REVIEW

Peptide models of protein β-sheets: design, folding and insights into stabilising weak interactions

Mark S. Searle

School of Chemistry, University of Nottingham, University Park, Nottingham, UK NG7 2RD. E-mail: mark.searle@nottingham.ac.uk; Fax: 0115 951 3564; Tel: 0115 951 3567

Received (in Cambridge, UK) 12th March 2001
First published as an Advance Article on the web 7th June 2001

Covering: 1996–2000

1 Introduction

The nature of the stereochemical code that determines whether a given peptide sequence is pre-disposed to adopt a specific conformation has engaged biological chemists for decades.[1] The observation that many proteins can readily fold and unfold reversibly in the test tube in the absence of other determinants shows unambiguously that all the information needed to specify the conformation of many proteins, and subsequently their biological function, is encoded within the primary sequence of amino acids. Cracking the code opens up possibilities for the re-engineering of proteins, for the rational design of novel structures with novel chemical and biochemical functions or structural architectures, and for probing the nature and strength of the weak interactions relevant to a wide range of molecular recognition phenomena in chemistry and biology.

From the mechanistic angle, the pathway by which the polypeptide chain assembles from a random coil "disordered" state to the final active folded protein has been the subject of intense investigation in recent times with development of new techniques for probing rapid kinetic processes at the early stages of the folding process.[2] It is generally accepted that proteins fold *via* a hierarchical mechanism in which the folding protein acquires ever increasing degrees of complexity.[3,4] Thus, folding initiation events occur at the local sequence level and involve residues close together in the amino acid sequence. These nuclei of structure promote interactions between different parts of the sequence leading ultimately to a co-operative rate-limiting step from which the native state emerges. Peptide models have proved extremely valuable in probing the relationship between local sequence information and folded conformation in the absence of the tertiary interactions found in the native state of proteins, allowing intrinsic secondary structure propensities to be investigated in isolation. While α-helical peptides have been studied in great detail, and considerable advances have been made in understanding the energetics of helix formation, in the rational design of α-helical proteins and in protein re-design,[5–13] in contrast, model β-sheets have until recently been neglected, partly because of their tendency to aggregate. However, numerous water-soluble, non-aggregating, monomeric peptides have now been identified and characterised,[14] and these model systems (involving largely natural amino acids) are the focus of this review. Rather than catalogue all the systems described, particular emphasis is placed on the progress being made to unravel the stereochemical code that dictates whether a particular sequence will fold into a monomeric β-sheet in aqueous solution. For reasons of design and chemical synthesis, these are exclusively anti-parallel β-sheets, although others have used non-natural linkers to engineer parallel strand alignment,[15–18] but these are beyond the scope of this review. The majority of work has described β-hairpin motifs, consisting of two anti-parallel β-strands linked by a β-turn sequence (Fig. 1). More recently a number of three-stranded anti-parallel β-sheet motifs have been described. Several factors are of key importance and will be addressed, including (i) the nature and role of the β-turn in promoting and stabilising anti-parallel β-sheet formation; (ii) the role of intrinsic backbone ϕ, ψ propensities of different residues in the pre-organisation of β-strands; (iii) the role of co-operativity in β-sheet folding and stability, and the validity of simple two-state models for folding; (iv) insights into the nature of the stabilising weak interactions and the thermodynamics of folding; (v) the nature and energetic contribution of side chain interactions to β-sheet stability; (vi) the role of co-operative interactions perpendicular to the β-strand direction and their role in the propagation of multi-stranded β-sheets.

2 Early results with β-hairpins derived from native protein sequences

The early focus on peptides excised from native protein structures provided the first insights into autonomously folding β-hairpins, preceding the more rational approach to β-sheet design. Peptides derived from tendamistat,[19] the B1 domain of protein G,[20,21] ubiquitin[22,23] and ferredoxin[24] showed that these sequences could exist in the monomeric form without aggregating, but that in most cases they showed a very limited tendency to fold in the absence of tertiary contacts. The use of organic co-solvents appeared to induce native-like conformation.[21,22,24] Study by Blanco *et al.*[20] of a peptide derived from the B1 domain of protein G provided the first example of native-like folding in water of a fully native peptide sequence. In contrast, the studies of hairpins isolated from the proteins ubiquitin and

DOI: 10.1039/b008074h

ferredoxin, which are structurally homologous to the G B1 domain (all form an α/β-roll fold), showed these to be unfolded in water.[22,24] Recent studies of the native ubiquitin peptide have now revealed evidence for a small population of the folded state in water.[25]

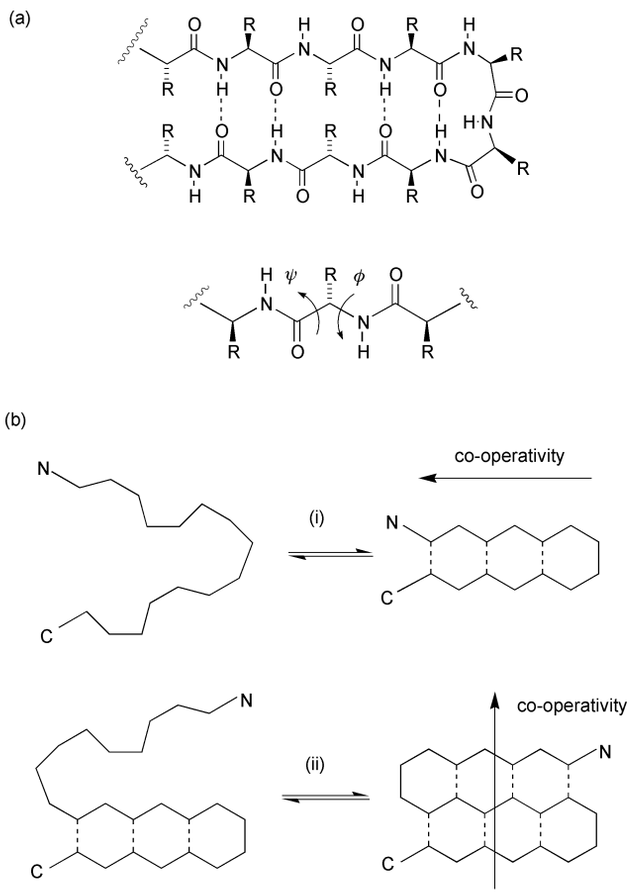

Fig. 1 (a) β-Strand alignment and interstrand hydrogen bonds in a β-hairpin peptide; R groups represent amino acid side chains, main chain ϕ and ψ angles are shown, (b) models for folding of β-hairpin and three-stranded β-sheet peptides demonstrating the possibility of co-operative interactions being propagated parallel (i) and perpendicular (ii) to the β-strand direction.

The apparent lack of evidence for folding of the peptide derived from residues 15–23 of tendamistat (YQSWRYSQA),[19] and from the N-terminal sequence of ubiquitin (MQIFVK-TLTGKTITLEV)[23] led to partial redesign of the sequence to enhance folding by modifying the β-turn sequence by introducing an NPDG type I turn which is the most common type I turn sequence in proteins. Thus, in the former peptide SWRY was replaced by NPDG,[19] and in the latter the G-bulged type I turn TLTGK was replaced by NPDG.[23] By introducing this tight two-residue loop across PD it was envisaged that both hairpins would be stabilised. This was certainly the case, however, the most striking observation was that both peptides folded into non-native conformations with a three-residue G-bulged type I turn (PDG) re-established across the turn (Fig. 2). These initial studies in rational re-design led to strikingly irrational results, prompting a much more systematic approach to β-hairpin design. The above results revealed that the β-turn sequence and its allowed geometry appears to be crucial in dictating the β-strand alignment, but also (particularly in the case of the re-designed ubiquitin peptide) that non-native side chain interactions can have a significant stabilising effect on β-hairpin conformation. From the protein folding view point, it is evident that one important role of the native turn sequence may be to preclude the formation of non-native conformations that may be incompatible with formation of the native state.

3 Rational approach to β-hairpin design

3.1 Role of the β-turn sequence in β-hairpin folding

The rational design of the turn regions of anti-parallel β-sheets has benefited greatly from the statistical analysis of protein structures in the protein data bank (PDB). The early systematic classification of β-turns reveals a wide variety of geometries and sizes of loop.[26–28] For the design of β-hairpins, the emphasis has been on incorporation of the smallest turn sequence possible to limit the entropic destabilisation effects. While two-residue type I and type II turns are generally common in β-turns, the backbone conformation (defined by ϕ and ψ torsion angles) of a two-residue type I or type II turn results in a local left-handed twist, which is not compatible with the

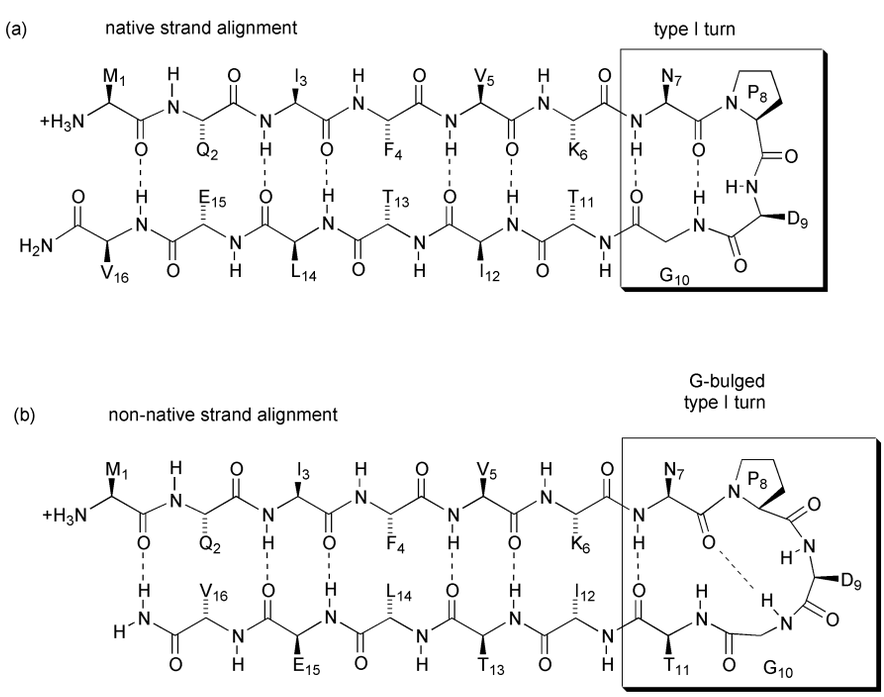

Fig. 2 N-terminal β-hairpin of ubiquitin in which the native TLTGK turn sequence has been replaced by the sequence NPDG. (a) Native strand alignment giving rise to a type I NPDG turn, (b) non-native strand alignment with a G-bulged type I turn (NPDGT).

right-handed twist found in protein β-sheets. However, the diastereomeric type I′ and II′ turns have the opposite ϕ and ψ angles to their type I and type II counterparts giving the former the right-handed twist complementary to the orientation of the β-strands. These conclusions appear to rationalise the above observations with β-turn modifications introduced into the β-hairpins of tendamistat and ubiquitin. The introduced NPDG type I turns are not compatible with the right-handed twist of the β-strands resulting in a refolding to a more flexible G-bulged type I turn.

The work of Haque and Gellman[29] systematically showed the importance of backbone ϕ, ψ angle preferences for the residues in the turn sequence by comparing the stability of a number of β-hairpin peptides derived from the ubiquitin sequence (MQIFVKSXXKTITLVKV) containing either XX = ^{L}Pro-X or ^{D}Pro-X. Replacing ^{L}Pro with ^{D}Pro switches the twist from left-handed (type I or II) to right-handed (type I′ or II′), making the latter compatible with the hairpin conformation. NMR data indicate that each of the ^{D}Pro containing peptides showed a significant degree of folding, whereas the ^{L}Pro analogues appeared to be unfolded. Similar conclusions were drawn from studies of a series of 12-mers containing XG turn sequences, with X = ^{L}Pro or ^{D}Pro.[30]

While replacement of ^{L}Pro with ^{D}Pro results in a stabilising two-residue turn, a number of natural L-amino acids are commonly found in the α_L region of conformational space, compatible with the type I′ or II′ turn conformation. Statistical analyses from a number of groups have identified X-Gly as a favoured type I′ turn. A number of studies have used the NG sequence to design β-hairpin motifs that showed a high population of the folded structure with the required turn conformation and strand alignment.[31–34] The work by de Alba *et al.* identified two possible conformations of the peptide ITSYNGKTYGR. The NOE data appear to be compatible with rapidly interconverting conformations involving a YNGK type I′ turn and an NGKT type II′ turn, each giving a distinct pattern of cross-strand NOEs. Ramirez-Alvarado *et al.*[31] also described an NG-containing 12-mer (RGITVXGKTYGR; X = N), which they subsequently extended to a series of hairpins to examine the correlation between hairpin stability and the database frequency of occurrence of residues in position X.[34] Using NMR and CD measurements they concluded that X = Asn > Asp > Gly > Ala > Ser in promoting hairpin folding, in agreement with the intrinsic ϕ, ψ preferences of these residues. Despite extensive analysis of NG type I′ turns in the PDB, it is still not entirely clear why Asn at the first position is so effective in promoting turn formation. There is no evidence for specific side chain to main chain hydrogen bonds that might stabilise the desired backbone conformation, although specific solvation effects cannot be ruled out.[33]

Evidence that the NG turn is able to nucleate folding in the absence of cross-strand interactions was demonstrated in a truncated analogue of one designed 16-residue β-hairpin sequence KKYTVSINGKKITVSI,[35,36] in which the sequence was truncated to SINGKKITVSI, lacking the N-terminal five residues.[33,37] Evidence from NOE data showed that the turn was significantly populated with interactions observed between Ser6 Hα↔Lys11 Hα and Ile7 NH↔Lys10 NH (Fig. 3) which are only compatible with a folded type I′ turn around NG. Two destabilised β-hairpin mutants (KKYTVSINGKKITKSK with electrostatic repulsion between the N- and C-terminal Lys residues, and KKATASINGKKITVSI with the loss of key hydrophobic residues in one strand) showed no evidence from NMR chemical shift data for cross-strand interactions, however, careful examination of NOE data revealed evidence for NG turn nucleation.[33] Titration with organic co-solvent showed both peptides to fold significantly, indicating that the turn sequence probably already predisposes the peptide to form a β-hairpin but that favourable cross-strand interactions are required for stability and not specificity.

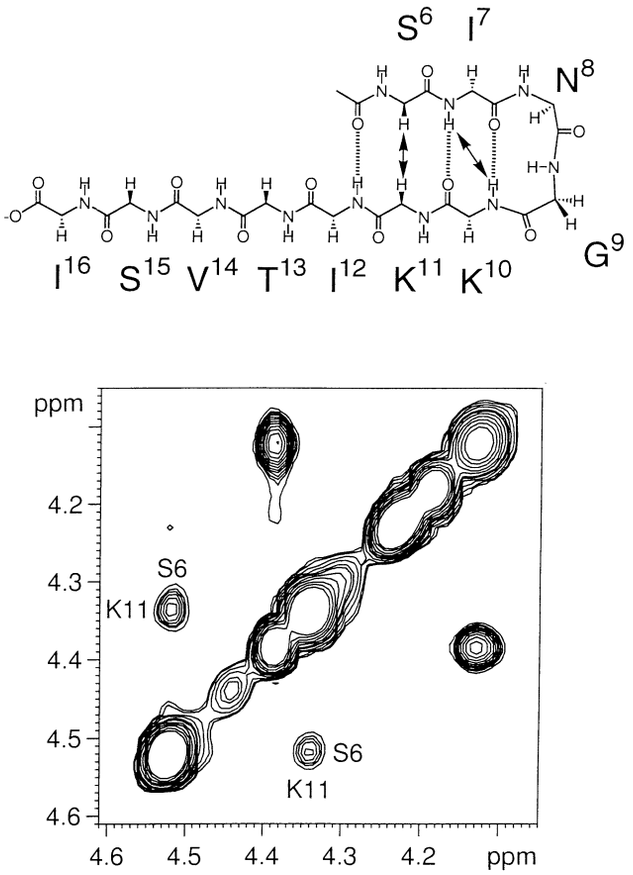

Fig. 3 Conformation of a truncated β-hairpin in which hydrophobic residue on the N-terminal β-strand have been removed. Interstrand NOEs (arrows) in the turn region show that the INGK sequence is adopting a type I′ NG turn. The Ser6 Hα↔Lys11 Hα NOE is illustrated in the portion of the NOESY spectrum of the peptide in aqueous solution. (Reproduced with permission from the *Journal of Molecular Biology*.[33])

This principle was convincingly illustrated by the work of de Alba *et al.*, in a series of six hairpin sequences (10-mers) where strand residues were conserved but turn sequences varied.[32] Using a number of NMR criteria they were able to show that changes in turn sequence could result in a variety of turn conformations including two residue 2 : 2 turns, 3 : 5 turns and 4 : 4 turns (see earlier nomenclature,[26–28]) with different pairings of amino acid side chains. As with the earlier examples cited with the turn modification described for hairpins derived from tendamistat and ubiquitin, the bulged-type I turn (3 : 5 turn) appears to be an intrinsically stable turn with the necessary right-handed twist. Together these data strongly support a model for hairpin folding in which the turn sequence strongly dictates the preferred conformation, and that strand alignment, cross-strand interactions and subsequently conformational stability are dictated by the specificity of the turn.

3.2 Intrinsic ϕ, ψ propensities of amino acids

In attempts to deconvolute the factors that contribute to a particular residue in a particular context adopting a particular conformation, intrinsic backbone ϕ, ψ preferences have been examined using statistically based analyses of high-resolution structures in the PDB. One particularly novel approach presented by Swindells *et al.*, was to determine ϕ, ψ propensities of different residues in coil regions of protein structures, that is, residues not in regular secondary structure (β-sheet or α-helix).[38] The advantage of this approach is that intrinsic conformational preferences can be identified free of the interactions associated with regular secondary structure. The striking observation is that in this context ϕ and ψ angles (see

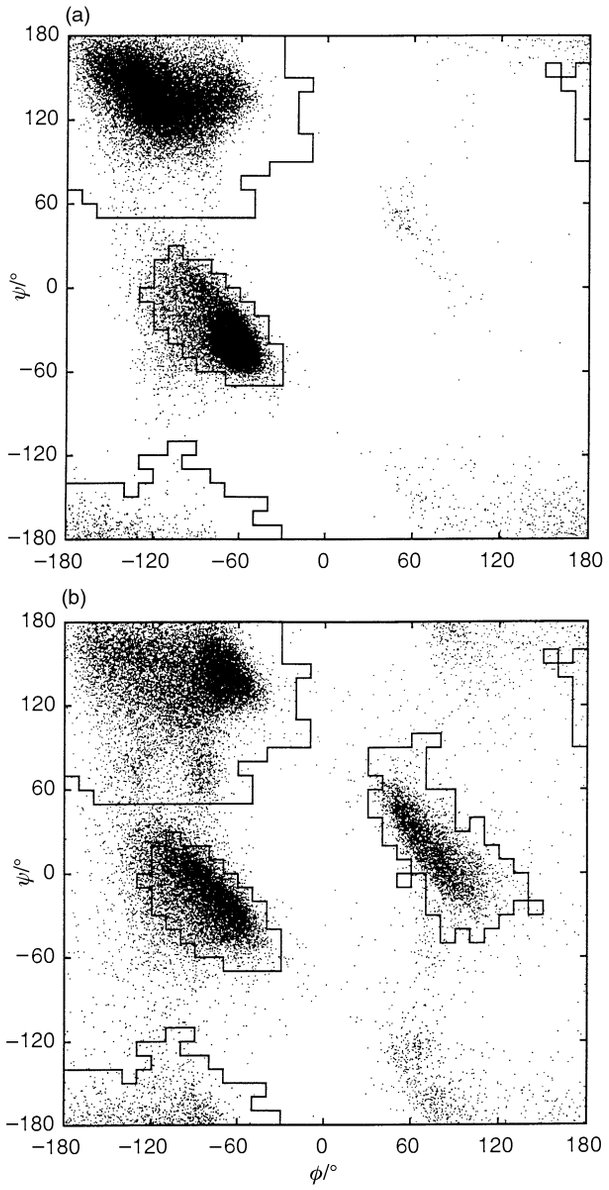

Fig. 4 Ramachandran plots of residue backbone ϕ and ψ angles taken from a database of 512 high resolution protein X-ray structures showing (a) residues in regular β-sheet (ϕ, ψ: −120°, 120°) and α-helix (ϕ, ψ: −60°, 60°), and (b) residues in the irregular coil regions of the same structures (ϕ, ψ: 60°, 0° is the α_L region of conformational space mainly occupied by Gly). The distribution in (b) shows that residues have a natural propensity to occupy the α- and β-regions of conformational space even when they are not involved in regular protein secondary structure. (Reproduced with permission from the *Journal of Molecular Biology*.[41])

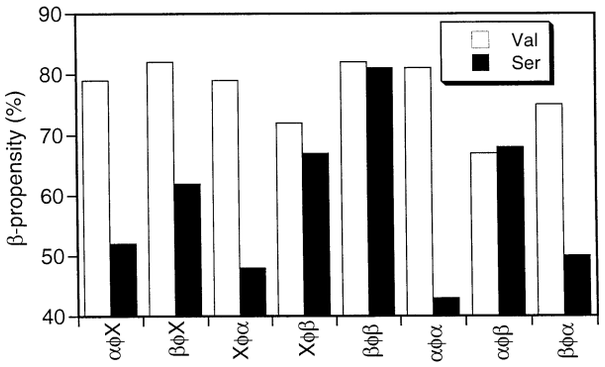

Fig. 5 β-Propensities of Ser and Val (ϕ = S or V) calculated from the data in Fig. 4(b). β-Propensity is a measure of the number of times a particular residue is found in the β-region of the Ramachandran plot as a fraction of the total distribution between α and β-space [β/(α + β)]. The context-dependence of the β-propensity is estimated by considering the nature of the neighbouring residue (X = any residue, α is a residue that prefers the α-helical region—Asp, Glu, Lys or Ser, β is a residue that prefers the β-sheet region—Ile, Val, Phe or Tyr). While Val is relatively insensitive to the nature of the flanking residues, Ser can be forced to adopt a higher β-propensity if it has bulky neighbours. Thus, the intrinsic β-propensity of a particular residue is highly context-dependent.

Fig. 1) are far from randomly distributed, but that most occupy regions of Ramachandran space associated with regular secondary structure (Fig. 4). The observed ϕ, ψ distributions for individual residues has been taken as representative of those found in denatured states of proteins providing the basis of a "random coil" state from which residue-specific NMR parameters ($^3J_{NH-H\alpha}$ and NOE intensities) can be derived as a reference state for folding studies.[39,40] While β-propensity is found to vary significantly from one residue to the next, context-dependent effects appear to play an important part.[41] While V, I, F and Y for example have a high intrinsic preference to be in the β-region of Ramachandran space, since this minimises steric repulsion with flanking residues, their conformation is relatively insensitive to the nature of the flanking residues (Fig. 5). In contrast, small or unbranched side chains have a higher preference for the α-helical conformation, however, this preference can be significantly modulated by its neighbours through a combination of steric and hydrophobic interactions, as well as both repulsive and attractive electrostatic interactions. The general effects for Ser and Val are illustrated in Fig. 5, showing all possible combinations of flanking residues. With Ser, for example, having bulky flanking residues either side with high β-propensity (βSβ), significantly increases the β-propensity of the Ser residue to minimise the steric repulsion between the flanking residues.[41] Thus, intrinsic structural propensities appear to be highly context-dependent.

This statistical framework has been extended to a number of experimental systems to examine the extent to which isolated β-strand sequences (in the absence of secondary structure interactions) are pre-disposed by the primary sequence to adopt an extended β-like conformation. The isolated 8-mer (GKKITVSI), corresponding to the C-terminal β-strand of the hairpin (KKYTVSINGKKITVSI), was examined by NMR analysis of $^3J_{NH-H\alpha}$ values and backbone NOE intensities. Surprisingly, many of these parameters are similar to those for the folded hairpin despite the monomeric nature of the 8-mer.[36,41] In an analogous study of the C-terminal strand of the ubiquitin hairpin described above,[25,42] similarly large deviations of coupling constants and NOE intensities from random coil values suggested that the isolated β-strands are partially pre-organised into an extended conformation supporting a model for hairpin folding which may not require a significant further organisation of the peptide backbone, a factor that may contribute significantly to hairpin stability, and be important to the mechanism of protein folding initiation by restricting the conformational search, leading to faster folding kinetics. Several studies of denaturated states of proteins have also highlighted the influence of neighbouring residues in modulating main chain conformational preferences.[43–46]

3.3 Side chain interactions in the stabilisation of β-hairpins

The origin of the specificity of β-hairpin folding has been discussed above and relates largely to the conformational preferences of the turn sequence. However, the origin of the stability of the folded state has been attributed to interstrand hydrogen bonding and/or hydrophobic interactions, though which dominates is still a matter of debate. Ramirez-Alvarado *et al.*[31] reported that the population of the folded state of the hairpin RGITVNGKTYGR was significantly diminished by replacing residues on the N-terminal strand, and then the C-terminal

strand, by Ala. The loss of stability was attributed firstly to a reduction in hydrophobic surface burial, but also due to the intrinsically lower β-propensity of Ala, the latter contributing through an adverse conformational entropy term. To compensate for this de Alba *et al.*[47] described a family of hairpins derived from the sequence IYSNSDGTWT. The effects of residue substitutions in the first three positions were examined while maintaining the overall β-character of the two strands. Several favourable cross-strand pair-wise interactions were identified that were apparent in earlier, and more recent, protein database analysis of β-sheet interactions.[48,49] For example, Thr–Thr and Tyr–Thr cross-strand pairs produced stabilising interactions, whilst Ile–Thr and Ile–Trp had a destabilising effect. Other studies of Ala substitution in one β-strand have similarly highlighted hydrophobic burial as a key factor in conformational stability,[33] while the observation of large numbers of side chain NOEs have been used as evidence for hydrophobic stabilisation in water.[23,29–31]

In contrast, Constantine *et al.*[50] have used molecular dynamics simulations coupled with NOE restraints to conclude that small β-hairpin peptides are unable to bury a sufficient amount of hydrophobic surface area for this to promote folding. Thus, transiently formed interstrand hydrogen bonds were concluded to play a significant role in stabilisation. This has been given more recent theoretical consideration. While hydrogen bonds in aqueous solution have long been associated with very small free energy contributions, on the basis that a peptide–peptide bond is isoenergetic with a peptide–water hydrogen bond, hydrogen bonds in the solvent-excluded core of proteins are thought to contribute significantly to stability, with a substantial cost associated with burying an unsatisfied hydrogen bond donor or acceptor.[51,52] Sippl *et al.*[53] concluded from an analysis of hydrogen bonding distances and geometries in proteins, that certain separations (around 4 Å) produce a high energy conformation which, despite the low thermodynamic stability of the hydrogen bond, may produce a large kinetic barrier to dissociation which may act in a highly co-operative sense to stabilise elements of secondary structure in proteins, but by implication, to a lesser extent in smaller peptides where the total number of interactions is smaller.

More recent examination of the effects of interstrand salt bridges between the C- and N-terminal residues of a β-hairpin has established a small but measurable effect, despite the solvent exposed nature of the interacting residues.[54] Others have similarly highlighted pH-dependent effects on hairpin stability.[55] In the hairpin KKYTVSINGKKITVSI (described above), the C-terminal carboxylate group of Ile16 was titrated between the free acid form and the conjugate base. This is the only carboxylate group in the molecule and allows the salt bridge between Lys1 and Ile16 to be selectively switched on and off. The pH switch, in combination with a Lys1→Gly1 mutation has enabled the electrostatic and hydrophobic contributions of Lys1 to be characterised from changes in the folded population of the β-hairpin *via* a thermodynamic cycle (Fig. 6). Although the overall contribution of the salt bridge is small (sum of electrostatic and hydrophobic contributions <2 kJ mol^{-1}), the system is sufficiently sensitive to be able to detect small changes in populations and free energies. These energies appear to be consistent with those of salt bridges in surface exposed sites in α-helical peptides and small proteins.[54]

3.4 Aggregation of peptide β-sheets

Numerous water-soluble, non-aggregating, monomeric peptides have now been identified and characterised,[14] as summarised in the above. The earlier difficulties with peptide aggregation and low solubility appear to have been partially overcome. It is difficult to point to hard and fast rules for avoiding problems with aggregation, however, examination of the successfully designed sequences so far reported suggests that incorporation of

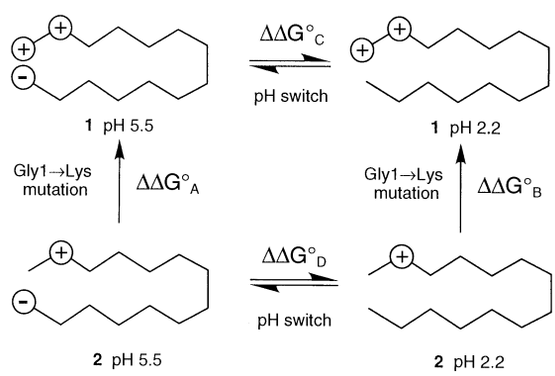

Fig. 6 Thermodynamic cycle used to estimate the energetic contribution of a salt bridge between the N-terminal Lys1 residue and the carboxylate group of the C-terminal Ile16 residue of the β-hairpin peptide KKYTVSINGKKITVSI (peptide **1**); peptide **2** is an analogue with the mutation Lys1→Gly1. Positive charges represent the side chains of Lys1 and Lys2, while the negative charge the carboxylate group of Ile16. The latter is switched between the carboxylate and the free acid by reducing the pH. The population of the folded state in each case enables the various free energy differences to be estimated ($\Delta\Delta G^\circ_A$, $\Delta\Delta G^\circ_B$, $\Delta\Delta G^\circ_C$ and $\Delta\Delta G^\circ_D$) from which the energetic contributions of the salt bridge between Lys1↔Ile16 and Lys2↔Ile16 have been estimated (≈−1 and ≈−0.5 kJ mol^{-1}, respectively). (Reproduced with permission from the *Journal of the American Chemical Society*.[54] Copyright 1999 American Chemical Society.)

solubilising charged residues that lead to mutual repulsion in the aggregated state, appears to be one strategy. Peptide aggregation problems appear to run hand-in-hand with stability of the folded β-sheet; successful design, leading to a well-folded system, invariably results in problems with intermolecular interactions at NMR concentrations.

4 Evidence for two-state co-operative folding of β-hairpin peptides

While the number of β-hairpin model systems is expanding rapidly, the number that have been amenable to quantitative analysis and thermodynamic characterisation remains small. The problems are several-fold: (i) the nature of the "folded" state is not entirely clear cut in terms of the dominant stabilising interactions, (ii) limiting (reference) parameters for the fully folded state are subject to some uncertainty making quantitative analysis of folded populations problematic, and (iii) folding models are based on an assumption of a two-state co-operative process which has not generally been clearly justified and may not be appropriate.

4.1 The nature of the folded state

It is unlikely that the folded state of a model β-hairpin peptide resembles a β-sheet in a native protein, with the former sampling a much larger number of conformations of similar energy stabilised by a fluctuating ensemble of transient interactions. Evidence to suggest that hydrogen bonding contributes significantly to folding in water has not yet been fully justified. IR analysis of the amide I band of a folded β-hairpin (as judged by NMR) does not show significant differences from data on a non-hydrogen bonded short reference peptide in the region expected for β-sheet formation (≈1630 cm^{-1}),[56,57] but this band does appear under aggregating conditions. Similarly, β-hairpins that appear to be well folded on the basis of various NMR criteria seem to be weakly folded by CD analysis.[58] This discrepancy has been attributed largely to weak interstrand hydrogen bonding interactions in the folded state. Indeed, molecular dynamics simulations using ensemble-averaging approaches or time-averaged NOEs tend to de-emphasise the role of hydrogen bonding between the peptide backbone of the two strands, but emphasise the role of hydrophobic side chain interactions.[33,50,59,60] In all cases described, side chain NOEs across the β-strands support such interactions, however, cross-strand

backbone NOEs only imply the possibility of hydrogen bonds since weakly populated folded states lead to only small NH↔ND protection factors in water.

4.2 Quantitative analysis of β-sheet stability

Quantification of the population of the folded structure in solution of β-sheet peptides still presents a challenge, largely as a consequence of uncertainties in limiting values for the fully folded state. Far UV-CD proves to be unreliable as a consequence of the complicating influence of the β-turn conformation and possibly aromatic residues, where present.[61] Added to this, the CD spectrum of β-sheet is intrinsically weak compared with α-helical secondary structure. The use of NMR parameters (Hα chemical shifts, $^3J_{NH-H\alpha}$ values and NOE intensities) to quantify folded populations has been discussed.[14,36,58,61] NMR offers the advantage that several independent parameters can be used in quantitative analysis to provide a consensus picture of the folded state. There still appear to be significant discrepancies between CD analysis and NMR, with peptides that appear to be significantly folded by NMR giving rise to a largely random coil CD spectrum. One interpretation of this observation is that in aqueous solution the peptides fold as a collapsed state with an ill-defined hydrogen bonding network dominated by side chain interactions. Interestingly, in many cases the addition of organic co-solvents changes the CD spectrum dramatically. The interpretation of co-solvent-induced folding is also subject to some uncertainty. Do 2,2,2-trifluoroethanol and methanol actually significantly perturb the equilibrium between the folded and unfolded states (induce folding), or do these solvents exert their influence by changing the nature of the folded state such as to stabilise interstrand hydrogen bonding interactions without significantly changing the folded population? The latter hypothesis would appear to more readily account for the observation of solvent-induced effects on the CD spectrum, and finds some support from studies of cyclic β-hairpin analogues where the folded population is fixed, but whose CD spectrum undergoes large solvent-induced changes.[58]

The use of cyclic β-hairpin analogues has been exploited in a number of studies to generate a fully folded NMR reference state for comparison with the folding of acyclic analogues.[62,63] Backbone cyclisation through amide bond formation, and through use of a disulfide bridge seems to work equally well. Such an approach has recently been used effectively to measure the thermodynamics of folding of a short hairpin carrying a motif of aromatic residues.[63] The cyclic analogues show a much higher stability than their acyclic counterparts, including significant protection from amide H–D exchange due to enforced interstrand hydrogen bonding. While peptide cyclisation seems a worthwhile approach to defining the fully folded state, the effects of conformational constraints (β-turns) at each end of the structure may introduce an overly restrained structure that may have its own limitations when comparing with acyclic analogues.

4.3 Co-operative interactions in β-sheet peptide folding

Folding kinetics for a β-hairpin derived from the C-terminus of the B1 domain of protein G have been described from measurements of tryptophan fluorescence following a laser-induced temperature-jump.[64,65] Kinetic analysis of this peptide (and a dansylated† analogue) reveals a single exponential relaxation process with time constant 3.7 ± 0.3 μs. The data indicate a single kinetic barrier separating folded and unfolded states, consistent with a two-state model for folding. Subsequently, the authors developed a statistical mechanical model based on these observations, describing the stability in terms of a minimal number of parameters: loss of conformational

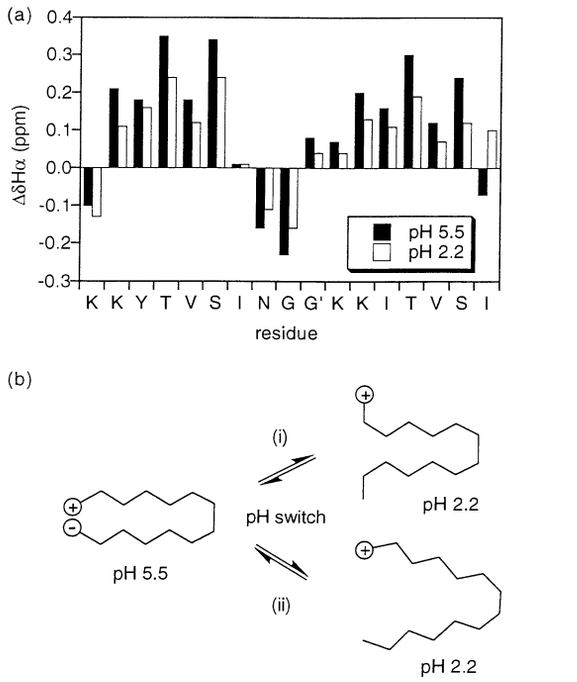

Fig. 7 (a) Hα chemical shift deviations from random coil values for the β-hairpin peptide KKYTVSINGKKITVSI at pH 5.5 and 2.2. At low pH the salt bridge between the C- and N-terminal residues is disrupted resulting in a reduction in the population of the folded state; (b) two scenarios for the effects of changing pH, (i) end-fraying produces local perturbations, (ii) co-operative unfolding affects all residues. (Reproduced with permission from the *Journal of Molecular Biology*.[33])

entropy, backbone stabilisation by hydrogen bonding and formation of a stabilising hydrophobic cluster between three key residues. This model seems sufficient to reproduce all of the features observed experimentally, with a rough, funnel-like energy landscape dominated by two global minima representing the folded and unfolded states. The formation of the hydrophobic cluster appears to be a key folding event. Nucleation by the turn seems most likely, consistent with experimental measurements of loop formation on the timescale of ≈1 μs.[66]

Several other studies have attempted to demonstrate, to at least a reasonable approximation, that β-sheet peptide folding has some of the co-operative features evident in native proteins. What are the criteria for demonstrating co-operative folding? The most readily applicable is that during the thermal melting transition all residues reflect the same melting process, that is, that the fraction of folded peptide reflected by each residue is the same. Several studies have attempted to demonstrate this.[36,54,62,63,67] Similarly, any other change in environmental conditions such as pH or solvent composition should also produce a uniform change in the population of the folded state reflected by each residue. At low pH a salt bridge between the C- and N-terminal residues of the β-hairpin **K**KYTVSING-KKITVSI–$\mathbf{CO_2}^-$ is switched-off, resulting in a small destabilisation of the folded hairpin.[34,54] This can be monitored from changes in Hα chemical shifts. Two scenarios can be envisaged (Fig. 7): (i) loss of the salt bridge results in partial fraying of the ends of the peptide resulting in local perturbations to Hα shifts, and (ii) co-operative unfolding occurs in which all residues (even those in the remote turn sequence) are affected by the interaction between the two ends. The experimental Hα shift changes are summarised in the histograms in Fig. 7 at pH 2.2 and 5.5, which demonstrate that all residues are affected by the pH switch, not just those adjacent to the interaction site, in favour of the co-operative model. Thus, in this model system both temperature- and pH-induced unfolding seems to support the two-state assumption. A key factor in the characterisation of these simple model systems is that the equilibrium constant

† The IUPAC name for dansyl is naphthyl-1-sulfonyl.

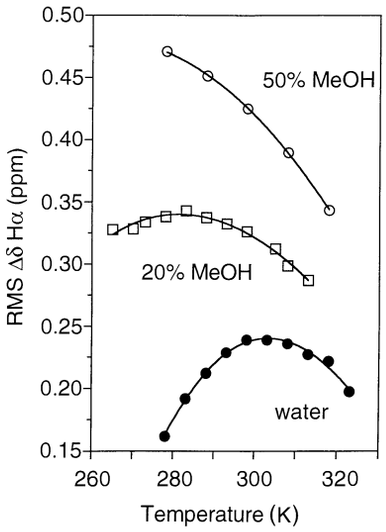

Fig. 8 Temperature-dependent stability of the β-hairpin peptide KKYTVSINGKKITVSI at pH 5.5 in water, 20% methanol and 50% methanol. Temperature is plotted against the RMS value for the deviation of Hα chemical shifts from random coil values, assuming a two-state folding model. In water the peptide shows "hot" and "cold" denaturation, but in 50% methanol the folded population increases at low temperature. The best fits to the three sets of data are shown; in water folding is slightly entropy-driven with a large negative ΔC°_p value, while in 50% methanol folding is strongly enthalpy driven with ΔC°_p close to zero. In 20% methanol the values are intermediate. (Reproduced with permission from the *Journal of the American Chemical Society*.[36] Copyright 1998 American Chemical Society.)

(*K*) for folding is close to unity such that small changes in stability are amplified into the largest possible changes in folded population detected by spectroscopic measurements. Only NMR gives the resolution at the individual residue-level to monitor unfolding at all sites.

5 Thermodynamics of hairpin folding

In the few cases where the two-state model has been justified, thermodynamic data on folding have also been presented based on studies of the temperature-dependence of the folded state. The β-hairpin KKYTVSINGKKITVSI (see above) shows the unique property of cold denaturation, with a maximum in the stability curve occurring at 298 K as judged by changes in Hα chemical shift (Fig. 8).[33,36] Such pronounced curvature is clear evidence for entropy-driven folding with a significant change in heat capacity. Both of these thermodynamic signatures are hallmarks of the hydrophobic effect contributing strongly to hairpin stability. Further examination of folding in the presence of methanol co-solvent shows that the signature changes such that folding becomes strongly enthalpy-driven, and that in 50% aqueous methanol a linear temperature–stability profile is indicative of the absence of any significant contribution of the hydrophobic effect to folding.[36] The population of the folded state appears to be enhanced by methanol, reflecting similar observations in helical peptides where the phenomenon has been attributed to the effects of the co-solvent destabilising the unfolded peptide chain so promoting hydrogen bonding interactions in the folded state.

In studies (both calorimetric and NMR) of the folding of the C-terminal hairpin from the B1 domain of protein G, enthalpy-driven folding is observed in water.[67] In contrast with the above data, where the stabilising hydrophobic interactions involved aliphatic side chains, here an aromatic cluster is responsible for folding. This begs the question whether the latter π–π interactions lead to stabilisation through fundamentally electrostatic interactions rather than through the hydrophobic effect. Studies of a designed β-hairpin system, also carrying the same motif of three aromatic residues, report qualitatively similiar enthalpy-

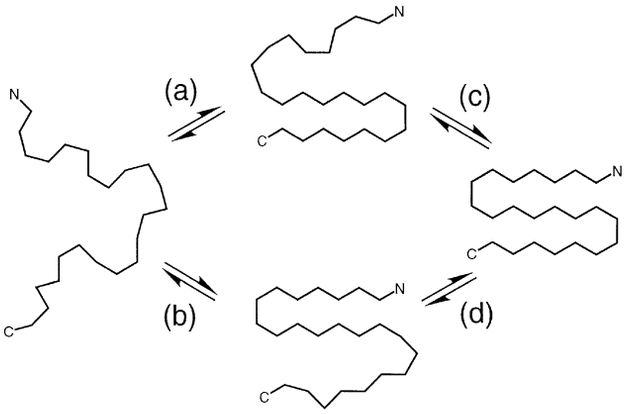

Fig. 9 Four-state model for the folding of a three-stranded anti-parallel β-sheet peptide: (a) folding of the C-terminal β-hairpin, and (b) folding of the N-terminal β-hairpin, (c) and (d) the preformed hairpins can act as templates for the folding of the third strand. (Reproduced with permission from the *Journal of the American Chemical Society*.[68] Copyright 2000 American Chemical Society.)

driven folding,[63] while the thermodynamics of the N-terminal hairpin component of a designed three-stranded anti-parallel β-sheet enables similar conclusions to be drawn[68] (see further below). Espinosa and Gellman[63] suggest that the difference in the thermodynamic signature for aliphatic *versus* aromatic side chain interactions may have its origins in enthalpy–entropy compensation effects. They propose, based on a hypothesis of Diederich *et al.*,[69] but also expounded by others,[70] that enthalpy-driven binding may be fundamentally a consequence of tighter interfacial interactions, giving rise to stronger electrostatic (enthalpic) interactions. The limited data available from such systems suggest that the thermodynamic driving force for β-hairpin folding is highly dependent on the nature of the side chain interactions involved. Further quantitative analysis of peptide folding is required to substantiate these hypotheses.

6 Triple-stranded β-sheet peptides

The logical extension of the earlier studies on β-hairpin peptides was to design three-stranded anti-parallel β-sheets. The design principles are largely those already discussed, focusing on the importance of turn sequence and favourable motifs of interacting side chains. An overriding question concerns the extent to which co-operative interactions perpendicular to the strand direction are important in stabilising these structures[58] [see Fig. 1(b)]; in other words, how good is a pre-organised β-hairpin motif at templating the interaction of a third strand?[71] Several studies have attempted to address this important question. The earliest study described a 24-residue peptide incorporating two NG turns (KKFTLSINGKKYTISNGKTYITGR) that showed little evidence for folding in water but was significantly stabilised in aqueous methanol solutions.[72,73] By comparison with a 16-residue β-hairpin analogue consisting of the same C-terminal sequence (GKKYTISNGKTYITGR), it was possible to show that the Hα shift perturbations for the C-terminal β-hairpin were greater in the presence of the interactions of the third strand. Subsequent design strategies, incorporating the ^{D}Pro-Gly loop, together with other NG-containing turn sequences have illustrated that it is possible to design structures that fold in water[74–76] that show some degree of co-operative interaction between strands. Several other studies have reported peptides that fold to three- (and four-) stranded β-sheets, but in organic solvents.[77]

The exact nature of the folded state in aqueous solution again remains uncertain. However, it would seem that folding is not highly co-operative and that the predominant folded species is in equilibrium with populations of the individual hairpins (see Fig. 9). The work of Kortemme *et al.*[75] showed a sigmoidal melting curve by monitoring Trp fluorescence, although

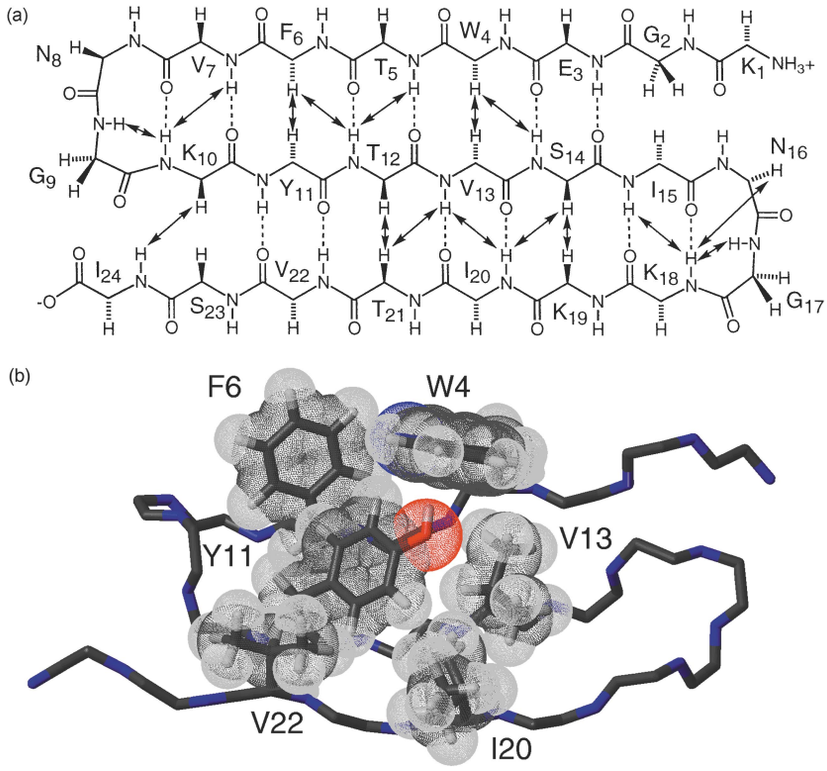

Fig. 10 (a) Amino acid sequence of a designed three-stranded anti-parallel β-sheet showing β-strand alignment and putative interstrand hydrogen bonds based on the pattern of observed NOEs; (b) calculated structure from NOE restraints showing the clustering of hydrophobic residues (particularly the aromatic residues W4, F6 and Y11) on one face of the sheet.

it is unclear whether this reflects a single two-state unfolding of their three-stranded sheet or the unfolding of one of the hairpin components. Schenck and Gellman[74] demonstrated co-operative interactions using their ^{D}ProGly to ^{L}ProGly switch, the latter destabilising one hairpin component selectively. From chemical shift analysis they showed that the individual β-hairpins are co-operatively stabilised by the presence of the third strand. de Alba *et al.*[76] also reported a designed β-sheet system, but were unable to find convincing evidence for co-operative stabilisation of either hairpin by the third strand.

A more recent quantitative study of co-operative interactions was based on a designed system incorporating an earlier β-hairpin with a third strand capable of forming a stabilising motif of interstrand interactions between aromatic residues (Fig. 10),[68] as described in earlier studies of native and designed β-hairpin sequences.[63] The earlier study of the isolated C-terminal β-hairpin showed cold denaturation,[36] approximating to two-state unfolding. In the designed three-stranded system, the same hairpin component shows the same cold denaturation profile, however, the N-terminal hairpin (sharing a common central strand) showed increased folding at low temperature. While the first process is characterised by entropy-driven folding, the latter is enthalpy-driven (Fig. 11). Clearly, two different thermodynamic profiles are not consistent with a single two-state folding model, but the data could be rationalised in terms of a four-state model in which the individual hairpins with an unfolded tail are also populated (Fig. 9). Examination of the folding of the isolated C-terminal hairpin, and comparison of the data with those of the three-stranded analogue, shows good evidence that the C-terminal hairpin is co-operatively stabilised by the interaction of the third strand, even though overall folding is not highly co-operative. The data in Fig. 11 show the temperature-dependence of the splitting of the Gly Hα resonances in the two NG turns. Larger values indicate a higher folded population. The splitting for G17 [see Fig. 10(a)] is greater in the three-stranded sheet than for the isolated C-terminal hairpin, while many Hα resonances are further downfield shifted than in the isolated hairpin. The

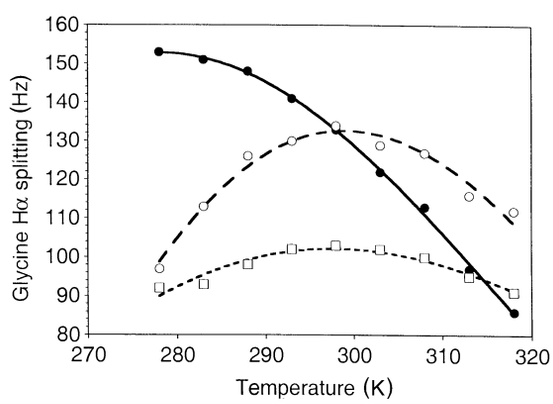

Fig. 11 Temperature-dependent stability profiles for the various β-hairpin components of the three-stranded sheet shown in Fig. 9 using the Hα splitting of Gly9 and Gly17 in the two β-turns: Gly9 solid line, Gly17 open circles, Gly17 (open squares) in the isolated C-terminal hairpin peptide corresponding to residues 9–24. A larger Gly Hα splitting indicates a higher population of the folded hairpin. Different profiles for Gly9 and Gly17 in the three-stranded sheet show that the peptide cannot be folding *via* a simple two-state model involving only random coil and fully folded peptide. The data have been fitted assuming the four-state model shown in Fig. 9. (Reproduced with permission from the *Journal of the American Chemical Society*.[68] Copyright 2000 American Chemical Society.)

temperature-dependence of the stability shows the C-terminal hairpin in both cases to undergo cold denaturation. The N-terminal hairpin carrying the aromatic motif of residues increases in population at low temperature; fitting the data shows the former to be entropy-driven and the latter enthalpy-driven.[68] Entropic effects seem to be the likely explanation for the small co-operative stabilisation effect on the C-terminal hairpin (<2 kJ mol^{-1}), with each hairpin providing a possible template against which the third strand can interact. With one strand pre-organised, the entropic cost of association of an additional strand is largely confined to the associating strand.[71,74] The nature of the folded state is unlikely to compare

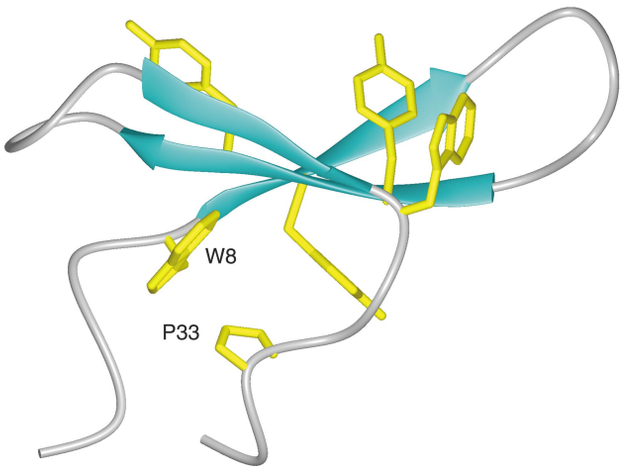

Fig. 12 NMR structure of the WW domain of the formin binding protein (protein data bank accession code: 1EOI), consisting of a three-stranded anti-parallel β-sheet motif; conserved residues are shown with tertiary contacts evident between W8 and P33.

with that of a β-sheet in a native protein, more likely, hydrophobic contacts between side chains stabilise a collapsed conformation where interstrand hydrogen bonds may play a minor stabilising role. These "loosely" defined interactions between side chains, rather than a native-like "crystalline" array of hydrogen bonds, may explain why co-operative interactions have only a small effect on overall stability because only a small energy barrier separates the folded from partially or fully unfolded states.

7 β-Sheet motifs in native proteins

One important characteristic of native proteins is their ability to fold in a highly co-operative fashion. One would expect that much simpler model systems such as β-sheet peptides, which lack defining tertiary interactions and which are held together by a much smaller number of weak interactions, are unlikely to show this property to any pronounced degree. The limited number of three-stranded β-sheet peptides so far described (typically <25 residues) confirms this, although some evidence for co-operative interactions has certainly been presented. Purely on entropic grounds, the addition of a third strand to an existing β-hairpin template would be expected to be less unfavourable than the folding of the β-hairpin where both strands require unfavourable conformational entropy changes to occur on folding. In all cases to date, it seems that designed three-stranded β-sheet structures are in equilibrium with their partially folded β-hairpin components (see Fig. 9: four-state model). This is an indication that the energy landscape is relatively flat unlike that for native proteins where there is a large energy difference between the single native state and other conformations. It is interesting to look at examples of the smallest known β-sheet proteins and make comparisons with the designed systems. The so-called WW domains[78–80] form a single folded motif consisting of three strands of anti-parallel β-sheet but with well-defined tertiary interactions arising from folding back of the N- and C-termini to form a small compact hydrophobic core that allows the protein to fold co-operatively despite its small size (some as small as 35 residues) (Fig. 12). In contrast, the outer surface protein A (OspA) is a considerably larger protein with an unusual structure defined by two β-sheet-rich globular domains linked by a single-layer anti-parallel β-sheet whcih is solvent exposed on both faces (Fig. 13).[81] This connecting region of β-sheet appears to be highly stable and formed co-operatively with the rest of the protein. Thus, the OspA system demonstrates that tertiary contacts are not the sole determinant of co-operative folding. Further, insertion of a duplicate hairpin into the native single-layer β-sheet structure results again in a highly stable, co-operatively folded protein

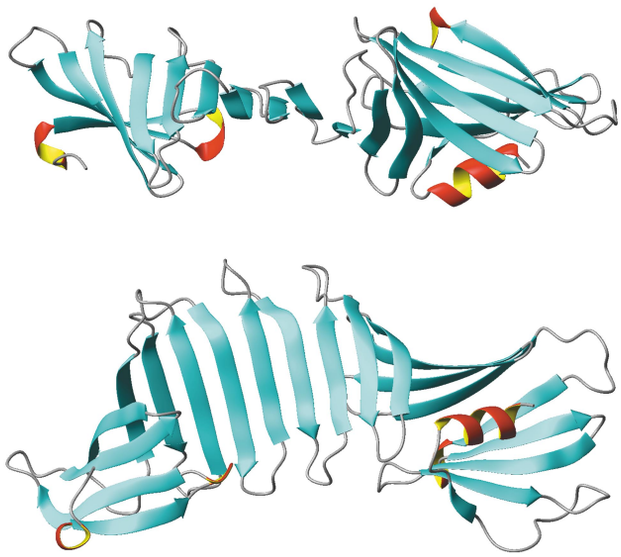

Fig. 13 X-Ray structure of the outer surface protein A (OspA) from two orientations showing the single layer β-sheet connecting two globular domains (PDB code: 1osp).

with the inserted β-strands lacking any tertiary contacts.[82] More recent studies have investigated the conformational properties of the β-hairpin peptide excised from the centre of the OspA β-sheet. The peptide is largely unfolded in solution but shows a time-dependent tendency to aggregate, ultimately forming amyloid-like fibrils.[83] Thus, the conclusion from this study is that the size of the array of weakly interacting β-strands has a significant impact on the overall stability of the sheet,[84] and that model systems of limited size, involving a limited set of weak interactions, and lacking tertiary contacts, are unlikely to deliver a single low energy folded conformation that shows the characteristics of these extended β-sheet structures.

8 Concluding remarks: weak interactions and stabilisation of peptide β-sheets

Apart from the desire to be able to design molecules to order with specific tailored properties, the model β-sheet systems described in this review have enabled us to make some progress in testing our understanding of the basic design principles helped in no small part by the statistical analysis of the ever-growing number of protein structures. This field is, however, lacking the rigorous quantitative approaches to measuring β-sheet stability that are now well-established in the field of α-helical model systems. Even the simplest of anti-parallel β-sheets (β-hairpins) contain several layers of complexity, requiring dissection of both the β-turn and β-strand contributions to the stability of the folded conformation, and an understanding of the nature of the stabilising interactions, the influence of intrinsic secondary structure propensities and their context-dependence. Some progress has been made in understanding the underlying stereochemical code that dictates one conformation and not another. The physical techniques, particularly NMR, are in place and are being used to probe folding in solution at the individual residue level, however, apart from a few notable exceptions, a thermodynamic description of the folded state is largely lacking. The whole area of β-sheet peptide design would benefit significantly from a more quantitative approach to understanding the nature of the stabilising weak interactions involved, and the thermodynamics and kinetics of β-sheet formation.

9 Acknowledgements

I thank all members of my research group, past and present, for their contributions to the research in my own laboratory. I am

grateful to the BBSRC and EPSRC of the UK for research funding.

References

1 B. Honig, *J. Mol. Biol.*, 1999, **293**, 283.
2 D. J. Brockwell, D. A. Smith and S. E. Radford, *Curr. Opin. Struct. Biol.*, 2000, **10**, 16.
3 R. L. Baldwin and G. Rose, *Trends Biochem. Sci.*, 1999, **24**, 26.
4 R. L. Baldwin and G. Rose, *Trends Biochem. Sci.*, 1999, **24**, 77.
5 A. Chakrabartty and R. Baldwin, *Adv. Protein Chem.*, 1995, **46**, 141.
6 V. Munoz and L. Serrano, *J. Mol. Biol.*, 1995, **245**, 275.
7 K. T. O'Neil and W. F. Degrado, *Science*, 1990, **250**, 646.
8 R. Hill and W. Degrado, *J. Am. Chem. Soc.*, 1998, **120**, 1138.
9 V. Munoz, P. Cronet, E. Lopez-Hernandez and L. Serrano, *Folding Des.*, 1996, **1**, 167.
10 R. T. Sauer and W. A. Lim, *Curr. Opin. Struct. Biol.*, 1992, **2**, 46.
11 A. Viguera, V. Villegas, F. Aviles and L. Serrano, *Folding Des.*, 1997, **2**, 23.
12 S. Dalal, S. Balasubramanian and L. Regan, *Nat. Struct. Biol.*, 1997, **4**, 548.
13 D. Jones, C. Moody, J. Uppenbrink, J. Viles, P. Doyle, C. Harris, L. Pearl, P. Sadler and J. Thornton, *Proteins: Struct.*, *Funct.*, *Genet.*, 1996, **24**, 502.
14 S. H. Gellman, *Curr. Opin. Chem. Biol.*, 1998, **2**, 717.
15 D. S. Kemp, B. R. Bowen and C. C. Meundel, *J. Org. Chem.*, 1990, **55**, 4650.
16 J. S. Nowick, E. M. Smith and M. Pairish, *Chem. Soc. Rev.*, 1996, 401.
17 H. Diaz, K. Y. Tsang, D. Choo, J. R. Espina and J. W. Kelly, *J. Am. Chem. Soc.*, 1993, **115**, 3790.
18 J. S. Nowick and S. Insaf, *J. Am. Chem. Soc.*, 1997, **119**, 10903.
19 F. J. Blanco, M. A. Jimenez, J. Herranz, M. Rico, J. Santoro and J. L. Nieto, *J. Am. Chem. Soc.*, 1993, **115**, 5887.
20 F. J. Blanco, G. Rivas and L. Serrano, *Nat. Struct. Biol.*, 1994, **1**, 584.
21 F. J. Blanco, M. A. Jimenez, A. Pineda, M. Rico, J. Santoro and J. L. Nieto, *Biochemistry*, 1994, **33**, 6004.
22 J. P. L. Cox, P. A. Evans, L. C. Packman, D. H. Williams and D. N. Woolfson, *J. Mol. Biol.*, 1993, **234**, 483.
23 M. S. Searle, D. H. Williams and L. C. Packman, *Nat. Struct. Biol.*, 1995, **2**, 999.
24 M. S. Searle, R. Zerella, D. H. Williams and L. C. Packman, *Protein Eng.*, 1996, **9**, 559.
25 R. Zerella, P. A. Evans, J. M. C. Ionides, L. C. Packman, B. W. Trotter, J. P. Mackay and D. H. Williams, *Protein Sci.*, 1999, **8**, 1320.
26 B. L. Sibanda and J. M. Thornton, *Nature*, 1986, **316**, 170.
27 B. L. Sibanda, T. L. Blundell and J. M. Thornton, *J. Mol. Biol.*, 1989, **206**, 759.
28 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207.
29 T. S. Haque and S. H. Gellman, *J. Am. Chem. Soc.*, 1997, **119**, 2303.
30 H. E. Stanger and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 4236. (PG *versus* NG.)
31 M. Ramirez-Alvarado, F. J. Blanco and L. Serrano, *Nat. Struct. Biol.*, 1996, **3**, 604.
32 E. de Alba, M. A. Jimenez and M. Rico, *J. Am. Chem. Soc.*, 1997, **119**, 175.
33 S. R. Griffiths-Jones, A. J. Maynard and M. S. Searle, *J. Mol. Biol.*, 1999, **292**, 1051.
34 M. Ramirez-Alvarado, F. J. Blanco, H. Niemann and L. Serrano, *J. Mol. Biol.*, 1997, **273**, 898.
35 A. J. Maynard and M. S. Searle, *Chem. Commun.*, 1997, 1297.
36 A. J. Maynard, G. J. Sharman and M. S. Searle, *J. Am. Chem. Soc.*, 1998, **120**, 1996.
37 S. R. Griffiths-Jones, A. J. Maynard, G. J. Sharman and M. S. Searle, *Chem. Commun.*, 1998, 789.
38 M. B. Swindells, M. W. Macarthur and J. M. Thornton, *Nat. Struct. Biol.*, 1995, **2**, 596.
39 K. M. Fiebig, H. Schwalbe, M. Buck, L. J. Smith and C. M. Dobson, *J. Phys. Chem.*, 1996, **100**, 2661.
40 L. J. Smith, K. A. Bolin, H. Schwalbe, M. W. Macarthur, J. M. Thornton and C. M. Dobson, *J. Mol. Biol.*, 1996, **255**, 494.
41 S. R. Griffiths-Jones, G. J. Sharman, A. J. Maynard and M. S. Searle, *J. Mol. Biol.*, 1998, 284, 1597.
42 M. Jourdan, S. R. Griffiths-Jones and M. S. Searle, *Eur. J. Biochem.*, 2000, **267**, 3539.
43 D. Neri, M. Billeter, G. Wider and K. Wuthrich, *Science*, 1992, **257**, 1559.
44 M. Frank, G. M. Clore and A. Gronenborn, *Protein Sci.*, 1995, **4**, 2605.
45 D. Buckler, E. Haas and H. Sheraga, *Biochemistry*, 1995, **34**, 15965.
46 D. Shortle, *FASEB J.*, 1996, **10**, 27.
47 E. de Alba, M. Rico and M. A. Jimenez, *Protein Sci.*, 1997, **6**, 2548.
48 M. A. Wouters and P. M. G. Curmi, *Proteins: Struct.*, *Funct.*, *Genet.*, 1995, **22**, 119.
49 E. G. Hutchinson, R. B. Sessions, J. M. Thornton and D. N. Woolfson, *Protein Sci.*, 1998, **7**, 2287.
50 K. L. Constantine, L. Mueller, N. H. Andersen, H. Tong, C. F. Wandler, M. S. Friedrichs and R. E. Bruccoleri, *J. Am. Chem. Soc.*, 1995, **117**, 10841.
51 F. Avbelj and J. Moult, *Biochemistry*, 1995, **34**, 755.
52 F. Avbelj and L. Mele, *J. Mol. Biol.*, 1998, **279**, 665.
53 M. Sippl, M. Ortner, M. Jaritz, P. Lackner and H. Flockner, *Folding Des.*, 1996, **1**, 289.
54 M. S. Searle, S. R. Griffiths-Jones and H. Skinner-Smith, *J. Am. Chem. Soc.*, 1999, **121**, 11615.
55 E. de Alba, F. J. Blanco, M. A. Jimenez, M. Rico and J. L. Nieto, *Eur. J. Biochem.*, 1995, **233**, 283.
56 C. S. Colley, S. R. Griffiths-Jones, M. W. George and M. S. Searle, *Chem. Commun.*, 2000, 593.
57 J. L. R. Arrondo, F. J. Blanco, L. Serrano and F. M. Goni, *FEBS Lett.*, 1996, **384**, 35.
58 E. Lacroix, T. Kortemme, M. Lopez de la Paz and L. Serrano, *Curr. Opin. Struct. Biol.*, 1999, 487.
59 B. Ma and R. Nussinov, *J. Mol. Biol.*, 2000, **296**, 1091.
60 H. Wang and S.-S. Sung, *Biopolymers*, 1999, **50**, 763.
61 M. Ramirez-Alvarado, T. Kortemme, F. J. Blanco and L. Serrano, *Bioorg. Med. Chem.*, 1999, **7**, 93.
62 F. A. Synd, J. F. Espinosa and S. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 11577.
63 J. F. Espinosa and S. H. Gellman, *Angew. Chem.*, *Int. Ed.*, 2000, **39**, 2330.
64 V. Munoz, P. A. Thompson, J. Hofrichter and W. A. Eaton, *Nature* 1997, **390**, 196.
65 V. Munoz, E. R. Henry, J. Hofrichter and W. A. Eaton, *Proc. Natl. Acad. Sci.*, *USA*, 1998, **95**, 5872.
66 S. J. Hagen, J. Hofrichter, A. Szabo and W. A. Eaton, *Proc. Natl. Acad. Sci.*, *USA*, 1996, **93**, 11615.
67 S. Honda, N. Kobayashi and E. Munekata, *J. Mol. Biol.*, 2000, **295**, 269.
68 S. R. Griffiths-Jones and M. S. Searle, *J. Am. Chem. Soc.*, 2000, **122**, 8350.
69 D. B. Smithrud, T. B. Wyman and F. Diederich, *J. Am. Chem. Soc.*, 1991, **113**, 5420; D. B. Smithrud and F. Diederich, *J. Am. Chem. Soc.*, 1990, **112**, 339.
70 G. J. Sharman, M. S. Searle, B. Benhamu, P. Groves and D. H. Williams, *Angew. Chem.*, *Int. Ed.*, 1995, **34**, 1483; M. S. Searle, G. J. Sharman, P. Groves, B. Benhamu, D. A. Beauregard, M. S. Westwell, R. J. Dancer, A. J. Maguire, A. C. Try and D. H. Williams, *J. Chem. Soc.*, *Perkin Trans. 1*, 1996, 2781.
71 A. V. Finkelstein, *Proteins*, 1991, **9**, 23.
72 G. J. Sharman and M. S. Searle, *Chem. Commun.*, 1997, 1955.
73 G. J. Sharman and M. S. Searle, *J. Am. Chem. Soc.*, 1998, **120**, 5291.
74 H. L. Schenck and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 4869.
75 T. Kortemme, M. Ramirez-Alvarado and L. Serrano, *Science*, 1998, **281**, 253.
76 E. de Alba, J. Santoro, M. Rico and M. A. Jimenez, *Protein Sci.*, 1999, **8**, 854.
77 C. Das, S. Raghothama and P. Balaram, *J. Am. Chem. Soc.*, 1998, **120**, 5812; C. Das, S. Raghothama and P. Balaram, *Chem. Commun.*, 1999, 967.
78 M. J. Macias, V. Gervais, C. Civera and H. Oschkinat, *Nat. Struct. Biol.*, 2000, **7**, 375.
79 E. K. Koepf, H. M. Petrassi, G. Ratnaswamy, M. E. Huff, M. Sudol and J. W. Kelly, *Biochemistry*, 1999, **38**, 14338.
80 E. K. Koepf, H. M. Petrassi, M. Sudol and J. W. Kelly, *Protein Sci.*, 1999, **8**, 841.
81 T.-N. Pham, A. Koide and S. Koide, *Nat. Struct. Biol.*, 1998, **5**, 115.
82 S. Koide, X. Huang, K. Link, A. Koide, Z. Bu and D. M. Engelman, *Nature*, 2000, **403**, 456.
83 S. Ohnishi, A. Koide and S. Koide, *J. Mol. Biol.*, 2000, **301**, 477.
84 M. S. Westwell, M. S. Searle and D. H. Williams, *J. Mol. Recogn.*, 1996, **9**, 88.