Design, construction and properties of peptide *N*-terminal cap templates devised to initiate α-helices. Part 1. Caps derived from *N*-(4-chlorobutyryl)-(2*S*)-Pro-(2*S*)-Pro-(2*S*)-Ala-OMe and *N*-[(2*S*)-2-chloropropionyl]-(2*S*)-Pro-(2*S*)-Pro-(2*S*,4*S*)-4-hydroxyPro-OMe

Arwel Lewis,[a] Martin D. Ryan[b] and David Gani*[a]

[a] *School of Chemistry and Centre for Biomolecular Sciences, The Purdie Building, The University, St. Andrews, Fife, UK KY16 9ST*
[b] *Department of Biochemistry, School of Biology and Medical Sciences, Irvine Building, The University, St. Andrews, Fife, UK KY16 9AL*

Received (in Cambridge) 31st July 1998, Accepted 22nd September 1998

The structural features required of *N*-terminal helix caps which might display high α-helix-forming propensities (Zimm–Bragg σ-values close to unity) were considered and two systems were identified as initial candidates for synthesis and evaluation. The first system was an eleven-membered cyclic all-*trans* triamide derived from *N*-(4-chlorobutyryl)-(2*S*)-Pro-(2*S*)-Pro-(2*S*)-Ala-OMe in which it was intended to remove the only amide NH proton to generate a better nucleophile, and then ring-close by displacing the 4-halogen atom as halide. The precursor was prepared in moderate overall yield and displayed conformational isomerism in chloroform solution. Some of the conformers appeared to contain one or more *cis* amide bonds. No suitable conditions could be found to effect macrocyclisation. The second system was a twelve-membered cyclic all-*trans* triamide derived from *N*-[(2*S*)-2-chloropionyl]-(2*S*)-Pro-(2*S*)-Pro-(2*S*,4*S*)-4-hydroxyPro-OMe in which it was intended to generate an alkoxide anion and then ring-close by displacing halide ion from the 2-chloropropionyl moiety. The precursor was prepared in moderate overall yield and populated one all-*trans* rotameric form in chloroform solution. This system also failed to macrocyclise under a variety of conditions. The major reason for the lack of reactivity in each system was analysed and appeared to be the generation of large molecular dipoles in the transition state for the cyclisation, originating from the requirement for the alignment of the carboxamide carbonyl groups. Useful modifications to the systems described here that were designed to overcome or minimise the effect of the dipole in destabilising the required conformation for helix cap generation are presented in the following two articles.

Introduction

The polypeptide chains of 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, the latter of which can occur in orientations where the polypeptide chains are aligned in either an antiparallel or a parallel fashion.

For polypeptide helices, there are two major structural types, both of which are right-handed. These differ in the number of amino acid residues that are required to make one turn of the helix and the number of atoms that are involved in the smallest hydrogen bonded network. 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** (Fig. 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** (Fig. 2). The 3_{10} helix is tighter and smaller in diameter than the α-helix and, in natural polypeptides, is also 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. Indeed, 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 constraints, the sum of the largely enthalpic advantages of forming H-bonds within the helix, together 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. 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,

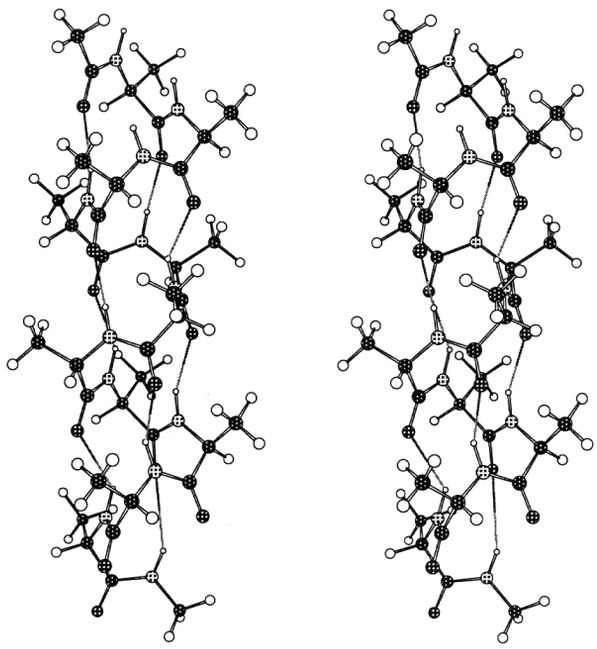


Fig. 1 Structure **1**: stereoview of α-helical arrangement of *N*-acetyl-poly-(2*S*)-alanine methylamide.

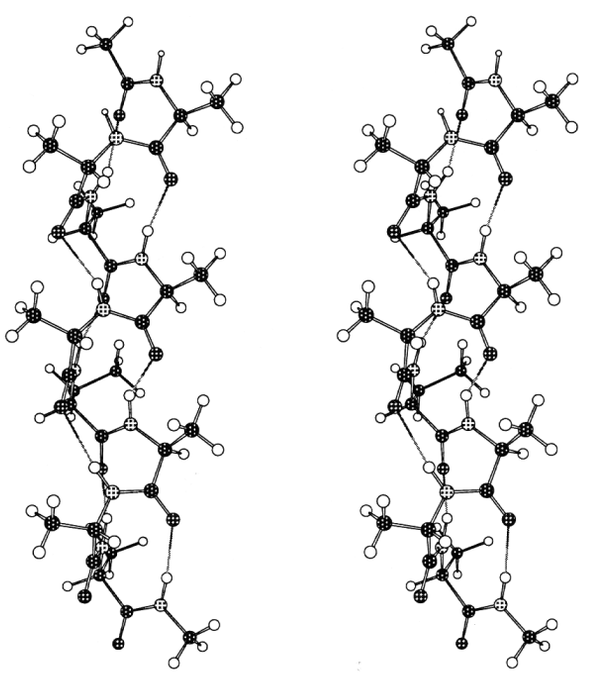


Fig. 2 Structure **2**: stereoview of 3_{10}-helical arrangement of *N*-acetyl-poly-(2*S*)-alanine methylamide.

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 helix-*breaking* (*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 amphilicity, 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*-terminals 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. 3, 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 conformations *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 (FMDV)

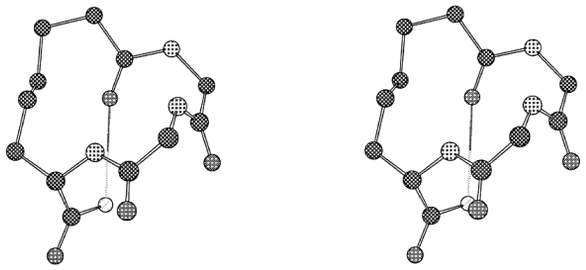


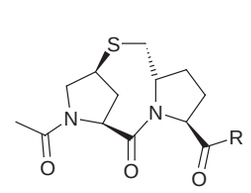
Fig. 3 Stereoview of an *N*-terminal template in an α-helical conformation. The carboxamide O-atoms are aligned downwards towards the *C*-terminal end and the position of the first potential H-bond for the first N-atom of appended peptides is shown. The structure shown was cut from an extended peptide α-helix bearing the displayed template, without altering the template coordinates.

polyprotein, where it is believed that the 19-amino acid residue 2A region forms an α-helix-type-VI reverse turn structure in the exit pore of eukaryotic ribosomes that allows it to catalyse self-cleavage during translation.[48]

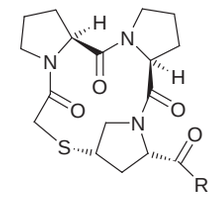
A major problem in all template designs (*e.g.* Fig. 3) is the requirement to align H-bond donors or acceptors in the correct geometry for helix initiation. The aligned carbonyl groups and NH moieties in α-helices and helix-like structures confer significant dipolar electric fields, positive at the *N*-terminus 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.

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 a 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**.

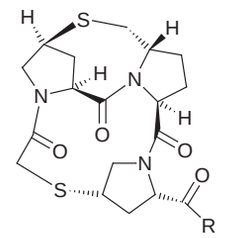
Macrocycle **4** can potentially supply three aligned carbonyl groups only if all three amide bonds within the ring possess the *trans* configuration (*ttt*). In solution, the macrocycle was found to exist as two conformers, the *cct* form and *ctt* form, and did not have useful helix-initiating properties.[37,38] 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]



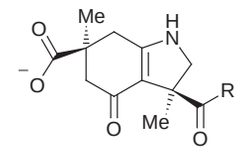
3



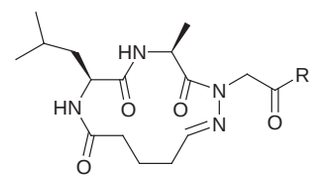
4



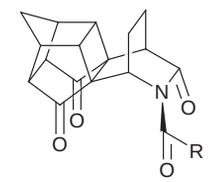
5



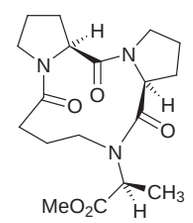
6



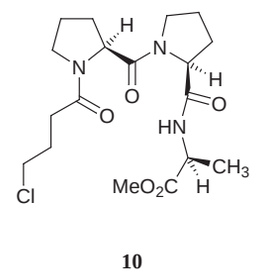
7



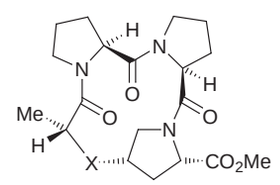
8



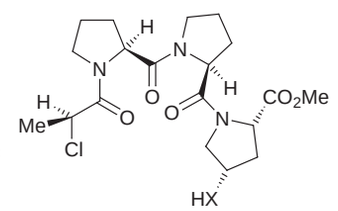
9



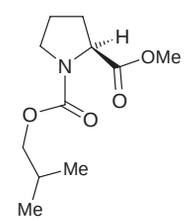
10



11



12



13

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 carboxy group on the six-membered ring to be deprotonated in order to function, and presumably works by forming a favour-

able 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 six 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. In this and the following two articles we describe the preliminary results of work in this area and demonstrate that such systems can be synthesised.

Results and discussion

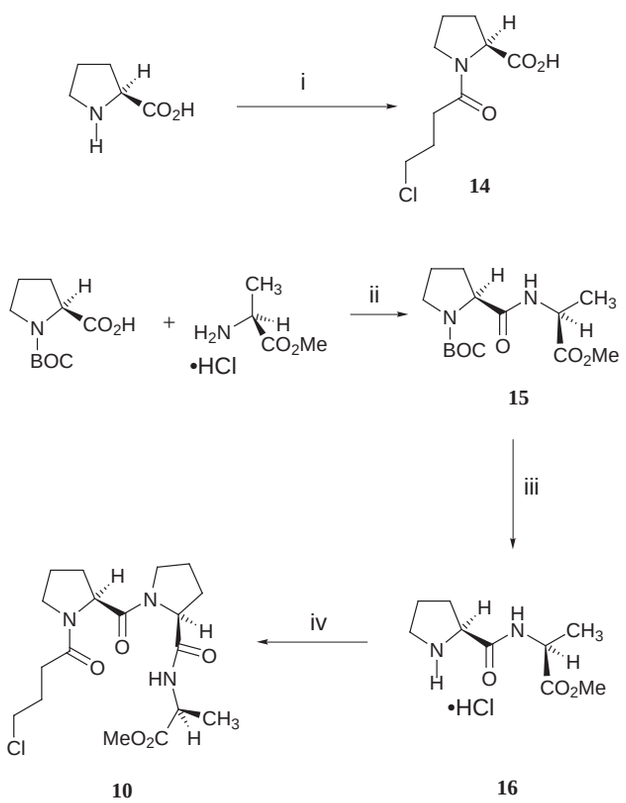
Identification of a suitable macrocyclic framework

At the start of our programme in this area, which stemmed from the need to assess structural features of 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, 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*-rotameric 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 on 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.

The synthesis of tripeptide **10** was achieved in four steps starting from commercially available materials. (2*S*)-Proline was directly acylated using 4-chlorobutyryl chloride to give the *N*-acyl derivative **14**. (2*S*)-*N*-Boc-proline was coupled to (2*S*)-alanine methyl ester using mixed anhydride methodology to give the protected dipeptide **15** in moderate yield (Scheme 1).



Scheme 1 *Reagents and conditions*: i, 4-chlorobutyryl chloride, EtOAc, reflux, 75 min, 94%; ii, IBCF, NMM, THF, DMF, −15 °C→rt, 2 h, 65%; iii, HCl(g), EtOAc, 0 °C, 70 min, 99%; iv, **14**, $SOCl_2$, pyridine, DMAP, CH_2Cl_2, 2 days, 39%.

Subsequent removal of the *N*-protecting group with HCl afforded dipeptide ester hydrochloride **16** in quantitative yield.

The reaction of the amine **16** with (2*S*)-*N*-(4-chlorobutyryl)-

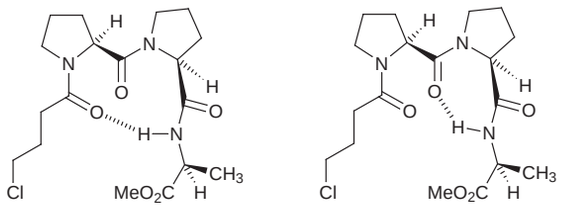


Fig. 4 Possible conformations available to peptide **10**.

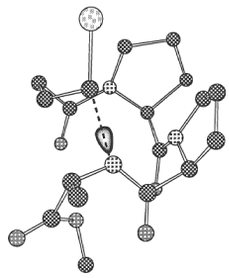


Fig. 5 Transition-state conformation required for the cyclisation of compound **10** calculated by modelling the best trajectory for opening the product macrocycle with chloride ion in the reverse reaction direction.

proline **14** requires an amide bond formation between two proline residues. This was known and was expected to be a difficult type of peptide coupling reaction, due to steric hindrance.[52] Standard mixed anhydride methodology and the use of the more powerful coupling reagents TBTU and pyBroP all failed to produce the required acylated tripeptide **10** but the use of thionyl chloride and pyridine methodology, which had already proven its utility for other hindered systems[53] afforded **10** in 25–-39% yield as an oil after chromatographic purification on flash silica. The compound eluted as a single band on analytical TLC plates in several solvents and gave the expected analytical data. 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. 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,[54] 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) (Fig. 4).

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 alklylations had been achieved using a variety of bases.[55–58] 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.

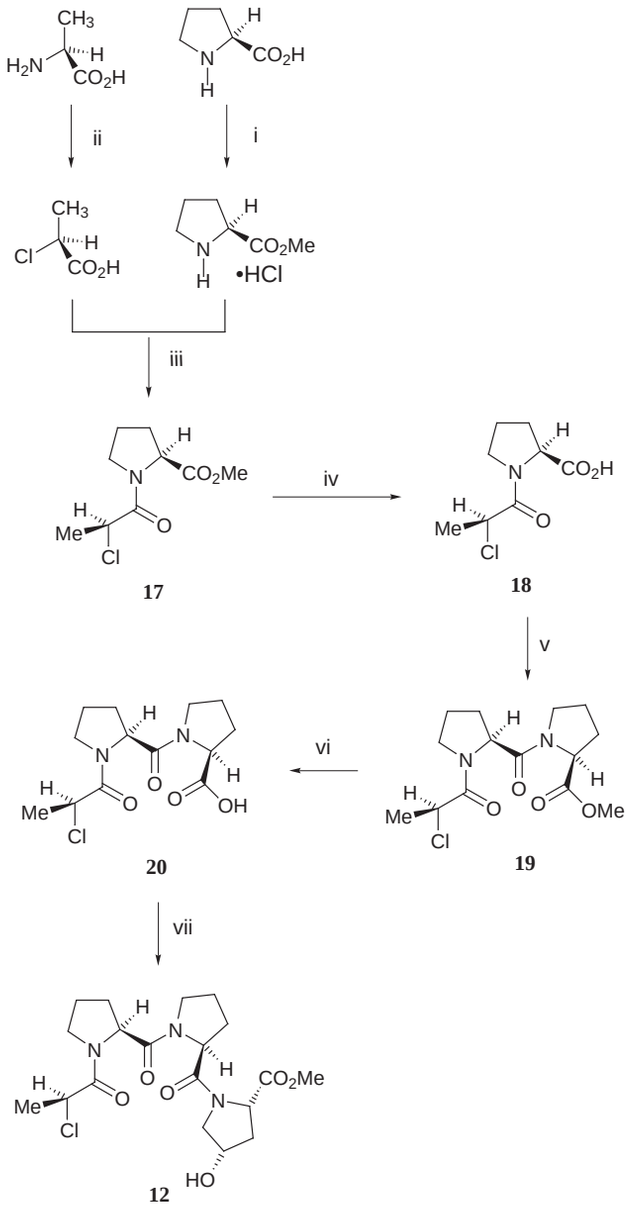
In rationalising the failure of the chloride **10** to react, it is expected that the cyclisation was hampered by the hindered nature of the nucleophile, and/or, the unfavourable conformations adopted by the acyclic compound in the presence of metal ions that can chelate the anion and other carbonyl 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-turns shown in Fig. 4, 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 tris-lactam **9**, all three carboxamide carbonyl groups 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, Fig. 5.

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*)-3-hydroxyproline residue [or a (2*S*,4*S*)-3-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 3-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. This was assessed by considering the stability of the *ctt* form of the linear precursor and the size of the steric interactions required for cyclisation. 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.[59]

The synthesis of linear precursor **12** was achieved in eight steps starting from the amino acids (2*S*)-proline, (2*S*,4*S*)-4-hydroxyproline and (2*S*)-alanine, Scheme 2. (2*S*)-Alanine was converted to (2*S*)-2-chloropropionic acid, with retention of configuration at C-2 *via* the established diazotisation–-chlorination protocol.[60] Condensation of the resulting acid with (2*S*)-proline methyl ester using standard mixed anhydride methodology gave 35–50% of the desired dipeptide **17**. The conversion efficiency was not improved by the use of the more highly activating reagents BOP-Cl or PyBOP. Despite its somewhat hindered nature, the amine component was not expected to present any real difficulties. Thus, the problem appeared to rest in the activation of the α-chloro acid, and indeed, in the case of the isobutyl chloroformate mediated reaction, significant amounts of a urethane **13** were formed. Variations in reaction conditions indicated that at least some of this material was derived through attack of the proline amine group on the carbonate carbonyl group.

The acylated proline methyl ester **17** was converted to the corresponding acid **18** through alkaline hydrolysis without disturbing the stereochemical integrity of the chiral centre at C-2. Note that while we had considered the possibility of using the more electrophilic 2-bromopropionyl analogue which after conversion to the 2-bromopropionyl analogue of **12** might have been easier to cyclise in the final step, we had shown that the C-2 position of the 2-bromopropionyl analogue of ester **17** was



Scheme 2 *Reagents and conditions*: i, $SOCl_2$, MeOH, reflux, 90 min, 100%; ii, $NaNO_2$, HCl, KCl, H_2O, 0–5 °C, 2 h, 73%; iii, IBCF, NMM, THF, DMF, −15 °C→rt, 18 h; iv, 1.0 mol dm^{-3} NaOH, MeOH, H_2O, 2 h, 36% (2 steps); v, (2*S*)-proline methyl ester hydrochloride, $SOCl_2$, pyridine, DMAP, CH_2Cl_2, 4 days, 33%; vi, 1.0 mol dm^{-3} NaOH, MeOH, H_2O, 90 min, 98%; vii, methyl (2*S*,4*S*)-4-hydroxyproline hydrochloride, $SOCl_2$, pyridine, DMAP, CH_2Cl_2, 2 days, 42%.

epimerised during saponification.[53] The 2-chloropropionyl proline **18** was coupled to (2*S*)-proline methyl ester using the acid chloride method to afford the tripeptide **19** in 33% yield and saponification of the methyl ester gave the acylated bis-proline **20** in almost quantitative yield. The final residue to be attached in order to complete the sequence was (2*S*,4*S*)-4-hydroxyproline methyl ester. This was achieved through activation of the acylated bis-proline **20** with thionyl chloride in the presence of pyridine and DMAP to afford the required triprolyl peptide alcohol **12** (X = O) in 42% yield. The compound was extremely polar but nevertheless displayed the expected analytical data and showed only one predominant set of signals in ^1H 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** 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** existed in the ground state.

The intended cyclisation of compound **12** 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 depronating 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 not 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. The synthesis and properties of sulfur analogues of compounds **11** and **12** are presented in the following two articles.

Experimental

NMR spectra were recorded on a Bruker AM-300 spectrometer (^1H, 300 MHz; ^{13}C, 75.4 MHz), a Varian Gemini spectrometer (^1H, 200 MHz; ^{13}C, 50.3 MHz), a Varian Gemini spectrometer (^1H, 300 MHz; ^{13}C, 74.76 MHz) and a Varian Unity Plus 500 spectrometer (^1H, 500 MHz; ^{13}C, 125.6 MHz). ^1H-NMR spectra were referenced internally to $(C^2H_3)_2SO$ (δ 2.47), ^2HOH (δ 4.68) or C^2HCl_3 (δ 7.27). ^{13}C-NMR were referenced to $(C^2H_3)_2SO$ (δ 39.70) or C^2HCl_3 (δ 77.5). *J* Values are given in Hz. Carbon and proton resonances of amino acids in NMR spectra are assigned as α, β, γ and δ according to normal convention. Where more than one conformational isomer is present due to the presence of a tertiary amide bond, these are assigned as *c* (*cis*) or *t* (*trans*), according to the isomeric state of the amide bond. If the isomeric states of the amide bonds are not known, the conformations are assigned as *A*, *B*, *C etc.* The three-letter codes Chp and Thp refer to (2*S*)-2-chloropropionyl and (2*R*)-2-thiopropionyl residues respectively.

Mass spectra and accurate mass measurements were recorded on a VG 70-250 SE, a Kratos MS-50 or by the SERC service at Swansea using a VG AZB-E. Fast atom bombardment spectra were recorded using glycerol as a matrix. Major fragments were given as percentages of the base peak intensity (100%). Infrared spectra were recorded using a Perkin-Elmer 1710 FT-IR spectrometer. The samples were prepared as Nujol

mulls or thin films between sodium chloride discs. Absorption maxima are given in wavenumbers (cm^{-1}) relative to a polystyrene standard. Melting points were measured using an electrothermal melting point apparatus and are uncorrected. Optical rotations were measured on an Optical Activity AA-1000 polarimeter using 10 cm path lengths cells at room temperature and are given in units of 10^{-1} deg cm^2 g^{-1}. Thionyl chloride was distilled over sulfur and the initial fractions were always discarded. *N*-Methylmorpholine was distilled over ninhydrin. All other reagents were used without further purification. Light petroleum refers to the fraction boiling at 40–60 °C and "ether" refers to diethyl ether.

All experiments were performed at room temperature (20–25 °C) unless otherwise stated. Flash chromatography was performed according to the method of Still *et al.*[62] using Fluka C60 (40–60 mm mesh) silica gel. Analytical thin layer chromatography was carried out on 0.25 mm precoated silica gel plates (Macherey-Nagel SIL g/UV254) and compounds were visualised using UV fluorescence, iodine vapour or ninhydrin.

Abbreviations

Boc, *tert*-butoxycarbonyl; BOP-Cl, *N*,*N*-bis(2-oxo-3-oxazolidinyl)phosphinic chloride; CD, circular dichroism; DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; DMAP, 4-(*N*,*N*-dimethylamino)pyridine; DMF, *N*,*N*-dimethylformamide; FMDV, foot and mouth disease virus; IBCF, isobutyl chloroformate; LDA, lithium diisopropylamide; NMM, *N*-methylmorpholine; PyBOP, benzotriazol-1-yloxytrispyrrolidinophosphonium hexafluorophosphate; pyBroP, bromotripyrrolidinophosphonium hexafluorophosphate; TBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate; TFE, 2,2,2-trifluoroethanol; THF, tetrahydrofuran.

(2*S*)-*N*-(4-Chlorobutyryl)proline 14

A suspension of (2*S*)-proline (4.0 g, 34.7 mmol) in ethyl acetate (100 cm^3) was treated with 4-chlorobutyryl chloride (6.0 cm^3, 53.5 mmol), and the mixture heated at reflux for 75 min. After cooling to room temperature, the solvent was removed under reduced pressure. The residual oil was washed with several portions of ethyl acetate–light petroleum (1:1) to remove unreacted acid chloride, affording (2*S*)-*N*-(4-chlorobutryryl)proline **14** as a pale yellow oil (7.2 g, 94%) (HRMS: Found M^+, 220.0740. $C_9H_{15}{}^{35}ClNO_3$ requires 220.0736); $[\alpha]_D$ −19.9 (*c* 1.0 in MeOH); ν_{max}(thin film)/cm^{-1} 2966 (OH), 1735 (acid CO) and 1605 (amide CO); δ_H(300 MHz; C^2HCl_3) 1.93–2.27 (6 H, m, βCH_2, γCH_2 and CH_2CH_2Cl), 2.47–2.55 (2 H, m, βCH_2 and CH_2Cl), 3.47–3.69 (4 H, m, δCH_2 and CH_2CO), 4.50 (1 H, m, αCH) and 10.14 (1 H, br s, CO_2H); δ_C(74.76 MHz; C^2HCl_3) 24.96 (γCH_2), 27.82 (CH_2CH_2Cl), 28.92 (βCH_2), 31.33 (CH_2Cl), 44.83 (CH_2CO), 48.03 (δCH_2), 59.78 (αCH), 173.34 (amide CO), 174.76 (acid CO); *m*/*z* (EI) 222 and 220 (2 and 5%, chlorine isotopes, M^+), 176 and 174 (13 and 25, chlorine isotopes, $[M - CO_2H - H]^+$), 157 (14, $[M - CH_2CH_2Cl]^+$), 140 (26, $[M - CO_2H - Cl]^+$), 107 and 105 [4 and 14, chlorine isotopes, $CO(CH_2)_2Cl^+$], 70 (100, $C_4H_8N^+$) and 41 (37, $C_3H_5^+$).

Methyl (2*S*)-*N*-[(2*S*)-*N*-(*tert*-butoxycarbonyl)prolyl]alaninate 15

A solution of (2*S*)-*N*-(*tert*-butoxycarbonyl)proline (2.5 g, 11.6 mmol) and *N*-methylmorpholine (1.3 cm^3, 11.6 mmol) in dry THF (20 cm^3) was cooled to −15 °C, and isobutyl chloroformate (1.55 cm^3, 11.4 mmol) was added with stirring. After 2 min, a solution of (2*S*)-alanine methyl ester hydrochloride (1.6 g, 11.6 mmol) and *N*-methylmorpholine (1.3 cm^3, 11.6 mmol) in THF (10 cm^3) and DMF (26 cm^3) was added in one portion to the stirred suspension. The reaction mixture was allowed to warm to room temperature over 2 h and left standing overnight. The hydrochloride salts were filtered off and the solvent removed under reduced pressure. The residual oil was redissolved in ethyl acetate (50 cm^3) and the solution washed with aqueous HCl (0.5 mol dm^{-3}, 2 × 30 cm^3), aqueous sodium hydrogen carbonate (5%, 2 × 30 cm^3) and brine (30 cm^3). The solution was dried ($MgSO_4$), and the solvent removed under reduced pressure to afford the product as a white solid which was recrystallised from ethyl acetate–light petroleum to yield the dipeptide **15** as colourless crystals (2.4 g, 65%), mp 78–79 °C (lit.,[63] 79–81 °C); $[\alpha]_D$ −94.8 (*c* 1.1 in MeOH) [lit.,[63] −92.3 (*c* 1.0 in MeOH)]; ν_{max}(Nujol)/cm^{-1} 3267 (NH), 1752 (ester CO), 1695 (urethane CO), 1658 (amide CO) and 1163 (C–O); δ_H(200 MHz; C^2HCl_3) 1.33 [3 H, d, *J* 7.2, βCH_3(Ala)], 1.41 [9 H, s, $C(CH_3)_3$], 1.73–1.98 [3 H, m, $\frac{1}{2}\beta CH_2$(Pro) and γCH_2(Pro)], 2.01–2.27 [1 H, m, $\frac{1}{2}\beta CH_2$(Pro)], 3.31–3.50 (2 H, m, δCH_2), 3.68 (3 H, s, CO_2CH_3), 4.23 [1 H, m, αCH(Ala)], 4.50 [1 H, m, αCH(Pro)], 6.61 and 7.34 (*c* and *t*, 1 H, br s, NH); δ_C(74.76 MHz; C^2HCl_3) 17.90 [βCH_3(Ala)], 23.64 [*c*, γCH_2(Pro)], 24.34 [*t*, γCH_2(Pro)], 28.18 [$C(CH_3)_3$ and *t*, βCH_2(Pro)], 30.88 [*c*, βCH_2(Pro)], 46.93 [δCH_2(Pro)], 47.88 [αCH(Ala)], 52.15 (CO_2CH_3), 59.71 and 60.59 [*c* and *t*, αCH(Pro)], 80.12 [$C(CH_3)_3$], 154.49 and 155.48 (*c* and *t*, urethane CO), 172.43 (amide CO) and 173.22 (ester CO); *m*/*z* (EI) 300 (2%, M^+), 227 (7, $[M - OC_4H_9]^+$), 199 (7, $[M - CO_2C_4H_9]^+$), 170 (31, $C_8H_{14}N_2)O_2^+$), 114 (71, $C_5H_{10}N_2O^+$), 70 (100, $C_4H_8N^+$) and 57 (72, $C_4H_9^+$).

Methyl (2*S*)-*N*-{(2*S*)-*N*-[(2*S*)-*N*-(4-chlorobutyryl)prolyl]prolyl}-alaninate 10

Dipeptide **15** (2.09 g, 7.0 mmol) was dissolved in ethyl acetate (75 cm^3) at 0 °C and dry HCl gas was bubbled through the cold solution for 70 min. The solution was allowed to stand at 0 °C for 90 min, after which time the solvent was removed under reduced pressure to yield the amine hydrochloride salt **16** as a clear oil (1.64 g, 99%).

A solution of (2*S*)-*N*-(4-chlorobutyryl)proline **14** (1.55 g, 7.1 mmol) and pyridine (0.89 cm^3, 11.0 mmol) in dichloromethane (40 cm^3) was treated dropwise with thionyl chloride (0.66 cm^3, 9.0 mmol). After 5 min, a solution of the amine hydrochloride **16** (1.20 g, 5.1 mmol) and DMAP (1.24 g, 10.2 mmol) in dichloromethane (35 cm^3) was added in one portion. The solution was allowed to stir for 2 days, after which time it was washed with aqueous HCl (0.5 mol dm^{-3}, 2 × 50 cm^3), aqueous sodium hydrogen carbonate (5%, 2 × 50 cm^3) and brine (50 cm^3). The solution was dried ($MgSO_4$), and the solvent removed under reduced pressure to afford a brown oil, which was purified by column chromatography using ethyl acetate–methanol (9:1) as the eluent to yield the peptide **10** as an oil (0.79 g, 39%); R_f 0.20 (HRMS: Found M^+, 401.1712. $C_{18}H_{28}{}^{35}ClN_3O_5$ requires 401.1717); $[\alpha]_D$ −84.5 (*c* 1.6 in MeOH); ν_{max}(thin film)/cm^{-1} 3329 (NH), 2955 (CH), 1739 (ester CO), 1652 (amide CO) and 1165 (C–O); δ_H(300 MHz; C^2HCl_3) (3 conformations) 1.32, 1.38 and 1.45 (3 H, d, *J* 7.2, βCH_3), 1.89–2.19 [10 H, m, 2 × βCH_2(Pro), 2 × γCH_2(Pro) and CH_2CH_2Cl], 2.45–2.49 (2 H, m, CH_2Cl), 3.50–3.71 [6 H, m, 2 × δCH_2(Pro) and CH_2CO], 3.63 (3 H, s, CO_2CH_3), 3.97 and 4.20–4.58 [3 H, m, 2 × αCH(Pro) and αCH(Ala)], 7.10, 7.55 and 7.68 (1 H, br d, NH); δ_C(74.76 MHz; C^2HCl_3) 16.35, 17.23 and 17.90 [βCH_3(Ala)], 22.41, 22.78, 24.22, 24.94, 25.27 and 25.43 [2 × γCH_2(Pro)], 27.58 and 28.01 (CH_2CH_2Cl), 28.53 and 28.95 [2 × βCH_2(Pro)], 31.26 and 31.54 (CH_2Cl), 44.51 and 44.73 (CH_2CO), 47.15, 47.34 and 47.62 [2 × δCH_2(Pro)], 48.25 and 48.89 [αCH(Ala)] 52.17 and 52.40 (CO_2CH_3), 57.91, 58.17 and 58.85 [αCH(Pro)], 59.82, 60.45 and 61.11 [2 × αCH(Pro)], 170.76, 171.50, 171.73 and 171.99 (amide CO), 173.18 and 173.52 (ester CO); *m*/*z* (EI) 403 and 401 (1 and 4%, chlorine isotopes, M^+), 204 and 202 (9 and 29, chlorine isotopes, $C_9H_{13}NO_2Cl^+$), 174 (51, $C_8H_{13}NOCl^+$) and 70 (100, $C_4H_8N^+$).

Methyl (2*S*,4*S*)-4-hydroxyprolinate hydrochloride

Thionyl chloride (0.35 cm^3, 4.1 mmol) was added dropwise to

methanol (10 cm^3) at 0 °C. (2*S*,4*S*)-4-Hydroxyproline (0.40 g, 2.3 mmol) was added and the resulting solution heated under reflux for 90 min. The solvents were removed under reduced pressure to afford the hydrochloride as a white solid in quantitative yield. A small portion was recrystallised from ethyl acetate–hexane to yield colourless needles, mp 79–80 °C (Found: C, 36.4; H, 6.75; N, 7.5. $C_6H_{11}NO_3{\cdot}HCl{\cdot}H_2O$ requires C, 36.3; H, 7.1; N, 7.05%); $[\alpha]_D$ −49.2 (*c* 0.5 in MeOH); ν_{max}(Nujol)/cm^{-1} 3327 (OH), 2921 (CH), 1732 (CO), 1252 (OH def.) and 1182 (C–O); δ_H(300 MHz; ^2H_2O) 2.27–2.39 (2 H, m, βCH_2), 3.34 (2 H, br s, δCH_2), 3.72 (3 H, s, CO_2CH_3) and 4.51–4.60 (2 H, m, αCH and γCH); δ_C(74.76 MHz; ^2H_2O) 36.76 (βCH_2), 53.55 (δCH_2), 53.93 (CO_2CH_3), 58.45 (αCH), 69.00 (γCH) and 170.93 (ester CO); *m*/*z* (EI) 145 (11%, $[M - HCl]^+$), 101 (13, $[M - HCl - CH_2NH - CH_3]^+$), 86 (100, $[M - HCl - CO_2CH_3]^+$), 68 (44, $C_4H_6N^+$) and 58 (16, $NHCH_2CHO^+$).

(2*S*)-2-Chloropropionic acid

A stirred suspension of (2*S*)-alanine (15.0 g, 0.17 mol) in saturated aqueous potassium chloride (36 cm^3) was treated dropwise with concentrated HCl (60 cm^3). The mixture was cooled to 0 °C, and sodium nitrite (23.3 g, 0.34 mol) was added in portions over 75 min. After completion of addition, the brown solution was stirred at 0–5 °C for a further 1 h, then allowed to reach room temperature overnight. The solution was extracted with ether (3 × 100 cm^3), the organic phase dried ($MgSO_4$), and the solvent removed under reduced pressure to yield the acid as a pale yellow oil (13.4 g, 73%), bp 80 °C/~15 mmHg (lit.,[60] 77 °C/10 mmHg); $[\alpha]_D$ −17.7 (*c* 0.8 in MeOH) (lit.,[60] −18.2); ν_{max}(thin film)/cm^{-1} 3110 (OH), 1735 (CO), 1457 (CH_3 def.), 1210 (C–O) and 668 (C–Cl); δ_H(300 MHz; C^2HCl_3) 1.73 (3 H, d, *J* 7.2, βCH_3), 4.45 (1 H, q, *J* 6.9, αCH) and 10.70 (1 H, br s, CO_2H); δ_C(74.76 MHz; C^2HCl_3) 21.71 (βCH_3), 52.63 (αCH) and 176.75 (acid CO); *m*/*z* (CI) 111 and 109 (3 and 11%, chlorine isotopes, $[M + H]^+$) and 57 (100, $CHCO_2^+$).

(2*S*)-*N*-[(2*S*)-2-Chloropropionyl)]proline 18

A stirred solution of (2*S*)-2-chloropropionic acid (6.0 g, 55.3 mmol) and *N*-methylmorpholine (6.2 cm^3, 55.3 mmol) in dry THF (100 cm^3) at −15 °C was treated dropwise with isobutyl chloroformate (7.40 cm^3, 54.5 mmol), followed after 2 min by a solution of (2*S*)-proline methyl ester hydrochloride (9.2 g, 55.5 mmol) and *N*-methylmorpholine (6.2 cm^3, 55.3 mmol) in DMF (30 cm^3). The mixture was allowed to reach room temperature and stirred for 18 h. The hydrochloride salts formed were filtered off and the solvents removed under reduced pressure to yield an orange oil, which was redissolved in dichloromethane (140 cm^3). The solution was washed with aqueous HCl (0.5 mol dm^{-3}, 2 × 100 cm^3), aqueous sodium hydrogen carbonate (5%, 2 × 100 cm^3) and brine (50 cm^3). The solvent was removed under reduced pressure to yield a pale yellow oil which was purified by gradient column chromatography using light petroleum–ethyl acetate as the eluent to give the methyl ester **17** as a pale yellow oil (4.9 g).

The above oil was dissolved in methanol (45 cm^3), and aqueous sodium hydroxide (1.0 mol dm^{-3}, 55 cm^3) was added. The mixture was stirred at room temperature for 2 h, after which time aqueous HCl (1.0 mol dm^{-3}, 55 cm^3) was added. Methanol was removed under reduced pressure, the solution further acidified to pH 1, and the solution extracted with ethyl acetate (3 × 80 cm^3). The combined organic extracts were dried ($MgSO_4$), and the solvent removed under reduced pressure to yield the acid **18** as a white solid (4.05 g, 36%). A small portion was recrystallised to give colourless crystals, mp 161–162 °C; $[\alpha]_D$ −59.2 (*c* 1.0 in MeOH); ν_{max}(Nujol)/cm^{-1} 3000 (OH), 1739 (acid CO), 1621 (amide CO) and 652 (C–Cl); δ_H(200 MHz; C^2HCl_3) 1.64 [3 H, d, *J* 6.6, βCH_3(Chp)], 1.97–2.29 [4 H, m, βCH_2(Pro) and γCH_2], 3.60–3.89 (2 H, δCH_2), 4.46–4.56 [2 H, m, αCH(Pro) and αCH(Chp)] and 10.05 (1 H, s, CO_2H); δ_C(50.31 MHz; C^2HCl_3) 20.89 [βCH_3(Chp)], 25.25 (γCH_2), 29.25 [βCH_2(Pro)], 47.69 (δCH_2), 51.53 [αCH(Chp)], 60.03 [αCH(Pro)], 169.13 (amide CO) and 175.78 (acid CO); *m*/*z* (EI) 208 and 206 (4 and 15%, chlorine isotopes, M^+), 170 (39, $[M - HCl]^+$), 162 and 160 (22 and 60, chlorine isotopes, $[M - CO_2H - H]^+$), 126 (49, $[M - CO_2H - Cl]^+$), 114 (51, $C_5H_8NO_2^+$), 70 (100, $C_4H_8N^+$) and 57 (29, $[CH_3CHCO + H]^+$).

Methyl (2*S*)-*N*-{(2*S*)-*N*-[(2*S*)-2-chloropropionyl]prolyl}prolinate 19

Thionyl chloride (0.40 cm^3, 5.5 mmol) was added dropwise to a stirred solution of (2*S*)-*N*-[(2*S*)-2-chloropropionyl]proline **18** (0.70 g, 3.4 mmol) and pyridine (0.53 cm^3, 6.6 mmol) in dichloromethane (20 cm^3). After 10 min, a solution of methyl (2*S*)-prolinate hydrochloride (0.51 g, 3.08 mmol) and DMAP (0.80 g, 6.6 mmol) in dichloromethane (20 cm^3) was added in one portion. The dark brown solution was stirred for 24 h, then washed with aqueous HCl (0.5 mol dm^{-3}, 2 × 35 cm^3), aqueous sodium hydrogen carbonate (5%, 2 × 35 cm^3) and brine (30 cm^3). The organic phase was dried ($MgSO_4$), and the solvent removed under reduced pressure to yield a brown oil which was purified by gradient column chromatography using light petroleum–ethyl acetate as the eluent to afford methyl ester **19** as colourless crystals (0.32 g, 33%), mp 126–128 °C; R_f 0.25 (ethyl acetate) (Found: C, 53.25; H, 6.8; N, 8.6. $C_{14}H_{21}ClN_2O_4$ requires C, 53.1; H, 6.7; N, 8.85%); $[\alpha]_D$ −131.0 (*c* 0.9 in MeOH); ν_{max}(Nujol)/cm^{-1} 1735 (ester CO) and 1652 (amide CO); δ_H(300 MHz; C^2HCl_3) 1.54 [3 H, d, *J* 6.6, βCH_3(Chp)], 1.90–1.98 [6 H, m, 2 × $\frac{1}{2}\beta CH_2$(Pro) and 2 × γCH_2(Pro)], 2.10–2.15 [2 H, m, 2 × $\frac{1}{2}\beta CH_2$(Pro)], 3.56–3.78 [4 H, m, 2 × δCH_2(Pro)], 3.62 (3 H, s, CO_2CH_3), 4.44–4.48 [2 H, m, αCH(Pro) and αCH(Chp)] and 4.53 [1 H, dd, J_1 7.8, J_2 3.9, αCH(Pro)]; δ_C(74.76 MHz; C^2HCl_3) 20.71 [βCH_3(Chp)], 24.93 and 25.13 [2 × γCH_2(Pro)], 28.49 and 28.99 [2 × βCH_2(Pro)], 46.89 and 47.44 [2 × δCH_2(Pro)], 51.49 [αCH(Chp)], 52.31 (CO_2CH_3), 58.52 and 58.91 [2 × αCH(Pro)], 167.90 [CO(Chp)], 170.70 (amide CO) and 173.06 (ester CO); *m*/*z* (EI) 318 and 316 (14 and 36%, chlorine isotopes, M^+), 280 (14, $[M - HCl]^+$), 225 {11, $[M - ClCH(CH_3)CO]^+$}, 190 and 188 [7 and 21, chlorine isotopes, $ClCH(CH_3)CONC_4H_7CO^+$], 162 and 160 [52 and 96, chlorine isotopes, $ClCH(CH_3)CONC_4H_7^+$], 128 (29, $NC_4H_7CO_2CH_3^+$) and 70 (100, $C_4H_8N^+$).

(2*S*)-*N*-{(2*S*)-*N*-[(2*S*)-2-Chloropropionyl]prolyl}proline 20

A solution of methyl ester **19** (0.30 g, 3.0 mmol) in methanol (2 cm^3) was treated with aqueous sodium hydroxide (1.0 mol dm^{-3}, 2 cm^3) and the mixture stirred for 90 min. The solution was treated with aqueous HCl (0.5 mol dm^{-3}, 2 cm^3) and the methanol was removed under reduced pressure. The solution was further acidified to pH 1–2 and extracted with ethyl acetate (5 × 20 cm^3). The combined organic extracts were dried ($MgSO_4$), and the solvent removed under reduced pressure to yield acid **20** as a white solid (0.28 g, 98%), mp 84–86 °C (Found: C, 50.9; H, 6.55; N, 8.65. $C_{13}H_{19}ClN_2O_4{\cdot}\frac{1}{4}H_2O$ requires C, 50.8; H, 6.4; N, 9.1%); $[\alpha]_D$ −104.5 (*c* 0.9 in MeOH); ν_{max}(Nujol)/cm^{-1} 2922 (CH), 1702 (acid CO) and 1642 (amide CO); δ_H(300 MHz; C^2HCl_3) 1.57 [3 H, d, *J* 6.6, βCH_3(Chp)], 1.97–2.08 [6 H, m, 2 × $\frac{1}{2}\beta CH_2$(Pro), 2 × γCH_2(Pro)], 2.11–2.22 [2 H, m, 2 × $\frac{1}{2}\beta CH_2$(Pro)], 3.52–3.59 and 3.63–3.84 [4 H, m, 2 × δCH_2(Pro)], 4.46–4.53 [2 H, m, αCH(Pro) and αCH(Chp)], 4.59–4.63 [1 H, m, αCH(Pro)] and 9.34 (1 H, br s, CO_2H); δ_C(74.76 MHz; C^2HCl_3) 2.60 [βCH_3(Chp)], 25.07 [2 × γCH_2(Pro)], 28.57 and 28.76 [2 × βCH_2(Pro)], 47.31 and 47.58 [2 × δCH_2(Pro)], 51.52 [αCH(Chp)], 58.68 and 59.31 [2 × αCH(Pro)], 168.47 [CO(Chp)], 171.77 (amide CO) and 174.23 (acid CO); *m*/*z* (EI) 304 and 302 (4 and 14%, chlorine isotopes, M^+), 266 (7, $[M - HCl]^+$), 260 and 258 (4 and 11, chlorine isotopes, $[M - CO_2]^+$), 162 and 160 [46 and 86, chlorine isotopes, $ClCH(CH_3)CONC_4H_7CO^+$], 128 (14, $NC_4H_7CO_2CH_3^+$) and 70 (100, $C_4H_8H^+$).

Methyl (2*S*,4*S*)-*N*-((2*S*)-*N*-{(2*S*)-*N*-[(2*S*)-2-chloropropionyl]-prolyl}prolyl)-4-hydroxyprolinate 12

This compound was prepared in a manner identical with that described for the peptide **19**, using acid **20** (0.23 g, 0.76 mmol) and methyl (2*S*,4*S*)-4-hydroxyprolinate hydrochloride (0.12 g, 0.66 mmol) to yield the peptide **12** as a brown oil (0.12 g, 42%) which defied further purification; (HRMS: Found M^+, 429.1662. $C_{19}H_{28}{}^{35}ClN_3O_6$ requires 429.1667); ν_{max}(thin film)/cm^{-1} 3441 (OH), 2981 (CH), 1746 (ester CO), 1649 (amide CO), 1435 (CH def.) and 1066 (C–O); δ_H(300 MHz; C^2HCl_3) 1.61 [3 H, d, *J* 6.3, βCH$_3$(Chp)], 1.91–2.38 [10 H, m, 2 × βCH$_2$(Pro), βCH$_2$(Hyp) and 2 × γCH$_2$(Pro)], 3.56–3.86 [6 H, m, 2 × δCH$_2$(Pro) and δCH$_2$(Hyp)], 3.68 (3 H, s, CO_2CH_3), 4.47–4.69 [4 H, m, αCH(Chp), 2 × αCH(Pro) and αCH(Hyp)] and 5.10–-5.21 [1 H, m, γCH(Hyp)]; δ_C(50.31 MHz; C^2HCl_3) 20.78 [βCH$_3$(Chp)], 25.08 and 25.17 [2 × γCH$_2$(Pro)], 28.36 and 28.52 [2 × βCH$_2$(Pro)], 37.04 [βCH$_2$(Hyp)], 47.37 and 47.61 [2 × δCH$_2$-(Pro)], 51.56 [αCH(Chp)], 52.94 (CO_2CH_3), 55.85 [δCH$_2$(Hyp)], 57.80, 58.21 and 58.77 [2 × αCH(Pro) and αCH(Hyp)], 71.41 [γCH(Hyp)], 168.08 [CO (Chp)], 170.56 and 171.56 (2 × amide CO) and 174.75 (ester CO); *m*/*z* (EI) 431 and 429 (1 and 3%, chlorine isotopes, M^+), 393 (6, $[M - HCl]^+$), 287 and 285 {4 and 7, chlorine isotopes, $[M - NCH_2CH(OH)CH_2CHCO_2CH_3 + H]^+$}, 259 and 257 {4 and 14, chlorine isotopes, $[M - CONCH_2CH(OH)CH_2CHCO_2CH_3 + H]^+$}, 190 and 188 (17 and 53, chlorine isotopes, $C_8H_{11}NO_2Cl^+$), 162 and 160 (34 and 87, chlorine isotopes, $C_7H_{11}NOCl^+$), 130 {69, $[NCH_2CH(OH)CH_2CHCO_2H + H]^+$}, 86 (84, $C_4H_8NO^+$) and 70 (100, $C_4H_8N^+$).

Acknowledgements

We thank Dr J. Wilkie, Dr T. Rutherford and Dr M. Akhtar for computational, NMR and scientific support respectively and the Wellcome Trust for a prize studentship to A. L.

References

1 J. S. Richardson, *Adv. Protein Chem.*, 1981, **34**, 167.
2 F. B. Sheinerman and C. L. Brooks III, *J. Am. Chem. Soc.*, 1995, **117**, 10098.
3 L. Pauling, R. B. Corey and H. R. Branson, *Proc. Natl. Acad. Sci. USA*, 1951, **37**, 205.
4 B. H. Zimm and J. K. Bragg, *J. Chem. Phys.*, 1959, **31**, 526.
5 S. Lifson and A. Roig, *J. Chem. Phys.*, 1961, **34**, 1963.
6 A. Yang and B. Honig, *J. Mol. Biol.*, 1995, **252**, 351.
7 N. E. Zhou, C. M. Kay, B. D. Sykes and R. S. Hodges, *Biochemistry*, 1993, **32**, 6190.
8 H. A. Scheraga, *Pure Appl. Chem.*, 1978, **50**, 315.
9 A. Chakrabartty, T. Kortemme and R. L. Baldwin, *Protein Sci.*, 1994, **3**, 843.
10 A. J. Doig and R. L. Baldwin, *Protein Sci.*, 1995, **4**, 1325.
11 J. J. Osterhout, R. L. Baldwin, E. J. York, J. M. Stewart, H. J. Dyson and P. E. Wright, *Biochemistry*, 1989, **28**, 7059.
12 P. S. Kim and R. L. Baldwin, *Nature*, 1984, **307**, 329.
13 J. W. Nelson and N. R. Kallenbach, *Biochemistry*, 1989, **28**, 5256.
14 K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart and R. L. Baldwin, *Nature*, 1987, **326**, 563.
15 A. Bierzynski, P. S. Kim and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 2470.
16 E. Bairaktari, D. F. Mierki, S. Mammi and E. Peggion, *J. Am. Chem. Soc.*, 1990, **112**, 5383.
17 N, Khan, A. Graslund, S. Ehrenberg and J. Schriver, *Biochemistry*, 1990, **29**, 5743.
18 E. T. Kaiser and F. J. Kézdy, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 1137.
19 E. T. Kaiser and F. J. Kézdy, *Science*, 1984, **223**, 249.
20 I. L. Karle, *Biopolymers*, 1996, **40**, 157.
21 S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 8898.
22 M. Reza Ghadiri and C. Choi, *J. Am. Chem. Soc.*, 1990, **112**, 1630.
23 F. Ruan, Y. Chen and P. B. Hopkins, *J. Am. Chem. Soc.*, 1990, **112**, 9403.
24 J. S. Albert, M. W. Peczuh and A. D. Hamilton, *Bioorg. Med. Chem.*, 1997, **5**, 1455.
25 D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391.
26 G. Osapay and J. W. Taylor, *J. Am. Chem. Soc.*, 1992, **114**, 6966.
27 C. Yu and J. W. Taylor, *Tetrahedron Lett.*, 1996, **37**, 1731.
28 J. C. Phelan, N. J. Skelton, A. C. Braisted and R. S. McDowell, *J. Am. Chem. Soc.*, 1997, **119**, 455.
29 G. R. Geier and T. Sasaki, *Tetrahedron Lett.*, 1997, **38**, 3821.
30 L. G. Presta and G. D. Rose, *Science*, 1988, **240**, 1632.
31 J. P. Schneider and J. W. Kelly, *Chem. Rev.*, 1995, **95**, 2169.
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.
33 D. S. Kemp and T. P. Curran, *Tetrahedron Lett.*, 1988, **29**, 4931.
34 D. S. Kemp and T. P. Curran, *Tetrahedron Lett.*, 1988, **29**, 4935.
35 D. S. Kemp, T. P. Curran, W. M. Davis, J. G. Boyd and C. Muendel, *J. Org. Chem.*, 1991, **56**, 6672.
36 D. S. Kemp, T. P. Curran, J. G. Boyd and T. J. Allen, *J. Org. Chem.*, 1991, **56**, 6683.
37 D. S. Kemp, J. H. Rothman, T. P. Curran and D. E. Blanchard, *Tetrahedron Lett.*, 1995, **36**, 3809.
38 D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 3813.
39 D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 4019.
40 D. S. Kemp and J. H. Rothman, *Tetrahedron Lett.*, 1995, **36**, 4023.
41 K. F. McClure, P. Renold and D. S. Kemp, *J. Org. Chem.*, 1995, **60**, 454.
42 D. S. Kemp, T. J. Allen and S. L. Oslick, *J. Am. Chem. Soc.*, 1995, **117**, 6641.
43 D. S. Kemp, T. J. Allen, S. L. Oslick and J. G. Boyd, *J. Am. Chem. Soc.*, 1996, **118**, 4240.
44 D. S. Kemp, S. L. Oslick and J. G. Boyd, *J. Am. Chem. Soc.*, 1996, **118**, 4249.
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.
46 T, Arrhenius and A. C. Satterthwait, *Peptides: Chemistry, Structure and Biology: Proceedings of the 11th American Peptide Symposium, 1989*, 1990, pp. 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, *Perspect. Med. Chem.*, 1993, pp. 513–531.
48 M. L. L. Donnelly, D. Gani, M. Flint, S. Monaghan and M. D. Ryan, *J. Gen. Virol.*, 1997, **78**, 13.
49 D. S. Kemp, J. G. Boyd and C. C. Muendel, *Nature*, 1991, **352**, 451.
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.
51 M. Narita, S. Isokawa, M. Doi and R. Wakita, *Bull. Chem. Soc. Jpn.*, 1986, **59**, 3547.
52 V. V. Ryakhovskii, S. V. Agafonov and Y. M. Kosyrev, *Russ. Chem. Rev.* (*Engl. Transl.*), 1991, **60**, 924.
53 M. M. Lenman, A. Lewis and D. Gani, *J. Chem. Soc.*, *Perkin Trans. 1*, 1997, 2297.
54 Y. V. Venkatachalapathi and P. Balaram, *Nature*, 1979, **281**, 83.
55 K. Sukata, *Bull. Chem. Soc. Jpn.*, 1985, **58**, 838.
56 J. D. Park, R. D. Englert and J. S. Meek, *J. Am. Chem. Soc.*, 1952, **74**, 1010.
57 A. Koziara, S. Zawadzki and A. Zwierzak, *Synth. Commun.*, 1979, 527.
58 N. Ono, T. Yoshimura, T. Saito, R. Tamura, R. Tanikaga and A. Kaji, *Bull. Chem. Soc. Jpn.*, 1979, **52**, 1716.
59 D. E. Dorman and F. A. Bovey, *J. Org. Chem.*, 1973, **38**, 2379.
60 S. J. Fu, S. M. Birnbaum and J. P. Greenstein, *J. Am. Chem. Soc.*, 1954, **76**, 6054.
61 M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.*, 1991, **218**, 397.
62 W. C. Still, M. Kahn and A. Mitra, *J. Org. Chem.*, 1978, **43**, 2923.
63 J. Halstrom, K. Kovács and K. Brunfeldt, *Acta Chem. Scand.*, 1973, **27**, 3085.

Paper 8/06028B