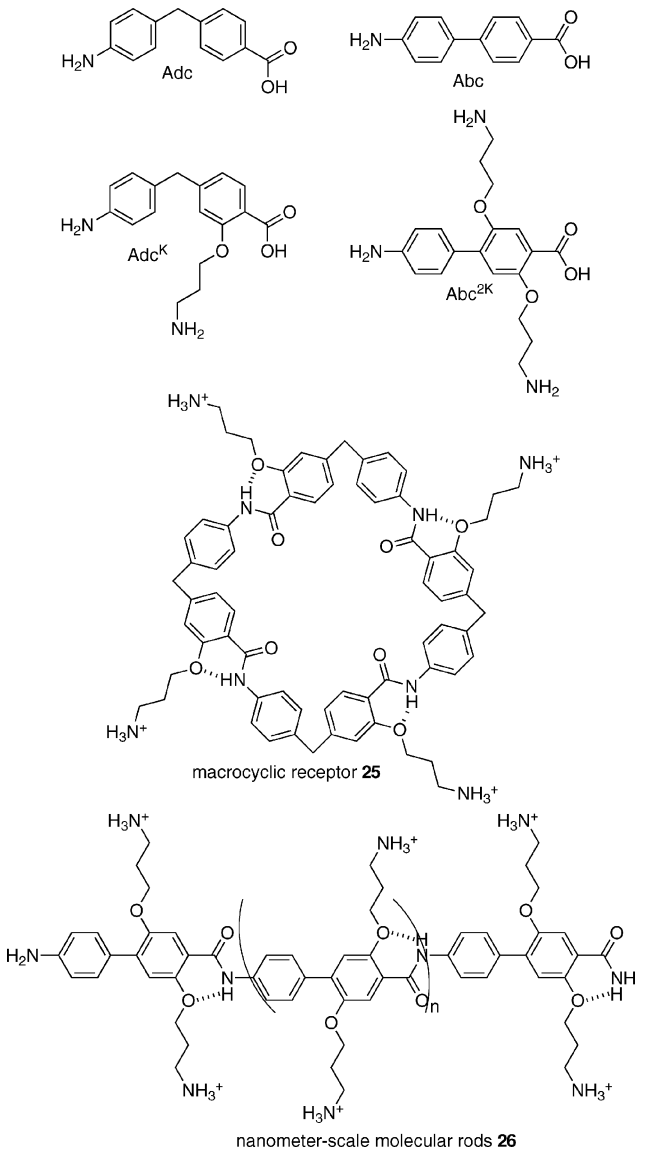


Fig. 15 Turn and hairpin structures.

The ease and reliability with which the Hao-family of amino acid building blocks can be assembled to create complex functional structures has inspired my students and me to pursue additional projects involving complex functional amino acids. One set of projects seeks to develop nanometre-scale amino acids for biomolecular nanotechnology. Graduate students Sang

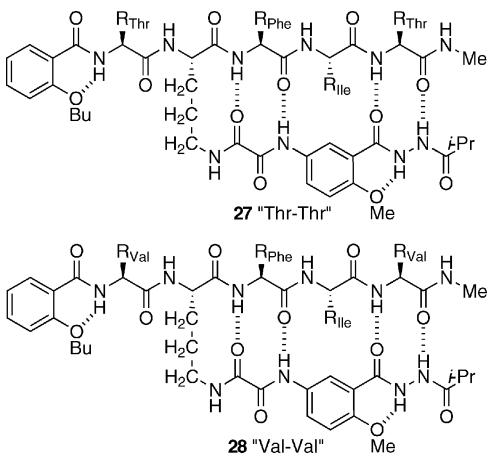
Woo Kang and Chris Gothard have now developed macrocyclic receptors and rods based on the nanometre-scale amino acid building blocks "Adc" (aminodiphenylmethanecarboxylic acid) and "Abc" (aminobiphenylcarboxylic acid) and their water-soluble analogues Adc^K and Abc^{2K}.[35,36] Using a suitably protected version of Adc^K and standard solid-phase peptide synthesis techniques, Sang has synthesized macrocyclic receptor **25**, which binds large hydrophobic molecules, such as sodium cholate, in aqueous solution. Chris has synthesized nanometre-scale rods **26**, in 3–10 nm lengths, also using standard solid-phase peptide synthesis techniques and a suitably protected version of Abc^{2K}. The ease of these syntheses has further reinforced my appreciation of the power of using designed modular building blocks to create complex structures inspired by biomolecules.



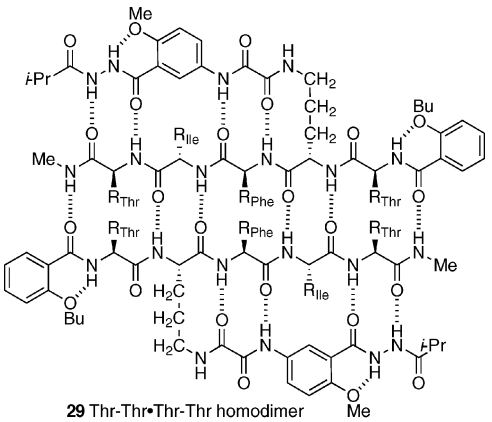
Of ^{1}H NMR spectroscopy and fast and slow exchange

An underappreciated factor that often distinguishes biomacromolecules from small molecules is that interactions between biomacromolecules often occur slowly, while interactions among small molecules generally occur rapidly. The stability of many protein and nucleic acid oligomers and complexes permits their manipulation and study as stable entities in a variety of electrophoretic, chromatographic, and spectroscopic techniques. As my students and I began working with larger and more complex molecules, we have observed that the stability of their complexes has increased. This increase in stability has provided new opportunities (and in some cases new challenges).

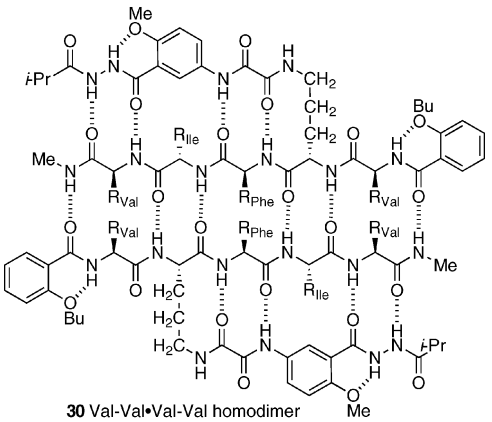
I began to appreciate this difference when graduate student Michael Chung began to tackle the question raised by Jim Tsai's observation of differences in stability among the homo- and heterodimers of his artificial β-sheets.[37,38] Michael envisioned using peptides containing Orn(i-PrCO–Hao), of the general structure introduced by Bill Kemnitzer and Santanu Maitra, to study the effects of different amino acid side chains upon interactions between β-sheets. Michael prepared peptides such as **27** ("Thr–Thr") and **28** ("Val–Val"), which present different amino acids along the exposed β-sheet edge. Unlike Jim's artificial β-sheets, these peptides form dimers that exchange slowly on the NMR time scale. The Thr–Thr·Thr–Thr and Val–Val·Val–Val homodimers **29** and **30**, for example, generate an equilibrium concentration of Thr–Thr·Val–Val heterodimer **31** when mixed. The slower exchange of these dimers reflects the larger number of hydrogen bonds among the peptide amide groups (six in **29**, **30**, and **31**, *vs.* four in **20**).



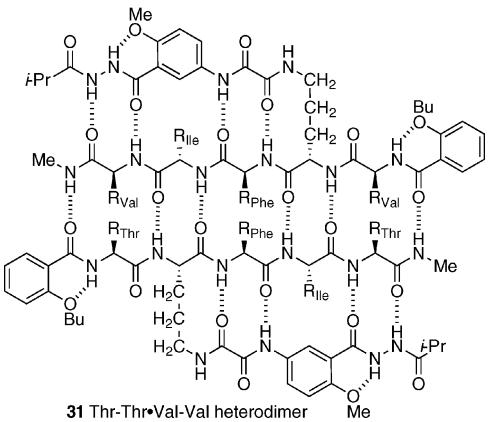
The slow exchange permits the separate observation of the homo- and heterodimers by ^{1}H NMR spectroscopy and allows them to be quantified and their relative stabilities to be determined (Fig. 16). Through studies of this sort, Michael established the relative thermodynamic stabilities of many amino acid pairings. For example, studies of the "Thr–Thr" and "Val–Val" peptides (**27** and **28**), in conjunction with studies of "Thr–Val" and "Val–Thr" peptides (not shown), established that an intermolecular Thr–Thr pair and an intermolecular Val–Val pair collectively are 0.6 kcal mol^{-1} more stable than two intermolecular Thr–Val pairs in 10% CD_3SOCD_3 in $CDCl_3$ solution.[37a] These studies have established that interactions between polar amino acids can impart sequence selectivity into the pairing of β-sheets. In other studies, Michael established the effects of chirality, aromatic interactions, and steric effects on the pairing of β-sheets.[37b,c,38,39]



29 Thr-Thr•Thr-Thr homodimer



30 Val-Val•Val-Val homodimer



31 Thr-Thr•Val-Val heterodimer

The benefit of being able to observe multiple species in slow chemical exchange by NMR spectroscopy comes with the challenge that there are more species to observe and characterize and the problem that overlapping resonances and chemical exchange between species considerably complicates the analyses of NMR spectra. The ever-present potential for overlapping resonances can lead to uncertainties about which crosspeaks in two-dimensional NMR spectra are associated with which species. This problem can be particularly vexing when the relative concentrations of the two species are very different (*e.g.*, a 5 : 1, 10 : 1, or 30 : 1 ratio) and identification of all of the resonances associated with the minor species is not possible. Chemical exchange that is slow on the millisecond timescale associated with the uncertainty principle but rapid on the hundred-millisecond to second timescale associated with proton relaxation also poses a special challenge, because

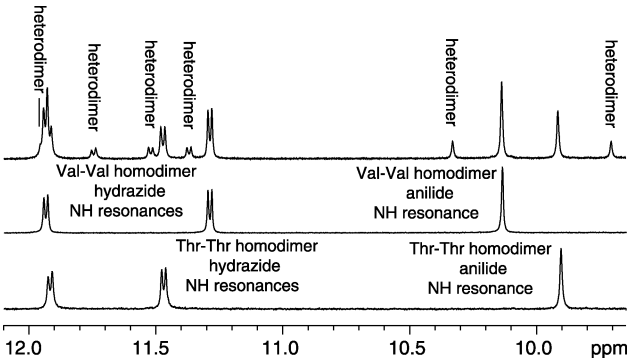


Fig. 16 ^{1}H NMR spectra of peptides **27** (Thr–Thr, lower), **28** (Val–Val, middle), and a mixture of the two peptides (upper) in 10% CD_3SOCD_3 in $CDCl_3$ solution at 253 K. The spectra of the pure peptides show homodimers (**29** and **30**); the spectrum of the mixture shows predominantly the homodimers and smaller quantities of the heterodimer (**31**).[37a]

crosspeaks in two-dimensional spectra may arise from interactions within either species.

We have found two-dimensional chemical exchange spectroscopy (EXSY) to be particularly valuable in dealing with the problems and challenges posed by multiple species in slow chemical exchange.[37b,40] By performing an EXSY experiment with a mixing time of hundreds of milliseconds and running the experiment at a temperature at which chemical exchange occurs at a comparable rate (*e.g.*, by warming the sample until slight broadening of the resonances occurs), it is possible to identify the two (or more) resonances associated with each proton in the molecule. The EXSY experiment is far more effective for complex molecules than a variable-temperature coalescence experiment and readily permits the assignment of overlapping resonances and small resonances associated with minor species. Fig. 17 provides an example of the application of this technique to establish that

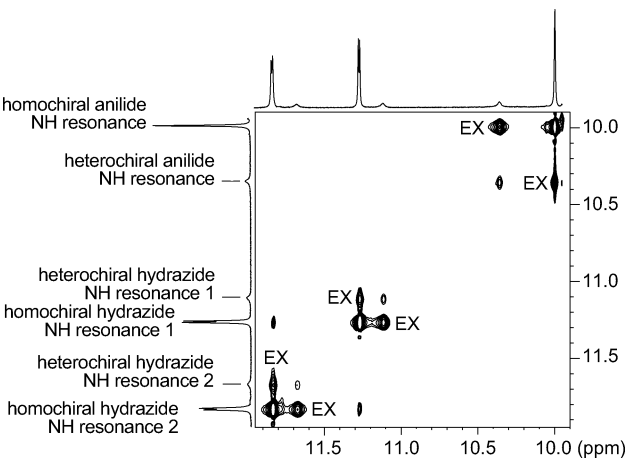
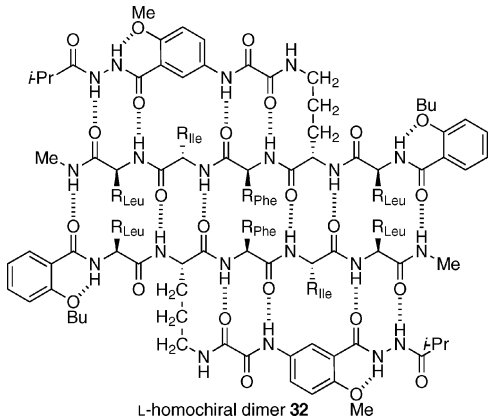
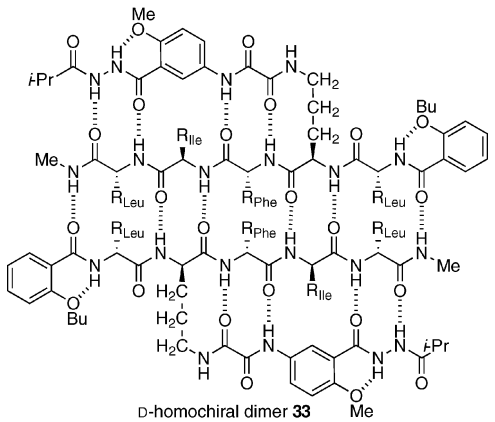


Fig. 17 2D EXSY spectrum of a mixture of L-homochiral dimer **32** and D-homochiral dimer **33**. The spectrum was recorded at 800 MHz in $CDCl_3$ at 308 K at 2.0 mM of each peptide using a 500 ms mixing time. EXSY crosspeaks resulting from chemical exchange with heterochiral dimer **34** are marked "EX".[37b]

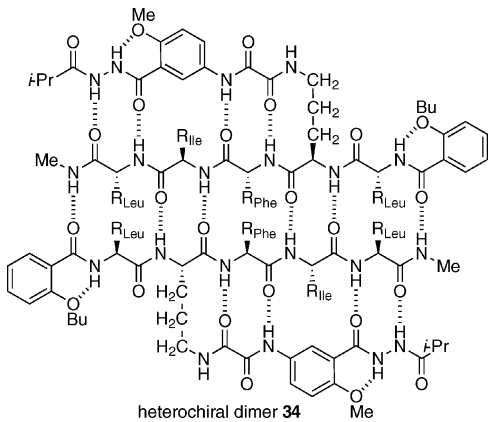
homochiral β-sheet dimers **32** and **33** are in equilibrium with a small population of heterochiral β-sheet dimer **34**.



L-homochiral dimer **32**



D-homochiral dimer **33**



heterochiral dimer **34**

The ease and success with which we have studied these β-sheet forming molecules has shaped and reinforced my understanding of many of the points I have raised in the preceding sections. The molecules are composed of modular building blocks and are easy to synthesize. Intramolecular hydrogen bonding helps create folded structures that form well-defined hydrogen-bonded dimers; all but one hydrogen-bond donor is satisfied either through intramolecular or intermolecular hydrogen bonding. The molecules contain spectroscopic probes for folding and dimerization and exhibit clear NOEs that reflect the folding and dimerization. Slow exchange among the dimers facilitates the identification and analysis of the species that are present.

The challenges and rewards of water-soluble structures

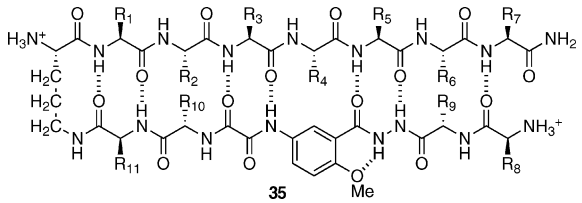
Water-soluble systems that mimic biomolecular structures and interactions in aqueous solution provide challenges and opportunities that are very different from systems that dissolve, fold, and interact in organic solvents. As we have already seen in the discussion of micellar supramolecular receptors, hydrogen bonding is difficult to achieve in water. For this reason, the design of systems that fold and interact through hydrogen bonding in aqueous solution is especially challenging. Water-soluble structures also require different methods of purification than those that dissolve in less polar organic solvents. The latter can often be purified on a scale of up to hundreds of milligrams or grams by "flash" chromatography on silica gel, while the former require reverse-phase high-performance liquid chromatography (RP-HPLC) and are limited to smaller scales (*e.g.*, tens of milligrams). On the other hand, RP-HPLC is often more effective than "flash" chromatography and generally affords a clean compound. An added benefit of RP-HPLC is that the removal of the aqueous phase by lyophilization often yields fluffy solids that are easy to manipulate even on a milligram scale.

The peptide chemistry community began to successfully tackle the thorny problem of creating small peptides that fold to form β-hairpin structures in aqueous solution in the mid-1990s and has now developed a variety of different structures that fold into these minimal two-stranded β-sheet structures.[41] What was, and continues to be, missing from the elegant structures that have been developed is *generality*—the ability of these structures to transcend specific peptide sequences and fold without regard to the individual amino acids present. The β-hairpin structures that have been developed are, for the most part, remarkably fragile and teeter on a knife-edge equilibrium between the folded and unfolded states. Alteration of one or two amino acids is often sufficient to tip the balance toward the unfolded state.[42] Also lacking from these structures is *intermolecularity*—the ability to interact in a controlled fashion. In contrast to the widespread interactions between β-sheets in protein quaternary structures, the β-hairpins that have been developed do not self-assemble into well-defined structures.

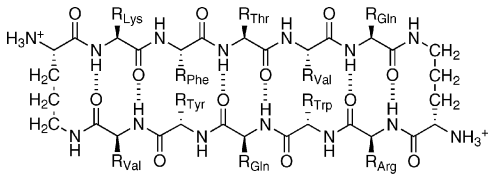
In 2001, my students and I began investigating the use of our δ-linked ornithine and Hao building blocks, with the goal of developing modular β-sheet systems that fold, self-assemble, and interact with proteins in aqueous solution. An innovation by postdoctoral student Nooshin Hashemzadeh led to our first water-soluble molecules that fold into well-defined structures in aqueous solution. We had struggled largely without success to design molecules based on Orn(i-PrCO–Hao), and other closely related systems, to adopt folded β-sheet structures in aqueous solution. Inspired by Kelly and coworkers' finding that hydrophobic cluster formation is important in β-hairpin formation, Nooshin envisioned separating the ornithine and Hao residues by pairs of hydrophobic amino acids.[43] We were delighted to discover that peptides **35** adopt folded β-sheet structures when the hydrophobic amino acids valine and leucine are incorporated at the R_1, R_2, R_{10}, and R_{11} positions.

A subsequent innovation led to structures that are more general, folding largely without regard to the residues that are present. In studying the δ-linked ornithine turn, postdoctoral student Justin Brower was inspired by earlier work by Gellman and

coworkers to prepare macrocyclic β-sheet **36** as a control.[34,44] We were, in hindsight, amazed with the facility with which the macrocyclization reaction to generate this 42-membered ring macrolactam occurred. With little special precaution or effort, other than to carry out the reaction under dilute conditions, the macrocyclization proceeded cleanly.

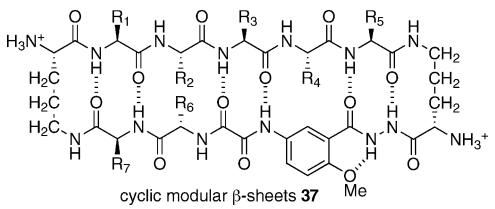


35



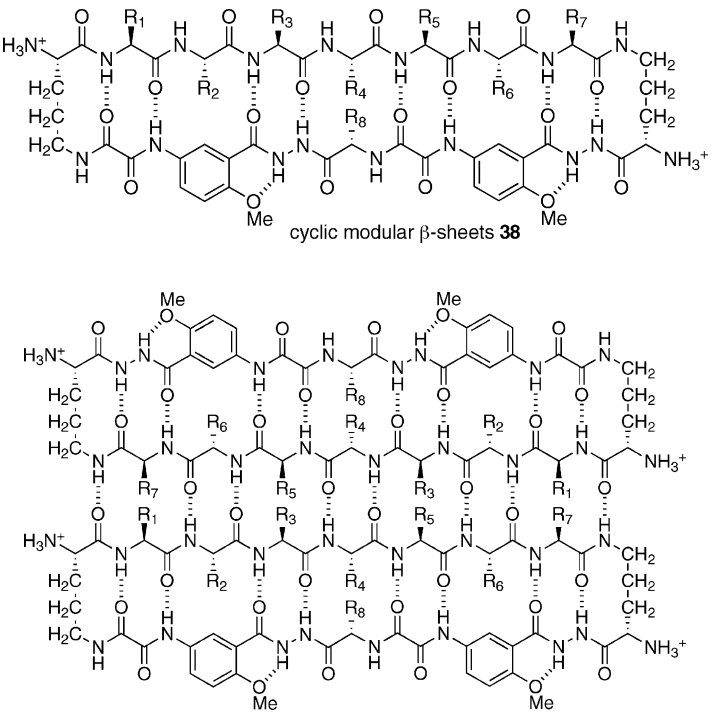
36

Drawing upon the ideas behind β-hairpin structures **35** and macrocycle **36** led us to conceive of macrocyclic β-sheets **37**, which incorporate features of both of these structures. A team effort by postdoctoral students Justin Brower, Wade Russu, and Jeremy Woods and graduate student Omid Khakshoor established that macrocyclic β-sheets **37** are easy to synthesize and generally exhibit good water solubility and folding properties, with little regard to the individual amino acids that are present. Further efforts established that the macrocyclic β-sheets can be linked together by the ornithine α-amino groups. We have termed peptides **37** *cyclic modular β-sheets* to reflect the generality with which they fold and the potential of these structures to be used as modular building blocks for the creation of more complex structures.



cyclic modular β-sheets **37**

An innovation by Omid Khakshoor has led to intermolecularity. Omid postulated that increasing the length of the extended peptide β-strand from a pentapeptide to a heptapeptide would facilitate edge-to-edge hydrogen-bonding interactions between β-sheets in aqueous solution. Upon preparing cyclic modular β-sheets **38**, he was delighted to find that the molecules fold, dimerize through edge-to-edge β-sheet interactions, and further self-assemble to form a tetramer consisting of a face-to-face dimer of dimers (Fig. 18 and 19). These structures are unique among small designed β-sheet structures, in that they achieve well-defined quaternary structure in addition to elements of secondary and tertiary structure. This system is reminiscent of the self-assembling supramolecular receptors with which my students and I had achieved molecular recognition of nucleic acid derivatives in water (Fig. 4), in that the hydrophobic face-to-face interactions involved in the self-assembly of the dimer of dimers facilitate the hydrogen-bonding edge-to-edge interactions between the β-sheets that comprise each dimer.



cyclic modular β-sheets **38**

Fig. 18 Hydrogen-bonded dimer formed by edge-to-edge interactions of cyclic modular β-sheets **38**.

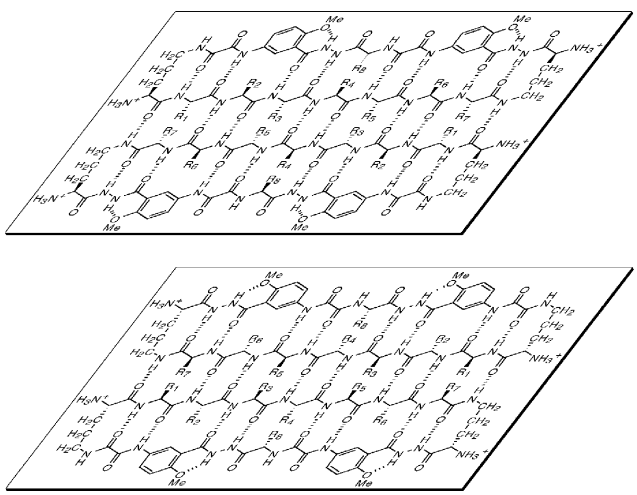


Fig. 19 Illustration of the β-sandwich tetramer of cyclic modular β-sheets **38**. The tetramer consists of two hydrogen-bonded dimers (Fig. 18) self-assembled through hydrophobic face-to-face interactions.

Our studies of the folding and interactions of these water-soluble β-sheet peptides have taught us that working with water-soluble structures can be remarkably rewarding. With good building blocks in hand, the systems are easy to synthesize and purify, and they form interesting, biologically relevant structures. Many of the principles that my students and I have learned are embedded in our current designs. Modularity helps make the molecules easy to synthesize, and judicious use of intramolecular hydrogen bonding and complementary hydrogen-bonding groups helps impart good solubility. Dimerization and higher-order

self-assembly allow us to create complex molecular architectures that mimic those found in proteins. ^{1}H NMR spectroscopy has proven critical to our understanding of these structures. Judicious use of spectroscopic probes that allow us to easily identify NOEs and other NMR properties associated with folding and self-association have allowed us to test our ideas. In studying these systems, pulsed-field gradient (PFG) NMR measurements of diffusion coefficients have proven particularly useful in establishing the oligomerization states of these species; these measurements promise to be a valuable addition to our arsenal of tools for characterizing the structures and interactions of our molecules.[45–47]

Conclusion

I never could have envisioned the directions that this series of studies would take when I embarked upon this research almost seventeen years ago. The learning that has come from first-hand experiences with each problem in biomolecular structure and interaction has directly influenced the problem selection and approach that has been taken in each subsequent study. Additional influence has come from the outstanding work of other members of the chemical community. I would like to think that the work can be described as creative. Insofar as this is the case, much of the creativity has come from having a general idea the problem upon which I wish to work, applying what my students and I have learned from our past experiences, and keeping our eyes open to what the molecules that we are studying can teach us. We are now trying to coax our molecules to interact with proteins and other biomolecular systems, and I look forward to learning from these studies in the upcoming years.

Acknowledgements

The importance of people. Almost everything that I have learned through these studies has been a direct result of the efforts of my talented undergraduate, graduate, and postdoctoral students. I am remarkably grateful to them for all of their hard work, as well as their keen insights. I have tried to acknowledge the principal contributors to each study in the text of this manuscript. Additional contributors, to whom I am also grateful, are cited in the references. I hope that this paper will help share with some of my past, current, and future students broader trends and insights that might not be apparent from working on individual projects. I also hope that these trends and insights will be of benefit to the general readership.

This work would not have been possible without the generous support over the years from the following granting agencies, foundations, companies, and intramural sources, to whom I offer my sincere thanks: National Institutes of Health (GM-49076); National Science Foundation; American Cancer Society; American Chemical Society Petroleum Research Fund; Camille and Henry Dreyfus Foundation; Alfred P. Sloan Foundation; Arnold and Mabel Beckman Foundation; Bill and Melinda Gates Foundation; American Chemical Society; Zeneca Pharmaceuticals group; Upjohn Company; Hoffman-La Roche Inc.; Pfizer, Inc.; University of California Cancer Research Coordinating Committee; University of California Systemwide Biotechnology Research and Education Program (UC BREP); UCI School of Physical Sciences Committee on Research; UCI Council on Research, Computing, and Library Resources; Institute for Brain Aging and Dementia (UCI), Cancer Research Institute (UCI), Institute for Genomics and Bioinformatics (UCI).

References

1 J. S. Nowick, N. A. Powell, E. J. Martinez, E. M. Smith and G. Noronha, *J. Org. Chem.*, 1992, **57**, 3763–3765.
2 S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173–180.
3 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore, *Chem. Rev.*, 2001, **101**, 3893–4011.
4 J. P. Fillers, K. G. Ravichandran, I. Abdalmuhdi, A. Tulinsky and C. K. Chang, *J. Am. Chem. Soc.*, 1986, **108**, 417–424.
5 (*a*) D. S. Kemp and D. R. Buckler, *Tetrahedron Lett.*, 1991, **32**, 3009–3012; (*b*) D. S. Kemp and D. R. Buckler, *Tetrahedron Lett.*, 1991, **32**, 3013–3016.
6 J. S. Nowick, P. Ballester, F. Ebmeyer and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1990, **112**, 8902–8906.
7 A Itai, Y. Toriumi, N. Tomioka, H. Kagechika, I. Azumaya and K. Shudo, *Tetrahedron Lett.*, 1989, **30**, 6177–6180.
8 D. L. Holmes and J. S. Nowick, unpublished results..
9 (*a*) J. S. Nowick, M. Abdi, K. A. Bellamo, J. A. Love, E. J. Martinez, G. Noronha, E. M. Smith and J. W. Ziller, *J. Am. Chem. Soc.*, 1995, **117**, 89–99; (*b*) J. S. Nowick, S. Mahrus, E. M. Smith and J. W. Ziller, *J. Am. Chem. Soc.*, 1996, **118**, 1066–1072; (*c*) D. L. Holmes, E. M. Smith and J. S. Nowick, *J. Am. Chem. Soc.*, 1997, **119**, 7665–7669.
10 E. M. Smith PhD Dissertation, University of California, Irvine, 1997.
11 (*a*) J. S. Nowick and J. S. Chen, *J. Am. Chem. Soc.*, 1992, **114**, 1107–1108; (*b*) J. S. Nowick, J. S. Chen and G. Noronha, *J. Am. Chem. Soc.*, 1993, **115**, 7636–7644; (*c*) J. S. Nowick, T. Cao and G. Noronha, *J. Am. Chem. Soc.*, 1994, **116**, 3285–3289.
12 L. Schechinger, PhD Dissertation, University of California, Irvine, 1999.
13 J. S. Nowick, V. Antonovich, G. Noronha and J. W. Ziller, *J. Org. Chem.*, 1995, **60**, 1888–1890.
14 (*a*) E. R. Jarvo and S. J. Miller, *Tetrahedron*, 2002, **58**, 2481–2495; (*b*) P. R. Schreiner, *Chem. Soc. Rev.*, 2003, **32**, 289–296; (*c*) P. I. Dalko and L. Moisan, *Angew. Chem., Int. Ed.*, 2004, **43**, 5138–5175; (*d*) J. Seayad and B. List, *Org. Biomol. Chem.*, 2005, **3**, 719–724.
15 J. S. Nowick, E. M. Smith and G. Noronha, *J. Org. Chem.*, 1995, **60**, 7386–7387.
16 (*a*) J. S. Nowick, D. L. Holmes, G. Mackin, G. Noronha, A. J. Shaka and E. M. Smith, *J. Am. Chem. Soc.*, 1996, **118**, 2764–2765; (*b*) E. M. Smith, D. L. Holmes, A. J. Shaka and J. S. Nowick, *J. Org. Chem.*, 1997, **62**, 7906–7907; (*c*) J. S. Nowick, M. Pairish, I. Q. Lee, D. L. Holmes and J.W. Ziller, *J. Am. Chem. Soc.*, 1997, **119**, 5413–5424; (*d*) J. H. Tsai, A. S. Waldman and J. S. Nowick, *Bioorg. Med. Chem.*, 1999, **7**, 29–38; (*e*) E. Junquera and J. S. Nowick, *J. Org. Chem.*, 1999, **64**, 2527–2531.
17 J. S. Nowick, E. M. Smith, J. W. Ziller and A. J. Shaka, *Tetrahedron*, 2002, **58**, 727–739.
18 J. S. Nowick, J. H. Tsai, Q.-C. D. Bui and S. Maitra, *J. Am. Chem. Soc.*, 1999, **121**, 8409–8410.
19 S. Maitra and J. S. Nowick, in β-Sheet Interactions Between Proteins, *The Amide Linkage: Structural Significance in Chemistry, Biochemistry, and Materials Science*, ed. A. Greenberg, C. M. Breneman and J. F. Liebman, Wiley, New York, 2000, ch. 15.
20 C. P. Hill, J. Yee, M. E. Selsted and D. Eisenberg, *Science*, 1991, **251**, 1481–1485.
21 This diagram was prepared using Molscript and Raster3D: (*a*) P. J. Kraulis, *J. Appl. Crystallogr.*, 1991, **24**, 946–950; (*b*) E. A. Merritt and D. J. Bacon, *Methods Enzymol.*, 1997, **277**, 505–524.
22 (*a*) A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio and R. Tycko, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 16742–16747; (*b*) S. Shivaprasad and R. Wetzel, *Biochemistry*, 2004, **43**, 15310–15317; (*c*) R. Nelson, M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, C. Riekel, R. Grothe and D. Eisenberg, *Nature*, 2005, **435**, 773–778; (*d*) T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 17342–17347; (*e*) A. T. Petkova, W.-M. Yau and R. Tycko, *Biochemistry*, 2006, **45**, 498–512.
23 Y. Dou, P.-F. Baisnée, G. Pollastri, Y. Pécout, J. S. Nowick and P. Baldi, *Bioinformatics*, 2004, **20**, 2767–2777.

24 http://www.igb.uci.edu/servers/icbs/.
25 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986, pp. 125–129.
26 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986, pp. 14–19.
27 (*a*) D. S. Wishart, B. D. Sykes and F. M. Richards, *J. Mol. Biol.*, 1991, **222**, 311–333; (*b*) D. S. Wishart, B. D. Sykes and F. M. Richards, *Biochemistry*, 1992, **31**, 1647–1651.
28 J. H. Tsai, PhD Dissertation, University of California, Irvine, 2000.
29 (*a*) J. S. Nowick, D. L. Holmes, G. Noronha, E. M. Smith, T. M. Nguyen and S.-L. Huang, *J. Org. Chem.*, 1996, **61**, 3929–3934; (*b*) J. S. Nowick, D. L. Holmes, G. Noronha, E. M. Smith, T. M. Nguyen, S.-L. Huang and E. H. Wang, *J. Org. Chem.*, 1998, **63**, 9144.
30 J. S. Nowick, D. M. Chung, K. Maitra, S. Maitra, K. D. Stigers and Y. Sun, *J. Am. Chem. Soc.*, 2000, **122**, 7654–7661.
31 K. D. Stigers, M. R. Koutroulis, D. M. Chung and J. S. Nowick, *J. Org. Chem.*, 2000, **65**, 3858–3860.
32 T. V. Khasanova and J. S. Nowick, Design and Synthesis of Functionalized Analogs of an Unnatural Amino Acid that Mimics a Tripeptide β-Strand, Presented at the 224th National Meeting of the American Chemical Society, Boston, MA, August 18, 2002, paper ORGN 16.
33 J. S. Nowick, K. S. Lam, T. V. Khasanova, W. E. Kemnitzer, S. Maitra, H. T. Mee and R. Liu, *J. Am. Chem. Soc.*, 2002, **124**, 4972–4973.
34 J. S. Nowick and J. O. Brower, *J. Am. Chem. Soc.*, 2003, **125**, 876–877.
35 J. S. Nowick, C. M. Gothard, S.-W. Kang and S. Maitra, in Nanometer-Scale Amino Acids for Biomolecular Nanotechnology, *Understanding Biology Using Peptides, Proceedings of the 19th American Peptide Symposium*, ed. S. E. Blondelle, American Peptide Society, San Diego, CA, 2006, in press.
36 (*a*) J. S. Nowick, C. Gothard, S.-W. Kang and S. Maitra, *iota*-Amino Acids For The Creation of Nanometer-Scale Molecular Architectures, Presented at the 226th National Meeting of the American Chemical Society, New York, NY, September 7, 2003, paper ORGN 65; (*b*) J. S. Nowick, C. Gothard, S.-W. Kang and S. Maitra, A Family of *iota*-Amino Acids For The Creation of Nanometer-Scale Molecular Architectures, Presented at the 38th National Organic Symposium, Bloomington, IN, June 8, 2003, paper A31; (*c*) J. S. Nowick and C. M. Gothard, Nanometer-Scale Molecular Rods, Presented at the 229th National Meeting of the American Chemical Society, San Diego, CA, March, 2005, paper ORGN 732; (*d*) J. S. Nowick, C. M. Gothard, S.-W. Kang and S. Maitra, Organic Approaches to Nanotechnology Through Nanometer-Scale Amino Acids, Presented at the 232nd National Meeting of the American Chemical Society, San Francisco, CA, September 12, 2006, paper ORGN 390.
37 (*a*) J. S. Nowick and D. M. Chung, *Angew. Chem., Int. Ed.*, 2003, **42**, 1765–1768; (*b*) D. M. Chung, Y. Dou, P. Baldi and J. S. Nowick, *J. Am. Chem. Soc.*, 2004, **126**, 3062–3063; (*c*) D. M. Chung and J. S. Nowick, *J. Am. Chem. Soc.*, 2005, **127**, 9998–9999.
38 D. M. Chung, PhD Dissertation, University of California, Irvine, 2004.
39 For examples of studies of side-chain interactions in β-sheets involving complementary approaches, see: (*a*) S. Lifson and C. Sander, *J. Mol. Biol.*, 1980, **139**, 627–639; (*b*) M. A. Wouters and P. M. G. Curmi, *Proteins: Struct., Funct., Genet.*, 1995, **22**, 119–131; (*c*) C. K. Smith and L. Regan, *Science*, 1995, **270**, 980–982; (*d*) E. G. Hutchinson, R. B. Sessions, J. M. Thornton and D. N. Woolfson, *Protein Sci.*, 1998, **7**, 2287–2300; (*e*) Y. Mandel-Gutfreund, S. M. Zaremba and L. M. Gregoret, *J. Mol. Biol.*, 2001, **305**, 1145–1159; (*f*) M. D. Distefano, A. Zhong and A. G. Cochran, *J. Mol. Biol.*, 2002, **322**, 179–188; (*g*) S. J. Russell, T. Blandl, N. J. Skelton and A. G. Cochran, *J. Am. Chem. Soc.*, 2003, **125**, 388–395; (*h*) S. E. Kiehna and M. L. Waters, *Protein Sci.*, 2003, **12**, 2657–2667; (*i*) M. L. Waters, *Biopolymers*, 2004, **76**, 435–445; (*j*) R. M. Hughes and M. L. Waters, *Curr. Opin. Struct. Biol.*, 2006, **16**, 514–524.
40 C. L. Perrin and T. J. Dwyer, *Chem. Rev.*, 1990, **90**, 935–967.
41 (*a*) C. K. Smith and L. Regan, *Acc. Chem. Res.*, 1997, **30**, 153–161; (*b*) F. Blanco, M. Ramírez-Alvarado and L. Serrano, *Curr. Opin. Struct. Biol.*, 1998, **8**, 107–111; (*c*) S. H. Gellman, *Curr. Opin. Chem. Biol.*, 1998, **2**, 717–725; (*d*) E. Lacroix, T. Kortemme, M. L. de la Paz and L. Serrano, *Curr. Opin. Struct. Biol.*, 1999, **9**, 487–493; (*e*) M. Ramírez-Alvarado, T. Kortemme, F. J. Blanco and L. Serrano, *Bioorg. Med. Chem.*, 1999, **7**, 93–103; (*f*) J. Venkatraman, S. C. Shankaramma and P. Balaram, *Chem. Rev.*, 2001, **101**, 3131–3152; (*g*) M. S. Searle, *J. Chem. Soc., Perkin Trans. 2*, 2002, 1011–1020; (*h*) M. S. Searle and B. Ciani, *Curr. Opin. Struct. Biol.*, 2004, **14**, 458–464; (*i*) R. M. Hughes and M. L. Waters, *Curr. Opin. Struct. Biol.*, 2006, **16**, 514–524.
42 Improved folding of β-hairpin structures often leads to uncontrolled aggregation through β-sheet interactions (ref. 41*g*).
43 K. Y. Tsang, H. Diaz, N. Graciani and J. W. Kelly, *J. Am. Chem. Soc.*, 1994, **116**, 3988–4005.
44 (*a*) F. A. Syud, J. F. Espinosa and S. H. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 11577–11578; (*b*) J. F. Espinosa and S. H. Gellman, *Angew. Chem., Int. Ed.*, 2000, **39**, 2330–2333.
45 A. S. Altieri, D. P. Hinton and R. A. Byrd, *J. Am. Chem. Soc.*, 1995, **117**, 7566–7567.
46 E. Ilyina, V. Roongta, H. Pan, C. Woodward and K. H. Mayo, *Biochemistry*, 1997, **36**, 3383–3388.
47 Y. Cohen, L Avram and L Frish, *Angew. Chem., Int. Ed.*, 2005, **44**, 520–554.