Pergamon

Tetrahedron 54 (1998) 15793–15819

TETRAHEDRON

Design, Synthesis, Structure and Properties of an α-Helix Cap Template derived from *N*-[(2*S*)-2-chloropropionyl]-(2*S*)-Pro-(2*R*)-Ala-(2*S*,4*S*)-4-thioPro-OMe which Initiates α-Helical Structures.

David Gani*, Arwel Lewis, Trevor Rutherford, John Wilkie, Iain Stirling, Thierry Jenn and Martin D. Ryan.

School of Chemistry and Centre for Biomolecular Science, The Purdie Building, The University, St. Andrews, Fife, KY16 9ST, UK

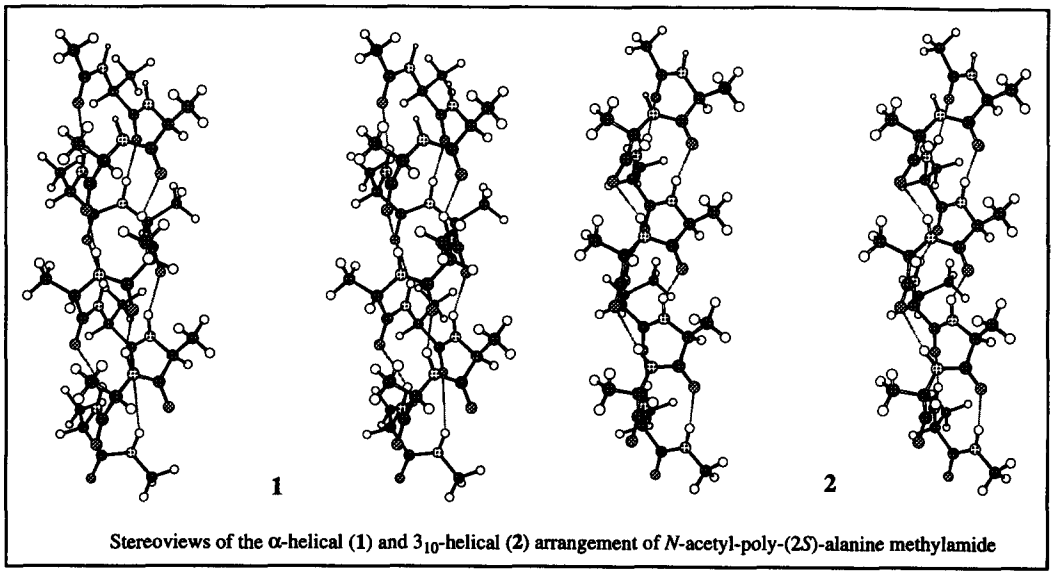
Received 7 August 1998; accepted 20 October 1998

Abstract: *A strategy based upon removing the requirement for all of the carbonyl dipoles to align at the same time in the transition state leading to the cyclisation of N-[(2S)-2-chloropropionyl]-(2S)-Pro-(2R)-Ala-(2S,4S)-4-thioPro-OMe to a Zimm-Bragg type α-helix peptide intitator template was successful. Each amide bond of the 12-membered macrocyclic template existed in the trans-rotomeric form. Derivatives of the template were prepared by extending the C-terminus and these were characterised by NMR spectroscopy and restrained simulated annealing. In deuterochloroform solution at low temperature, separate sets of NMR signals were observed for two rapidly interconverting helical conformational isomers of the thioether macrocycle which possessed an appended trialkylammonium ion. A similar time-averaged conformation was also observed in aqueous solution. At –80 °C in d_2-dichloromethane the rate of conformational exchange was slowed sufficiently to obtain resonance assignments and NOE data separately for each isomer. In the minor isomer (40%), the four carbonyl oxygen hydrogen-bond acceptors of the template are aligned in an α-helical conformation and in the major conformer the Pro² carbonyl dipole was anti-aligned with the other three dipoles. Thus, the conformers differ in the orientation of one carbonyl group. Molecular modelling calculations showed that the minor isomer was stabilised by coulombic interactions between the trialkylammonium salt and the carbonyl group dipole moments.*

Proteins are organised into secondary structural elements through hydrogen bonding networks in several different manners.[1] The two most common arrangements are helices and sheets.For the helices, there are two major structural types both of which are right-handed. For the more common α-helices there are 3.6 residues per turn, the pitch is 5.4 Å and hydrogen bonding occurs in networks forming 13-membered rings between the carboxamide O-atom of residue *i* and the NH moiety of residue (*i* + 4). Thus, the α-helix is also referred to as the 3.6_{13} helix **1**. The other helical form is derived from the type III bend in which hydrogen bonding occurs in 10-membered rings between the carboxamide O-atom of residue *i* and the NH moiety of residue (*i* + 3). If this type III bend is propagated, the resulting structure is called a 3_{10} helix **2**. The 3_{10} helix is tighter, less stable and occurs in proteins ten times less frequently.[2] However, in isolated short polypeptides, the 3_{10} helix can be the most stable form of secondary structure and it is frequently observed at the *C*-terminal end of an α-helix, see below. Whilst Pauling proposed that 3.6_{13} helices should exist no less than sixty years ago,[3] it remains a fact that, in general, the α-helix is not a stable conformation for short polypeptides. It is now evident that without additional contraints, the sum of the largely enthalpic advantages of forming H-bonds within the helix, together

0040-4020/98/$ - see front matter
PII: S0040-4020(98)00991-0

with favourable intra-helix side-chain interactions, is normally smaller than the entropy loss in forming the helix combined with the enthalpic loss associated with dipole-dipole repulsions.



Stereoviews of the α-helical (1) and 3_{10}-helical (2) arrangement of *N*-acetyl-poly-(2*S*)-alanine methylamide

Thus, peptides of fewer than about 20 residues usually exist as random coils and, in aqueous solution, amide groups in the backbone form hydrogen bonds with the solvent. Clearly, the situation within long polypeptides and proteins is very different where large areas of hydrophobic surface on the exterior of two or more helices can come together, or the dipoles of helices can be stabilised by charges or external H-bond formations at the termini.

In order to assess α-helix stability relative to the random coil state in peptides, Zimm and Bragg in 1959[4] (and Lifson and Roig shortly afterwards)[5] defined two parameters, the helix initiation constant sigma (σ), and the propagation constant *s*. The initiation constant reflects the probability of aligning the first three residues in an α-helical conformation. Without stabilising hydrogen bonds, this is a highly disfavoured process, due to repulsive interactions between the aligned dipoles and the loss of entropy.[6] Thus, σ generally has very low values of $<10^{-3}$, where a value of 1.0 would indicate that the first three residues were fixed in an α-helical conformation. The propagation constant *s* reflects the likelihood of an amino acid residue adopting α-helical torsion angles [for the N-C^{α} and C^{α}-CO bonds, respectively (φ and ψ values)] when added to the end of a pre-existing helix. There is a different *s* value for each amino acid, and the value changes depending on its position within the helix and which solvent, if any, is present in the study system. Note that some residues are more likely to be found at the helix termini than in central regions. Values for *s* of >1.0 indicate that the residue is a helix-*making* amino acid (*e.g.* alanine) and values <1.0 indicate that the residue is a helix-*breaking* amino acid (*e.g.* proline). Values of *s* have been determined experimentally,[7,8] theoretically[6] and by the analysis of protein crystal structure databases.[9,10] Together, these Zimm-Bragg parameters σ and *s* give a useful *indication* of the helix-forming propensities of each amino acid.

Although peptides of fewer than 20 residues usually exist as random coils, there are, nevertheless, several examples of short peptides which do show α-helical structure in aqueous solution. The first two of these to be

studied in depth were the S-peptide and C-peptide of an α-helical segment of the enzyme ribonuclease A[11-13] Here it was shown by Baldwin *et al.* that the structures were stabilised by charge-dipole interactions[14] and salt bridges[15] respectively. Bombolitin I, a 17-residue peptide isolated from bumble-bee venom, is also predominantly α-helical in the presence of micelles,[16] and molitin, a 22-residue gastrointestinal peptide hormone, is stabilised as an α-helix by charge-dipole interactions and salt bridges.[17] A variety of other peptide hormones also show α-helical structure during interactions with receptors. These peptides rely on amphiphilicity for their structural stability when associated with non-polar surfaces.[18,19] In fact, virtually all short peptides which show observable helical structure in aqueous solution owe this feature to stabilisation contributed by amphiphilicity, charge-dipole interactions, salt bridges or the involvement of helix-fortifying residues.[20] These factors have been used in the *de novo* design of helical peptides, *e.g.* the 17-residue peptide of Marqusee and Baldwin which contains three (Glu-Lys) salt bridges, charged groups of appropriate sign at both termini (see below), and a high proportion of helix-stabilising alanine residues.[21] The 18-residue peptide of Zhou *et al.* combines charge-dipole effects, salt bridges, high alanine content and amphiphilicity to generate helicity.[7] Many other less common types of conformational constraint have also been used to force peptides into α-helical conformations. For example, the metal ion coordination of the side-chains of the i and $(i + 4)$ residues bearing acid, or, amine and thiol groups has been used to achieve helicity.[22,23] Similarly, aspartic acid residues at positions i and $(i + 4)$ have been shown to bind to a positively charged guanidinium ion containing receptor, enforcing an α-helical conformation.[24] Other strategies have used covalent linkages to encourage α-helical conformations in peptides. For example, connection of the i to $(i + 7)$ residue side-chains through a disulfide bond,[25] connection of the i to $(i + 4)$ or i to $(i + 7)$ residue side-chains[26,27] through lactam bridges, or connection of the i to $(i + 7)$ residues through an alkyl bridge all generated helices.[28] In one case, the i and $(i + 11)$ side-chains were connected to a porphyrin macrocycle, and the structure of the peptide between the attachment points was found to be an α-helix.[29]

In untethered helical polypeptides of the type described above, the *N*- and/or *C*-terminal ends show much more motion than the central regions. This phenomenon, often referred to as helix fraying, is due to the absence of constraint through H-bond formation beyond the end of the helix. Such behaviour is easy to rationalise because it is evident that a residue positioned in the central region of a helix forms H-bonds with both its NH moiety and its carboxamide O-atom to other helix residues above and below the residue, respectively. However, the first four NH moieties of the first four residues and, also, the last four carbonyl O-atoms of the last four residues have no H-bonding partners and are, therefore, much less restrained. Hence, to stabilise helical structures, residues near the termini are preferred if their side-chains are able to supply hydrogen bond partners for unpaired main-chain NH and CO groups,[30] as noted above. Such considerations tend to suggest that if a rigid template possessing a σ value of 1.0 could be prepared which aligned NH moieties or carbonyl groups in the correct position required for helix propagation, Fig. 1, then attached polypeptides could self-assemble into helical structures. In essence, this approach has been followed by many groups with varying degrees of success where the objective has been to entrain natural polypeptides into α-helical conformation *without* modifying the peptide residues themselves.[31-47]

Indeed, this has been our own objective where systems have been sought that would allow the abilities of short sequences to adopt an α-helical conformation to be tested, in the absence of the stabilisation conferred by the macroscopic environment of the biomolecular assembly. An example of this is found in the 2A region of the foot and mouth disease virus polyprotein, where it is believed that the 19-amino acid residue 2A region

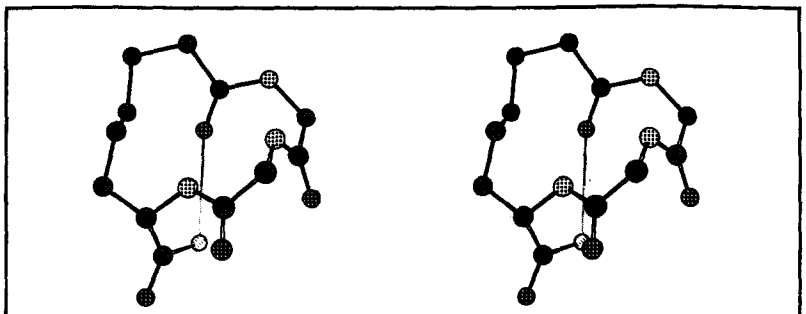
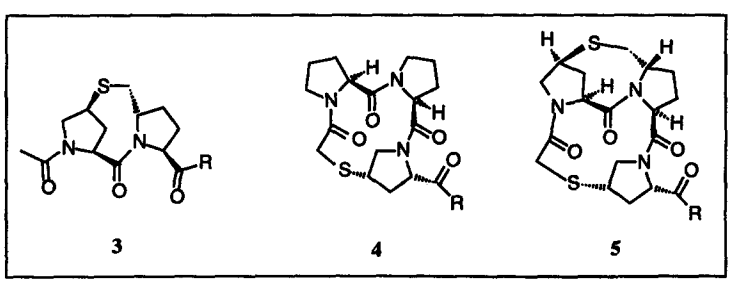


Figure 1 Stereoview of an *N*-terminal template in an α-helical conformation. The carboxamide O-atoms are aligned downwards towards the *C*-terminal end and the position of the first potential H-bond for the first N-atom of appended peptides is shown.

forms an α-helix-type-VI reverse turn structure in the exit pore of eukarotic ribosomes that allows it to catalyse self-cleavage during translation.[48]

A major problem in all template designs (*e.g.* Fig. 1) is the requirement to align H-bond donors or acceptors in the correct geometry for helix initiation. The aligned carbonyl groups and NH moieties in α-helices and helix-like structures confer significant dipolar electric fields, positive at the *N*-teminus and negative at the *C*-terminus, which destabilise the overall structure. In the central portions of helices, the formation of H-bonds offsets such destabilisation, as noted above, but at the termini, there is no compensating enthalpic gain. Thus, the alignment of carboxamide moieties or equivalent polar moieties within the template itself is a significant challenge. The design of Kemp's proline containing cap **3**[32-44] serves to illustrate the problem.

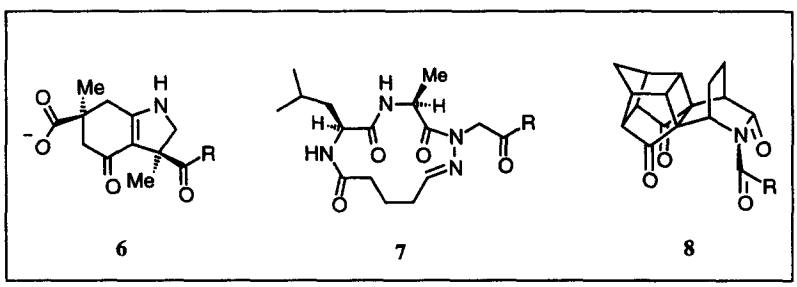


3 **4** **5**

Ideally, the three carbonyl groups present in system **3** should be aligned, and thus able to generate α-helical structure in an attached peptide. In the event, the macrocycle **3** shows a degree of flexibility, interconverting between three conformations, only one of which is able to nucleate a helix.[42] The amount of template present in this nucleating conformation, as determined by NMR spectroscopy, correlates with the degree of helicity of the attached peptide, and as such is a "reporter" on the conformation of the peptide. This has proved useful for the quantification of amino acid *s* values,[44,49] and for the study of the helix-stabilising effects of side-chains[50] and the use of trifluoroethanol as a solvent.[32] However, the lack of rigidity significantly decreases the efficiency of the template as an α-helix initiator and σ was estimated to be ~0.15 in water, as compared to the value of 1.0 required for an exactly α-helical array of carbonyl groups.[44] In an attempt to overcome this limitation, Kemp has also studied the triproline analogues **4** and **5**.

Macrocycle **4** can potentially supply three aligned carbonyl groups only if all three amide bonds within the ring possess the *trans* configuration (*ttt*). In solution, the macrocycle was found to exist as two conformers, the *cct* form and *ctt* form, and did not have useful helix-initiating properties.[37,38] The designed introduction of a 4,5'-methylenethio bridge between the first two prolines of structure **4** was expected to prevent the adoption of the *cis* configuration of the Pro^2-Pro^3 amide bond and gave macrocycle **5** as a preparative target. However, when the molecule was synthesised the desired *ttt* conformer did not form and the macrocycle populated the *ctt* form. Preliminary results suggested that the template could generate a 3_{10}-helix which may convert to an α-helix further along the peptide sequence, but further information on the system is required to complete the picture. Whilst template **5** apparently possesses greater initiating power than the previous two templates **3** and **4**, the cap again lacks the required number of aligned carbonyl groups for α-helix induction, presumably due to dipole repulsion, and demands a difficult 26-step synthesis.[39,40]

In addition to Kemp's work, three other groups have constructed templates potentially capable of initiating α-helices. Bartlett *et al.* recently reported on the hexahydroindol-4-one **6** and described the system as an "aglet" designed to prevent the *N*-terminal fraying of helices.[45]



The template requires the carboxyl group on the six-membered ring to be deprotonated in order to function, and presumably works by forming a favourable charge-helix dipole interaction which can stabilise the helix. The peptides must be linked *via* a lactate residue to the tertiary carboxylic acid and, if these criteria are met, the template is able to generate an α-helix, as has been demonstrated through CD and NMR spectroscopic analysis. The actual degree of helicity induced in the peptide is unclear and while standard methods suggested around 40% helicity, these may be inappropriate for the analysis of the short 6 residue peptide sequences involved. Whilst the template clearly does have a stabilising influence over α-helical conformations, the extent of this stabilisation is ambiguous, since the peptide appended to the template (-Glu-Ala-Leu-Ala-Lys-Ala-NH_2) already possesses a high α-helical propensity, consisting largely of alanine residues, and incorporates a Glu-Lys salt bridge.

Arrhenius and Satterthwait reported on an α-helix template in the form of the 13-membered macrocycle **7**.[46] However, the compound was only effective as a helix template in TFE, and the helix appeared to fray rapidly in moving away from the template. Müller *et al.* constructed several polyketone lactam derivatives which were expected to function as helix initiators, of which the cage compound **8** is representative.[47] Peptide conjugates of these compounds displayed significant helicity compared to untemplated peptides, but once again the experiments were performed in the presence of TFE and the peptide sequences themselves incorporated additional stabilising features.

All of the templates described above have been shown to enhance helical peptide conformations, however, in each case, the actual gain in stability of the helical form is difficult to assess because other helix-inducing influences were present. For example, the use of TFE as a solvent, intra-helix salt bridges, or sets of residues with high *s* values. Our objectives were to design, prepare and evaluate the efficiencies of high σ-value templates based on almost natural cap sequences and use these to probe, on one hand, α-helix propagation (*s*) values for individual residues, without bias, and on the other hand, the effects of helix length on the molecular dipole as measured by changes in the pK_a value of ionisable groups placed at the helix termini. Here we describe the preliminary results of work in this area and demonstrate that such systems can be synthesised.

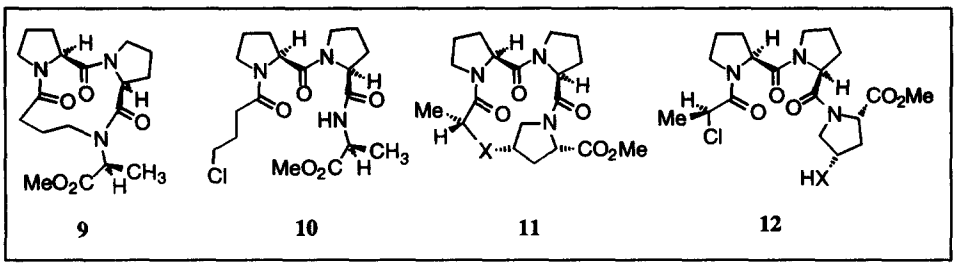
Results and discussion

Identification of a suitable macrocyclic framework

At the start of the programme in this area, which stemmed from the need to assess structural features of the self-cleaving sequence of the 2A region of FMDV, much of the work on templates described above had not yet appeared. Therefore, 3-D structures were identified that would fulfil the requirements of the perfect α-helix initiator, starting from scratch. Given our need to probe the putative α-helix-type-VI reverse turn structure of the FMDV-2A polypeptide,[48] we needed to initiate the helix from the *N*-terminal end and preserve the unadulterated natural structure of the *C*-terminus. Thus, we searched for template cap structures capable of initiating helices from the *N*-terminus. Five criteria were applied to constrain the search and these were: a) the template must provide three correctly oriented carbonyl groups; b) it should be conformationally rigid; c) it should lack hydrogen bond donors of its own to preclude the formation of other hydrogen bonded conformations; d) its structural and electronic properties should very closely resemble part of a natural helix, and; e) its synthesis should be reasonably concise, preferably modular, and amenable to modification to provide structural variants.

With its restricted dihedral angles and side-chain alkylated amino group, proline, as a modular unit, potentially fulfils many of the criteria more effectively than any other proteinogenic amino acid.[51] Therefore, if macrocycles constructed largely from proline residues could be appropriately linked together, they might assume structures able to nucleate helices. Accordingly, the first template design was the 11-membered macrocycle **9**, which consists of a Pro-Pro-Ala tripeptide sequence cyclised *via* a butyryl bridge between the Pro^2 and Ala^4 nitrogen atoms. In order to emulate an α-helical structure, each of the three amide bonds would need to be in the *trans* configuration. Molecular modelling suggested that this arrangement would be energetically accessible, but it was noted that the tertiary nature of each amide bond would increase the possibility of populating *cis*-rotomeric forms which would be ineffective templates. If the template **9** could be prepared in the correct conformation, it was reasoned that saponification of the methyl ester would allow activation of the *C*-terminal carboxylic acid for the subsequent attachment of amino acids or peptides with which the helix initiating properties of the system could be tested.

Disconnection of macrocycle **9** to give the 4-chlorobutyramide **10** was considered to furnish the most simple, viable and informative synthetic protocol. This was because cyclisation *via* S_N2 attack of the 4-chloride by a nitrogen nucleophile was expected to be facile, if the correct conformations could be populated, and because analysis of the population of conformers of **10** would provide useful structural information on the system, given that all of the amide bonds in **10** were already in place.



9 10 11 12

The synthesis of tripeptide **10** was achieved in four steps starting from commercially available materials in which the difficult Pro-Pro coupling step[52] was achieved using the thionyl chloride and pyridine methodology, which had already proven its utility for other hindered systems.[53] The compound migrated as a single band on analytical TLC plates in several solvents and gave the expected analytical data.[54] However, compound **10** displayed three distinct conformations in its ^{1}H- and ^{13}C-NMR spectra recorded in deuterochloroform, in a ratio of approximately 2:1:1.[54] The complexity of the system, *i.e.* six sets of signals for proline residues, precluded a full structural analysis. Nevertheless, a similar peptide studied by Venkatachalapathi and Balaram displayed comparable conformational isomerism by NMR spectroscopy,[55] and it seemed very likely that one or more of the conformations of compound **10** contained a *cis* acylproline amide bond. It was also noted that hydrogen bonding between the Ala[4] NH and one of the carboxamide O-atoms could be a factor in stabilising certain conformations and that this would form a 10-membered ring (β-turn/3_{10}-helical turn) or 7-membered ring (γ-turn).

The cyclisation of compound **10** to give the template structure **9** required the deprotonation of the Ala NH for two reasons. First the amide nitrogen itself lacked sufficient nucleophilicity, and, second deprotonation would remove the only potential the molecule possessed to form H-bonds. Analogous amide alkylations had been achieved using a variety of bases.[56-59] However, when we attempted to convert the chloride **10** to the template structure **9** at high dilution using either DBU, sodium hydride, potassium *tert*-butoxide or LDA as the base, none of the desired macrocycle **9** could be obtained. Indeed, after an aqueous acidic work-up, only starting material was recovered along with a very small amount of polymerised material.

In rationalising the failure of the chloride **10** to react, it is expected that the cyclisation was hampered by the hindered nature of the nucleophile, and/or, the unfavourable conformations adopted by the acyclic compound in the presence of metal ions that can chelate the anion and other carbonyl group O-atoms, in the cases where metal salts were used as the bases. Indeed, it seems likely the very conformational structures available to the all-*trans* peptide, the β- and γ-turn, that we had hoped to avoid by removal of the H-bonding potential from the alanine NH moiety, might be more highly populated by chelation to a metal ion. Furthermore, we were aware that in order to achieve the cyclisation to the trislactam **9**, all three carboxamide O-atoms in the all-*trans* precursor **10** would need to be aligned and, simultaneously, the chloromethyl and carboxamide groups of the 4-chlorobutyramide moiety would need to be almost eclipsed in the transition state.

While it was accepted that it would always be difficult to align the three carboxamide O-atoms in the transition state for the cyclisation, there appeared to be no obvious alternative strategy, and thus attention was focussed on the entropic problems associated with eclipsing chloromethyl and carboxamide groups of the 4-chlorobutyramide moiety. It was reasoned that re-design of the system such that the nucleophile was both constrained and placed closer to a more ordered electrophilic halogenoacylproline would potentially remove such problems. Thus, the

(2*S*)-alanine residue could be replaced by a (2*S*,4*S*)-4-hydroxyproline residue [or a (2*S*,4*S*)-4-thioproline residue] where the hydroxy group (or thio group) would serve as the nucleophile, such that the chain length of the halogenoacyl moiety could be shortened. The ring size would need to be expanded from 11-membered to 12-membered to allow for the pseudoequatorial preference for the 4-substituent on the new proline residue, which would allow a 2-chloroacyl group to be used in place of the 4-chlorobutyryl group in system **10**.

These modifications resulted in the new template target **11**, X = O or S and its immediate precursor **12**, X = O or S. The increased reactivity of the 2-halogenoacylamide in structure **12** was also expected to facilitate cyclisation but a 2-chloropropionyl group was chosen in preference to the 2-chloroacetyl group because, unlike in the previous system **10**, there appeared to be a low energy pathway for the formation of a *ctt* macrocycle. It was reasoned that the bulky secondary acyl chloride should destabilise the *cis* form of the first amide bond such that it would prefer to populate the *trans* rotomeric form, as is well established for acyl prolines.[60]

The synthesis of linear precursor **12** (X = O) was achieved in eight steps starting from the amino acids (2*S*)-proline, (2*S*,4*S*)-4-hydroxyproline and (2*S*)-alanine.[54] The compound was extremely polar but nevertheless displayed the expected analytical data and showed only one predominant set of signals in ^{1}H- and ^{13}C-NMR spectra recorded in deuterochloroform. Further analysis indicated that the only detectable species was the required *ttt* rotomeric form, an immensely pleasing result because it seemed that cyclisation would give the required *ttt* template **11** (X = O) as the predominant product. Note that in the case of the triproline system, incorrectly folded forms of the target macrocycle **11** which possessed one or more *cis*-amide bonds were not expected to be able to isomerise to more stable structures due to the very high activation energies required to distort the macrocycle. Thus, it was particularly important that at least some of the *ttt* form of precursor **12** (X = O) existed in the ground state.

The intended cyclisation of compound **11** (X = O) is a variant of the Williamson ether synthesis, which involves reaction of an alkoxide with an alkyl halide. The weaker amine bases pyridine and DBU failed to promote cyclisation even upon use of large excesses and, prolonged reaction times. The presence of a full negative charge on the 4-O-atom of Pro^{4} was expected to help to preorganise the transition state for the cyclisation by locating the negative charge towards the *N*-terminus of the forming helical structure and, in doing so off-setting some of the energy penalty associated with aligning the carbonyl groups. However, a similar lack of reactivity was observed with stronger bases capable of fully deprotonating the 4-OH group of Pro^{4}; the use of sodium methoxide, potassium *tert*-butoxide, sodium hydride and potassium hexamethyldisilazide (KHMDS) in a variety of solvents either at ambient or elevated temperatures failed to effect cyclisation.

It was rapidly becoming evident that the alignment of the carbonyl groups in the transition state required for helical cap formation was a very serious problem. Although it was not expected that the problems of metal ion chelation would be such a serious problem as for the cyclisation of compound **10** to template **9**, the possibility remained and a metal ion associated with the 4-O-atom of Pro^{4} in the *ttt* form of compound **12** certainly could also bind to the first and second carbonyl O-atom of the sequence. Furthermore, we could not rule out the possibility that the nucleophilic and electrophilic sites, both of which are located in branched secondary positions, were simply too crowded to react. It is also well established that proline oligomers tend to adopt an extended conformation due to steric interactions present between the pyrrolidine rings.[61]

In order to try to circumvent the problems associated with aligning carboxamide dipoles, which we suspected we would have to confront, the system was modified such that the 4-OH group of Pro^{4} in structure **12** was replaced by a softer and more nucleophilic thiol group.

While this work was in progress, Kemp's group published the structure of the 12-membered macrocyclic triproline template **4** designed to initiate an α-helix that was very similar in structure to our own targets, compound **11**.[54] Although the template was synthesised *via* a completely different route, the 12-membered macrocycle had been prepared and differed only in that a thioether linkage was used in place of the oxygen ether in system **11** (X = O) and that the sulfur atom was connected to an achiral acetyl moiety rather than to the (2*R*)-centre of a propionyl residue, as in our own system **11**. Effectively the difference here amounts to the use of a glycine residue analogue by Kemp in structure **4** and a (2*R*)-alanine residue in structure **11** where, in the required all-*trans* amide forms of the templates, the heteroatom would reside in a position occupied by the amino group of the first residue in a natural α-helix. This latter difference in the structure of the first residue is of considerable importance because we had introduced the extra steric bulk specifically to destabilise the *cis* rotomeric form of the first amide linkage. Moreover, Kemp actually prepared the *ctt* and *cct* forms of the template rather than the desired *ttt* form. Given that it is extremely difficult for the amide rotomers to equilibrate in the closed cycle, Kemp's system must have possessed a significant population of the first amide bond in the *cis*-form prior to cyclisation. As such, the carbonyl groups would not have been aligned in the transition state for cyclisation, using our own macrocyclisation protocols, Fig. 2.

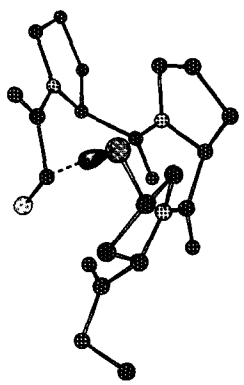


Figure 2
Transition-state conformation required for the cyclisation to *ctt* rotomer of Kemp's template **4**

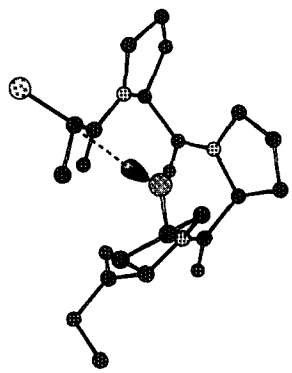


Figure 3
Transition-state conformation required for the cyclisation of *ttt* rotomer of compound **12**

Thus, it seemed very likely that the reason compound **12** could not be converted to the requisite macrocycle was exactly because all of the amide groups were *trans*-oriented for significantly more time, as was determined by NMR spectroscopy,[54] and therefore, that all of the carbonyl groups were aligned to give a large dipole moment and, hence, a high energy transition state for cyclisation, Fig. 3.

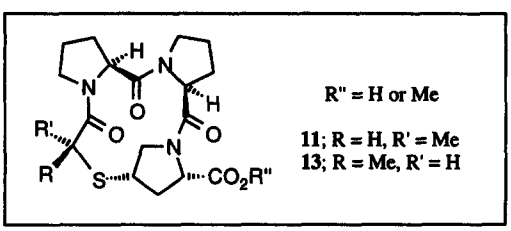
In view of the results of this analysis, it seemed worthwhile to investigate the system further by first preparing the thioether version of compound **11**. The replacement of the O-atom by sulfur was expected to facilitate the formation of an anionic nucleophile and slightly ease crowding in the cyclised material, due to the increased length of C-S bonds. It was determined that a modification of our previous synthetic strategy would furnish the maximum amount of information because the rotomeric forms of each amide bond could be assessed in the precursor, by NMR spectroscopy, as before. We, therefore, decided that a last step should be the displacement of chloride by a Pro[4]-thiolate ion derived from a template precursor such as compound **12** (X = S). Given that it

was expected that the synthesis of the precursor **12** (X = S) would not present a formidable challenge in its own right, it was of interest to examine the possible mechanism for its cyclisation to the *ttt*-form of the template, structure **11** (X = O), in more detail.

We knew from our earlier work that the precursor **12** should exist primarily in its *ttt*-form. We also suspected from our analysis so far, that the alignment of carbonyl dipoles in the ground state of the required product increased its internal energy relative to the ground state of the precursor. Thus, as the ground state *ttt*-form of the precursor **12** began to approach the transition state for cyclisation, the energy of the system would increase, even before the concerted S_N2 displacement of the chloride ion became an important contributor to the energy of the system. However, the ground state of the precursor **12** is actually in equilibrium with a number of other rotomeric forms including the *ctt* and the *cct* forms, both of which could cyclise to form the corresponding macrocycles. Indeed, *N*-acetyl analogues of these latter compounds were the predominant products of Kemp's attempted synthesis of template **4**.[37,38] Thus, it seemed appropriate to obtain some measure of the energy penalty associated with aligning the carbonyl dipoles into a helical structure.

Application of the Hammond postulate for the endothermic cyclisation of precursors such as **12** (X = S) to template structures indicates that the structures of the transition states should resemble the structures of the products. Furthermore, the relative energies of the transition states should relate to the energies of the products. It would be possible to obtain relative energies for the cyclised products by performing molecular mechanics calculations on each of the possible conformers of the template **11** (X = S). It is also reasonable to assume that the energetics for bond breaking and making in the actual displacement of chloride by thiolate in each transition state would be equal, such that for comparative purposes, these contributions to the transition state energies cancel. Thus, from transition state theory, the rates of formation of given cyclic products would be related to the concentration of each of the transition state complexes for which the relative energies should correspond to those calculated for the various conformational forms of the cyclised template. By compensating for the actual ground state concentrations of the various rotomeric forms of the starting material, it should, therefore, be possible to predict the kinetic product of a given cyclisation reaction and then go on to identify, and then test, improvements in the system that would lead to the formation of a high σ value template.

Accordingly, the energies of all of the conformations of the macrocyclic thioether template systems **4**, **11** and **13** containing, respectively acetyl-, (2*R*)-propionyl- and (2*S*)-propionyl- moieties in the first position, were calculated.[62]



The results are summarised in Table 1, where the fifteen computed lowest energy conformations for each macrocycle are shown.

Table 1 *Molecular mechanics energy calculations (in kJ mol^{-1}) for 4, 11 and 13*

Acid 4			Acid 11			Acid 13		
Code	**State**	**Energy**	**Code**	**State**	**Energy**	**Code**	**State**	**Energy**
G	*ctt*	299.8	*A*	*cct*	392.9	*G*	*ctt*	388.8
B	*cct*	300.6	*M*	*ttt*	405.4	*AE*	*ctt*	397.9
AE	*ctt*	307.0	*AA*	*tct*	413.1	*AG*	*ctt*	400.2
R	*ctt*	311.0	*S**	*ttt*	413.9	*R*	*ctt*	401.7
K	*ctt*	312.0	*I*	*ttt*	417.8	*AB*	*ctt*	414.4
M	*ttt*	317.5	*U**	*ttt*	419.3	*A*	*cct*	418.9
*U**	*ttt*	318.7	*O**	*ttt*	421.9	*U**	*ttt*	423.6
*S**	*ttt*	323.9	*Q*	*tct*	422.8	*D*	*ctt*	424.2
AA	*tct*	324.4	*F*	*ttt*	424.1	*I*	*ttt*	425.3
AB	*ctt*	325.1	*N*	*cct*	424.4	*M*	*ttt*	425.4
*O**	*ttt*	325.9	*G*	*ctt*	425.8	*Q*	*tct*	428.6
I	*ttt*	327.0	*Y*	*tcc*	426.8	*O**	*ttt*	431.4
N	*cct*	330.1	*Z*	*ttt*	431.3	*AA*	*tct*	432.0
A	*cct*	330.1	*P*	*tct*	433.8	*S**	*ttt*	433.3
Y	*tcc*	332.1	*AG*	*ctt*	436.5	*P*	*tct*	433.3

Conformational energies were calculated using the AMBER all-atoms molecular mechanics force-field[63] and the Insight II program.[64] Energies are relative to an arbitrary minimum and are given in kJ mol^{-1}. "State" describes the configurations of the three amide bonds in the 12-membered ring and conformers with the correct set of dihedral angles for α-helix initiation are marked with an asterisk. Note that for torsion about the CO-N bond, the torsional angle, ω, must be close to 0° (*cis*) or 180° (*trans*).

Cis-trans isomerisation at each of the 3 tertiary amide bonds generates 8 possible configurational isomers where only the *ttt* form is potentially capable of α-helix initiation. Each of these 8 "parent" conformations of defined ω values possesses a set of sub-states. These sub-states differ in the φ (N-C$^{\alpha}$ bond torsion) and ψ (C$^{\alpha}$-CO bond torsion) angles for each proline residue and for the α-thio acid residue, and in the torsional angle about the Pro^{4}-C$_{4}$-S bond of the thioether bridge. Some of these sub-states which exist in the *ttt* form, for example, structures *M* and *I*, do not contain aligned carbonyl groups and, therefore, are not capable of α-helix initiation.

The relative conformational energies of Kemp's macrocycle **4** indicate that 6 isomeric forms exist which are of greater stability than the most likely α-helix-initiating *ttt* sub-state ***U***. The five lowest energy forms of these possess *cis* stereochemistry for the first amide bond. Note that this site is adjacent to the site of structural variation

in the target compounds **11** and **13**. The energy difference between the most stable forms ***G*** and ***B*** and the most stable α-helix-initiating *ttt* sub-state ***U*** is about 19 kJ mol^{-1}. While it is accepted that such calculated energy differences are not accurate, they are useful, especially when large in magnitude, and in this case the modelling study correctly predicts the structure of compound **4**, which was experimentally determined by Kemp *et al.* to assume the *ctt* and *cct* forms.[37,38]

Compared to compound **4**, structure **11** possesses an extra methyl group in the (*R*)-configuration at C-2 in the first residue. The calculations show that, relative to the situation for compound **4**, the *ctt* forms of compound **11** are significantly destabilised in favour of the *ttt* forms. Indeed, 6 of the 9 most stable sub-states for compound **11** are *ttt* forms and none of these 9 most stable sub-states are *ctt* forms. The most stable form is a *cct* sub-state, ***A***, in which the extra steric size of the methyl group is repealed by the two adjacent *cis*-amide bonds which place it in an equatorial position on the periphery of the macrocycle. Interestingly, the orientation of the carbonyl dipoles in sub-state ***A*** are also randomised, Fig. 4.

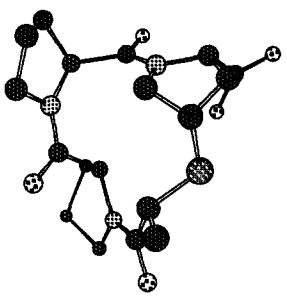


Figure 4
The most stable form of **6**, *cct* sub-state, ***A***

Given that the *cct*-rotomeric form of the precursor **12** (X = S) is not likely to be highly populated, since the equivalent *cct* rotomeric form of **12** (X = O) was not,[54] this consideration alone would not mean that ***A*** would be the kinetic product of any attempted macrocyclisation. However, the finding that the ***M*** sub-state of the *ttt* form was the most stable of the *ttt* forms, and 8.5 kJ mol^{-1} more stable than the lowest energy helix nucleating form, ***S****, was of more concern. This is because sub-state ***M*** possesses non-aligned carbonyl groups, in which the second carbonyl group O-atom is twisted up and in above the 5-methylene group of Pro^4, so that it is about 140° out of alignment with the first carbonyl group (*i.e.* nearly anti-aligned) Fig. 5.

Furthermore, peptide extended analogues of the ***M*** sub-state of compound **11** could not be expected to untwist at accessible temperatures to give peptide extended analogues of the ***S**** form. The conformational reorganisation would require two isomerisations of the second amide bond *via* the *tct* form, ***AA***, both of which are thermodynamically up-hill and both of which require large activation energies of 75 kJ mol^{-1} or more, for overcoming amide resonance and reorganising the macrocycle. These isomerisations would need to occur without the single isomerisation of the first amide bond to the *cis* form, either directly from the ***M*** sub-state (to give the species ***G***) or, from the intermediate ***AA*** which would give the most stable species, ***A***. Thus, it appeared that if cyclisation did occur from a predominantly all-*trans* form of the precursor **7**, the ***M*** sub-state and possibly also, the ***A*** sub-state, would be the predominant products.

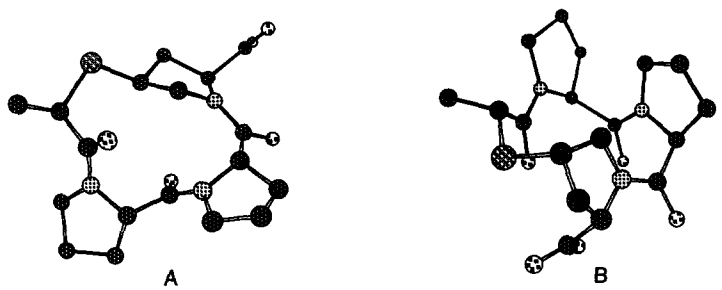


Figure 5 The *M* sub-state (**A**) and the *S** sub-state (**B**) of **11**.

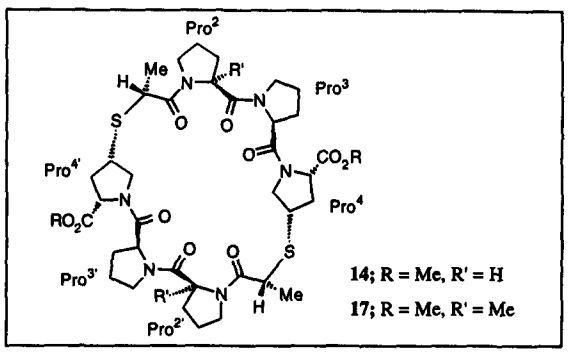
Methyl substitution with (*S*)-stereochemistry at C-2 of the first residue gives compound **13**. In this system the 6 most stable sub-states possess *cis* stereochemistry in the first amide bond and 5 of these are *ctt* sub-states. The difference in energy of 34.8 kJ mol^{-1} between the most stable sub-state, ***G***, and the lowest energy helix nucleating form, ***U****, in this case, is larger than that for the diastereomer **11**, 21 kJ mol^{-1} (for ***A*** isomerising to ***S****). Indeed, the situation is worse for compound **13** than for the unsubstituted system **4** largely because the helix nucleating forms of **11** possess an adverse 1,3-diaxial interaction between the methyl group and the 5-methylene group of Pro^2. It is interesting to note that the non-helix nucleating *ttt* form, sub-state ***M***, is also destabilised by the same diaxial interactions. Thus, compound **13**, it appears, is not likely to cyclise to give an all *trans* form of the template.

From the comparison of the conformational energies for each compound, it is apparent that macrocycle **6** is the most likely to afford a sub-state capable of displaying α-helix-initiating properties (for example, conformer ***S***). However, it is also apparent that for the parent acid derivative conformer ***A*** with a *cct* arrangement of amide bonds and conformer ***M*** a twisted *ttt* sub-state which possesses no net dipole are the likely kinetic products. In order to validate the analysis experimentally and to test the effect of introducing modifications into structure **11**, preparative routes to the parent system **11** (X = S) were devised building upon the protocols employed in the attempted synthesis of the oxygen-containing analogue **11** (X = O).[54]

The synthesis of the precursor **12** (X = S) is described in detail elsewhere.[62] The thiol **12** displayed the required analytical properties and showed one predominant set of signals in ^{1}H- and ^{13}C-NMR spectra corresponding to the *ttt* rotomeric form.[62]

In order to effect cyclisation, solutions of thiol **12** in a range of different solvents were treated with a variety of bases. The use of DBU, sodium methoxide and potassium *tert*-butoxide each failed to mediate the formation of a new product. However, when solutions of thiol **12** were heated in the presence of sodium hydride, significant changes were observed in the NMR spectra of the crude product, indicating that a reaction had occurred. Column chromatography of the crude material allowed isolation of a single product in 15–25% yield and its NMR spectra were in accord with those expected for a single conformer of macrocycle **11**. The ^{1}H-NMR spectrum displayed a striking 0.3 ppm upfield shift for the doublet attributed to the propionyl methyl group at the stereogenic centre, in keeping with a change of substituent from chlorine to sulfur at the C-2 position. The ^{13}C-NMR spectrum also showed two upfield shifts; one of 3 ppm for the methyl group carbon, and one of 10 ppm for the stereogenic

carbon atom at C-2. The single set of resonances in the ^{13}C NMR spectrum ruled out acyclic oligomers and polymers and the product lacked an S–H stretch in the infra-red spectrum further supporting the NMR evidence that a thioether was present. However, mass spectra of the compound failed to show the desired molecular ion (M^+ = 409 Da) and, instead, a peak of mass 819 Da of 16% intensity was observed together with a more intense peak at 841 Da, corresponding to the sodium ion complex, and a peak at 857 Da, corresponding to the potassium ion complex. Thus, the reaction product was actually the cyclic dimer **14** containing a 24-membered ring. The dimeric product **14** was indistinguishable from monomeric compound **11** by any of the other analytical techniques employed because it possesses C_2 symmetry. Conformational analysis of the product and its free acid form were conducted using high-field 2D NMR experiments. From these experiments it was evident that all of the amide bonds were of *trans*-stereochemistry. The set of ROESY cross-peaks observed for the system were consistent with the structure of the dimer **14**.

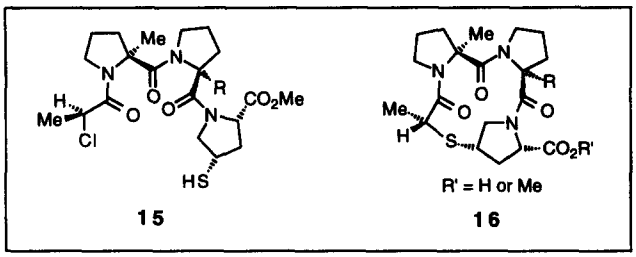


The result is highly informative and indicates that the energies inherent in the folded conformations of the predominantly all-*trans* precursor thiol **12** which are required for intramolecular cyclisation are too high. However, because the thiolate group and the α-chloroacyl moiety are inherently reactive enough, intermolecular thiol alkylation takes place, to give an acyclic dimer which has sufficient flexibility to react further to give the 24-membered macrocycle **14**. Given that the precursor thiol **12** does not populate the *cct* rotomeric form significantly, this outcome is in accord with the modelling studies and provides a firm basis for the informed structural alteration of the target structure **11**.

α-Alkyl amino acids such as 2-aminoisobutyric acid (Aib) are known to favour folded conformations,[65-67] and to promote cyclisation reactions *via* the *gem*-dimethyl effect.[68] It was reasoned that if such an α-alkyl residue could be incorporated into the thiol **12**, the moiety would appreciably increase the time that such a molecule spends in folded conformations conducive to monomeric cyclisation. The minimum change which could be envisaged is the substitution of a (2*S*)-α-methylproline residue for one of the proline residues in macrocycle **6**.

α-Methylproline (Pro^{Me}) has received little attention to date. Those studies which have appeared indicate that the amide bond preceding an α-methylproline residue is exclusively *trans*.[69,70] Likewise, the *cis* configuration of the following amide bond is also destabilised. Substitution of Pro for Pro^{Me} at position (i + 1) in tetrapeptide sequences by Robinson and co-workers was found to restrict the conformational space available to the peptide so that it occupied turn-like conformations for longer periods of time,[71] exactly the type of effect that was needed to coax the precursor **11** into a helical transition state. The substitution of a (2*S*)-Pro^{Me} residue for Pro^2 in the

triprolyl peptide **11** defines the new precursor target **15** which could be tested for cyclisation to the template structure **16**.



Molecular mechanics simulations preformed on the *ttt* substates of potential template **16** (R = R' = H) were in accord with expectations based upon the Hammond postulate which we had invoked to link the structures and relative energies of the products to those of the transition states for their formation, see above. The results for acid **16** (R = R' = H) indicated that the helix-initiating *S* substate would be of *lower* energy than the non-productive *M* substate in contrast to the original thioether macrocycle **11** (Table 1).

It was envisaged that the preparative route employed for the synthesis of precursor **12** could be adopted for the preparation of the methyl homologue **15**.[62] Thus, a supply of (2*S*)-Pro^{Me} was prepared folowing the method of Seebach[72] and the material was converted to the required thiol precursor **15**. In the case of the cyclisation of thiol **15**, caesium carbonate proved to be slightly more effective than sodium hydride and afforded a cyclic product in 30% yield. NMR spectroscopic analysis showed several of the expected chemical shift changes and, once again, mass spectral analysis confirmed that the structure was a cyclic dimer **17**. It was evident that the incorporation of a Pro^{Me} residue failed to promote sufficiently the required folded forms of the acyclic precursor to allow formation of the template structure **16** (R = H)

Further promotion of folded forms of the linear precursor was clearly necessary to enable cyclisation to take place. Continuing with a strategy based upon destabilising unwanted conformers in the transition state and extending the principles applied in design of the Pro^{Me} macrocycle **16** leads to a new target which incorporates two Pro^{Me} residues. Molecular mechanics simulations for acid **16** (R = Me, R' = H) predicted further advantage in the form of an increased preference for the potential helix-initiating *S ttt* sub-state over the *M* sub-state, as expected. Also, the introduction of a second Pro^{Me} residue would exert an additional destabilising effect on the *cis* configuration of the second and third amide bonds in the precursor. Hence, it appeared that if the 12-membered macrocycle could be formed, its most stable conformation would be one that could initiate an α-helical conformation in attached peptides, for example, the *S** sub-state. Unfortunately, however, the presence of two Pro^{Me} residues posed very serious synthetic problems and it was not possible to synthesise the precursor **15** (R = Me).[62]

Effect of Hydrogen Bonding on the Conformations of Linear Precursors

Given that the results of studies thus far had established that adverse steric effects alone are insufficient to promote the necessary dipole alignment in the folding of the triprolyl sequence into a template of type **11** of the correct conformation, the potential stabilising effect of hydrogen bonding in the transition state was examined. In nature, polypeptide folding is stabilised by hydrogen bonding and we wished to test whether the inclusion of

one H-bond would significantly alter the energies of the pathways leading to 12-membered template formation and 24-membered macrocyclisation in favour of the template structure.

Conversion of the methyl ester of the original linear precursor **12** to a methylamide gives triprolyl peptide **18**. The potential now exists for formation of intramolecular hydrogen bonds which could stabilise folded states approximating to the desired cyclic structure, thus assisting cyclisation. Possible hydrogen bonding schemes involve the terminal methyl amide in $i \rightarrow (i-5)$ (α-helical turn), $i \rightarrow (i-4)$ (3_{10}-helical/β-turn) or $i \rightarrow (i-3)$ (γ-turn) interactions with preceding carbonyl groups (Fig. 6).

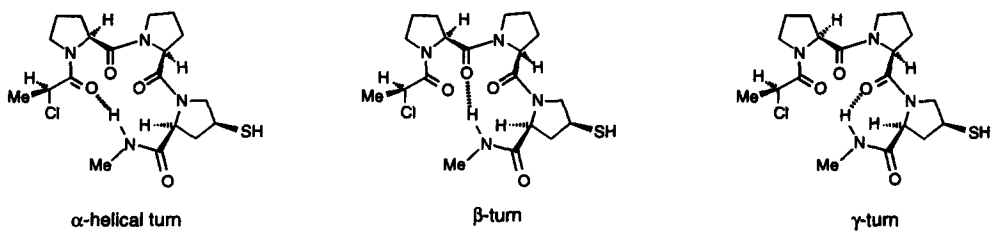


Figure 6 Different types of turn possible for compound **18**

Model studies on other peptide sequences indicate that the β-turn seems the most stable of the three in the absence of other factors.[73,74] These conformations may allow favourable geometry for intramolecular reaction, particularly if the α-helical turn makes a prominent contribution to the secondary structure of the linear precursor. Hydrogen bonding has found uses in favouring cyclisations in the past, for example in Wenger's synthesis of cyclosporin A.[75]

Hydrogen bonding of this sort provides an enthalpic bias towards certain folded conformations, but the actual geometry adopted depends on a fine balance between this favourable effect and other forces which act against compact structures, such as loss of entropy upon formation of hydrogen bonded rings, torsional strain, and repulsive dipolar interactions.[73,76,77] It remained to be seen whether the enthalpic gain of a hydrogen bond would be sufficient to overcome the repulsive steric and dipolar interactions which prohibited cyclisation of earlier linear precursors, **12** and **15**. Hydrogen bonding interactions in Pro-Pro-Ala tripeptides had been demonstrated to promote folded forms of this sequence[55] and we hoped that this might extend to the similar Pro-Pro-Pro sequence.

Accordingly, compound **18** was prepared in a several step synthesis starting from *t*BOC-(2*S*)-Pro-(2*S*,4*R*)-4-hydroxyProOMe through an initial direct aminolysis with methylamine. The methyl ester **14** displayed two conformations in NMR spectra recorded in C^2HCl_3, the *trans trans* (*tt*) and *cis trans* (*ct*), in 55% and 45% abundance respectively. *Cis* and *trans* isomers of urethane bonds are of very similar energy due to the similar size of the carbonyl oxygen and the alkoxy substituent. The slight preference for the *trans* configuration may be due to an O–H·····O=C 10-membered ring hydrogen bond which can form between the Pro^2 γ-hydroxy substituent and the urethane carbonyl oxygen. No trace of the *cis* configuration about the Pro-Pro amide bond could be found, indicating that it is of significantly higher energy than the *trans* form in the absence of external stabilising forces.

In contrast, the methyl amide displayed *three* conformational isomers. Since the only change is incorporation of a hydrogen bond donor, the additional conformation must be one stabilised by a hydrogen bond. The NMR spectra revealed that the *tc*, *tt*, and *ct* configurations of the urethane and Pro-Pro amide respectively, were present, in relative abundance 14:62:24.

The *tc* state is not detected in the methyl ester and must therefore be stabilised by an N–H·····O=C hydrogen bond involving the methyl amide NH and the urethane carbonyl oxygen. The presence of a hydrogen bond is further supported by a 0.4 ppm upfield shift of the amide proton resonance relative to the non-hydrogen bonded form; infra-red spectra in dichloromethane solution also show two N–H stretch bands at 3448 and 3342 cm^{-1}, which correspond to non-hydrogen bonded and intramolecularly hydrogen bonded N–H groups respectively.

Subsequent removal of the *N*-protecting group and reaction with activated (2*R*)-2-chloropropionyl-(2*S*)-Pro gave the 4-hyoxydroPro containing precursor which was converted to the thiolacetate **19**.[62] NMR spectra revealed the presence of two conformations, corresponding to the *ttt* (45%) and *ttc* (55%) forms respectively. This finding accords with the re-establishment of *cis-trans* isomerisation about the Pro^2-Pro^3 amide bond (Fig. 7).

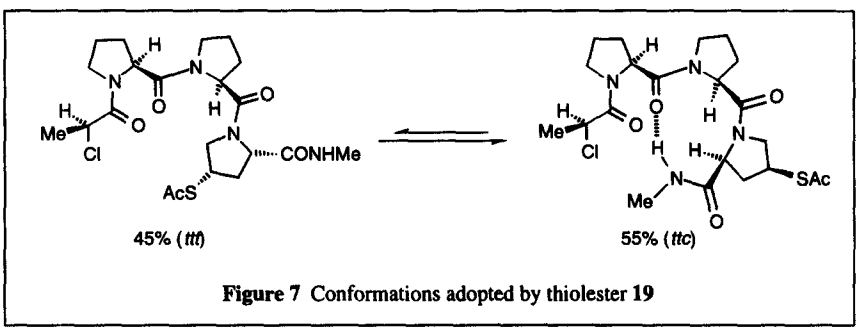


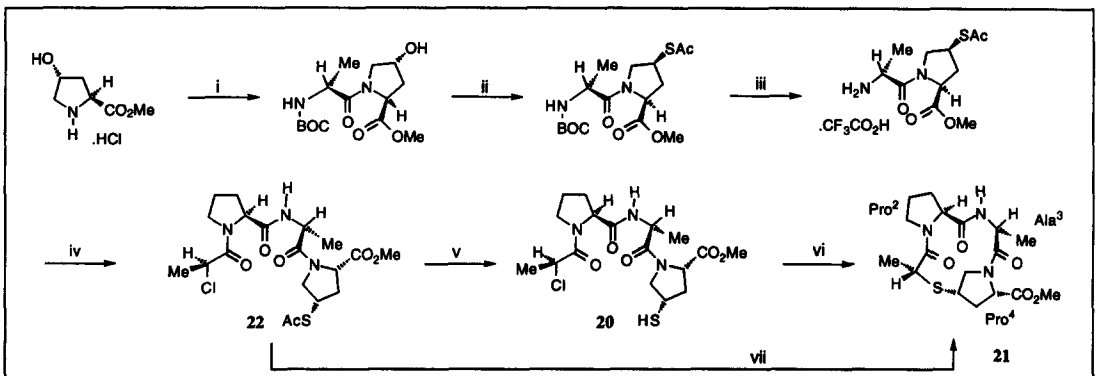
Figure 7 Conformations adopted by thiolester **19**

Conversion of thiolester **19** to the free thiol **18** did not change the conformational equilibrium adopted in chloroform solution and the *ttt* and *ttc* forms existed in a ratio of approximately 1:1. Spectra recorded in ^2H_2O show only the *ttt* state, presumably because the intramolecular hydrogen bond found in the *ttc* form in chloroform is replaced by intermolecular solvent-peptide hydrogen bonds. Thus, it appears that addition of a hydrogen bonding group at the *C*-terminus will not be able to assist the peptide thiol **18** to adopt the required conformations for close approach of the nucleophile, the thiol moiety itself, to the electrophile. A hydrogen bond is present in approximately 50% of the population of molecules in chloroform solution and to some degree the hydrogen bond allows the molecule to overcome repulsive interactions which would otherwise prohibit the folded form. However, the stabilised conformation is one in which the thiol group is held away from the intended site of reaction, and as such is of no use in assisting the cyclisation process relative to the situation for the methyl ester **12**. Hence, the hydrogen bond stabilises the ground state such that any beneficial hydrogen bonding interactions between the methylamide proton and the first carbonyl group in the transition state for cyclisation would be largely cancelled.

In summary, it is evident that using methods to force the acyclic template **12** into a helical structure by destabilising undesirable conformations in the transition state are insufficient on their own to overcome the problems of aligning the carbonyl group dipoles. It is also evident that the second strategy of attempting to stabilise the conformation of an α-helical transition state by providing a mechanism to form the 13-membered hydrogen-bonding network typical of α-helices, is fraught with problems due to the stabilisation of usually rare

ground-state conformations. A third strategy for the preparation of nearly natural α-helical cap templates, based on the notion that it is too difficult to align at least three carboxamide group dipoles simultaneously, was also pursued. The relaxation of the highly restrained C^{α}-N bond torsion for Pro^3 in the acyclic precursor **12**, through its substitution for a (2*R*)-alanine residue, effectively decouples the motion of the second carboxamide group from the C^{α}-N bond torsion and allows the second carboxamide group to rotate. It was reasoned that this would allow the modified acyclic precursor **20** [to macrocylic template structures (*e.g.* **21**)] possessing *ttt* amide bond stereochemistry to form macrocyclic transition states which were of lower energy because the Pro-Ala amide bond dipole would not need to align with the other carboxide group dipoles.

The synthesis of compound **20** followed analogous routes to those developed for the preparation of the acyclic precursor **12**[62] and all intermediates were fully characterised and displayed the expected properties, Scheme 1.



Scheme 1 *Reagents and conditions*: i, (2*R*)-*N*-(tert-butoxycarbonyl)-alanine, BOP-Cl, DIEA, CH_2Cl_2, 0-5 °C, 5 h then rt, 4 days, 55%; ii, DIAD, PPh_3, AcSH, THF, 0 °C, 2.5 h then rt, 15 h, 92%; iii, TFA, CH_2Cl_2, 0 °C, 90 min; iv, N-[2*S*)-2-chloropropionyl]-(2*S*)-proline, BOP-Cl, DIEA, CH_2Cl_2, 0-5 °C, 3 h then rt, 2 days, 68% (2 steps); v, KOH, MeOH, H_2O, 2.5 h; vi, Cs_2CO_3, DMF, 4 days, 25%; vii, KOH, MeOH, H_2O, 65 °C, 2.5 h, 69%

Surprisingly, mild alkaline hydrolysis of the thiolester **22** to the free thiol **20** did not give a pure product and apparently some thiolate alkylation had already taken place to give the template **21**. The impure material **20** was taken straight on and subjected to cyclisation using caesium carbonate in DMF under high dilution conditions at room temperature over 4 days, Scheme 1. Column chromatography of the crude product allowed the isolation of a compound displaying NMR spectra corresponding to those expected of the 12-membered macrocycle **21**. Gratifyingly, mass spectrometric analysis revealed this to be the desired *monomeric* cyclic compound **21** ($[M + H]^+$ = 384). Some of the cyclic dimer was also formed, but was easily separated by chromatography. This methodology allowed a 25% conversion of thiolacetate **22** to the macrocyclic compound **21**. Optimisation of the process established that compound **21** could be produced directly from the thiolacetate **22** *via* action of a dilute solution of potassium hydroxide in methanol–water at 65 °C for 2.5 h in an excellent 69% yield.

Conformational Studies

Macrocycle **21** was crystallised as a hemihydrate from ethyl acetate–hexane for analysis by X-ray diffraction. In the solid-state structure all amide bonds exist in the *trans* configuration, fulfilling the first objective. However, the

ttt sub-state adopted is apparently incapable of helix initiation, as the second and third carboxamide carbonyl group lack the required alignment. In fact, the molecule lacks any significant dipole and although the first carboxamide group (in the mercaptopropionyl bridge) possesses an approximately α-helical geometry, the Pro^2 carbonyl group has roughly opposite alignment, and the Ala^3 carbonyl is splayed out. Of course, this is the expected situation in the absence of any H-bond donors and the structure vindicates our original concerns regarding the alignment of carboxamides dipoles in synthesising the macrocycle, *vide supra*. Thus, it now appears certain that the reason why the acyclic precursor can cyclise is because the Pro^2 carbonyl group can 'about turn' in the transition state and minimise the repulsive forces associated with bringing together aligned carbonyl dipoles. Note that in Kemp's system **4** a similar result was observed, but here, because the system was so rigid, one or more amide bonds needed to rotate to the *cis* form in order to minimise the transition state energy for cyclisation.

Before elaborating the system it was necessary to verify that the conformer present in the solid-state was not merely a consequence of crystal packing. The solution state conformation was characterised by NMR spectroscopy and restrained simulated annealing. In $^2H_6$-DMSO solution, the macrocycle **21** displayed only one conformation, which corresponded almost exactly to that found in the solid state. The observation of one set of resonance lines for the macrocycle indicated that there was no significant conformational exchange that was slow on the NMR timescale, and any minor conformation was below the limit of detection (~11% population). The diagnostic NOE's which indicate the "up" alignment (*i.e.* anti-alignment) of the Pro^2 carbonyl group all involve the NH proton of the Ala^3 residue. This proton shows NOE interactions with the α-CH hydrogen of Pro^2 and both of the δCH_2 protons of Pro^4 (Fig. 8), which also characterises a *trans* configuration for the amide bond.

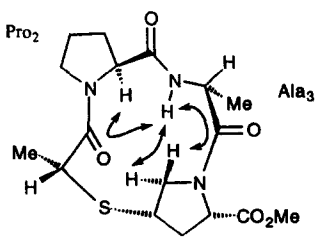


Figure 8
NOE's observed for macrcocycle **21** in $^2H_6$-DMSO solution

The absence of NOE's to either Pro^2 βCH_2 proton argues against any significant population of the α-helical target conformation with the Pro^2 carbonyl aligned "down".

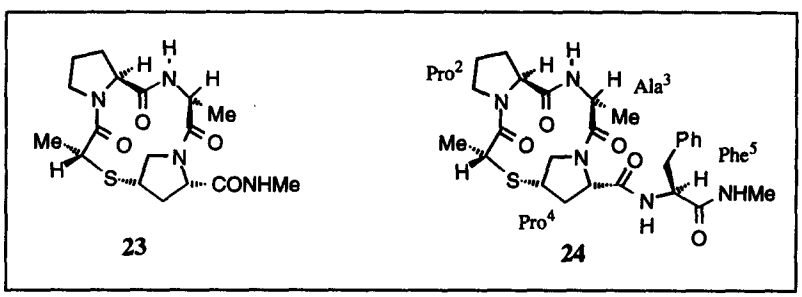
Evaluation of Macrocycle 21 as an α-Helix Initiator

At first glance, the arrangement of carbonyl groups in macrocycle **21** may not appear to be favourable for initiating an α-helix. The Pro^2 carboxamide group has the opposite alignment to that required, and the Ala^3 carbonyl is also appreciably displaced from the position expected for an α-helix. However, the first carboxamide is in approximately the correct position to accept a hydrogen bond from the first amide NH group of an attached peptide and, thus, induce an α-helical geometry in at least the first appended residue. Note that the formation of such an H-bond requires the movement of the Ala^3 carbonyl O-atom into the correct position and that, as a result, the entire cap structure would become α-helical, except the Pro^2 carbonyl group would stay anti-aligned in the

absence of its own H-bond donor, to minimise the molecular dipole. The next (second) NH group of an attached peptide could be in the correct position to hydrogen bond to the Pro^2 carbonyl group, if this carbonyl group is able to flip alignment, and there is no steric reason why it should not do so. If it does align correctly, the Ala^3 and Pro^4 carbonyl groups would be in the desired position, as defined by the constrain in the pyrrolidine ring of Pro^4. Indeed, subsequent H-bond formation would be to carbonyl groups that are already positioned correctly through serving as NH H-bond donors. Thus, after the formation of two H-bonds, the system should behave as a perfect template for α-helix propagation. This "conformational co-operativity" is favoured by the formation of hydrogen bonds, but disfavoured on entropic grounds and by repulsive dipolar interactions. For now then, we should take heed of the fact that peptide helices display 'fraying' at their termini and, thus, for short added sequences we should expect to observe signicant motion at the *C*-terminus.

From the analysis above, it is evident that a rotation of the Pro^2 carboxamide (about the dihedral angles ψ of Pro^2 and ϕ of Ala^3) would be essential to bring the carbonyl group into the correct alignment for helix propagation. Therefore, it was an important objective to assess the magnitude of activation energy required for its re-alignment and how the template might be modified to stabilise the correctly aligned rotomer. In the study of the feasibility of α-helix induction by template **21**, which would require the synthesis of derivatives with appended peptide chains, it was decided to address the question of Pro^2 carboxamide dipole re-alignment first through the sequential introduction of H-bond donors and/or positive charges.

One hydrogen bond donor was introduced through conversion of the methyl ester **21** to the methylamide **23** *via* direct aminolysis. The required amide product was recovered in quantitative yield and displayed the expected analytical and spectroscopic properties.



23 **24**

Conformational analysis of compound **23** in H_2O–^2H_2O (9:1) indicated that there was no change in the conformation of the macrocycle upon addition of the hydrogen bond donor. The Pro^2 carbonyl group did not flip to form a 3_{10}-helical hydrogen bonding network with the methylamide NH proton. Experiments to establish whether there was an α-helical 13-membered hydrogen bonding network connecting the methylamide NH to the mercaptopropionyl carbonyl group were inconclusive. Only weak NOE cross peaks were observed to support the notion that the methylamide group may spend some time folded underneath the macrocycle. This was the expected result since a large movement of the Pro^4 carboxamide moiety and the proper alignment of the Ala^3 carbonyl dipole would be required to form a 13-membered H-bonding interaction.

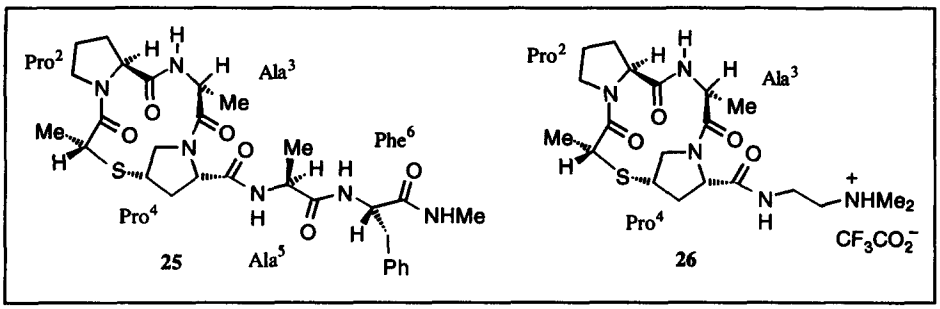
In order to attach more than one hydrogen bond donor site to the template it was neccessary to hydrolyse the methyl ester **21** and perform subsequent peptide coupling steps. The ester was saponified to afford the acid, as colourless crystals in 83% yield. The acid displayed the expected analytical and spectroscopic properties and

would be a precursor for several derivatives. The template cap acid was coupled to (2*S*)-phenylalanine methylamide using BOP-Cl activation to give the template-peptide conjugate **24** in 62% yield which possessed the expected analytical properties.

In C^2HCl_3 solution, the macrocyclic portion of compound **24** showed no change in conformational preference relative to that in the methyl ester **21**. However, several NOEs indicate that there was some interaction between the lower face of the macrocycle and the attached peptide. This implied that the Phe^5 residue spent some time folded underneath the macrocycle in a position conducive to the formation of a hydrogen bond between the Phe NH and the mercaptopropionyl carbonyl group. The existence of such a hydrogen bonding interaction requires the Phe^5 residue to adopt an α-helical position. This notion is supported by a 0.2 ppm upfield shift for the methyl protons of the mercaptopropionyl residue, which is consistent with a ring current effect caused by the close proximity of the phenyl ring. There was also a 0.8 ppm downfield shift of the Phe^5 NH proton relative to its position in a model linear peptide in which it is known intramolecular hydrogen bonding is absent. While the folded conformation is evidently populated to some extent, other evidence indicates that the attached peptide is extremely mobile. For example, some NOE cross-peaks are present which are not possible for the folded conformation, and the $^3J_{HN\alpha}$ value for the Phe^5 residue is 8.0 Hz, which is not consistent with α-helical geometry.[79] Thus, as expected, the *C*-terminus shows the properties of fraying[36] and a good deal of mobility.

In order to provide hydrogen bond donors for each acceptor carboxamide group in the macrocycle **21**, one further derivative containing an attached dipeptide amide was required, compound **25**. The peptide component (2*S*)-Ala-(2*S*)-Phe-NHMe was synthesised by activating BOC-(2*S*)-Ala-OH using the mixed anhydride method and adding (2*S*)-phenylalanine methylamide and was coupled to the template cap acid to give the required conjugate **25** in 33% yield.

NMR studies performed on the template conjugate **25** in C^2HCl_3 solution provided convincing evidence for the presence of a 13-membered intramolecular hydrogen bond network between the NH of the appended (2*S*)-Ala^5 residue and the first carbonyl group. There was a relatively strong set of NOEs similar to those obtained for compound **24**, indicative of the peptide folding underneath the macrocyclic template. The $^3J_{HN\alpha}$ value for the attached alanine residue was 3.9±0.4 Hz, from which we estimate a ϕ dihedral angle of −57°, which is consistent with α-helical geometry.[79] These data therefore indicate that the macrocycle possesses α-helical character and that the first attached amino acid residue, Ala^5, is part of the helical structure, but not the subsequent residues. The less mobile, tighter α-helical conformation obtained for the conjugate **25** over derivative **24** could reflect the higher intrinsic α-helical propensity of Ala compared with Phe,[9,80] or the fact that the fraying *C*-terminus had been moved further away by one residue. Whatever the cause, it was evident that the Pro^2 carbonyl group in conjugate **25** was still aligned antiparallel to the aligned mercaptopropionyl and Ala^3 carbonyl groups. From this analysis we must conclude that the enthalpic benefit of hydrogen bonding between short appended peptides and the template in the "helical" conformation appears unable to counterbalance the repulsive dipolar forces and entropic losses associated with this conformation.



Hydrogen bonds are not the only stabilising forces acting upon peptide and protein α-helices in nature. Another common stabilising feature is the presence of an associated counter charge near one of the helix termini which can interact favourably with the helix dipole.[14,81,82] The desired alignment of carbonyl groups in the parent macrocycle is destabilised by repulsive dipole-dipole interactions but this arrangement could be stabilised by a positively charged group located on the helix axis at the *C*-terminus. The latter effect was demonstrated by Gellman *et al.* who appended a tetraalkylammonium group at the *C*-terminus of a depsipeptide with the resulting induction of an α-helical conformation.[74] The introduction of an anionic group at the *N*-terminus should exert a similar effect, but this is synthetically challenging and an on-going objective in our laboratory. However, since the conjugates **24** and **25** were showing helical characteristics, it was clearly of interest to further investigate the properties of the parent macrocycle **21** by placing a positive charge at the *C*-terminus. The question of whether the Pro^2 carboxamide group in conjugates of **21** could be aligned with the other template derived carboxamide groups needed to be addressed.

We chose to attach dimethylethylenediamine at the *C*-terminus of the macrocycle to give conjugate **26** since this strategy would introduce a hydrogen bond donor, and, after protonation of the tertiary amine, a suitably disposed cationic group. The amide bond formation was achieved by reacting the mixed anhydride of template acid directly with *N,N*-dimethylethylenediamine to afford the conjugate **26** as colourless crystals in 78% yield. The compound displayed the expected analytical and spectroscopic properties and was directly protonated with trifluoroacetic acid in DCM to give the required trialkylammonium salt.

In deuterochloroform solution the ^{1}H and ^{13}C NMR spectra of conjugate **26** showed one set of resonances, but with significant broadening of the resonance lines for the Pro^2 Cα and CO ($\Delta\nu_{1/2}$=30 Hz for Cα at 30 °C). NOESY assignments for the Ala^3 amide proton indicated several short H-H distances that could not be realised simultaneously in a single conformation. On cooling, the sample exhibited two distinct sets of resonances. The free energy of activation at coalescence ($\Delta G_c^{\ddagger}$) can be estimated from the following expression:[83]

$$\Delta G_c^{\ddagger} = 4.57\, T_c \left[9.97 + \log\left(T_c / \Delta\nu\right)\right]$$

where T_c is the coalescence temperature, $\Delta\nu$ is the difference in resonance frequency of the two spins in the slow exchange limit. For the Ala^3 NH, with a coalescence temperature of 243±5 K and $\Delta\nu$ of 87 Hz, the calculated free energy of activation for the transition is 48 kJ/mol. The corresponding exchange rate at coalescence was 190 s^{-1}.

At −55 °C conformational exchange rates were sufficiently low to obtain full ^{1}H and ^{13}C resonance assignments for each of two conformational isomers, with relative abundance 3:2 (estimated from peak integrals

of resolved amide proton resonances). For a two-site conformational exchange with relative populations in the ratio 3:2 the difference in free energy of the two sites is 1.0 kJ/mol. However, the conformational exchange was too rapid to separate the NOE's for each isomer.

In d_2-dichloromethane solution at −80 °C the full assignment of NOE spectra was still complicated by residual exchange between two conformers with almost degenerate chemical shifts. Conformational exchange also precluded the separation of spin systems by NMR filtration techniques requiring more than a few milliseconds mixing time. Only one NOE between two non-exchangeable protons, the NOE between Ala^3 Hα and Pro^4 Hδ, differs significantly in the two conformers (Table 2). NOE assignments for conjugate **26** at −80 °C were accomplished only for the NOE's involving NH protons (Table 2), but both conformers could be characterised sufficiently from the assignment of amide proton NOEs alone.

Table 2 *Modelling restraints, and comparison of experimental vs. modelled internuclear distances for amide proton signals of macrocycle* ***26*** *in $C^2H_2Cl_2$ solution at −80 °C.*

		Major Isomer			*Minor Isomer*		
From	*To*	*NOE*[a]	r_{noe}[b] (Å)	r_{model}[c] (Å)	*NOE*[a]	r_{noe}[b] (Å)	r_{model}[c] (Å)
Ala^3 NH	Ala^3 Hα	weak	2.7	3.0	strong	2.6	2.3
Ala^3 NH	Ala^3 Hβ	weak	2.7	3.1	weak	3.7	3.3
Ala^3 NH	Pro^2 Hα	strong	2.2	2.2	-	-	3.6
Ala^3 NH	Pro^2 Hβ	-	-	4.5	weak	3.4	2.4
Ala^3 NH	Pro^2 Hγ	-	-	5.9	weak	2.9	4.1
Ala^3 NH	Pro^2 Hδ	-	-	4.3	weak	3.3	2.7
Ala^3 NH	Pro^2 Hδ	weak	2.7	3.1	strong	2.7	2.4
TAA^5 NH	Pro^2 Hα	weak	3.4	4.4	-	-	6.2
TAA^5 NH	Pro^4 Hα	medium	2.7	3.4	weak	4.0	2.3
TAA^5 NH	Pro^4 Hδ	weak	3.1	3.6	weak	3.7	5.1
TAA^5 NH	TAA^5 Hα	strong	2.4	2.9	medium	3.2	2.4
TAA^5 NH	TAA^5 Hα'	strong	2.2	3.0	medium	2.9	3.0
TAA^5 NH	TAA^5 Hβ	weak	2.6	2.7	weak	3.7	3.7
TAA^5 NH	TAA^5 Hδ	weak	2.9	3.3	-	-	3.1
TAA^5 NH	Ala^3 Hβ	weak	3.8	4.9	-	-	4.8
TAA^5 Hγ	TAA^5 Hα	medium	2.9	2.7	-	-	3.2
TAA^5 Hγ	TAA^5 Hα'	medium	2.7	3.5	-	-	3.8
TAA^5 Hγ	TAA^5 Hβ	medium	2.5	2.6	-	-	2.5
TAA^5 Hγ	TAA^5 Hβ'	medium	2.6	2.9	-	3.3	3.0
TAA^5 Hγ	TAA^5 Hδ	strong	2.2	2.4	medium	2.9	2.4
TAA^5 Hγ	TAA^5 Hδ'	strong	2.2	2.4	-	-	2.4
TAA^5 Hγ	Pro^4 Hδ	-	-	4.7	weak	3.0	5.0
TAA^5 Hγ	Ala^1 Hβ	weak	3.2	4.0	-		5.9

a. defined as weak (r>3.0 Å), medium (2.5 < r < 3.0 Å), strong (r<2.5 Å); TAA=trialkylammonium.

b. effective time-averaged internuclear distance calculated from relative volume of NOESY cross-peaks, using the isolated spin-pair approximation (i.e. assuming a rigid molecule with isotropic reorientation).

c. internuclear distance in lowest energy model from simulated annealing.

In the major conformation the NOE between Ala^3 NH and Pro^2 Hα was strong, whilst that between Ala^3 NH and Ala^3 Hα was weak. In the minor conformer these relative intensities were reversed, indicating that the conformational difference was the orientation of the Pro^2–Ala^3 amide group. The variation was not, however, an amide bond isomerisation; a *cis* amide bond would give no NOE between Ala^3 NH and any Pro^2 protons, whereas NOE's were observed from Ala^3 NH to Pro^2 Hα in the major isomer and to Pro^2 Hβ and Pro^2 Hδ in the minor isomer. Relative NOE's were quantified using cross-peak volumes from a NOESY experiment, with 200 ms mixing time, for use as modelling restraints during simulated annealing, *vide infra*.

It was hoped that the helix cap would be capable of templating helices in neat aqueous solution, without additional co-solvents. In aqueous solution (95% H_2O, 5% ^2H_2O, 50 mM Tris-d_{11} buffer, pH 7.5) the ^{13}C spectra of compound **26** showed one set of resonance lines. Acetone-d_6 (15% v/v) was added to facilitate low temperature solution studies, and had an insignificant effect on ^{13}C chemical shifts, from which we infer that a low concentration of acetone did not alter the conformational preference. When cooled to −15 °C, some resonances of Pro^2 broadened appreciably; $\Delta\nu_{1/2}$ for Cβ and the carbonyl carbon were ~40 Hz and >100 Hz, respectively, and the Cα resonance was similarly broadened (but not resolved). The line broadening was consistent with the observations in deuterochloroform solution. Additionally, there was broadening of the *N*-methyl carbon resonances. Intensities of NOE's (measured from a NOESY spectrum with 200 ms mixing time) were characteristic of the two rapidly exchanging conformers observed previously in dichloromethane.

Models for each isomer were generated by restrained simulated annealing, using the distance restraints derived from the dichloromethane solution at −80 °C, as shown in Table 2. Two sets of simulations were performed, applying the constraints from the two isomers separately. AM1 *in vacuo* semi-empirical energy minimizations were performed on the lowest energy annealed model of each isomer, to give the optimised structures depicted in Fig. 9.

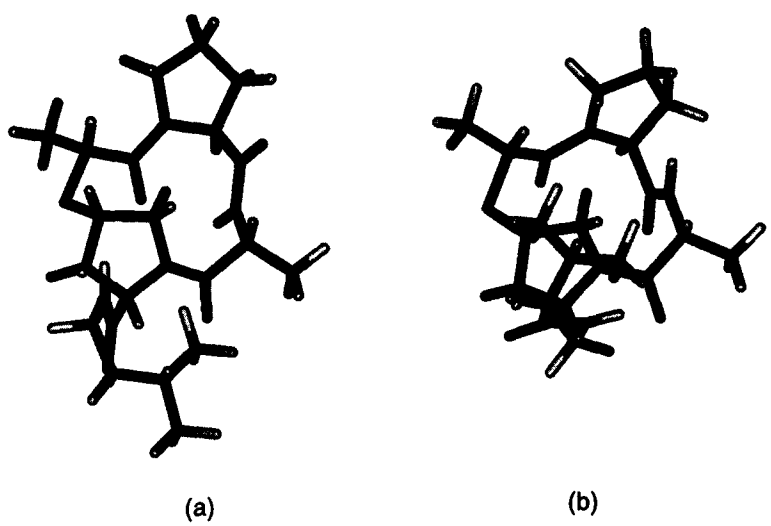


Figure 9 AM1-optimised models of (a) the major isomer, and (b) the minor isomer of the cap macrocycle with attached trialkylammounium salt **26**

In the AM1 optimised model of the target conformation (Fig. 9b) the ammonium proton is situated equidistant (2.2 Å) from the carbonyl oxygens of Pro^2 and the thiopropionyl moiety, suggestive of possible hydrogen bonding, but had no interactions with Ala^3.

Restrained and unrestrained molecular dynamics simulations for the minor (40%) conformation using a dielectric constant of 4 (approximating to a non-polar environment) showed no evidence of hydrogen bonding between the macrocycle and the pendant chain. Instead, the trialkylammonium cation was positioned beneath the centre of the macrocycle, allowing maximum charge-dipole interaction with the aligned carbonyl groups. When the dielectric constant was increased (to 80) to simulate an aqueous environment, the trialkylammonium group was found to drift away from the position directly beneath the macrocycle.

This striking change in macrocycle conformation shows that the generation of potential α-helix initiating templates is possible by judicious use of charge-dipole interactions. Modelling studies suggest that these initiating forms are stabilised solely by the charged ammonium group effect without contribution from hydrogen bonding. It should be noted that both structures are α-helical and that the only difference concerns the orientation of the dipole of the Pro^2 carbonyl group.

Future work is directed towards placing negatively charged groups at the *N*-terminal (top) of the template so that the effect of the entire range of amino acids on the stability of α-helical conformations can be assessed. The results of the study described here will also allow the conformation of biologically active sequences to be probed, for example, the activity of the 18-amino acid 2A fragment of the foot and mouth disease virus which it appears catalyses its own cleavage during ribosomal processing.[48,84]

Acknowledgements

We thank the Wellcome Trust for grant 040331, Dr. M. Akhtar for scientific support, Dr P. Lightfoot for providing the X-ray crystal structure of compound **21**, the Wellcome Trust for a studentship to AL and the BBSRC for grants B03817, B05162 and B06902 to DG.

References

1. J. S. Richardson, *Adv. Protein Chem.*, 1981, **34**, 167-339.
2. F. B. Sheinerman and C. L. Brooks III, *J. Am. Chem. Soc.*, 1995, **117**, 10098-10103.
3. L. Pauling, R. B. Corey and H. R. Branson, *Proc. Natl. Acad. Sci. USA*, 1951, **37**, 205-211.
4. B. H. Zimm and J. K. Bragg, *J. Chem. Phys.*, 1959, **31**, 526-535.
5. S. Lifson and A. Roig, *J. Chem. Phys.*, 1961, **34**, 1963-1974.
6. A. Yang and B. Honig, *J. Mol. Biol.*, 1995, **252**, 351-365.
7. N. E. Zhou, C. M. Kay, B. D. Sykes and R. S. Hodges, *Biochemistry*, 1993, **32**, 6190-6197.
8. H. A. Scheraga, *Pure Appl. Chem.*, 1978, **50**, 315-314.
9. A. Chakrabartty, T. Kortemme and R. L. Baldwin, *Protein Science*, 1994, **3**, 843-852.
10. A. J. Doig and R. L. Baldwin, *Protein Science*, 1995, **4**, 1325-1336.
11. J. J. Osterhout, R. L. Baldwin, E. J. York, J. M. Stewart, H. J. Dyson and P. E. Wright, *Biochemistry*, 1989, **28**, 7059-7064.
12. P. S. Kim and R. L. Baldwin, *Nature*, 1984, **307**, 329-334.

13. J. W. Nelson and N. R. Kallenbach, *Biochemistry*, 1989, **28**, 5256-5261.
14. K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart and R. L. Baldwin, *Nature*, 1987, **326**, 563-567
15. A. Bierzynski, P. S. Kim and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 2470-2474.
16. E. Bairaktari, D. F. Mierki, S. Mammi and E. Peggion, *J. Am. Chem. Soc.*, 1990, **112**, 5383-5384.
17. N. Khan, A. Graslund, A. Ehrenberg and J. Shriver, *Biochemistry*, 1990, **29**, 5743-5751.
18. E. T. Kaiser and F. J. Kézdy, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 1137-1143.
19. E. T. Kaiser and F. J. Kézdy, *Science*, 1984, **223**, 249-255.
20. I. L. Karle, *Biopolymers*, 1996, **40**, 157-180.
21. S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 8898-8902.
22. M. Reza Ghadiri and C. Choi, *J. Am. Chem. Soc.*, 1990, **112**, 1630-1632.
23. F. Ruan, Y. Chen and P. B. Hopkins, *J. Am. Chem. Soc.*, 1990, **112**, 9403-9404.
24. J. S. Albert, M. W. Peczuh and A. D. Hamilton, *Bioorg. Med. Chem.*, 1997, **5**, 1455-1467.
25. D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391-9392.
26. G. Osapay and J. W. Taylor, *J. Am. Chem. Soc.*, 1992, **114**, 6966-6973.
27. C. Yu and J. W. Taylor, *Tetrahedron Lett.*, 1996, **37**, 1731-1734.
28. J. C. Phelan, N. J. Skelton, A. C. Braisted and R. S. McDowell, *J. Am. Chem. Soc.*, 1997, **119**, 455-460.
29. G. R. Geier and T. Sasaki, *Tetrahedron Lett.*, 1997, **38**, 3821-3824.
30. L. G. Presta and G. D. Rose, *Science*, 1988, **240**, 1632-1641.
31. J. P. Schneider and J. W. Kelly, *Chem. Rev.*, 1995, **95**, 2169-2187.
32. A. Cammers-Goodwin, T. J. Allen, S. L. Oslick, K. F. McClure, J. H. Lee and D. S. Kemp, *J. Org. Chem.*, 1996, **118**, 3082-3090.
33. D. S. Kemp and T. P. Curran, *Tetrahedron Lett.*, 1988, **29**, 4931-4934.
34. D. S. Kemp and T. P. Curran, *Tetrahedron Lett.*, 1988, **29**, 4935-4938.
35. D. S. Kemp, T. P. Curran, W. M. Davis, J. G. Boyd and C. Muendel, *J. Org. Chem.*, 1991, **56**, 6672-6682.
36. D. S. Kemp, T. P. Curran, J. G. Boyd and T. J. Allen, *J. Org. Chem.*, 1991, **56**, 6683-6697.
37. D. S. Kemp, J. H. Rothman, T. P. Curran and D. E. Blanchard, *Tetrahedron Lett.*, 1995, **36**, 3809-3812.
38. D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 3813-3816.
39. D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 4019-4022.
40. D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 4023-4026.
41. K. F. McClure, P. Renold and D. S. Kemp, *J. Org. Chem.*, 1995, **60**, 454-457.
42. D. S. Kemp, T. J. Allen and S. L. Oslick, *J. Am. Chem. Soc.*, 1995, **117**, 6641-6657.
43. D. S. Kemp, T. J. Allen, S. L. Oslick and J. G. Boyd, *J. Am. Chem. Soc.*, 1996, **118**, 4240-4248.
44. D. S. Kemp, S. L. Oslick and J. G. Boyd, *J. Am. Chem. Soc.*, 1996, **118**, 4249-4255.
45. R. E. Austin, R. A. Maplestone, A. M. Sefler, K. Liu, W. N. Hruzewicz, C. W. Liu, H. S. Cho, D. E. Wemmer and P. A. Bartlett, *J. Am. Chem. Soc.*, 1997, **119**, 6461-6472.
46. T. Arrhenius and A. C. Satterthwait, *Peptides: Chemistry, Stucture and Biology: Proceedings of the 11th American Peptide Symposium, 1989*, 1990, 870-872.
47. K. Müller, D. Obrecht, A. Knierzinger, C. Stankovic, C. Spiegler, W. Bannwarth, A. Trzeciak, G. Englert, A. M. Labhardt and P. Schönholzer, *Perspectives in Medicinal Chemistry*, 1993, 513-531.
48. M. L. L. Donnelly, D. Gani, M. Flint, S. Monaghan and M. D. Ryan, *J. Gen. Virol.*, 1997, **78**, 13-21.
49. D. S. Kemp, J. G. Boyd and C. C. Muendel, *Nature*, 1991, **352**, 451-454.
50. K. Groebke, P. Renold, K. Y. Tsang, T. J. Allen, K. F. McClure and D. S. Kemp, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 4025-4029.
51. M. Narita, S. Isokawa, M. Doi and R. Wakita, *Bull. Chem. Soc. Jpn.*, 1986, **59**, 3547-3552.

52. V. V. Ryakhovskii, S. V. Agafonov and Y. M. Kosyrev, *Russ. Chem. Rev. (Engl. Transl.)*, 1991, **60**, 924-933.
53. M. M. Lenman, A. Lewis and D. Gani, *J. Chem. Soc., Perkin Trans. 1*, 1997, 2297-2311.
54. A. Lewis, M. D. Ryan and D. Gani, *J. Chem. Soc., Perkin Trans. 1*, 1998, Part 1-submitted
55. Y. V. Venkatachalapathi and P. Balaram, *Nature*, 1979, **281**, 83-84.
56. K. Sukata, *Bull. Chem. Soc. Jpn.*, 1985, **58**, 838-843.
57. J. D. Park, R. D. Englert and J. S. Meek, *J. Am. Chem. Soc.*, 1952, **74**, 1010-1012.
58. A. Koziara, S. Zawadzki and A. Zwierzak, *Synth. Commun.*, 1979, 527-529.
59. N. Ono, T. Yoshimura, T. Saito, R. Tamura, R. Tanikaga and A. Kaji, *Bull. Chem. Soc. Jpn.*, 1979, **52**, 1716-1719.
60. D. E. Dorman and F. A. Bovey, *J. Org. Chem.*, 1973, **38**, 2379-2383.
61. M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.*, 1991, **218**, 397-412.
62. A. Lewis, J. Wilkie, T. J. Rutherford and D. Gani, *J. Chem. Soc., Perkin Trans. 1*, Part 2, 1998-submitted
63. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner, *J. Am. Chem. Soc.*, 1984, **106**, 765-784.
64. Insight II, 9685 Scranton Road, San Diego, CA 92121-2777, USA
65. E. Benedetti, *Biopolymers*, 1996, **40**, 3-44.
66. R. Nagaraj, N. Shamala and P. Balaram, *J. Am. Chem. Soc.*, 1979, **101**, 16-20.
67. S. K. Awasthi, S. R. Raghothama and P. Balaram, *J. Chem. Soc., Perkin Trans. 2*, 1996, 2701-2706.
68. P. G. Sammes and D. J. Weller, *Synthesis*, 1995, 1205-1222.
69. N. G. Delaney and V. Madison, *J. Am. Chem. Soc.*, 1982, **104**, 6635-6641.
70. J. L. Flippen-Anderson, R. Gilardi, V. Madison and N. G. Delaney, *J. Am. Chem. Soc.*, 1983, **105**, 6609-6614.
71. C. Bisang, C. Weber, J. Inglis, C. A. Schiffer, W. F. van Gunsteren, I. Jelesarov, H. R. Bosshard and J. A. Robinson, *J. Am. Chem. Soc.*, 1995, **117**, 7904-7915.
72. D. Seebach, M. Boes, R. Naef and W. B. Schweizer, *J. Am. Chem. Soc.*, 1983, **105**, 5390-5398.
73. E. A. Gallo and S. H. Gellman, *J. Am. Chem. Soc.*, 1993, **115**, 9774-9788.
74. E. A. Gallo and S. H. Gellman, *J. Am. Chem. Soc.*, 1994, **116**, 11560-11561.
75. R. M. Wenger, *Helv. Chim. Acta*, 1984, **67**, 502-525.
76. S. H. Gellman, G. P. Dado, G. Liang and B. R. Adams, *J. Am. Chem. Soc.*, 1991, **113**, 1164-1173.
77. M. Oka, G. T. Montelione and H. A. Scheraga, *J. Am. Chem. Soc.*, 1984, **106**, 7959-7969.
78. A. Lewis, J. Wilkie, T. J. Rutherford and D. Gani, *J. Chem. Soc., Perkin Trans. 1*, Part 3, 1998-submitted
79. A. Pardi, M. Billeter and K. Wüthrich, *J. Mol. Biol.*, 1984, **180**, 741-751.
80. M. Blaber, X. Zhang and B. W. Matthews, *Science*, 1993, **260**, 1637-1640.
81. D. Sali, M. Bycroft and A. R. Fersht, *Nature*, 1988, **335**, 740-743.
82. S. Dasgupta and J. A. Bell, *Int. J. Pept. Protein Res.*, 1993, **41**, 499-511.
83. J. M. Lehn, F. G. Riddell, B. J. Price and I. O. Sutherland, *J. Chem. Soc., B*, 1967, 387-390.
84. M. D. Ryan, M. L. L. Donnelly, A. Lewis, A. P. Mehrotra, J. Wilkie and D. Gani, *Bioorg. Chem.* 1998, in the press.