

Synthesis and Characterization of Template-Assembled Three- and Four-Helix Bundle Proteins

by

Ashley Scott Causton

B.Sc. (Honours), University of Greenwich, 1992

M.Sc., University of British Columbia, 1994

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Chemistry)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April, 2001

© Ashley Scott Causton, 2001

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

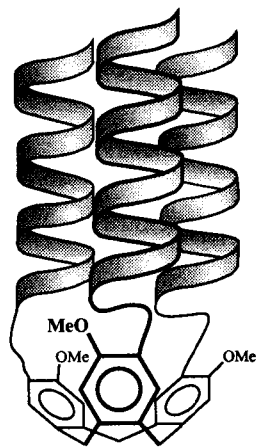
Department of CHEMISTRY

The University of British Columbia
Vancouver, Canada

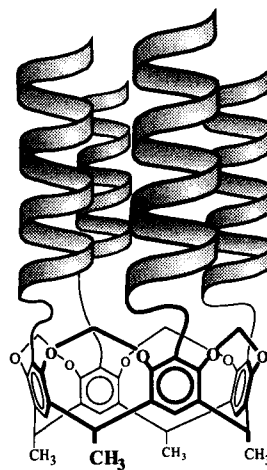
Date 25th JUNE 2001

Abstract

The unique three-dimensional structure of a protein is the result of a multitude of non-covalent interactions within the polypeptide backbone, and between the side chains of its constituent amino acids. One method that can be used to simplify and study the non-covalent forces present in protein structure is to use a template assembled synthetic protein (TASP): This method employs peptide strands that are attached to a template which directs them to fold into a pre-determined folding pattern.



Three-Helix Bundle TASP
(Cyclotribenzylene (CTB) Template)



Four-Helix Bundle TASP
(Cavitand Bowl Template)

This thesis studies three- and four-helix bundles using the TASP concept. The templates are cyclotribenzylene ("CTB") and cavitand bowl ("bowl") macrocycles, which have three- and four-fold symmetry, respectively. Amphiphilic peptides (for example, CEELLKKLEELLKKG) which were designed to form helical bundles, were attached to the

templates via disulfide bonds. The CTB and bowl TASP were synthesized by first activating the side chain of a cysteine residue in the peptide strand, and then reacting it with a benzyl thiol moiety on the template. TASP containing various peptide sequences were investigated in terms of their structure and stability.

The CTB and bowl TASP were found to have helical structure, and enhanced stability to chemical denaturation, exhibiting the co-operative unfolding that is a characteristic of natural proteins. The initial CTB and bowl TASP contained peptides with the template-bound cysteine residue adjacent to the “helix” sequence, and were found to undergo self-association. This was attributed to incorrect bundling of the constituent helices, resulting from a template-to-helix linker that was too short. Therefore, flexible glycine residues were introduced between the template-bound cysteine residue and the helix in the attached polypeptides. A three-glycine “spacer” (*i.e.* CGGGEELLKKLEELLKKG) was found to be optimal for reducing TASP self-association.

The ultimate goal of this research is to produce TASP with “native-like” conformational specificity. This first generation of CTB and bowl TASP appeared to have molten globule structure, which is consistent with other *de novo* designed proteins that contained a degenerate series of amino acids in their core; this is a good starting point for future designs. By varying the size and shape of the central residue in the TASP’s helices, some difference in their stability was observed. These studies demonstrated the validity of these TASP models for investigating protein structure.

Contents

Abstract.....	ii
Contents.....	iv
List of Figures.....	viii
List of Tables.....	xii
List of Abbreviations.....	xiii
Acknowledgements.....	xv

Chapter One: Introduction

1.0 Rationale Behind Thesis.....	1
1.1 Proteins and the Protein Folding Problem.....	3
1.2 Protein Structure.....	5
1.2.1 Primary Structure.....	5
1.2.2 Secondary Structure: The α -Helix.....	6
1.2.3 Supersecondary structure and Protein domains: Helix Bundles in Proteins...8	
1.2.4 Tertiary Structure: Helix Bundles as Proteins.....	9
1.2.5 Quaternary structure: Coiled-Coils.....	10
1.3 Non-covalent Forces Involved in Protein Structure and Protein Folding.....	11
1.3.1 Significant Forces in Protein Structure.....	11
1.3.2 Mechanism of Protein Folding.....	13
1.3.3 The Molten Globule State.....	13
1.3.4 Protein Folding with Chaperones and Co-Factors.....	14
1.4 Decoding the Protein Folding Problem.....	15
1.4.1 Studies of Secondary Structure: Important Forces in α -Helices.....	16
1.4.1.1 Helix Propensities.....	17
1.4.1.2 Helix Chain Length.....	18
1.4.1.3 Helix Macrodipole.....	18
1.4.1.4 Side Chain Interactions Within The Helix.....	19
1.4.1.5 Capping of the Helix.....	19
1.4.2 Investigation of Tertiary Structure: Helix Bundling in Proteins.....	20
1.4.2.1 Side Chain Packing: Inter-Helical Packing.....	20
1.4.2.2 Alteration of Natural Protein Sequences.....	22
1.4.3 <i>De Novo</i> Protein Design.....	25
1.4.4 Template Assembled Synthetic Proteins.....	30
1.4.4.1 Peptides as Templates.....	32
1.4.4.2 Porphyrin Templates.....	32

1.4.4.3 Other Divergent TASP.s.....	34
1.4.4.4 Templat.ion by Metal Ions.....	36
1.4.4.5 Other Convergent Templates.....	38
1.5 Beyond the Protein Folding Problem: Designer Proteins.....	39
1.6 Thesis Goals and Summary of Thesis.....	40

Chapter Two: Synthesis of Templates

2.0 Choice of Template.....	45
2.1 Cyclotribenzylene Macrocycle.....	46
2.1.1 Synthesis of CTB Tris-Benzylthiol 1.....	46
2.2 Cavitand Bowl Macrocycle.....	49
2.2.1 Synthesis of Cavitand Bowls 2a and 2b.....	50
2.3 Experimental.....	52

Chapter Three: Design and Synthesis of TASP.s

3.0 Introduction.....	61
3.1 Design of CTB and Cavitand Based TASP.s.....	62
3.1.1 Design and Nomenclature of Peptides and TASP.s.....	62
3.2 Methods for the Synthesis of TASP.s.....	69
3.2.1 TASP.s Constructed using Protected Peptides and Amino Acid Side Chains..	69
3.2.2 Synthesis of TASP.s using Unprotected Peptides.....	72
3.2.2.1 The Thioether Bond.....	72
3.2.2.2 Thioesters, Oximes, Amides and Amines.....	75
3.3 Synthesis of CTB and Cavitand Bowl TASP.s.....	77
3.3.1 Synthesis of Peptides.....	77
3.3.2 Cavitand Bowl TASP.s Synthesized via a Thioether Bond.....	77
3.3.3 TASP.s Synthesized via a Disulfide Bond.....	78
3.3.3.1 Activated Template Thiols plus a Cysteine Containing Peptide.....	80
3.3.3.2 Activated Cysteine Containing Peptide + Benzyl Thiol Template...	83
3.4 Conclusions for Chapter 3.....	85
3.5 Experimental.....	86

Chapter Four: Initial Three- and Four-Helix Bundle TASP.s

4.0 Introduction.....	97
4.1 Methods for Characterizing the Global Structure/Stability of Proteins.....	98
4.1.1 Circular Dichroism Spectroscopy.....	98
4.1.2 Effect of Varying pH. Addition of Salt, or Trifluoroethanol.....	99
4.1.3 Stability to Chemical Denaturants and Temperature.....	99
4.2 Methods for Determining Association State of Proteins.....	100

4.2.1	Concentration Dependent Circular Dichroism.....	101
4.2.2	Concentration Dependent GnHCl-Induced Unfolding Curves.....	101
4.2.3	Size Exclusion Chromatography.....	102
4.2.4	Electrospray Mass Spectrometry.....	103
4.2.5	Analytical Ultracentrifugation.....	103
4.3	Methods for Determining the Conformational Specificity of Proteins.....	104
4.3.1	ANS Binding.....	104
4.3.2	One Dimensional ¹ H NMR Spectra.....	105
4.4	Introduction to Initial CTB and Bowl TASP.....	106
4.5	CTB/N1/L, CTB/EGG/N1/L, and CTB/H-EGG/N1/L.....	108
4.6	CTB/N2/L.....	117
4.7	Bowl/N1/L, Bowl/EGG/N1/L, and Bowl/H-EGG/N1/L.....	125
4.8	Bowl/N2/L.....	129
4.9	Bowl/N2/AAAAA.....	135
4.10	Bowl/e/N2/L.....	140
4.11	Conclusions for Chapter 4.....	144
4.12	Experimental.....	150
4.12.1	CD studies.....	150
4.12.2	Global Conformational Stability of TASP: Calculated by Analysis of GnHCl-Induced Denaturation.....	152
4.12.3	Sedimentation Equilibrium Ultracentrifugation.....	153
4.12.4	NMR Spectroscopy.....	164

Chapter Five: Effect of Linker Length on CTB and Bowl TASP

5.0	Introduction.....	165
5.1	CTB/N2-GG/L and CTB/N2-GGG/L.....	168
5.2	CTB/C2-GGG/L.....	174
5.3	Bowl/N2-GG/L and Bowl/N2-GGG/L.....	178
5.4	Bowl/C2-GGG/L.....	186
5.5	Conclusions for Chapter 5.....	190
5.6	Experimental.....	196

Chapter Six: Effect of Changing the Central Leucine Amino Acid in the helices of CTB and Bowl TASP

6.0	Introduction.....	206
6.1	CTB/N2-GGG/I, CTB/N2-GGG/V, and CTB/N2-GGG/Nle.....	208
6.2	Bowl/N2-GGG/I, Bowl/N2-GGG/V, and Bowl/N2-GGG/Nle.....	214
6.3	Further Investigation of TASP Structure.....	219
6.4	Conclusions for Chapter 6.....	224
6.5	Experimental.....	229

Chapter Seven: Summary, Conclusions, and Future Work

8.0 Introduction.....	230
8.1 Synthesis of TASP.....	230
8.2 Oligomeric State of TASP.....	232
8.3 Structure and Stability of TASP.....	236
8.4 Conformational Specificity of TASP.....	238

References.....	240
------------------------	------------

Appendix A: The one- and three-letter codes for the amino acids mentioned in this thesis, and their side chains.....	262
---	------------

List of Figures

Figure 1.1:	(a) A polypeptide chain folds to form a protein. Lines represent unstructured polypeptide chains, ribbons represent α -helices. (b) α -Helical peptide strands on a template gives a template assembled synthetic protein that mimics protein structure.....	2
Figure 1.2:	The formation of a polypeptide via condensation of the constituent amino acids. The dotted square represents the plane formed by the peptide amide bond.....	6
Figure 1.3:	The right handed α -helix. The first and last four amino acids in the helix backbone have "unsatisfied" hydrogen bonding moieties.....	7
Figure 1.4:	The helices in square bundles are splayed out from their closest point of contact.....	9
Figure 1.5:	A two stranded coiled-coil.....	10
Figure 1.6:	The order of preference of the twenty natural amino acids for being in an α -helix.....	17
Figure 1.7:	Interhelical side chain packing. (a) Knobs into holes. (b) Ridges into grooves.....	21
Figure 1.8:	Helical wheel showing the heptad repeat, and the interactions seen in a two stranded coiled-coil.....	24
Figure 1.9:	Helical Wheel diagram showing the hydrophobic and hydrophilic faces of the peptide Ac-GELEELLKKLKELLKG-NH ₂ when it is folded into an α -helix.....	27
Figure 1.10:	Iterative design of a four-helix bundle.....	28
Figure 1.11:	Designed "clockwise" folding of a three helix bundle.....	29
Figure 1.12:	Schematic showing how a template assembled protein (TASP) mimics protein structure....	30
Figure 1.13:	One example of Mutter's peptide template.....	32
Figure 1.14:	Porphyrin Templates: (a) Sasaki/Kaiser. (b) DeGrado, Nishino.....	33
Figure 1.15:	Fairlie's organic templates.....	34
Figure 1.16:	Other divergent organic templates: (a) Sasaki. (b) Nishino. (c) Inoue. (d) Sherman....	35
Figure 1.17:	(a) Ghadiri's four-helix bundle induced via the addition of Ru(II). (b) Ghadiri's and Sasaki's three-helix bundles induced via the addition of iron (II), nickel (II), copper (II) or ruthenium (II). (c) Schepartz's two helix bundles induces via the addition of iron (II). (d) Minoura's two helix bundle induced via the addition of potassium (I).....	37
Figure 1.18:	Morii's "template" is formed by the adamantane-cyclodextrin inclusion complex.....	38
Figure 1.19:	Schematic representations of four-helix bundle TASPs: (a) The hydrophobic template may lead to TASP self-association in water. (b) Charged groups on the template may prevent self-association. (c) Charged groups on the peptide sequence may prevent self-association.....	42
Figure 1.20:	Optimal length of template-to-helix linker is important to the bundling of the constituent helices. (a) Ideal linker. (b) Too short a linker. (c) Too long a linker.....	43
Figure 2.1:	The organic macrocycles that we used as templates for three- and four-helix bundles.....	45
Figure 2.2:	The CTB skeleton has been made using various "rim" (X, Y) and "linking" (Z) groups....	46
Figure 2.3:	Synthesis of 4-bromo-3-methoxybenzyl alcohol 7, the monomer used for the cyclotrimerization reaction.....	47
Figure 2.4:	Synthesis of CTB tris-benzylthiol 1 from benzylalcohol monomer 7.....	48
Figure 2.5:	The resorcin[4]arene skeleton, which can have various "rim" (Y) and "feet" (X).....	49
Figure 2.6:	Cram's cavitand bowl, for which the "rim" (Y) and "feet" (X) can be varied.....	50
Figure 2.7:	Synthesis of Cavitand bowls 2a and 2b.....	51
Figure 3.1:	Helix web diagrams of (a) three- and (b) four-helix bundles, for the sequence XEKLKELKELLEKG, where X is the "linker" to the template.....	65
Figure 3.2:	Codes and identities of peptides used in this thesis.....	66
Figure 3.3:	Schematic preparation of TASPs using protected amino acids and peptides. (a) Mutter's sequential addition of amino acids onto a resin-bound peptide template using standard solid state peptide synthesis methods. (b) Sasaki and Kaiser's addition of peptide strands with protected side chains and a free amino terminus, to a porphyrin template with	

Figure 3.4:	TASPs constructed using thioether bonds. (a) DeGrado's porphyrin TASP. (b) Fairlie's various aromatic TASPs. (c) Sherman's cavitand bowl TASP, Futakai's and Haehnel's peptide strand TASPs. (d) Tuchscherer's peptide strand TASP.....	74
Figure 3.5:	TASPs constructed using unprotected peptide strands with linkers other than thioethers. (a) Kent's thioester linked peptide strand template. (b) Rose's and Tuchscherer's oxime linked peptide strand templates. (c) Futakai's amide linked peptide strand template. (d) Sasaki's amine linked macrocyclic template.....	76
Figure 3.6:	The reaction of benzyl bromide bowl 2a and a cysteine containing peptide.....	78
Figure 3.7:	(a) A mixture of cysteine-containing peptides under reversible redox conditions will result in a mixture of disulfide linked products (assuming they have equal stability). (b) Reacting the thiol template with a cysteine-containing peptide resulted solely in the homo-dimer peptide as product.....	79
Figure 3.8:	Schematic synthesis of hetero-dimeric disulfide bonds by first activating one thiol of the pair.....	80
Figure 3.9:	Synthesis of TASPs via activated benzyl thiol groups on a template, and subsequent reaction with a cysteine containing peptide.....	82
Figure 3.10:	Synthesis of TASPs by first activating the peptide with an S-pyridyl group, and subsequent addition of this activated peptide to the benzyl thiol containing CTB 1 or cavitand bowl 2b template.....	84
Figure 4.1:	(a) Potential inter-macrocycle self-association in aqueous solution. (b) TASP containing charged, pre-cysteine residues to prevent inter-macrocycle self-association by charge-charge repulsion.....	107
Figure 4.2:	(a) Full (190-300 nm) and (b) expanded (240-300 nm) CD spectra of TASPs CTB/N1/L and CTB/EGG/N1/L. (c) Full (190-300 nm) and (d) expanded (240-300 nm) CD spectra of TASPs CTB-f/H-EGG/N1/L, CTB-s/H-EGG/N1/L and peptide H-EGG/N1/L	110
Figure 4.3:	Effect of GnHCl on the helicity of (a) CTB/N1/L, (b) CTB/EGG/N1/L, (c) CTB-f/H-EGG/N1/L, and (d) CTB-s/H-EGG/N1/L.....	113
Figure 4.4:	GnHCl-induced unfolding curves of (a) CTB/N1/L versus peptide N1(Spy)/L, and (b) CTB/EGG/N1/L versus CTB-f/H-EGG/N1/L, and CTB-s/H-EGG/N1/L.....	114
Figure 4.5:	CD spectrum of CTB/N2/L, compared to CTB/N1/L.....	117
Figure 4.6:	GnHCl-induced denaturation curves of (a) CTB/N2/L, and (b) CTB/N2/L versus CTB/N1/L.....	118
Figure 4.7:	The peptide sequences N1/L and N2/L. Examples of the favourable (<i>i, i+3</i> and <i>i, i+4</i>) and the unfavourable (<i>i, i+1</i> and <i>i, i+2</i>) interactions between oppositely charged residues in an α -helix are highlighted.....	121
Figure 4.8:	NMR spectra of CTB/N2/L compared to peptide Ac-EELLKKLEELLKKG-NH ₂	124
Figure 4.9:	(a) Full and (b) Near-UV, CD spectra of Bowl/N1/L, Bowl/EGG/N1/L, and Bowl/H-EGG/N1/L.....	126
Figure 4.10:	GnHCl-induced unfolding curves for; (a) Bowl/N1/L. (b) Bowl/EGG/N1/L. (c) Bowl/H-EGG/N1/L. (d) Bowl/N1/L, Bowl/EGG/N1/L, Bowl/H-EGG/N1/L, compared to peptide N1(Spy)/L.....	127
Figure 4.11:	CD spectrum of Bowl/N2/L, compared to Bowl/N1/L.....	129
Figure 4.12:	GnHCl-induced denaturation curves of (a) Bowl/N2/L, and (b) Bowl/N2/L.....	130
Figure 4.13:	Structures of ArBowl/N2/L (peptide linked via N-terminus of polypeptide backbone), and Bowl/N2/L (peptide linked via N-terminal cysteine side chain).....	132
Figure 4.14:	NMR spectra of Bowl/N2/L compared to peptide Ac-EELLKKLEELLKKG-NH ₂	134
Figure 4.15:	CD spectrum of Bowl/N2/AAAAA, compared to Bowl/N2/L.....	136
Figure 4.16:	GnHCl-induced denaturation curves of Bowl/N2/AAAAA.....	137
Figure 4.17:	NMR spectra of Bowl/N2/AAAAA compared to Bowl/N2/L.....	139
Figure 4.18:	CD spectrum of Bowl/e/N2/L, compared to Bowl/N2/L.....	141
Figure 4.19:	GnHCl-induced denaturation curves of (a) Bowl/e/N2/L, and (b) Bowl/e/N2/L versus Bowl/N2/L.....	142

Figure 4.20:	(a) Schematic diagram of a cavitand TASP with too short a linker that results in steric crowding of the helices around the template. (b) The dipolar interactions between the template and linker. (c) The preferred disulfide dihedral angle of $\pm 90^\circ$. (d) Self-association of two TASPs by complementary pairing of oppositely charged side chains....	147
Figure 4.21:	Sedimentation equilibrium analysis of CTB/N2/L	160
Figure 4.22:	Sedimentation equilibrium analysis of Bowl/N2/L	161
Figure 4.23:	Sedimentation equilibrium analysis of Bowl/N2/AAAAA	162
Figure 5.1:	Experimental "average" molecular weights (according to sedimentation equilibrium studies) of bowl TASP that differ by their template-to-helix linker.....	166
Figure 5.2:	(a) Full CD, and (b) near-UV CD spectra of CTB/N2-GG/L , CTB/N2-GGG/L , and CTB/N2/L	168
Figure 5.3:	GnHCl-induced denaturation curves for (a) CTB/N2-GG/L , (b) CTB/N2-GGG/L , (c) CTB/N2-GG/L , CTB/N2-GG/L , compared to CTB/N2/L	170
Figure 5.4:	NMR spectra of CTB/N2-GG/L and CTB/N2-GGG/L , compared to CTB/N2/L	173
Figure 5.5:	CD spectra of CTB/C2-GGG/L , compared to CTB/N2-GGG/L	175
Figure 5.6:	GnHCl-induced denaturation curves for (a) CTB/C2-GGG/L , (b) CTB/C2-GGG/L , compared to CTB/N2-GGG/L	176
Figure 5.7:	NMR spectra of CTB/C2-GGG/L , compared to CTB/N2-GGG/L	177
Figure 5.8:	(a) Full, and (b) Near-UV CD spectra of Bowl/N2-GG/L , and Bowl/N2-GGG/L , compared to Bowl/N2/L	179
Figure 5.9:	GnHCl-induced denaturation curves of (a) Bowl/N2-GG/L , (b) Bowl/N2-GGG/L (c) Comparison between Bowl/N2-GG/L , Bowl/N2-GGG/L , and Bowl/N2/L	180
Figure 5.10:	Structures of ArBowl/GGG-N2/L (peptide linked via the peptide backbone), and Bowl/N2-GGG/L (peptide linked via cysteine side chain).....	182
Figure 5.11:	NMR spectra of Bowl/N2-GG/L and Bowl/N2-GGG/L , compared to Bowl/N2/L , and ArBowl/N2-GGG/L	185
Figure 5.12:	CD spectrum of Bowl/C2-GGG/L compared to Bowl/N2-GGG/L	186
Figure 5.13:	GnHCl-induced denaturation curves of (a) Bowl/C2-GGG/L , and (b) Comparison between Bowl/C2-GGG/L and Bowl/N2-GGG/L	188
Figure 5.14:	NMR spectra of Bowl/C2-GGG/L , compared to Bowl/N2-GGG/L	189
Figure 5.15:	Diagram showing the positions of amino acid residues on a helix on (a) Bowl/N2-GG/L ; and (b) Bowl/N2-GGG/L	191
Figure 5.16:	Proposed helix nucleating hydrogen bond between an "unsatisfied" helix backbone N-H, and lone pair of electrons of an acetal oxygen on bowl template.....	195
Figure 5.17:	Sedimentation equilibrium analysis of CTB/N2-GG/L	200
Figure 5.18:	Sedimentation equilibrium analysis of CTB/N2-GGG/L	201
Figure 5.19:	Sedimentation equilibrium analysis of CTB/N2-GGG/L	202
Figure 5.20:	Sedimentation equilibrium analysis of Bowl/N2-GG/L	203
Figure 5.21:	Sedimentation equilibrium analysis of Bowl/N2-GGG/L	204
Figure 5.22:	Sedimentation equilibrium analysis of Bowl/N2-GGG/L	205
Figure 6.1:	CD spectra of CTB/N2-GGG/I , CTB/N2-GGG/V , and CTB/N2-GGG/Nle , compared to CTB/N2-GGG/L	209
Figure 6.2:	GnHCl-induced denaturation curves for (a) CTB/N2-GGG/I , (b) CTB/N2-GGG/V (c) CTB/N2-GG/Nle , (d) CTB/N2-GGG/I , CTB/N2-GGG/V , CTB/N2-GGG/Nle , compared to CTB/N2-GGG/L	211
Figure 6.3:	NMR spectra of CTB/N2-GGG/L , CTB/N2-GGG/V , and CTB/N2-GGG/Nle , compared to CTB/N2-GGG/L	213
Figure 6.4:	CD spectra of Bowl/N2-GGG/I , Bowl/N2-GGG/V , and Bowl/N2-GGG/Nle , compared to Bowl/N2-GGG/L	215
Figure 6.5:	GnHCl-induced denaturation curves for (a) Bowl/N2-GGG/I , (b) Bowl/N2-GGG/V , (c) Bowl/N2-GGG/Nle , (d) Bowl/N2-GGG/I , Bowl/N2-GGG/V , Bowl/N2-GGG/Nle compared to Bowl/N2-GGG/L	216
Figure 6.6:	NMR spectra of Bowl/N2-GGG/L , Bowl/N2-GGG/V , and Bowl/N2-GGG/Nle , compared to Bowl/N2-GGG/L	218
Figure 6.7:	NMR spectra of Bowl/N2-GGG/I at 277, 298 and 358 K.....	222
Figure 6.8:	NMR spectra of Bowl/N2-GGG/I in pH 7.0 phosphatebuffer and in 1:9 TFE:buffer.....	223

Figure 6.9:	Simplified diagram showing the proposed packing, and interactions in the three- and four-helix bundle TASP.....	228
Figure 7.1:	Suggested method for the synthesis of TASP containing different peptide sequences.....	232
Figure 7.2:	(a) A heptad repeat representation showing a “helix” (-EQLKKLEQLKKG) with an overall charge, which may discourage TASP self-association by charge-charge repulsion.	
	(b) A heptad repeat representation of a “helix” (-IAAIESKIAAIESKG) that is more suited to forming a three-helix bundle.....	235

List of Tables

Table 1.1:	The effect of altering in the core a and d positions on the GCN4-p1 sequence.....	24
Table 3.1:	LSIMS characterization of peptides.....	88
Table 3.2:	% Yields and ESMS characterization of thioether linked cavitand bowl TASP.....	89
Table 3.3:	% Yields and ESMS characterization of disulfide linked cavitand bowl TASP, made via an "activated" cavitand bowl template.....	90
Table 3.4:	% Yields and LSIMS characterization of S-pyridyl activated peptides.....	92
Table 3.5:	% Yields and ESMS characterization of disulfide linked CTB TASP, made via "activated" peptides.....	94
Table 3.6:	% Yields and ESMS characterization of disulfide linked cavitand bowl TASP, made via "activated" peptides.....	96
Table 4.1:	Thermodynamic evaluation of CTB/N1/L variants, calculated from their GnHCl-induced denaturation.....	116
Table 4.2:	Thermodynamic evaluation of CTB/N2/L (compared to CTB/N1/L), calculated from its GnHCl-induced.....	120
Table 4.3:	Thermodynamic evaluation of Bowl/N1/L variants, calculated from their GnHCl-induced denaturation.....	128
Table 4.4:	Thermodynamic evaluation of Bowl/N2/L compared to Bowl/N1/L and ArBowl/N2/L , calculated from their GnHCl-induced denaturation.....	132
Table 4.5:	Thermodynamic evaluation of Bowl/e/N2/L (compared to Bowl/N2/L), calculated from its GnHCl-induced denaturation.....	143
Table 4.6:	Mean residual ellipticity at 222 nm of CTB and bowl TASP presented in Chapter 4.....	151
Table 4.7:	Data used for, and obtained from, linear fit of sedimentation equilibrium data.....	163
Table 5.1:	Experimentally estimated molecular weights for CTB/N2-GG/L , and CTB/N2-GGG/L determined by sedimentation equilibrium ultracentrifugation.....	171
Table 5.2:	Thermodynamic evaluation of CTB/N2-GG/L , CTB/N2-GGG/L , compared to CTB/N2/L , calculated from their GnHCl-induced denaturation.....	172
Table 5.3:	Thermodynamic evaluation of CTB/C2-GGG/L , compared to CTB/N2-GGG/L , calculated from their GnHCl-induced denaturation.....	176
Table 5.4:	Experimentally estimated molecular weights for TASP Bowl/N2-GG/L , and Bowl/N2-GGG/L determined by sedimentation equilibrium ultracentrifugation.....	181
Table 5.5:	Thermodynamic evaluation of Bowl/N2-GG/L and Bowl/N2-GGG/L , compared to Bowl/N2/L and ArBowl/N2-GGG/L , calculated from their GnHCl-induced denaturation curves.....	183
Table 5.6:	Thermodynamic evaluation of Bowl/C2-GGG/L (compared to Bowl/N2-GGG/L), calculated from their GnHCl-induced denaturation curves.....	188
Table 5.7:	Mean residual ellipticity at 222 nm of CTB and bowl TASP presented in Chapter 5.....	197
Table 5.8:	Data and numerical constants used for, and obtained from, linear least-squares fit of sedimentation equilibrium data.....	198
Table 5.9:	Data and numerical constants used for, and obtained from, linear least-squares fit of sedimentation equilibrium data.....	199
Table 6.1:	Thermodynamic evaluation of CTB/N2-GGG/I , CTB/N2-GGG/V , CTB/N2-GGG/Nle , compared to CTB/N2-GGG/L , calculated from their GnHCl-induced denaturation.....	212
Table 6.2:	Thermodynamic evaluation of Bowl/N2-GGG/I , Bowl/N2-GGG/V , Bowl/N2-GGG/Nle , compared to Bowl/N2-GGG/L , calculated from their GnHCl-induced denaturation.....	217
Table 6.3:	Comparison of the global conformational stability of bowl and CTB TASP with different central amino acids.....	226
Table 6.4:	Mean residual ellipticity at 222 nm of CTB and bowl TASP presented in Chapter 6.....	229

Abbreviations

One- and three-letter codes for the amino acids mentioned in this thesis can be found in Appendix A.

ANS	1-anilinonaphthalene-8-sulfonate
Bowl	cavitand bowl (template)
Bz	benzoyl
CD	circular dichroism
CTB	cyclotribenzylene (template)
δ	chemical shift
d	day(s)
DCM	dichloromethane
DIPEA	diisopropylethylamine
DMA	<i>N,N</i> -dimethylacetamide
DMF	<i>N,N</i> -dimethylformamide
DMS	dimethyl sulfate
DMSO	dimethyl sulfoxide
DPDS	2,2'-dipyridyl disulfide
ESMS	electrospray ionization mass spectrometry (or spectrum)
EtOAc	ethyl acetate
EtOH	ethanol
GnHCl	guanidine hydrochloride
h	hour(s)
HOBt	1-hydroxybenzotriazole
HPLC	high performance (pressure) liquid chromatography
<i>J</i>	coupling constant
LSIMS	liquid secondary ionization mass spectrometry or spectrum
<i>m/z</i>	mass-to-charge ratio

M	moles per litre
MeOH	methanol
min	minutes
mp	melting point
MS	mass spectrometry or spectrum
MW	molecular weight
NBS	<i>N</i> -bromosuccinimide
NMR	nuclear magnetic resonance (spectroscopy)
1d	one-dimensional
ppm	parts per million
rpm	revolutions per minute
rt	room temperature
s	seconds
Spy	S-pyridyl group
[θ]	molar ellipticity per residue
TASP	template assembled synthetic protein
TFA	2,2,2-trifluoroacetic acid
TFE	trifluoroethanol
THF	tetrahydrofuran

Acknowledgements

I wish to extend my sincere gratitude to the following individuals:

My supervisor, Professor John Sherman, for his enthusiasm, high standards, support, and guidance during my “journey” through graduate school. Additionally, I’d like to thank him for “learning me to write proper...like what real scientists do!”

Past and present members of the Sherman group, especially those working in the peptide area, Adam Mezo and Diana Wallhorn, for their “team” spirit in sharing instrument time.

Professor Ron Reid’s group (Patrick, Scott, and Ted) in Pharmaceutical Sciences, for the use of their peptide synthesizer when ours was “sick”. Professor Mike Fryzuk, for the loan of his HPLC. The manifold support services associated with UBC chemistry. Shouming He (Professor Steve Withers) for running the electrospray mass spectrum. Dr. Nick Burlinson, Professor Lawrence McIntosh, and Professor Les Burtnick, for their expertise in NMR spectroscopy, and sedimentation equilibrium ultracentrifugation. Christoph Naumann for his enthusiasm and help running the NMR experiments. Diana Wallhorn for proofreading this thesis. Dr Adam Mezo (Professor Barbara Imperiali, M.I.T.) for running the analytical ultracentrifuge.

My family and friends for their support, and Marisa for everything.

Introduction

1.0 Rationale Behind Thesis

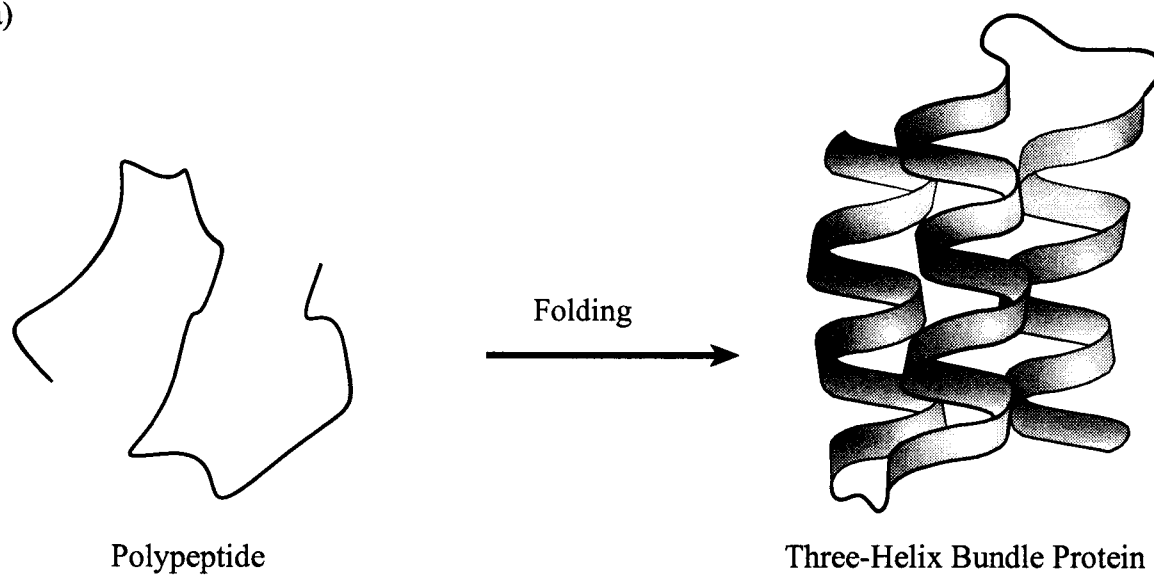
The three-dimensional shape that is unique for every protein results from the specific folding of its polypeptide chain (shown schematically in Figure 1.1a). How does this come about? What features are important to this process? Researchers into this, the protein folding problem, have been trying to answer these questions since it was proposed that the information required for a protein to fold is encoded by its amino acid sequence.

The goal of this thesis is to investigate protein structure, which we simplify by linking peptide strands onto rigid organic macrocycles that act as templates (Figure 1.1b). These templates preorganize the peptide strands and direct them into predetermined folding patterns. We use macrocycles with three- and four-fold symmetry to investigate three- and four-helix bundles, two common motifs seen in proteins.

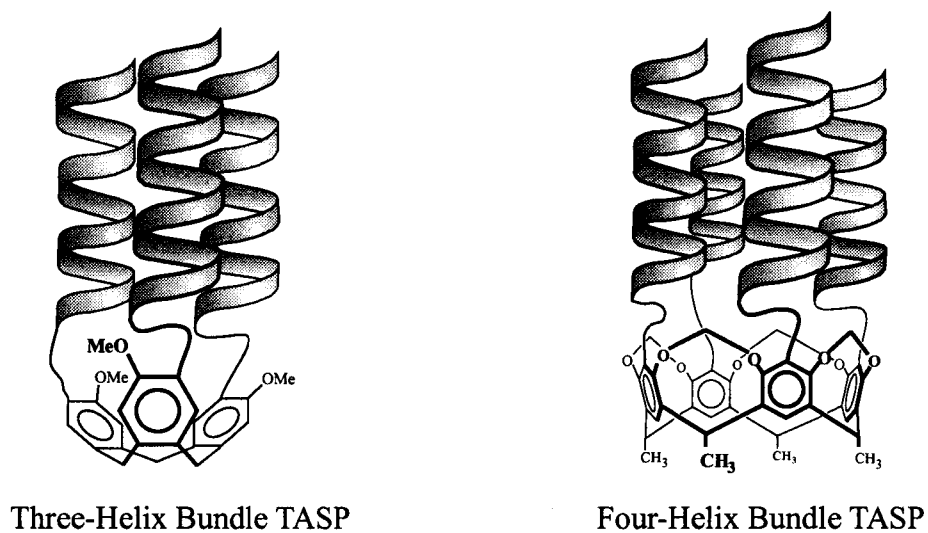
This chapter introduces the protein folding problem and the non-covalent forces that are important in a protein's three dimensional structure. It also covers some of the ways these forces have been investigated, including the analysis and modification of protein structure, as well as the design of *de novo* proteins and template assembled synthetic proteins. An outline of the thesis research goals concludes the chapter.

Figure 1.1: (a) A polypeptide chain folds to form a protein. Lines represent unstructured polypeptide chains, and ribbons represent α -helices. (b) α -Helical peptides on a template give a template assembled synthetic protein (TASP) that mimics protein structure.

(a)



(b)



1.1 Proteins and the Protein Folding Problem

Proteins are essential for life, and have many diverse functions in biological systems, ranging from the physical durability found in connective tissue to the high specificity found in enzymes.¹ In the early 1950s Sanger proved a protein to be a linear amino acid polymer, constructed from a combination of the 20 proteinogenic amino acids.² Furthermore, he determined that the specific sequence of these amino acids was unique to a given protein. It was another two decades before Anfinsen revealed the significance of reversible denaturation / renaturation to a protein's biological properties, in relation to its folded state.³ Anfinsen concluded that the primary amino acid sequence alone contained all the information necessary to define three dimensional structure and thus the function of a protein. Since then researchers have been trying to solve “the protein folding problem”: To determine how three-dimensional structure arises from a primary amino acid sequence.

The three dimensional architecture of these biopolymers has a relatively fixed structure.^{4,5} This results from a multitude of non-covalent interactions involving the local environment, protein backbone and the various amino acid side chains, which may be polar, non-polar or charged.⁵ The ultimate goal for researchers probing the protein folding problem is to understand how three-dimensional structure arises in terms of the chemistry and physics of interactions within the biopolymer.⁶⁻⁸ The solution to this problem has important implications for the biotechnology industry where there is an ever-growing interest in new catalysts and molecular machines. Thus, the design of proteins with novel properties from first principles, *i.e. de novo* design, is a major goal in this area.⁶

Recent advances in protein engineering have had a great impact on this research by making larger quantities of desired native and mutant proteins available for study.^{6,9} This, coupled with advances in X-ray crystallography and protein NMR, has made structural determination more expedient, increasing the knowledge base available and fueling interest in this area. In 1999 alone about 1000 new protein structures were solved.¹⁰ The three dimensional architectures of natural proteins have revealed some common features, with a predicted 1000 different substructural motifs.¹⁰ Ultimately, computer prediction of a protein's structure from its amino acid sequence will use the information from both solved crystal and NMR structures, and solution experiments. There has been some progress towards the prediction of simple structures, but the protein folding problem remains far from being solved.¹¹⁻¹⁷ IBM recently announced that they're building a "super-computer" with the specific task of solving the protein folding problem.^{18,19}

There is a small total energy difference between the folded and unfolded state of proteins, which is easy to measure but currently impossible to calculate from first principles. This difference is 5-20 kcal mol⁻¹ in a protein for which the sum of all non-covalent interactions in the folded or random states is on the order of 200-400 kcal mol⁻¹.^{6,9,20} Folding of a protein is a directed process leading to a structure with highly specific interactions, and the calculation of a lowest free energy structure alone may not give the correct structure (see Section 1.3.2).⁴ Thus, the prediction of a protein's three dimensional structure remains one of the major unsolved problems in molecular biology.

1.2 Protein Structure

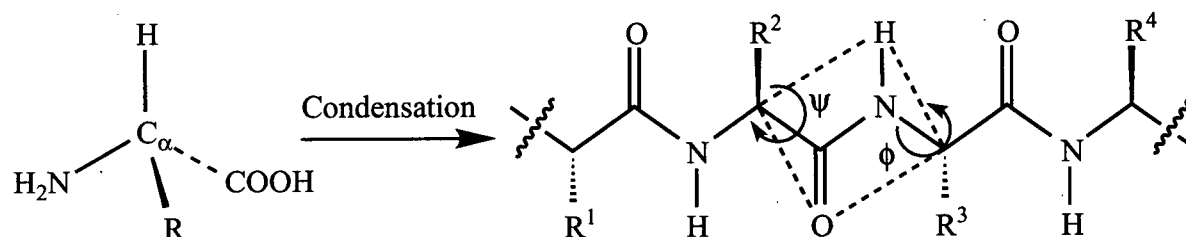
Before one can understand the three dimensional structure of a protein, an understanding of the lower levels of structure must be developed. The common themes seen in the lower levels of structure can occur with different amino acid sequences.^{21,22} Thus, the folding code is degenerate. If we can understand the forces involved with each constituent amino acid, and its contribution to the final structure, we can solve the protein folding problem.²³ Proteins are classed according to their structure and substructure,^{5,9,24} and this thesis is concerned with the forces present in three- and four-helix bundles.

1.2.1 Primary Structure

The primary structure refers to the specific amino acid sequence of a protein and is dictated genetically or by design. The protein biopolymer is formed by condensation between the carboxyl and amino termini of sequential L-amino acids (Figure 1.2).^{5,9} By convention, the amino acid sequence is written (left to right) from the amino to the carboxy end of the polypeptide. The three- and one-letter abbreviations for the twenty common amino acids are shown in Appendix A. The amide bonds between adjacent amino acid residues have significant double bond character and exist almost exclusively in an s-trans configuration.²⁵ Therefore the only two bonds around which rotation can readily occur in the backbone are the C-C_α bond (psi, ψ angle) and the N-C_α bond (phi, ϕ angle) (Figure 1.2).

Only certain combinations of these two angles occur due to steric interactions between the side chains of these chiral L-amino acid residues and the peptide backbone.^{5,9}

Figure 1.2: The formation of a polypeptide via condensation reactions between sequential amino acids. The dotted square represents the plane formed by the peptide amide bond.



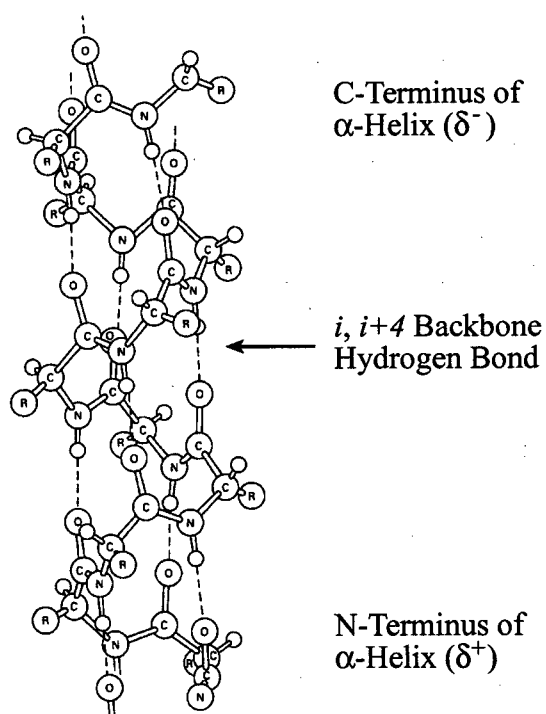
R = Side Chain

1.2.2 Secondary Structure: The α -Helix

Secondary structure refers to the local geometry of the polypeptide chain. Repetition of suitable combinations of phi and psi angles results in the formation of recognizable secondary structural features such as α -helices and β -sheets.^{5,9} These were both predicted in the 1950's by Corey and Pauling,²⁶ who combined basic stereochemical principles with the restrictions imposed by the planarity of the peptide bond. Hairpin or reverse turns occur between anti-parallel elements of α -helix or β -sheet. So called "random coil" (e.g. a fully unfolded protein) occurs where there is no recognizable secondary structure.

Repetition of phi (-57°) and psi (-47°) angles^{5,9} along a polypeptide chain results in the formation of α -helices²⁷ (Figure 1.3). The α -helix is the most common secondary structure, making up 33% of known protein structure.²⁸ It is characterized by regular hydrogen bonding along its backbone between the carbonyl of residue, i , and the N-hydrogen four residues along the polypeptide, $i+4$ (from the N- to C-terminus). Thus, the peptide backbone of an α -helix has a uniform right handed twist with 3.6 residues per turn, each turn being separated by a 5.4 \AA . The backbone hydrogen bonds, which each have a dipole moment, all point in the same direction along the helical axis. This results in a net dipole for the helix, with a partial positive charge at the amino terminus and a partial negative charge at the carboxy end.^{29,30}

Figure 1.3: The right handed α -helix. The first and last four amino acids in the helix backbone have “unsatisfied” hydrogen bonding moieties.



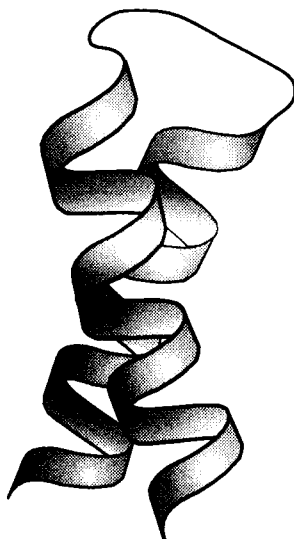
In proteins, helices can vary in length from eight to 37 residues in length,²⁷ with a typical length of 12 residues in globular proteins.³¹

1.2.3 Supersecondary Structure and Protein Domains: Helix Bundles in Proteins

From studies of the known protein structures it has been observed that the secondary structures can pack into simple geometric arrangements. These motifs can occur in different protein structures and are known as super-secondary structure. Two common motifs seen in natural proteins are three- and four-helix bundles, where the helices have a near parallel (or anti-parallel) arrangement (Figure 1.4 shows two helices in this arrangement).³² The helices in these bundles are commonly tilted at 20° to each other, diverging from their closest point of contact, and are known as “square bundles”, a term which differentiates them from coiled-coils (Section 1.2.5).

A protein domain also arises from the association of lower levels of structure. It can exist independently of the whole protein. Again, examples of protein domains include three-^{33,34} and four-helix bundles.³⁵

Figure 1.4: The helices in “square bundles” are splayed out from their closest point of contact.



1.2.4 Tertiary Structure: Helix Bundles as Proteins

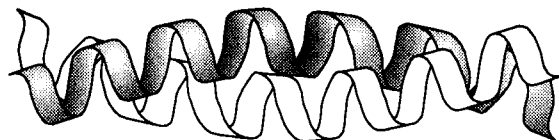
The overall arrangement of all lower levels of structure leads to the tertiary structure of a protein.^{5,9} The native state of a protein refers to its natural, folded structure. Proteins consisting of three-³⁶⁻³⁸ and four-helix^{39,40} bundles (Figure 1.1a shows a schematic representation of a three-helix bundle) have been extensively studied because of their relative simplicity. They are splayed out from their closest point of contact, resulting in a binding pocket between the helices which is often occupied by a prosthetic group.^{5,9,41}

1.2.5 Quaternary structure: Coiled-Coils

Quaternary structure arises through the non-covalent association of two or more proteins.^{5,9} These non-covalent forces are the same as those present in the lower levels of structure.

One common motif that promotes quaternary structure is the coiled-coil.⁴² Over 200 proteins have been identified to associate via this domain.⁴³⁻⁴⁹ A coiled-coil is formed when two or more right handed α -helices line up along their helical axes, resulting in a structure with an overall left handed superhelix, such that there is continuous contact with the adjacent helix (Figure 1.5).⁵⁰ Within a protein, the polypeptide strands that result in this motif are often unstructured as monomers. Coiled-coils have been extensively studied and reveal some important features to protein folding; their structure will be discussed further in Section 1.4.2.2.

Figure 1.5: A two stranded coiled-coil.



1.3 Non-covalent Forces Involved in Protein Structure and Protein Folding

1.3.1 Significant Forces in Protein Structure

An understanding of the dominant forces in protein folding has come from knowledge of the physical properties of molecules, coupled with the trends that are observed in protein structure. The unique shape of a protein results from a balance of forces that favour the native folded state over an unfolded, or misfolded state. This balance of forces often may be tipped in one direction by varying the surrounding environmental conditions, or by substituting one or more amino acid(s) in the sequence.

The dominant driving force for protein folding is the hydrophobic effect,^{5,8,51-56} usually considered to be the aversion of water for non-polar amino acid side chains, which tend to pack the core of the protein.⁵⁷ The hydrophobic effect is an entropically driven process, where the freedom given to water that no longer has to solvate a hydrophobic moiety largely outweighs the unfavourable entropy effect associated with restricting the polypeptide backbone.⁵⁸ Thus, the hydrophilic polar and charged amino acids tend to be found at the surface of the protein. The interior of the protein tends to be packed as tightly as possible (maximizing van der Waals interactions) within stereoelectronic considerations;

the presence of any cavities (or an “over-packed” core) is known to destabilize the folded state of the protein.⁵⁹⁻⁶² A protein's overall architecture can be further stabilized by hydrogen bonding^{63,64} within the polypeptide backbone, which gives rise to secondary structure (see Section 1.2.2). Interaction between suitably paired polar, charged or aromatic amino acid side-chains can also contribute to the stability and conformational specificity of the protein.^{5,8,65} Another example of side-chain interactions that can contribute to protein stability occurs covalently, when two cysteine residues that are spatially close in the overall architecture can oxidize to form a disulfide bond.^{5,9,66}

The forces outlined above can contribute to the global stability of the protein, but more important are the specific interactions between the side chains, especially within the core, that are crucial in defining the immobile, native structure of the protein. Many *de novo* designed proteins were found to be stable to denaturation, but lacked the native-like structure that is the hallmark of natural proteins (Section 1.4.3). Therefore, the global stability of a protein should not be confused with the conformational specificity associated with its “native-like” structure. Protein design relies on successful manipulation of the forces outlined above. The importance of these forces to α -helices and helical bundles will be expanded upon in Section 1.4.

1.3.2 Mechanism of Protein Folding

The native structure of a protein is not necessarily the conformation with the minimum thermodynamic energy.⁶⁷⁻⁶⁹ The folding process is thought to be under kinetic control because the final conformation of the protein can sometimes depend on the refolding conditions.^{68,69} Therefore, the native state of a protein is thought to be the lowest energy state that is kinetically accessible. Folding is thought to follow several different pathways, some of which may form common intermediates before the native structure is reached.⁷⁰⁻⁷⁸ The importance of hydrophobic collapse and secondary structure formation as early folding events is a subject of much study.^{75,79-83}

1.3.3 The Molten Globule State

One important intermediate in protein folding is called the molten globule state, which exists between native and unfolded protein structure.^{4, 84-86} Not all proteins exhibit this folding intermediate. A protein has only one native structure, with relatively immobile side chains that are packed into its core. The molten globule state is recognized as having the elements of secondary structure, but not the specific side chain interactions associated with the native state. Thus, it has a multitude of low energy conformations.⁵ *De novo* protein design (Section 1.4.3) has often resulted in structures that are stable to denaturation but existed as molten globules. Currently, an understanding of the features important for

native-like structures is a major goal of researchers into *de novo* protein design, and therefore a major goal of this thesis. Characterization of molten globules and native-like proteins will be discussed in Chapter 4.

1.3.4 Protein Folding with Chaperones and Co-Factors

Although all the information for a protein to fold is “encoded” in its primary amino acid sequence, other conditions must sometimes be met for a protein to fold correctly. To reach their native state, some proteins require prosthetic groups,^{87,88} carbohydrates,⁸⁹ or metal ions.^{87,90} Some proteins require the presence of other proteins, called chaperones, to fold properly.⁹¹⁻⁹³ Chaperones are thought to help organize the partially folded protein so it can achieve its final structure without misfolding.^{5,94}

1.4 Decoding the Protein Folding Problem

Researchers into the protein folding problem investigate the contribution made by every amino acid and element of secondary structure to the overall protein fold. Elements of secondary structure (Section 1.4.1 focuses on α -helices) are often studied because of the sheer complexity of interactions within a whole protein. Information gleaned from isolated secondary structure often can be applied to a bigger protein, but not always; the folding of fragments of polypeptide chain often can be context dependent.⁹⁵⁻¹⁰⁰

Observation of higher levels of structure in natural proteins (Section 1.4.2.1 deals with helix bundles) can give an insight into significant interactions between individual amino acids, and also between elements of secondary structure. The contribution of an amino acid to the overall stability can be probed by altering a known protein structure (Section 1.4.2.2). This approach has not always proved reliable for more complex proteins, where each amino acid may be involved in a multitude of interactions.¹⁰¹ Thus, it is difficult to interpret the alteration of one specific interaction within the protein, because there may be a structural rearrangement elsewhere in the protein: Proteins fold as a result of numerous co-operative interactions, and therefore it is often difficult to interpret the cause/effect for any alteration in a sequence. However, the interpretation of changes to the relatively simple coiled-coil motif (introduced in Section 1.2.5) has added much knowledge to the area of protein folding (Section 1.4.2.2).

A good way to test our knowledge of the rules governing protein folding is to design proteins “*de novo*”, from first principles (Section 1.4.3). This is an attractive approach for investigating protein folding because we can design and study our own simple structures. Template assembled synthetic proteins (TASPs) are a class of *de novo* proteins that can also simplify the tertiary interactions in proteins (Section 1.4.4). These involve a template that preorganizes the attached peptide strands, which fold into a predetermined folding topology.

1.4.1 Studies of Secondary Structure: Important Forces in α -Helices

The importance of secondary structure stability to the higher levels of protein structure has been demonstrated. For example, in the homo-dimer protein GCN4, whose quaternary structure is formed via a coiled-coil domain, the intrinsic stability of the isolated α -helix that takes part in the coiled-coil domain is thought to be a major determinant of the protein's folding rate.¹⁰² The α -helix is an important component of many protein structures and understanding its stability and structure is important to this thesis, as we want to produce three- and four-helix bundle TASPs. An understanding of the important features of this secondary structure has come from both the observation of α -helices in proteins, and the study of isolated α -helices. The following sections present some of the important factors in the stabilization of α -helices¹⁰³ (Sections 1.4.1.1 to 1.4.1.5).

1.4.1.1 Helix Propensities

Statistical analysis of protein structure shows that some amino acids occur more often in α -helices and others in β -sheets.¹⁰⁴⁻¹⁰⁷ These trends have been investigated by substitution with each of the 20 common amino acids into both isolated α -helices,¹⁰⁸⁻¹¹¹ and α -helices in proteins.¹¹²⁻¹¹⁷ A few minor differences exist between the systems,¹¹⁸⁻¹²² but the trends are similar (Figure 1.6).

Figure 1.6: The order of preference of the twenty natural amino acids for being in an α -helix.

(best) Ala>Arg>Leu>Lys>Met>Gln>Glu>Ile>Phe>Trp>

Ser>Tyr>His>Asp>Cys>Asn>Val>Thr>Gly>Pro (worst)

Alanine,¹⁰⁸ and amino acids with long hydrophobic side chains, or long side chains¹¹² with polar or charged groups at their termini,¹¹⁸ tend to be favoured in α -helices. β -Branched side chains,¹²³ and those with polar or charged groups only one carbon removed from the α -carbon are less favoured. Glycine favours the random coil due to its inherent flexibility. The side chain of proline is inflexible, which induces strain into the helix backbone. Proline also lacks an N-H to participate in any backbone hydrogen bond, making it a “helix-breaker”. The energy difference between alanine and glycine in a helix is about 1 kcal mol⁻¹, a small but significant energy difference that is manifested in the folded structure

of proteins. Within the context of three dimensional structure, the amino acid side chains will also contribute to the overall protein fold according to their size and hydrophobicity (or hydrophilicity).¹²⁴

1.4.1.2 Helix Chain Length

Typically helices are eight (which is the minimum number of amino acids required to form an isolated α -helix) to twelve amino acids long in globular proteins,³¹ generally the longer the helix the higher its stability (this observation can be related to the “unsatisfied” carbonyl and N-H hydrogen bonding moieties at the termini of the helix).^{125,126}

1.4.1.3 Helix Macro-dipole

A result of the directionality of the backbone hydrogen bonding in helices (see Figure 1.3) is that a macro-dipole is formed along the length of the helix.^{29,30} This electrostatic effect^{29,30,127-131} is estimated to put a half positive charge at the N-terminal end, and a half negative charge at the C-terminal end of the helix. Therefore, the helix is stabilized when amino acids with a negatively charged side chain are located near the N-terminal end, and those with a positively charged side chain near the C-terminus.¹¹⁸

1.4.1.4 Side Chain Interactions Within The Helix

The α -helix typically contains 3.6 residues per turn, therefore residues at either the i , $i+3$ or at the i , $i+4$ positions reside on the same side of the helix and can interact. The most common stabilizing interaction is a salt bridge between oppositely charged residues¹³²⁻¹³⁴ (typically Glu-Lys, which contributes about $0.5 \text{ kcal mol}^{-1}$ to the helix stability). However, having oppositely charged residues at either the i , $i+1$ or the i , $i+2$ position, favours the random coil, and therefore destabilize the helix. Stabilizing hydrogen bonds^{118,135}, non-polar,¹³⁶ and aromatic interactions¹³⁷ between side chains in a helix have also been studied.

1.4.1.5 Capping of the Helix

In protein α -helices, certain residues are seen to occur more frequently at, or close to the C- or N-termini.¹³⁸⁻¹⁴⁰ The effect of residues at the helix termini have been investigated in both peptides and proteins.¹⁴¹⁻¹⁴³ In the helix backbone, the first and last four hydrogen bonding moieties are unpaired. Residues that stabilize the termini have been shown to hydrogen bond with the “unsatisfied” backbone hydrogen bonds, stabilize the helix macrodipole (see Section 1.4.1.3), or aid in the burial of hydrophobic surface. Suitable capping of the N-terminus can contribute up to 2 kcal mol^{-1} to the helix stability.¹⁴⁴⁻¹⁵⁴ Due to its small size, glycine is thought to act as a good N-cap by not restricting the solvation of the unpaired N-Hs in the helix backbone.¹⁵⁵ Capping of the C-terminus (which is less

common in proteins) is thought to only contribute up to 0.4 kcal/mol, and is generally attributed to amino acids that stabilize the helix macrodipole.¹⁴⁰

1.4.2 Investigation of Higher Levels of Protein Structure: Helix Bundling in Proteins

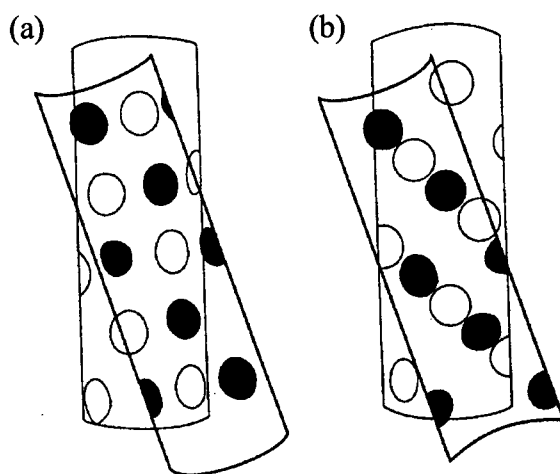
Within proteins and peptides, side chain interactions can stabilize secondary structure, but more importantly the specific side chain interactions dictate the tertiary structure. Helix bundles are formed via complementary interactions between their inter-helical side chains. The following section will look at some of the research on helical proteins, which is a combination of observation of crystal and NMR structures of natural proteins (Section 1.4.2.1) and solution experiments (Section 1.4.2.2).

1.4.2.1 Side Chain Packing: Inter-Helical Packing

Helices can interact at different angles with respect to the helical axis, giving rise to bundles with various topologies. We're interested in three- and four-helix bundles that pack against each other at approximately 20° with respect to the helical axis. Several theories have been proposed to explain the packing in these common motifs. Nearly fifty years ago, Crick proposed the “knobs into holes” model^{50,156} (Figure 1.7a), where the side chain,

“knob”, of one helix, points into the “hole” made by the side chains of the adjacent helix at positions i , $i+3$, $i+4$ and $i+7$. After about thirty years, this explanation of inter-helical packing fell out of favour as more protein x-ray crystal structures were revealed. Chothia proposed the “ridges into grooves” model¹⁵⁷⁻¹⁵⁹ to better explained different inter-helical packing angles (Figure 1.7b). Here, the side chains are arranged in a row along the surface of the helix, and form ridges that are separated by grooves. The ridges can be formed by side chains that are separated either by three (i , $i+3$, $i+6$...) or four (i , $i+4$, $i+8$...) residues, and they pack into grooves formed by an adjacent helix. Recently Bowie suggested that these “two-dimensional” models are only relevant to helices that pack directly “face-to-face”.¹⁶⁰⁻¹⁶⁵ He proposed that the flexible side chains can adjust to accommodate any variety of inter-helical packing angles.

Figure 1.7: Inter-helical side chain packing: (a) Knobs into holes. (b) Ridges into grooves.



Helices in bundles tend to be oriented anti-parallel rather than parallel as a result of interaction between their helix macrodipoles. However, the energy difference between the parallel and anti-parallel arrangements of helices is thought to be less than 1 kcal mol⁻¹.

1.4.2.2 Alteration of Natural Protein Sequences

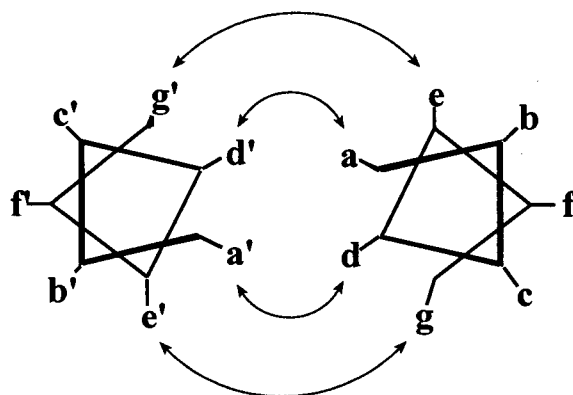
Attempts at altering a single amino acid in a large protein and observing any change in stability/structure can only be limited to simple proteins for reliable interpretation. Coiled-coil domains were introduced in Section 1.2.5, and are known to form dimeric,¹⁶⁶ trimeric,^{135,167-171} tetrameric,¹⁷²⁻¹⁷⁹ and pentameric¹⁸⁰ structures. Their relatively simple structure has prompted much investigation. The interacting α -helices in this structure result in a left handed supercoil with 3.5 residues per helical turn, as opposed to the 3.6 residues per turn that is typical of α -helices. Thus, the constituent α -helices in a coiled-coil can be characterized by what is called the heptad repeat: Each seven residues in a peptide sequence are represented by the repeating [abcdefg]_n unit, where every seventh residue is in an equivalent position. Typically, the **a** and **d** positions are occupied by hydrophobic amino acids that are completely buried in the protein's core and are largely responsible for holding the structure together (Figure 1.8a shows a two-stranded coiled-coil). Residues at positions **e** and **g** are partially exposed to the solvent and can also interact with adjacent helices. The core amino acids are known to interact via Crick's "knobs-into-holes" packing, and substituting them can lead to a change in the oligomeric state of the coiled-coil.^{28,181-183}

One example of this is a study on the polypeptide representing the N-terminal coiled-coil domain of the homo-dimeric DNA binding protein GCN4, the sequence is called GCN4-p1 (shown in Figure 1.8b).¹⁸⁴⁻¹⁸⁶ This native sequence contains mainly valines at the **a** positions and all leucines at the **d** positions of the heptad repeats. However, one of the core **a** positions contains the hydrogen bonding asparagine residue. Complementary pairing of these hydrophilic residues leads to a hydrogen bond that is buried in the core of the protein, this is thought to be responsible for specifying the parallel arrangement of the coiled-coil dimer. Replacement of this core asparagine with a hydrophobic residue (e.g. valine) results in a multitude of oligomeric states for the coiled coil. Therefore, by favouring the dimer, this buried hydrogen bonding pair also destabilizes any higher order aggregates, and imparts structural specificity to the protein.

Another study on GCN4-p1 that illustrates the subtle complexity of the protein folding problem involved varying the hydrophobic residues at the **a** and **d** positions of the helices in the GCN4-p1 coiled-coil homodimer protein (Table 1.1).¹⁸⁴ In this study the **a** and **d** positions of the constituent helices were replaced with I, L or V residues (e.g. GCN4-pVL had V occupying all of its **a** positions, and L at all the **d** positions). The GCN4-pIL, -pII and -pLI mutants were found to exist as dimers, trimers and tetramers respectively. Whereas, the GCN4-pVI, -pVL, -pLV, and -pLL variants existed in multiple oligomeric states. Thus, the shapes of the buried side chains in coiled-coils are essential determinants of the global protein fold. This work also demonstrated that the patterning of polar and hydrophobic amino acids in a protein sequence is insufficient to determine its three dimensional structure.

Figure 1.8: (a) Helical wheel showing the heptad repeat, and the interactions seen in a two stranded coiled-coil. Residues at positions **a** and **d** form the hydrophobic core of the protein. Residues at positions **e** and **g** can also interact. “Knobs-into-holes” packing occurs when, for example, the **a** residue (the “knob”) packs into the “hole” made by **a'**, **g'**, and two **d'** residues. (b) The GCN4-p1 amino acid sequence.

(a)



(b)

a b c d e f g a b c d e f g a b c d e f g a b c d e f g a b c
R M K Q L E D K V E E L L S K N Y H L E N E V A R L K K L V G E R

Table 1.1: The effect of altering in the core **a** and **d** positions on the GCN4-p1 sequence.

NAME	"a" residue	"d" residue	Oligomeric State
GCN4-pLL	L	L	Multiple
GCN4-pLI	L	I	Tetramer
GCN4-pLV	L	V	Multiple
GCN4-pIL	I	L	Dimer
GCN4-pII	I	I	Trimer
GCN4-pVL	V	L	Multiple
GCN4-pVI	V	I	Multiple

Coiled-coil dimers have been used to investigate the effect of changing single amino acids in the core,^{172,187-191} capping residues,^{192,193} chain length of the constituent helices,^{125,126,194} helix propensities of the amino acids,^{116,195-197} positional effects of disulfide bonds between the helices,¹⁹⁸ and the strength of inter-helical electrostatic interactions.¹⁹⁹ Their well understood structure has led to the *de novo* design of coiled-coils with different oligomerization states,^{169,200} that have metal induced folding (a molecular switch),²⁰¹ and that form heterodimers,²⁰²⁻²⁰⁴ or heterotrimers.²⁰⁵ Another extension of this work is where turn motifs have been added between the coiled coils, leading to α -helical bundle proteins constructed from a single polypeptide chain (see Section 1.4.3).^{35,206}

1.4.3 *De Novo* Protein Design²⁰⁷⁻²¹⁰

The rules of protein folding are sufficiently understood to design proteins that can adopt simple folded conformations. These *de novo* proteins must undergo controlled monomeric folding to form the designed structure; ideally any potential alternate folds must be destabilized by design. One problem with these man-made proteins is that they often result in molten globules (introduced in Section 1.3.3). However, researchers now have a better understanding of how to design a native-like fold.^{21,212} The literature cites examples of *de novo* designed β -sheet proteins,^{187,192,194,195,200-202,213,214} mixed α - β structures,^{14,215-218} and helical bundles.^{32,205,219-223}

The design of these proteins is done by an iterative process: The information from one design is used in the redesign of a better structure. An elegant example of this process is the design of an anti-parallel four-helix bundle protein. Initially this involved a 16-residue peptide strand, which was designed to form an amphiphilic helix with both a hydrophobic and hydrophilic face (Figure 1.9).²²⁴⁻²²⁶ In water, this peptide associated into a four-helix bundle via the hydrophobic face on each helix (Figure 1.10). The stepwise addition of suitable reverse turn fragments led to a 74 residue polypeptide that successfully folded into a four-helix bundle.²²⁷ Initially a molten globule, the structure of this protein has been further refined to adopt a more native-like structure.²²⁸⁻²³⁰

When designing a protein tertiary structure, not only should interactions be included to stabilize the desired fold, but also interactions that destabilize alternate folds. One example of the principle is Dutton's design of a three-helix bundle (Figure 1.11).¹⁸⁵ The design for this protein involved helices that were shortened versions of a known three-stranded coiled-coil, onto which were added reverse turn motifs. The authors knew that to produce a native-like structure the specificity of the side chain interactions, especially in the hydrophobic core, were more important than the global stability of the protein. The three helices in this protein adopted the "clockwise" orientation (Figure 1.11a) due to the favourable inter-helix electrostatic interactions and hydrophobic core side chain specificity. The alternate fold (Figure 1.11b) would involve less specific core interactions and electrostatic repulsion along two of the faces. Working independently, DeGrado reported an almost identical three-helix bundle design at the same time.^{231,232}

Figure 1.9: Helical Wheel diagram showing the hydrophobic and hydrophilic faces of the peptide Ac-GELEELLKKLKELLKG-NH₂ when it is folded into an α -helix. NOTE: Only residues 2 to 15 are included in the diagram, which is viewed from the N- to the C-terminus. The hydrophobic face is the driving force for self-association to form a four-helix bundle. The helix design also includes stabilizing intra-helical $i, i+4$ salt bridges, and suitably charged residues towards the N- and C-termini to minimize any unfavourable helix macrodipole effects. Additionally the helix includes charged residues that may form inter-helical salt bridges in the designed anti-parallel four-helix bundle.

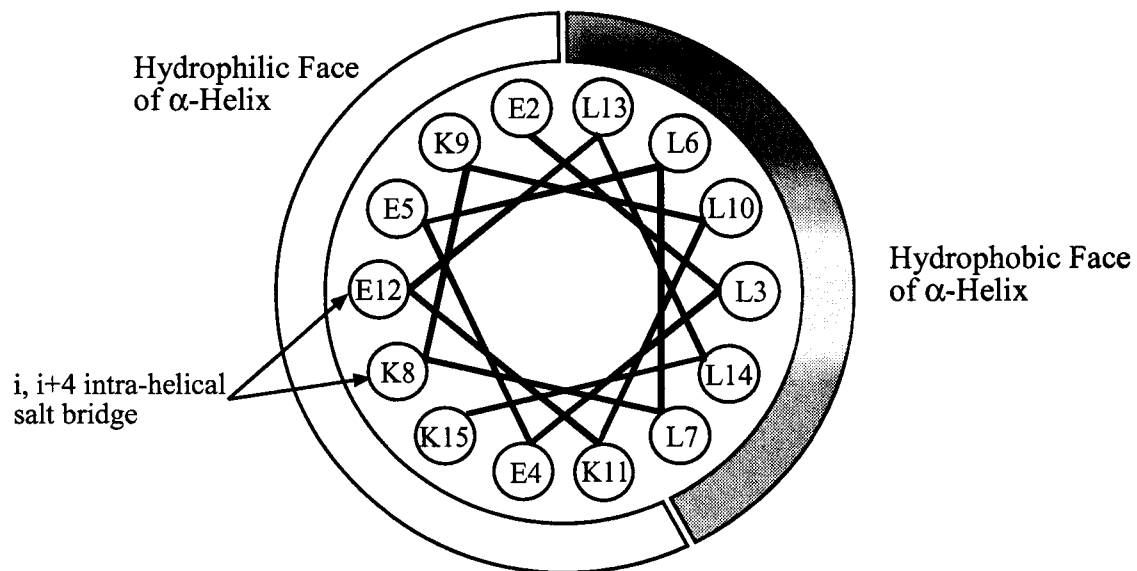
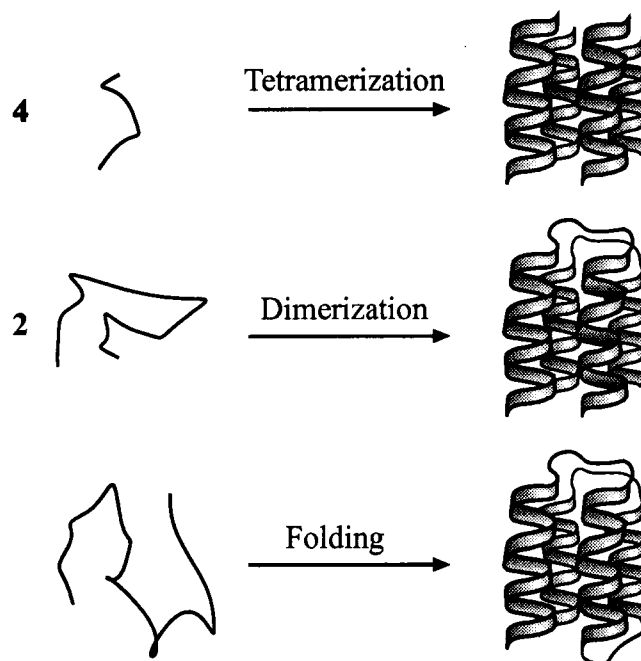


Figure 1.10: Iterative design of a four-helix bundle. The helical portion of the strands had the sequence Ac-GELEELLKKLKELLKG-NH₂ and the turn motifs contained –PRR–.

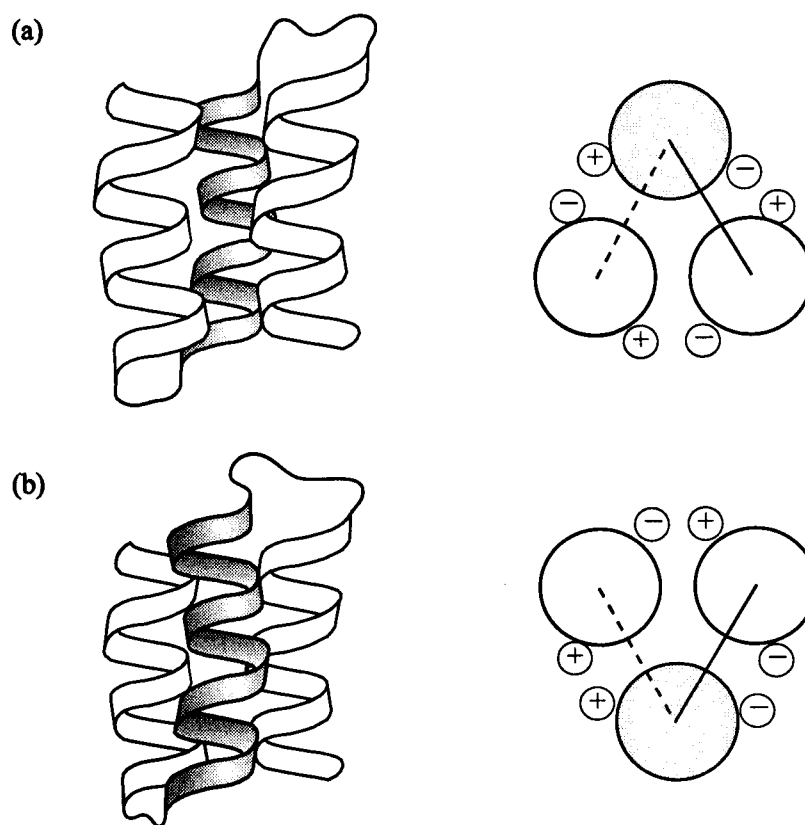


One of the major problems in designing *de novo* proteins, is to produce a “native-like”, rather than a molten globule-like structure. For helical bundles, the inter-helix side chain specificity is known to be important. Some progress has been made towards mimicking native-like structure using clusters of aromatic residues,²³³ hydrogen bond specificity,²³⁴ or by the suitable patterning of hydrophobic residues.^{235,236}

Other workers have concentrated on designing helical proteins based solely on the predicted position of hydrophilic, or hydrophobic residues in the folded structure.^{222,237,238} They then used a combinatorial approach to produce proteins with a sequence that only specifies hydrophobic or hydrophilic residue at a certain position. Most of these sequences

resulted in proteins that were monomeric, α -helical, and water soluble.²³⁷ Redesign of their best sequences from this combinatorial set led to a protein with native-like properties.²³⁸

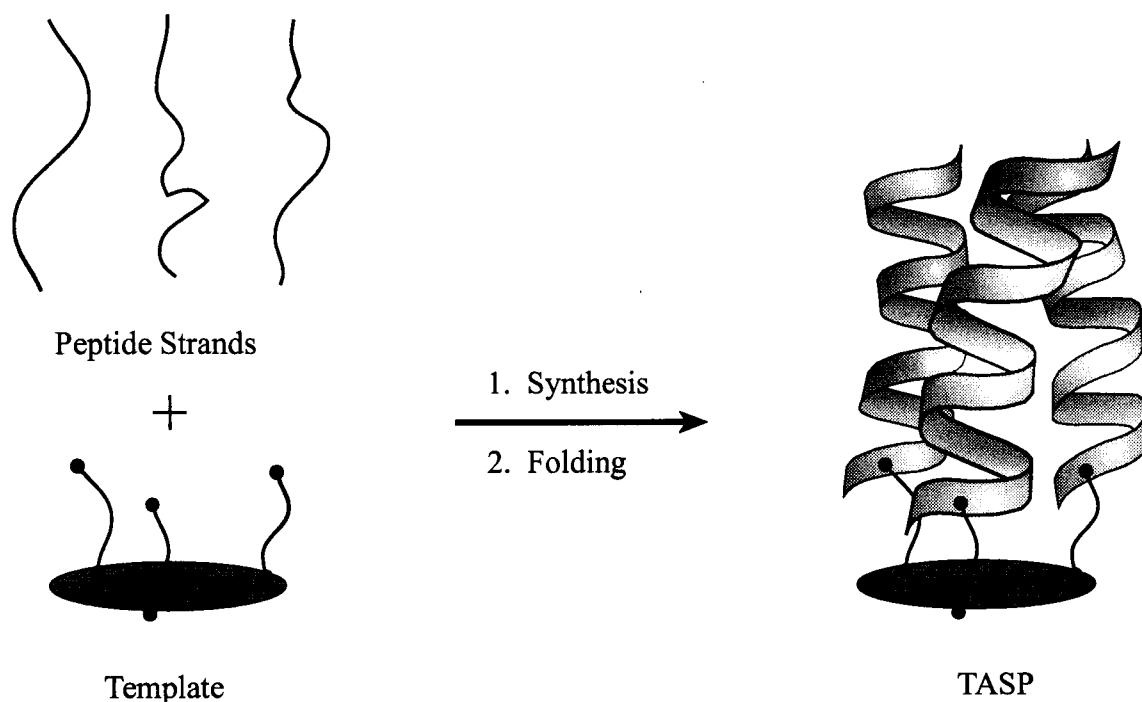
Figure 1.11: Designed “clockwise” folding of a three-helix bundle. The peptide sequence is Ac-R-VKALEEK-VKALEEK-VKAL-GGGG-R-IEELKKK-WEELKKK-IEEL-GGGG-E-VKKCEEE-VKKLEEE-IKKL-NH₂ (the **a** and **d** hydrophobic positions in each heptad repeat are shown in bold). The core contains primarily valine and leucine residues, a combination that has been shown to favour three stranded coiled-coils. The structures on the right hand side of the figure show the charged residues that occupy the **e** and **g** positions of the heptad. Turn motifs (-GGGG-) have been shown to be less important to the final topology of the bundle,^{239,240} but may effect the rate of folding.^{240,241}



1.4.4 Template Assembled Synthetic Proteins

The potentially complex folding patterns of a polypeptide can be overcome by attaching peptide strands to a template molecule. The template is responsible for directing the peptide strands into a pre-determined folding topology. Thus, a template assembled synthetic protein (TASP)^{242,243} can help to overcome the unfavourable loss of entropy associated with restricting the polypeptide during folding. For the protein sub-structure under investigation, a suitable template separates the peptide strands by distances similar to those seen in nature. Thus, by reducing the degrees of freedom afforded to the peptide, the potential for intra-molecular interaction is enhanced, which facilitates the desired folding pattern (Figure 1.12).

Figure 1.12: Schematic diagram showing how a template assembled protein (TASP) mimics protein structure.

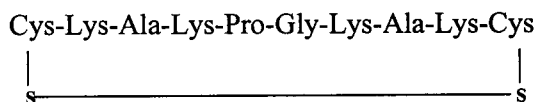


Templates have been used to promote structure in isolated α -helices,²⁴⁴⁻²⁴⁷ β -sheets,²⁴⁸⁻²⁵⁰ and α -helical bundles (described below – Section 1.4.4.1-1.4.4.5). TASP s can be constructed using either a divergent (Sections 1.4.4.1-1.4.4.3), or convergent (Sections 1.4.4.4-1.4.4.5) approach. Methods for covalently linking peptide strands to template molecules will be covered in more detail in Chapter 2. The main feature of all the TASP s outlined below is that the spacing of the constituent peptides approximates the inter-helical distance seen in nature of 6-18 Å.²⁵¹

1.4.4.1 Peptides as Templates

Mutter pioneered the TASP concept using a linear peptide as template onto which a variety of protein motifs were attached via the constituent lysine residues.^{242,243} To decrease its flexibility, this template was later cyclized via a disulfide bond between two cysteine residues in its sequence (Figure 1.13).^{252,253} Variants of this basic template^{254,255} have been used as scaffolds in the construction of TASPs which have bio-electric properties,²⁵⁶⁻²⁵⁸ or act as ion channels.^{259,260}

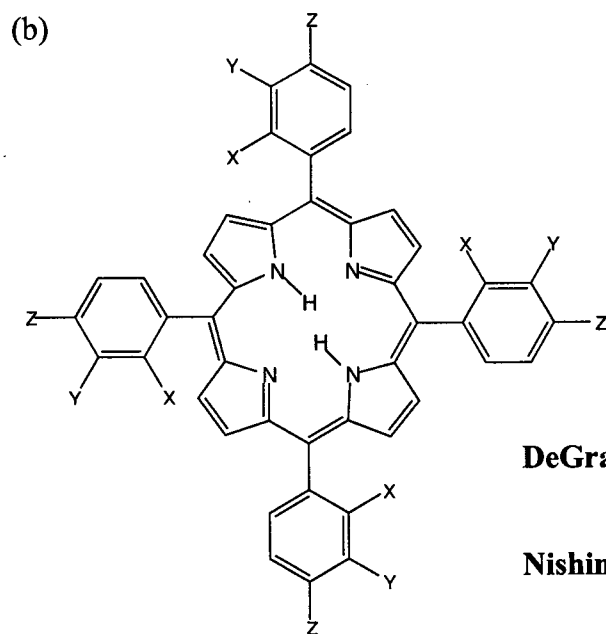
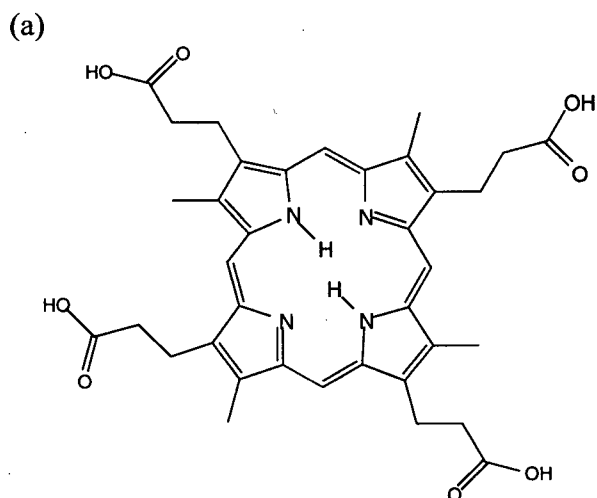
Figure 1.13: One example of Mutter's peptide template. A four-helix bundle TASP was made by linking peptides (EALEKALKEALAKLG) via the lysine side chains of the template.



1.4.4.2 Porphyrin Templates

Sasaki and Kaiser were the first to prepare a four-helix bundle metalloprotein based on a porphyrin molecule (Figure 1.14a).^{261,262} DeGrado synthesized a metalloporphyrin (Figure 1.14b) TASP that has proton channel activity.²⁶³ Nishino investigated two porphyrin TASPs (Figure 1.14b) which were soluble in lipid bilayer membranes.²⁶⁴ The ortho-substituted analogue had a higher helicity, and was better incorporated into phospholipid bilayers, characteristics that were attributed to a better template.

Figure 1.14: Porphyrin Templates: (a) Sasaki/Kaiser made a four-helix bundle TASP by linking peptides (AEQLLQEAQELLQEL) via the acid moieties on the template. (b) DeGrado and Nishino made four-helix bundle TASPs by linking peptides [(LS-Aib-LSL)₃, and A(ALLELLA)₃ respectively] via the acid moieties on the template.



DeGrado X & Z = H, Y = COOH

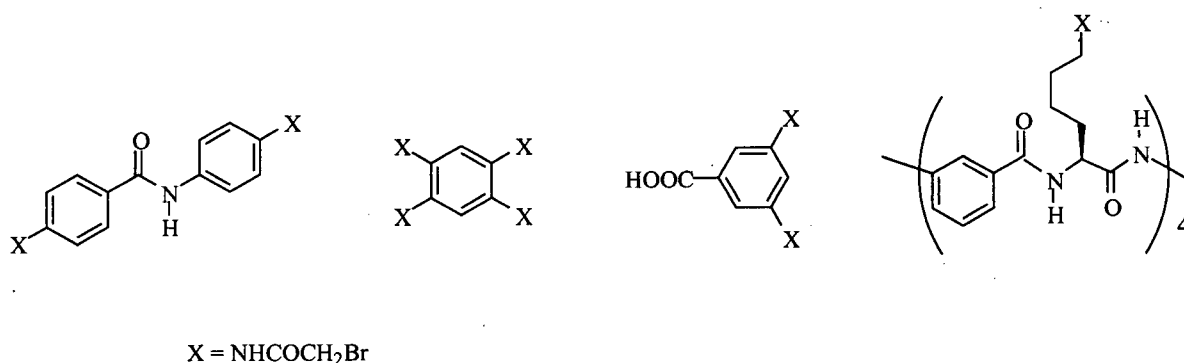
Nishino X = COOH, Y & Z = H

X & Y = H, Z = COOH

1.4.4.3 Other Divergent TASPs

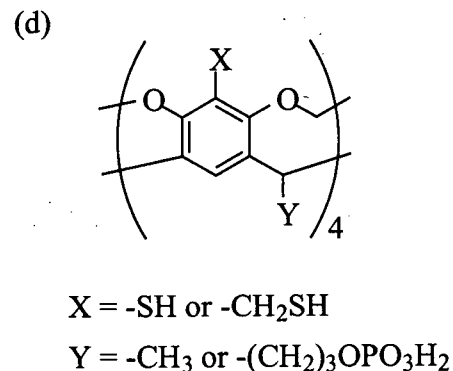
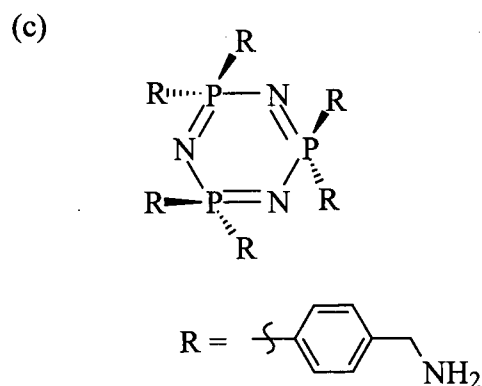
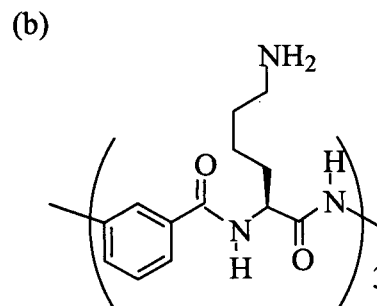
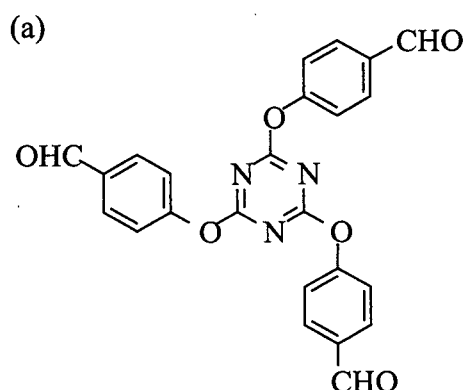
Fairlie investigated several organic templates (Figure 1.15), and proposed that the shape, size or directionality of a template is not crucial to the formation of a four-helix bundle TASP, provided that the linker to the peptide fragment is flexible enough.²⁶⁵

Figure 1.15: Fairlie's organic templates. Two-, and four-helix bundle TASPs were made by linking peptides (DAATALANALKKL) via the alkyl halide side chains on the template.



Three-helix bundles were modeled by Sasaki using an organic trialdehyde²⁶⁶ (Figure 1.16a), and by Nishino using a cyclic pseudopeptide template²⁶⁷ (Figure 1.16b). Inoue used a cyclic phosphazene²⁶⁸ (Figure 1.16c) to model a pair of three-helix bundles, which sat above and below the plane of the ring. Especially relevant to this thesis is the earlier work done in our (Sherman) group which modeled four-helix bundles on a cavitand template²⁶⁹⁻²⁷¹ (Figure 1.16d). This work demonstrated the importance of the template-to-helix linker in terms of both the association state, and conformational specificity (native-like structure) of the resultant TASP (work which will be covered in more detail elsewhere in this thesis).

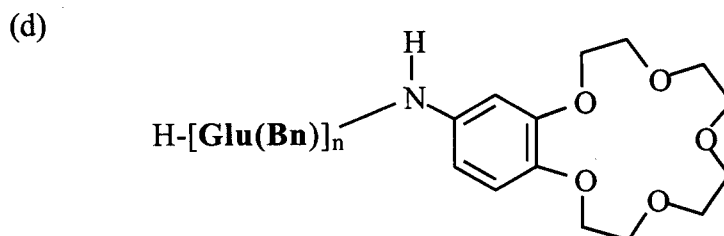
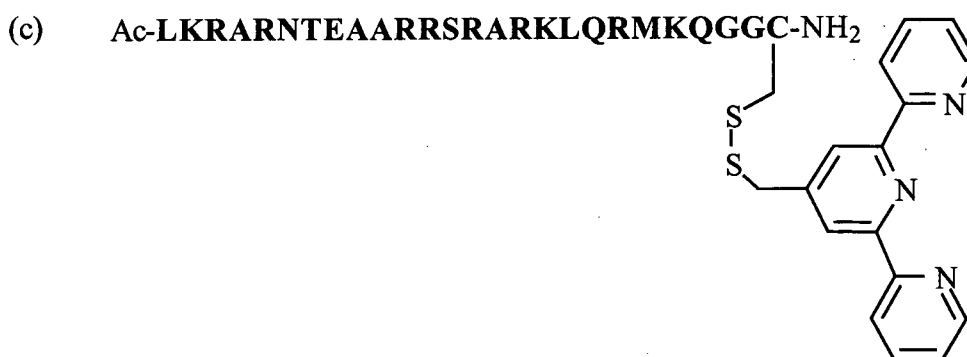
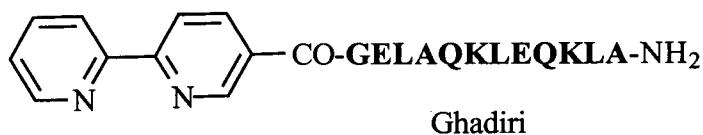
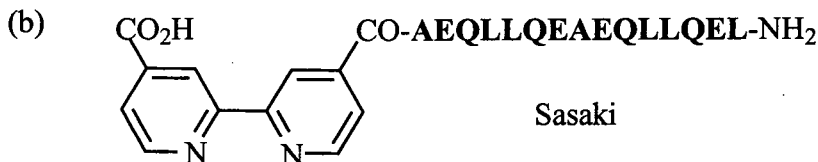
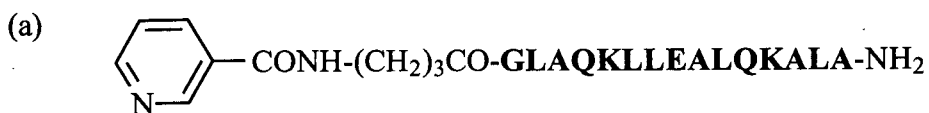
Figure 1.16: Other divergent organic templates: (a) Sasaki made three-helix bundle TASPs by linking peptides (AEQLLQEAELQLQEL) via the aldehyde moieties on the template. (b) Nishino made three-helix bundle TASPs by linking peptides (PE-Cha-LKA-Bpa-AEL-Cha-K) via the amine moieties on the template. (c) Inoue made a pair of three-helix bundle TASPs by linking peptides [poly-E(Bn)] via the amine moieties on the template. (d) Sherman made four-helix bundle TASPs by linking peptides (EELLKKLEELLKKG) via the thiol moieties on the template.



1.4.4.4 Templatation by Metal Ions

The addition of a suitable chelating moiety to the peptide strand can lead to assembly around a metal ion template. Four-helix bundles have been modeled by Ghadiri using a ruthenium (II) with four pyridyl-modified peptides²⁷² (Figure 1.17a). Three-helix bundles have been modeled by Sasaki using iron and bis-pyridyl moities,²⁷³⁻²⁷⁵ and independently by Ghadiri using nickel (II), copper (II) or Ruthenium (II) with bis-pyridyl moieties^{276,277} (Figure 1.17b). Schepartz produced a metal ion induced two-helix bundle TASP using a terpyridyl modified peptide and iron (II) metal²⁷⁸ (Figure 1.17c). This TASP dimer was found to selectively bind DNA. Finally, Minoura used two crown ether containing peptides (Figure 1.17d), which were assembled into a dimer upon the addition of potassium (I).²⁷⁹

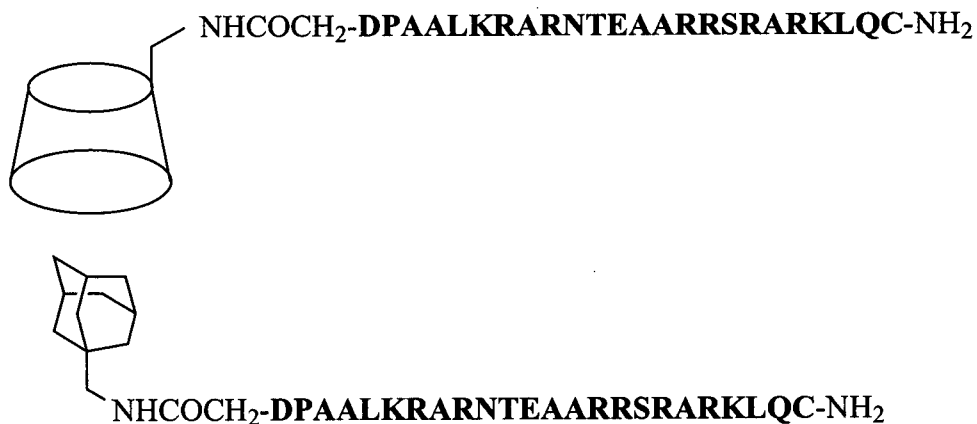
Figure 1.17: TASPs are formed via addition of a metal ion which induces chelation of the modified peptide strands: (a) Ghadiri's four-helix bundle induced via the addition of ruthenium (II). (b) Ghadiri's and Sasaki's three-helix bundles induced via the addition of iron (II), nickel (II), copper (II) or ruthenium (II). (c) Schepartz's two-helix bundle induced via the addition of iron (II). (d) Minoura's two-helix bundle induced via the addition of potassium (I).



1.4.4.5 Other Convergent Templates

Morii templated a two-helix bundle structure using a cyclodextrin-guest inclusion complex²⁸⁰ (Figure 1.18). This involves peptide strands that contain either a cyclodextrin or an adamantane at their termini. This inclusion dimer was found to bind DNA more tightly than the underivatized peptides.

Figure 1.18: Morii's "template" is formed by the adamantane-cyclodextrin inclusion complex.



1.5 Beyond the Protein Folding Problem: Designer Proteins

Currently, the tertiary structure of a protein is only available via x-ray crystallography or a detailed NMR investigation. Only with information from these structures, and from the investigation of simpler structures can we assess how the tertiary structure relates to the primary amino acid sequence. By building a computer databank of the common motifs seen in protein structures, it may one day be possible to predict the structure of a protein from its primary amino acid sequence alone. Beyond this goal, researchers are trying to design new proteins with novel properties,^{236,281-285} which may incorporate any natural or non-natural amino acids,²⁸⁶ small molecules,^{174,235,287-289} or ions.^{230,290-293} These new proteins could find a use as designer catalysts,^{181,183,294,295} or in the rapidly expanding "nanotechnology" industry as molecular machines, receptors,^{176,296} or ion channels.^{297,298} A polypeptide is an attractive target because of its rigid backbone, diverse side chain functionalities and synthetic availability. Site directed mutagenesis has opened the door to the production of large quantities of proteins that are made up of natural amino acids. Methods for the introduction of non-natural amino acids into a protein sequence has received much attention recently.^{299,300}

1.6 Thesis Goals and Summary of Thesis

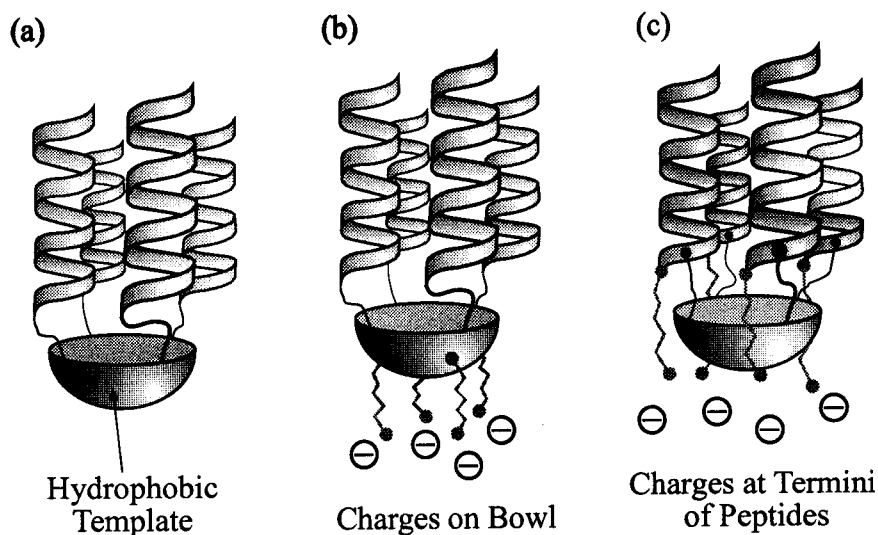
Listed below are the four major goals of this thesis, followed by a brief description of the work towards these goals (and the rationale behind it). More details are given in the individual chapters and at the end of the thesis. The four major goals of this thesis were:

1. TASP synthesis: Find an efficient method to covalently link unprotected peptide strands to the templates.
2. Produce TASPs that are monomers in solution.
3. Investigate the effect of the template on the peptide strands (in terms of structure and stability), and on the degree of inter-TASP association.
4. Produce TASPs that have “native-like” structure.

Chapter two gives background on the two templates used in this thesis, and outlines their syntheses. Chapter three is concerned with methods used for the syntheses of the TASPs, specifically the methods used for linking the peptides to the template. Nearly all the TASPs used in this thesis were made using disulfide bonds between a cysteine residue on the constituent peptides, and thiol moieties on the template.

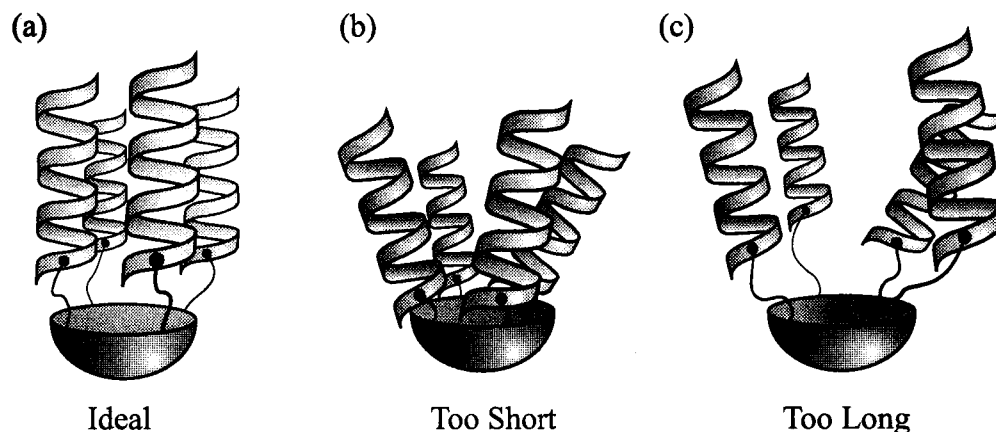
Chapter four introduces methods used for studying protein structure, then presents the characterization of the prototypical three- and four-helix bundle TASPs. A portion of this work has been published.³⁰¹ These initial TASPs were found to self-associate in solution, and the hydrophobic template macrocycles were recognized as a potential source of this self-association (Figure 1.19a). The addition of charged groups onto the underside of the organic macrocyclic template (Figure 1.19b) is one method that has been used in overcoming this potential source of TASP self-association.^{270,271} Theoretically, the templates are kept apart by charge-charge repulsion. In Chapter four, a more universal approach to overcoming this potential association of the hydrophobic templates was to make the environment surrounding the template charged. This was achieved by using an amino acid sequence which contained a charged group that resides in the vicinity of the macrocycle: Therefore, it was hoped that charge-charge repulsion between the “dangling” peptides on each TASP would prevent self-association via the templates (Figure 1.19c). However, this approach did not prevent TASP self-association, which appeared to originate from non-optimal packing of the template-bound helices.

Figure 1.19: Schematic representations of four-helix bundle TASPs: (a) The hydrophobic template may lead to TASP self-association in water. (b) Charged groups on the template may prevent self-association. (c) Charged groups on the peptide sequence may prevent self-association.



Chapter five reports a more successful approach to solving the self-association of the TASPs: Namely, to insert a flexible “linker” between the template and the helix. This was achieved using glycine residues as “spacers” between the template-bound cysteine residue and the “helix”. A suitable “spacer” would allow optimal interactions between the constituent helices on a TASP (Figure 1.20a). If the linker is too short (Figure 1.20b), this may lead to steric crowding of the helices at the template, and result in them pointing away from each other (possibly leading to association between the helices of two TASPs). Too long a template-to-helix linker may negate the directing ability of the template (Figure 1.20c).

Figure 1.20: Optimal length of template-to-helix linker is important to the bundling of the constituent helices. (a) Ideal linker. (b) Too short a linker. (c) Too long a linker.



Chapter five also compares the effect of linking the helices via their opposite termini, in order to investigate the potential interaction between the template and the different ends of the α -helices: Thus, any “end-capping” of the helices by the template, may result in TASPs with different stability and structure (when comparing the C- and N-linked helices).

Chapter six focuses on the final goal of this thesis, the production of TASPs with “native-like” structural properties. This goal was not achieved: All of the three- and four-helix bundle TASPs made in this thesis manifested structure akin to a “molten globule” (see Section 1.3.3). In this chapter some of the core amino acids of the peptides were altered, but displayed no increase in their conformational specificity. However, some differences were observed in the stability of these TASP analogues, these differences were attributed to the size and shape of the core amino acids.

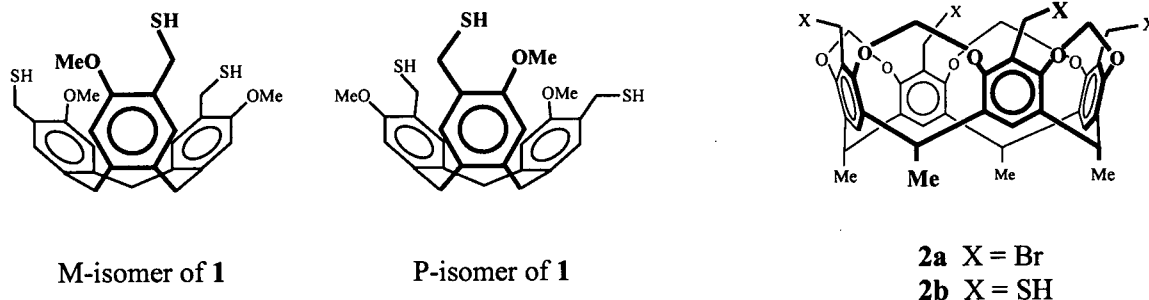
Chapter seven concludes the thesis, outlining the major findings of this research, and suggesting further experiments and directions for the research. Also I will compare the research presented here to other work from our group, as well as work from the literature.

Synthesis of Templates

2.0 Choice of Templates

The organic macrocycles that were used as scaffolds for the three- and four-helix bundles were cyclotribenzylene (CTB) **1** and cavitand bowl **2**, respectively (Figure 2.1). These rigid, bowl-shaped organic macrocycles were selected because they are synthetically available, and will result in inter-helix distances that approximate those seen in nature. The functional groups that served as synthetic handles for attaching the peptide strands were either benzyl bromide, or benzyl thiol moieties. Another feature of these templates is the enforced cavity, which has been shown to bind hydrophobic molecules in aqueous solution.^{302,303}

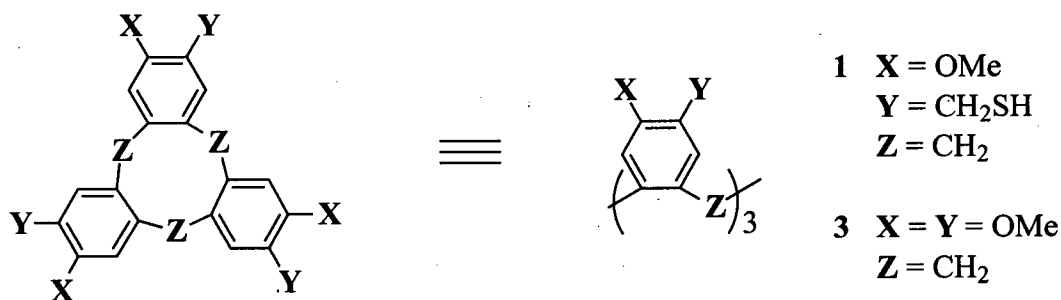
Figure 2.1: The organic macrocycles that were used as templates for three- and four-helix bundles. The CTB macrocycle **1** exists as a pair of enantiomers (known as the M- and P-isomers).³⁰⁴ Cavitand bowl macrocycle **2** contained either benzyl bromide **2a**, or benzyl thiol **2b** moieties at its “rim” positions.



2.1 Cyclotribenzylene Macrocycle

The cyclotribenzylene (CTB) skeleton was first synthesized in 1915 by Mrs. G.M. Robinson, who reported that an identical compound was produced from the acid catalyzed reaction either between formaldehyde and veratrole (1,2-dimethoxy benzene), or of veratryl alcohol (3,4-dimethoxybenzyl alcohol) alone.³⁰⁵ The bowl-like structure of this product **3** (Figure 2.2), was not identified until 1963, when the name cyclotrimeratrylene was adopted for macrocycles of this type.³⁰⁶ The CTB skeleton has been synthesized with various “rim” and “linking” groups.³⁰⁷ We chose to use the CTB **1** with methylene “linking” groups, and benzyl thiols at the “rim” positions to serve as the synthetic handles for the attachment of peptide strands (see Chapter 3).

Figure 2.2: The CTB skeleton has been made using various “rim” (X, Y) and “linking” (Z) groups.

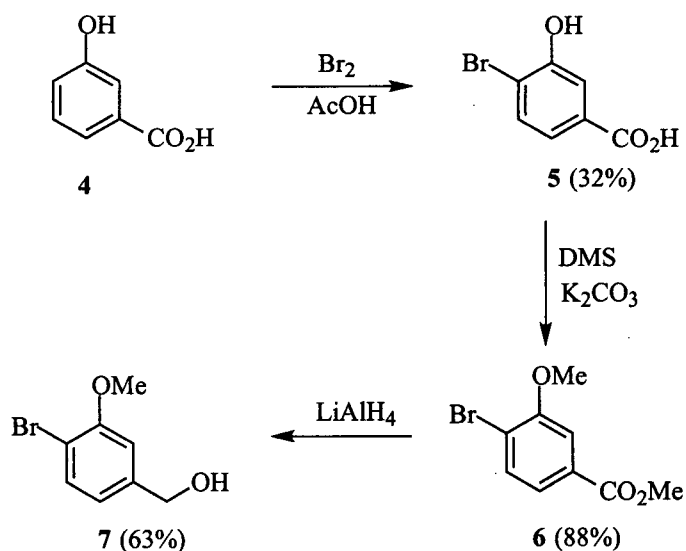


2.1.1 Synthesis of CTB Tris-Benzylthiol **1**

The CTB tris-benzylthiol **1** was first synthesized by Cram.³⁰⁸ The key step in the synthesis of CTB's is the cyclotrimerization reaction, which requires a suitable benzyl alcohol

monomer (see Figure 2.3 for synthesis of monomer). The first step in the synthesis of the monomer used in the synthesis of CTB 1 began with bromination of 3-hydroxybenzoic acid **4**, to yield 4-bromo-3-hydroxybenzoic acid **5**. Methylation of both the hydroxyl and acid functionalities of **5** was achieved using dimethyl sulfate, resulting in methyl 4-bromo-3-methoxybenzoate **6**. Reduction of the ester **6** with lithium aluminum hydride results in the 4-bromo-3-methoxybenzyl alcohol **7**, which was the starting material for the cyclization reaction.

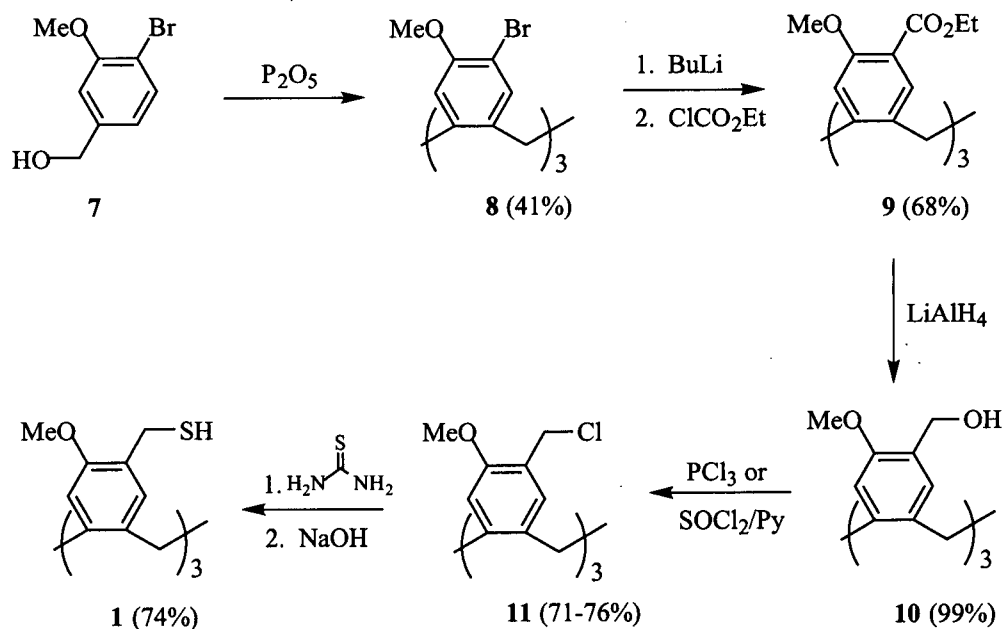
Figure 2.3: Synthesis of 4-bromo-3-methoxybenzyl alcohol **7**, the monomer used for the cyclotrimerization reaction to make the CTB macrocycle. NOTE: The yields reported below are my yields and are comparable to those reported in the literature.



The next step in the synthesis of CTB **1** was the cyclotrimerization of the benzylalcohol monomer **7**, which was achieved using phosphorous pentoxide (Figure 2.4). Treatment of the three aryl-bromide moieties of CTB **8** with butyllithium, followed by the addition of ethyl chloroformate yielded CTB tris-ester **9**. Reduction of CTB tris-ester **9** with

lithium aluminumhydride gave CTB tris-benzylalcohol **10**. The next step in the published synthesis reports the use of a N-chlorosuccinimide/triphenylphosphine mixture to convert CTB tris-benzyl alcohol **10** into CTB tris-benzylchloride **11**. We were unable to reproduce this reaction, but found that either thionyl chloride in the presence of pyridine, or phosphorous trichloride gave the required CTB tris-benzylchloride **11**. Subsequent reaction of CTB tris-benzylchloride **11** (which is highly insoluble) with thiourea, and hydrolysis of the intermediate with NaOH gave the required CTB tris-benzyl thiol **1**.

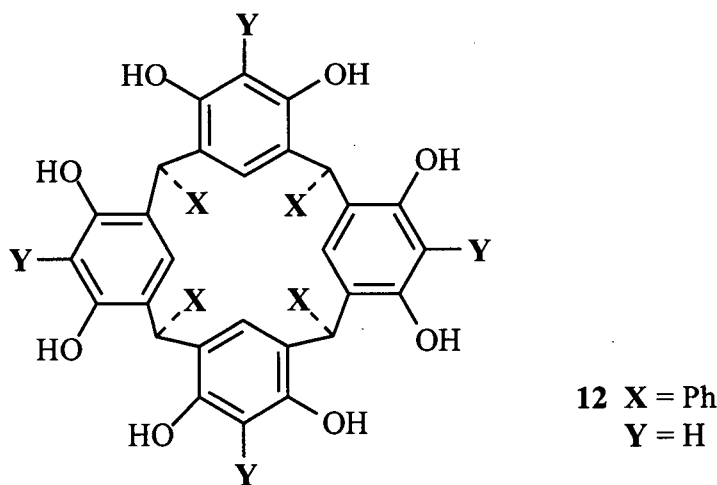
Figure 2.4: Synthesis of CTB tris-benzylthiol **1** from benzylalcohol monomer **7**. NOTE: The yields reported below are my yields, and generally lower than those reported in the literature. The conversion of CTB tris-benzylalcohol **10** to CTB tris-benzylchloride **11** employed a different procedure to that reported (see text for details).



2.2 Cavitand Bowl Macrocycle

In 1872 Alfred von Bayer reported a crystalline product from the acid catalysed reaction of benzaldehyde and resorcinol (1,3-dihydroxybenzene).³⁰⁹ However, it was not until 1968 that the structure was confirmed as the cis-cis-cis isomer **12** (Figure 2.5).³¹⁰ This macrocyclic skeleton, commonly called a resorcin[4]arene, is the thermodynamically most stable product from the reversible condensation reaction. It is possible to vary the “feet” or “rim” moieties of this macrocycle by changing the aldehyde or resorcinol derivative in the acid catalyzed reaction.³⁰⁹

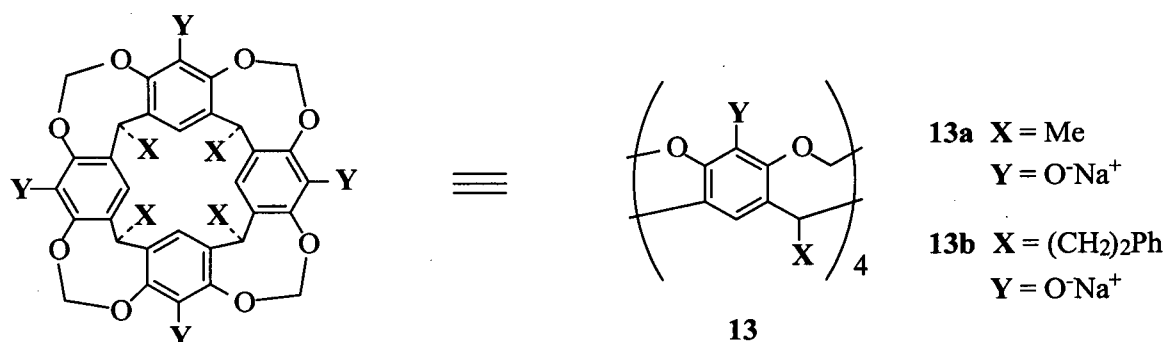
Figure 2.5: The resorcin[4]arene skeleton, which can have various functional groups at the “rim” (Y) and/or “feet” (X) positions. The conformation of this macrocycle exists in equilibrium between the “saddle” (major) and “cone” (minor) conformations.



Cram added methylene bridges between the hydroxyl moieties to yield the rigid bowl-like structure **13** (Figure 2.6), which he called a cavitand (a term used to describe a molecule with an enforced cavity with a convergent binding site). We chose to use a cavitand bowl

with either benzylthiol **2b**, or benzylbromide **2a** groups at the “rim” positions, and methyl groups as the “feet”. The “rim” positions were chosen as synthetic handles for the addition of peptide strands. Methyl “feet” were chosen because cavitand **13a** (a tetra sodium salt) which had methyl “feet”, was shown to exhibit concentration independent NMR spectra in D₂O, consistent with a non-aggregating species, whereas the analogous phenethyl footed cavitand **13b** exhibited concentration dependent NMR spectra, indicating aggregation in water.^{311,312}

Figure 2.6: Cram’s cavitand bowl, for which the “rim” (Y) and “feet” (X) positions can be varied.

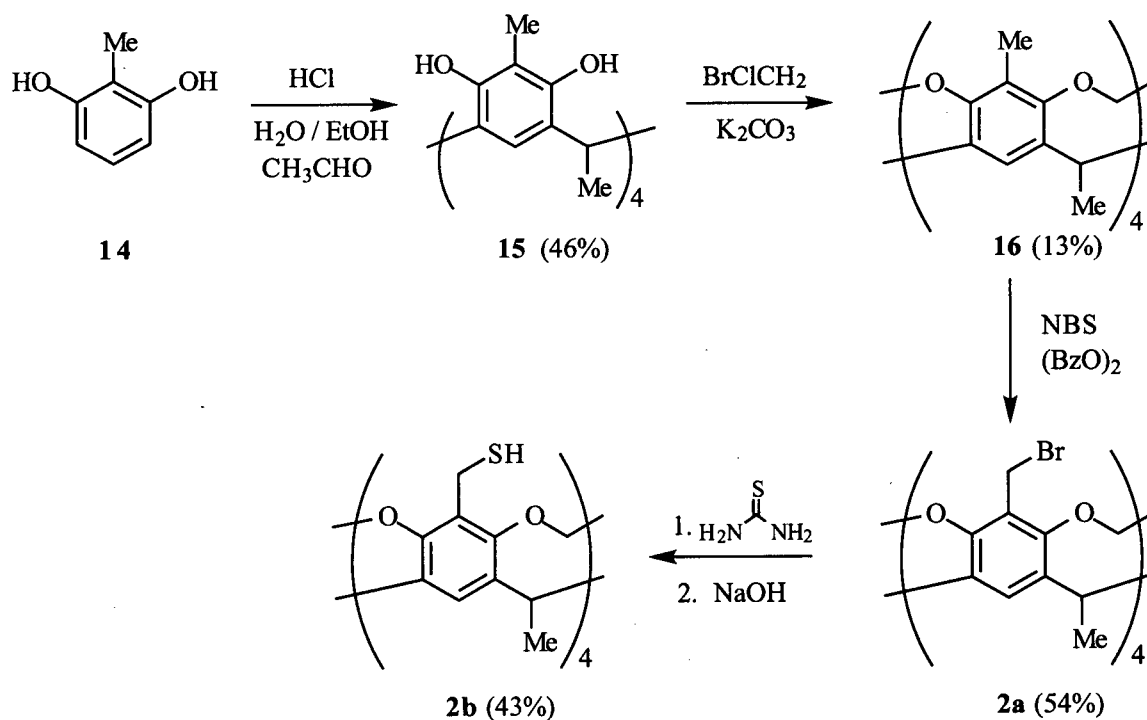


2.2.1 Synthesis of Cavitand Bowls 2a and 2b

Benzylbromide cavitand **2a** was prepared according to the literature method (Figure 2.7).³¹³ The first step in the synthesis was the acid catalyzed macrocyclization reaction between methylresorcinol (1,3-dihydroxy-2-methylbenzene) **14** and acetaldehyde, to produce tetra-methylresorcin[4]arene **15**.³¹⁴ Tetra-methyl cavitand bowl **16** was made by bridging the inter-aryl hydroxyl moieties using bromochloromethane in the presence of potassium carbonate as base.³¹⁴ Tetra-benzylbromide bowl **2a** was then made by treatment of tetra-methyl cavitand bowl **16** with N-bromosuccinimide.³¹³ Finally, tetra-benzylthiol cavitand

bowl **2b** was made via treatment of the tetra-benzylbromide bowl **2a** with thiourea, followed by hydrolysis of the intermediate with aqueous NaOH. Cram also synthesized the benzylthiol cavitand **2b** using a method that parallels the one used in the production of CTB tris-benzylthiol **1**.³¹⁵

Figure 2.7: Synthesis of Cavitand bowls **2a** and **2b**. **NOTE:** Each step of the synthesis was performed only once. Yields reported below are my yields and are comparable to those reported in the literature, except for a lower than expected yield from the bridging reaction to produce tetra-methyl cavitand bowl **16**.



2.3 Experimental

General:

All chemicals were reagent grade (Aldrich or BDH). Tetrahydrofuran (THF) was distilled under N₂ from sodium benzophenone ketyl. Diethylether, dimethylformamide (DMF), dimethylsulfoxide (DMSO) and *N,N*-dimethylacetamide (DMA) were dried over 4 Å molecular sieves. Liquid secondary ionization mass spectra (LSIMS) were recorded on a Kratos Concept IIH32 using various matrices as noted. All ¹H NMR spectra were recorded on a Bruker AC-200E spectrometer at ambient temperature using residual ¹H signals from deuterated solvents as a reference (CDCl₃, 7.24 ppm; DMSO-*d*₆, 2.49 ppm). Silica gel (230-400 mesh, BDH) was used for column chromatography and silica gel glass-backed analytical plates (0.2 mm, Aldrich) were used for thin layer chromatography, with UV detection. All compounds have been fully characterized in the literature. Spectral analyses obtained here agreed with that reported in the literature.

(5): 4-Bromo-3-hydroxybenzoic acid:³¹⁶

Bromine (22 g, 275 mmol) in acetic acid (70 mL) was added dropwise to a stirred suspension of 3-hydroxybenzoic acid (18.2 g, 132 mmol) in glacial acetic acid (180 mL). The reaction mixture was stirred at room temperature for 2 h, during which time the solid went into solution. The reaction mixture was reduced to about 100 mL volume *in vacuo*, and placed in the fridge (4 °C) overnight. The solid that formed was filtered and washed

with cold water. Multiple recrystallizations from boiling water yielded 9.1 g (32 %) of a white solid.

¹H NMR (DMSO-*d*₆, 200 MHz): δ 7.61-6.82 (m, 3H, ArH), 5.86 (s, 1H, ArOH).

(6): Methyl 4-Bromo-3-hydroxybenzoate:³¹⁷

A mixture of dimethyl sulfate (48.90 g, 390 mmol, HIGHLY TOXIC), 4-Bromo-3-hydroxybenzoic acid **5** (21.14 g, 97 mmol), K₂CO₃ (100.70 g, 730 mmol) in acetone (200 mL) was refluxed for 12 h. The reaction mixture was filtered, and the filtrate was washed with acetone and diethyl ether. The combined organic phases were evaporated, and an ice/ammonia solution mixture (100 g/25 mL) was poured onto the residue (to decompose any remaining dimethyl sulfate). The solid was filtered and recrystallized from 40-60 petroleum ether to produce 21.02g (88 %) of a pale yellow solid after drying *in vacuo*.

¹H NMR (CDCl₃, 200 MHz): δ 7.59-6.79 (m, 3H, ArH), 3.86 (s, 3H, ArOCH₃), 2.02 (s, 3H, CO₂CH₃).

(7): 4-Bromo-3-hydroxybenzyl alcohol:³¹⁷

The ester **6** (20.86 g, 85 mmol) in 100 mL dry THF was added dropwise to a suspension of LiAlH₄ (3.24 g, 85 mmol) in THF (100 mL) under a N₂ atmosphere. The reaction mixture was stirred for 3 h at room temperature before the careful addition of ethyl acetate (15 mL) and 2 M HCl (15 mL). The reaction mixture was filtered through celite, and the layers were separated. The aqueous phase was extracted with diethyl ether (2 x 50 mL), and the

combined organic phases were washed with saturated NaCl solution (100 mL), dried over MgSO_4 , and evaporated. The resulting dark brown oil was purified by silica gel column chromatography using 1:1 40-60 petroleum ether: diethyl ether as eluent to produce a pale brown oil 11.64 g (63 %).

^1H NMR (CDCl_3 , 200 MHz): δ 7.43-6.70 (m, 3H, ArH), 4.53 (s, 2H, CH_2O), 3.82 (s, 3H, OCH_3), 2.73 (br s, 1H, OH).

(8): Tris-bromo CTB:

10,15-Dihydro-2,7,12-tribromo-3,8,13-trimethoxy-5H-tribenzo[a,d,g] cyclononene:³¹⁸

The benzyl alcohol 7 (4.33 g, 20 mmol) was added to a stirring mixture of P_2O_5 (8.66 g, 60 mmol) in diethyl ether (125 mL) under a N_2 atmosphere. The mixture was refluxed for 2 d. The ether was evaporated and the residue dissolved in DCM (100 mL) and passed through a short column of silica gel using DCM as eluent. The volume of DCM was reduced to 100 mL before being poured onto 400 mL of diethyl ether. This solution was left for 16 h at 4 °C. The solid was filtered and washed with ice cold 4:1 diethyl ether: DCM before being dried *in vacuo* to produce 1.65 g (41 %) of a white solid.

^1H NMR (CDCl_3 , 200 MHz): δ 7.50 (s, 3H, Ar-H); 6.84 (s, 3H, Ar-H); 4.64 (d, 3H, Ar_2CH_2 , axial H, $J = 14$ Hz); 3.86 (s, 9H, OCH_3); (d, 3H, Ar_2CH_2 , equatorial H, $J = 14$ Hz).

MS (LSIMS, thioglycerol): 594, 596, 598, 600 ($\text{M}+\text{H}$)⁺

(9): CTB tris-ester:

Triethyl-10,15-dihydro-3,8,13-trimethoxy-5H-tribenzo[a,d,g]cyclononene-2,7,12-tricarboxylate:³⁰⁸

The tris-bromo CTB 8 (1.70 g, 2.8 mmol) was dissolved (with warming) in THF (300 mL) under a N₂ atmosphere. This solution was cooled to -78 °C (dry ice / acetone bath) and n-butyllithium (7.9 mL, 1.6 M in hexanes, 13 mmol) was added dropwise, forming a white precipitate. After 15 min, ethyl chloroformate (9.11 g, 84 mmol) was added by syringe, and the reaction mixture was allowed to warm to room temperature (1.5 h). The solvent was evaporated, and the residue partitioned between ethyl acetate (100 mL) and water (100 mL). The aqueous phase was further extracted with ethyl acetate (2 x 100 mL). The organic phases were combined, dried over MgSO₄, and evaporated. Recrystallization of the residue (using chloroform / ethyl acetate) resulted in 1.10 g (68 %) of a white solid.

¹H NMR (CDCl₃, 200 MHz): δ 7.81 (s, 3H, Ar-H); 6.95 (s, 3H, Ar-H); 4.68 (d, 3H, Ar₂CH₂ axial H, J = 13.5 Hz); 4.30 (q, 6H, CO₂CH₂, J = 7 Hz); 3.86 (s, 9 Hz, OCH₃); 3.67 (d, 3H, Ar₂-CH₂ equatorial H, J = 13.5 Hz); 1.22 (t, 9H, CH₃, J = 7 Hz).

MS (LSIMS, thioglycerol): 576 (M+H)⁺

(10): CTB tris-benzylalcohol:

10,15-Dihydro-3,8,13-trimethoxy-5H-tribenzo[a,d,g]cyclononene-2,7,12-trimethanol:³⁰⁸

The CTB tris-ester 9 (658 mg, 1.1 mmol) was dissolved in THF (100 mL, with warming), and added dropwise to a rapidly stirring suspension of LiAlH₄ (418 mg, 11 mmol) in THF (20 mL) under a N₂ atmosphere. The mixture was stirred at room temperature for 1 h. Ethyl

acetate (25 mL) was added dropwise, followed by 0.2 M HCl (10 mL). The reaction mixture was filtered and the solid washed with ethanol (50 mL) and acetone (50 mL). All the organic phases were combined and poured onto 100 mL water. This solution was evaporated to 100 mL volume *in vacuo*; the solution now contained a white solid which was filtered and dried *in vacuo* to yield 488 mg (99 %) of product.

¹H NMR (DMSO-*d*6): δ 7.42 (s, 3H, Ar-H); 6.91 (s, 3H, Ar-H); 4.82 (t, 3H, OH, J = 6 Hz); 4.83 (d, 3H, Ar₂CH₂ axial H, J = 13 Hz); 4.37 (d, 6H, ArCH₂O, J = 6 Hz); 3.73 (s, 9H, OCH₃); 3.59 (d, 3H, Ar₂CH₂ equatorial H, J = 13 Hz)

MS (LSIMS, thioglycerol): 450 (M+H)⁺

(11): CTB tris-benzylchloride:

10,15-Dihydro-2,7,12-trimethoxy-3,8,13-tris-(chloromethyl)-5H-tribenzo[a,d,g]cyclononene:

(Method 1) CTB tris -benzylalcohol **10** (40 mg, 89 μ mol) in DMF (1.5 mL) was added to a solution of thionyl chloride (106 mg, 890 μ mol) in DMF (1.5 mL) at 0 °C. Pyridine (71 mg, 890 μ mol) was added, and the reaction was stirred at 0 °C for 2 h, during which time a white precipitate was seen to form. The solid was filtered, washed with water, acetone and ethyl acetate, and dried overnight under high vacuum at 79 °C to produce 34 mg (76 %) of a white solid.

¹H NMR (nitrobenzene-*d*5): δ 7.32 (s, 3H, Ar-H); 6.72 (s, 3H, Ar-H); 4.41 (d, 3H, Ar₂CH₂ axial H, J = 14 Hz); 4.35 (broad s, 6H, ArCH₂Cl); 3.54 (s, 9H, OCH₃); 3.33 (d, 3H, Ar₂CH₂ equatorial H, J = 14 Hz).

MS (LSIMS, thioglycerol:DMSO): 506, 508 (M+H)⁺

(Method 2) Phosphorous trichloride (62 mg, 450 μ mol) was added to the CTB tris-benzylalcohol **10** (45 mg, 100 μ mol) in DMF (5 mL) under a N₂ atmosphere. The mixture was stirred at room temperature for 3 h, during which time a white precipitate was seen to appear. The solid was filtered, washed with DMF, water and acetone, before being dried *in vacuo* to produce 36 mg (71 %) of a white solid. See above for characterization.

(1): CTB tris-benzylthiol:

10,15-Dihydro-3,8,13-trimethoxy-5H-tribenzo[a,d,g]cyclononene-2,7,12-trimethanethiol:³⁰⁸

A mixture of CTB tris-benzylchloride **11** (29 mg, 57 μ mol), thiourea (14 mg, 190 μ mol) and degassed DMSO (5 mL) was stirred at 65 °C under a nitrogen atmosphere for 3.5 h. At room temperature, the reaction was poured onto 1 M NaOH (15 mL). This solution was stirred for 2 min before being acidified to pH 2 (litmus paper) with 2 M HCl. The resultant solid was filtered, washed with water and dried *in vacuo* at 80 °C for 72 h to produce 21 mg. (74 %) of a white solid.

¹H NMR (DMSO-*d*6): δ 7.42 (s, 3H, Ar-H); 7.00 (s, 3H, Ar-H); 4.75 (d, 3H, Ar₂CH₂ axial H, J = 13 Hz); 3.78 (s, 9H, OCH₃); 3.56 (d, 3H, Ar₂CH₂ equatorial H, J = 13 Hz); 3.56 (broad s, 6H, ArCH₂S).

MS (LSIMS, thioglycerol): 498 (M+H)⁺

(15): Tetra-methyl resorc[4]inarene:

2,5,8,11,14,17,20,23-Octamethylpentacyclo[19.3.1.1^{3,7}.1^{9,13}.1^{25,19}]octacosa-

1(25),3,5,7,(28),9,11,13(27),15,17,19(26),21,23-dodecaen-4,6,10,12,16,18,22,24-octol

Stereoisomer:³¹⁴

2-Methylresocinol **14** (1,3-dihydroxy-2-methylbenzene) (25.0 g, 200 mmol) was dissolved in a mixture of H₂O (75 mL), EtOH (150 mL) and concentrated HCl (75 mL). Nitrogen was bubbled through the mixture (5 min), which was then cooled to 0 °C (ice / water bath) before the dropwise addition of acetaldehyde (8.87 g, 200 mmol). The reaction was stirred at room temperature for a further 2 h. The oily purple solid was filtered using a fine frit, and washed with water, before being dried under high vacuum for 3 d. This produced 14.0 g (46 %) of a pale purple solid.

¹H NMR (DMSO-*d*₆, 200 MHz): δ 8.70 (s, 8H, OH), 7.41 (s, 4H, Ar-H), 4.46 (q, 4H, Ar₂CH, J = 7 Hz), 2.00 (s, 12H, ArCH₃), 1.70 (d, 12H, CH₃CH, J = 7 Hz)

MS (LSIMS, thioglycerol): 648 (M+H)⁺

(16): Tetra-methyl bowl:

1,7,11,15,21,23,25,28-Octamethyl-2,20:3,19-dimetheno-1*H*,21*H*,23*H*,25*H*-

bis[1,3]dioxcino[5,4-*i*:5',4'-*i'*]benzo[1,2-*d*:5,4-*d'*]-bis[1,3]benzodioxocin Stereoisomer:³¹⁴

Over 3 d, tetra-methyl resorcin[4]arene **15** (13.80 g, 23 mmol) was added to a mixture of bromochloromethane (14.84 g, 120 mmol) and K₂CO₃ (31.74 g, 230 mmol) in DMA (350 mL) under a N₂ atmosphere. The mixture was stirred for 2 d at room temperature before the

addition of a second portion of bromochloromethane (14.84 g, 120 mmol). The mixture was then heated at 60 °C for 24 h, before the addition of a further portion of bromochloromethane (14.84 g, 120 mmol). After a further 24 h heating at 60 °C, the solvent was evaporated and the product extracted from the gummy residue using DCM (350 mL). The DCM extract was evaporated to a low volume (35 mL) and filtered through a short plug of silica gel using DCM as eluent. After drying over MgSO₄, the DCM fractions that contained the product were evaporated *in vacuo* to yield 1.90 g (13 %) of a white solid.

¹H NMR (CDCl₃, 200 MHz): δ 7.14 (s, 4H, ArH), 5.92 (d, 4H, H_{out}, J = 7 Hz), 5.20 (q, 4H, CH₃CH, J = 8 Hz), 4.29 (d, 4H, H_{in}, J = 7 Hz), 2.00 (s, 12H, ArCH₃), 1.73 (d, 12H, CH₃CH, J = 8 Hz)

MS (LSIMS, thioglycerol): 648 (M+H)⁺

(2a): Tetra-benzylbromide cavitand bowl:

7,11,15,28-Tetrakis(bromomethyl)-1,21,23,25-tetramethyl-2,20:3,19-dimetheno-1H,21H,23H,25H-bis[1,3]dioxocino[5,4-*i*:5',4',*i'*]benzo[1,2-*d*:5,4-*d'*]bis[1,3]benzodioxocin Stereoisomer:³¹³

N-Bromosuccinimide (401 mg, 2.3 mmol) was added to a mixture of tetra-methyl cavitand bowl **16** (324 mg, 0.5 mmol) and benzoyl peroxide (37 mg, 0.15 mmol) in CCl₄ (15 mL) under N₂ at room temperature. This mixture was refluxed for 24 h. After evaporation of the solvent (which was saved for recycling), the residue was passed through a short silica gel column using CHCl₃ as eluent. Isolation of the product resulted in 260 mg (54 %) of a white solid after drying overnight under vacuum.

¹H NMR (CDCl₃, 200 MHz): δ 7.26 (s, 4H, ArH), 6.04 (d, 4H, H_{out}, J = 7 Hz), 5.02 (q, 4H, Ar₂CH, J = 7 Hz), 4.57 (d, 4H, H_{in}, J = 7 Hz), 4.42 (s 8H, CH₂Br), 1.75 (d, 12H, CH₃, J = 7 Hz)

MS (LSIMS, thioglycerol): 964 (M+H)⁺

(2b): Tetra-benzylthiol cavitand bowl:

1,21,23,25-2,20:3,19-dimetheno-1*H*,21*H*,23*H*,25*H*-bis-[1,3]dioxocino[5,4-*i*:5',4'-*i'*]benzo[1,2-*d*:5',4'-*d'*]bis[1,3]benzodioxocin-7,11,15,28-tetramethanethiol

Stereoisomer³¹⁵

A mixture of tetra-benzylbromide cavitand bowl **2a** (148 mg, 0.15 mmol) and thiourea (57 mg, 0.75 mmol) in DMF (25 mL) were stirred at room temperature for 2 h. The reaction mixture was then poured onto aqueous 1 M NaOH (150 mL). This solution was stirred for 10 min before being acidified to pH 2 (litmus paper) with concentrated HCl. The resultant precipitate was filtered, and washed with water, before being dried overnight at 79 °C *in vacuo*. The resultant brownish-yellow solid was filtered through a short column of silica gel using DCM as eluent, to produce 50 mg (43 %) of a white solid after vacuum drying.

¹H NMR (CDCl₃, 200 MHz): δ 7.18 (s, 4H, ArH), 5.95 (d, 4H, H_{out}, J = 7 Hz), 4.98 (q, 4H, Ar₂CH, J = 7 Hz), 4.49 (d, 4H, H_{in}, J = 7 Hz), 3.57 (d, 8H, CH₂S, J = 8 Hz), 1.89 (t, 4H, SH, J = 8 Hz), 1.73 (d, 12H, CH₃C, J = 7 Hz).

MS (LSIMS, thioglycerol): 776 (M+H)⁺

Design and Synthesis of TASP

3.0 Introduction

Careful design of the TASP is crucial for its success: For our three- and four-helix bundle TASP

s to fold correctly in water, the peptide must contain a specific pattern of amino acids in its sequence (Section 3.1).

The key step in the synthesis of our TASP

s is linking the templates to the unprotected peptide stands. Many of the amino acid side chains contain reactive functional groups, however, peptide strands with protected side chains are notorious for their insolubility: Therefore, we needed to find a coupling method that was both efficient and selective. Reagents for crosslinking proteins to various functionalities, including peptides and other proteins, have been known for a long time,³¹⁹⁻³²¹ and some of these methods have been used for covalently linking peptides to templates (Section 3.2). We chose to synthesize our TASPs using either disulfide or thioether bonds to link the template to the peptides (Section 3.3).

3.1 Design of CTB and Cavitand Based TASP

We decided to link both the CTB and cavitand bowl templates to peptide strands that contain a cysteine residue in their sequence. The thiol moiety of this cysteine residue was coupled to a template that possessed either a benzyl bromide group, to create a thioether linkage (Section 3.3.2), or a benzyl thiol moiety, to form a disulfide bond (Section 3.3.3). Theoretically, a cysteine residue can be placed anywhere in a peptide sequence: A peptide that is linked via its polyamide backbone is restricted to being linked via its N- or C-terminus only.

3.1.1 Design and Nomenclature of Peptides and TASP

The templates were chosen because they are both rigid macrocycles, which is an important characteristic for pre-organizing the peptide strands. Also, the functional groups on the rim of these macrocycles are 8-14 Å apart, which will result in the appropriate inter-helical distance found for three- and four-helix bundles in nature.⁵ Additionally, both macrocycles have an enforced cavity with potential for binding apolar substrates,³⁰³ which may be important if these TASP are to be used as substrate binders, catalysts, or drug delivery devices. Finally, CTB 1 is a mixture of two enantiomers (designated as the M- and P-isomers using IUPAC nomenclature – see Chapter 2),³⁰⁴ which upon addition of three chiral peptide strands would form two diastereomers, which are in principle separable, and may manifest different physical properties. Indeed, subtle interactions between helices

could potentially be probed via analysis of such diastereomers. Also, before we can obtain a crystal structure of the CTB TASP it would be essential to first separate the diastereomers.

The peptides were all designed around a sequence that was designed to fold into an amphiphilic α -helix; the template-enhanced bundling of these helices leads to the desired folding motif (Figure 3.1 shows helical wheel diagrams for the proposed three- and four-helix bundle TASPs). The specific sequences and shorthand of the peptides and TASPs used in this thesis are detailed in Figure 3.2 (the code is rather complicated because there are many parts of the TASP that were varied).

The individual helices used in our TASPs incorporate many of the favourable design features introduced in Chapter 1. The amino acids chosen for the TASPs had a high propensity for being in an α -helix. The pattern of the peptide sequence results in non-polar (usually leucine) residues occupying one face of the helices, resulting in inter-helical bundling via the hydrophobic effect in water. The hydrophilic face of the helix contains lysine and glutamic acid residues which reside on the exterior of the helix bundle, thus imparting water solubility. The $i, i+4$ spacing of these oppositely charged residues is optimal for intra-helix stabilization by salt bridges. The charge distribution within the helix helps to stabilize the helix macrodipole, for example, the negatively charged glutamic acid residues (when compared to the positively charged lysine residues) tend to be towards the electropositive N-terminus of the peptide strand. Favourable inter-helical salt bridges, such as those designed into coiled coils,²¹¹ are yet to be formally included in our model, but there is a possible E to K inter-helix salt bridge (see Chapter 4). A glycine residue is included as a

C-cap for unattached backbone hydrogen bonding moieties (to diminish fraying of the helices). A cysteine residue (usually towards the N-terminal end of the peptide) is the synthetic handle for coupling to the template, either via a disulfide or thioether bond.

The “pre-cysteine” N-terminal three residue sequences of peptides **H-EGG/N1/L** and **EGG/N1/L** (see Figure 3.2) aid in separation of the diastomeric CTB TASP_s (see Chapter 4). All peptides were amidated at their C-terminus and acetylated at their N-terminus to prevent the termini from being charged and thereby interacting unfavourably with the helix macrodipole³²² (the exception was the peptide **H-EGG/N1/L** which had a free amine at its N-terminus).

Figure 3.1: Helix web diagrams of: (a) Three- and (b) Four-helix bundles, for the sequence XEKLLKELKELLEKG, where X is linked to the template. The helices are viewed down the helical axis from the N- to C-termini. The thick circles represent hydrophilic amino acids.

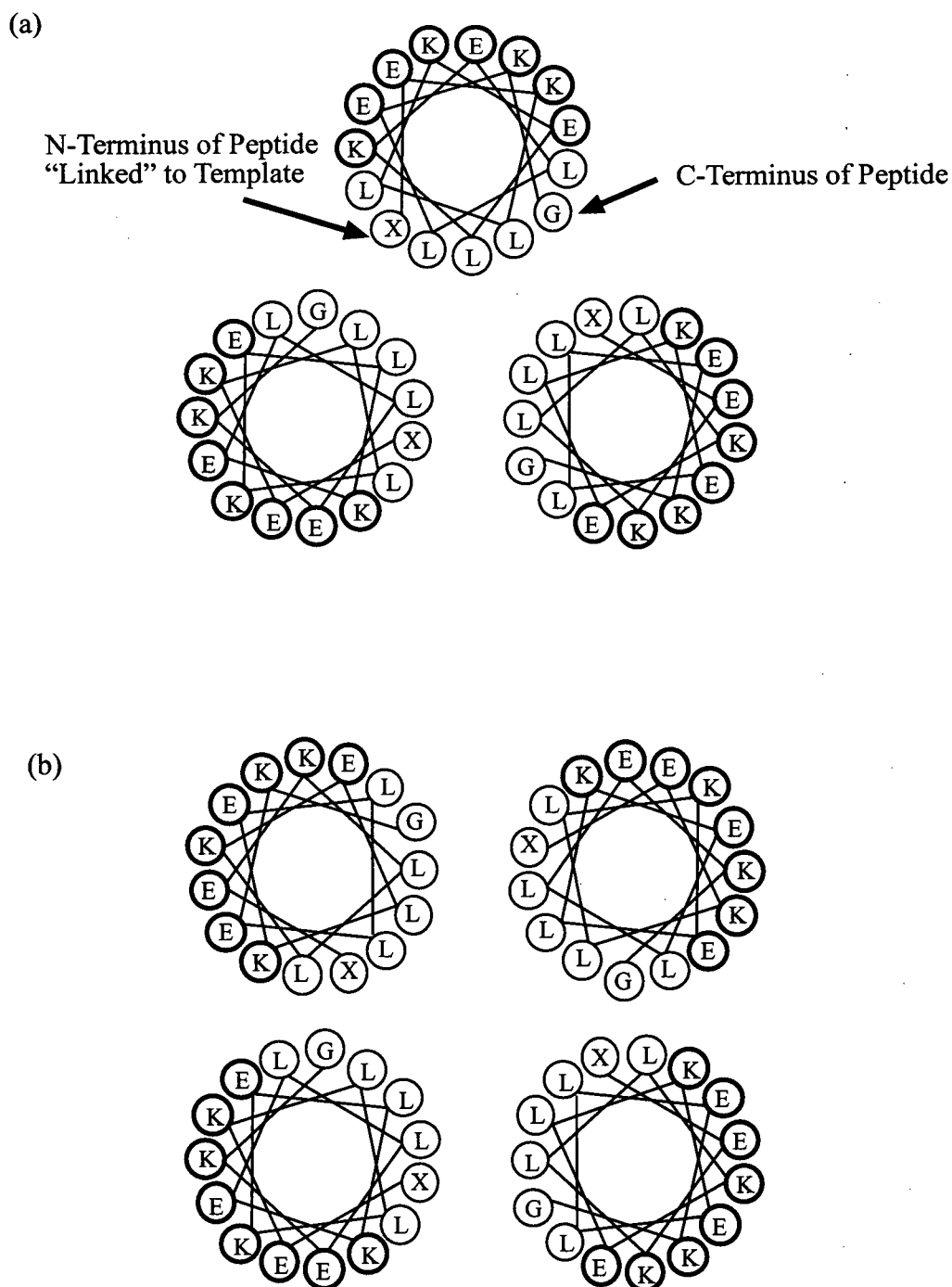


Figure 3.2: Codes and identities of peptides used in this thesis. All peptide C-termini were amidated, and N-termini were acetylated (the one exception of a peptide with a free N-terminus is designated by “H-”).

Code	Peptide Sequence
N1/L	<u>CEKLLKELKELLEKG</u>
H-EGG/N1/L	<u>H-EGGCEKLLKELKELLEKG</u>
EGG/N1/L	<u>EGGCEKLLKELKELLEKG</u>
EG/N1/L	<u>EGCEKLLKELKELLEKG</u>
EEG/N1/L	<u>EEGCEKLLKELKELLEKG</u>
N2/L	<u>CEELLKKLEELLKKG</u>
N2/AAAAA	<u>CEEAAKKAEEAAKKG</u>
N2-GG/L	<u>CGGEELLKKLEELLKKG</u>
N2-GGG/L	<u>CGGGEELLKKLEELLKKG</u>
N2-GGG/V	<u>CGGGEELLKKVEELLKKG</u>
N2-GGG/I	<u>CGGGEELLKKIEELLKKG</u>
N2-GGG/Nle	<u>CGGGEELLKKNleEELLKKG</u>
C2-GGG/L	<u>GEELLKKLEELLKKGGGC</u>

Code Key:

- (i) The letter C or N in the code refers to whether the cysteine residue of the peptide was located at the C- or N- terminus.

(ii) The number 1 or 2 following the C or N refers to the patterning of the hydrophilic E and K amino acids. NOTE: The “/L” refers to the “default” all leucine residues at the hydrophobic positions in the peptide sequence (see note “vi” below for more details).

e.g.

N1/L = CEKLLKELKELLEKG The E and K hydrophilic residues are “unpaired”

N2/L = CEELLKKLEELLKKG The E and K hydrophilic residues are “paired”

(iii) In the case of a peptide with its cysteine residue activated by an S-pyridyl group (see Section 3.3.3.2), the designation (Spy) was added onto the code.

e.g.

N1(Spy)/L = C(Spy)EKLLKELKELLEKG

(iv) Peptides with glycine spacers between the cysteine and the “helix” sequence (see Chapter 5) are given the letter(s) G according to their number.

e.g.

N2-GGG/L = CGGGEELLKKLEELLKKG

(v) Peptides with “pre-cysteine” amino acids are designated with this sequence before the “helix” part of the code.

e.g.

H-EGG/N1/L = H-EGGCEKLLKELKELLEKG

(vi) Changes from the “default” central lysine residue(s) in the helix are designated by the residue(s) that were substituted. These amino acids were either V, I, Nle, or five As.

e.g.

N2-GGG/V = CGGGEELLKKYEELLKKG

N2/AAAAA = CEEAAKKAEEAAKKG

(vii) When attached to a template the designation **CTB** or **Bowl** preceded the code, this also indicates that the template contained three or four peptides, respectively.

e.g.

CTB/H-EGG/N1/L = Three H-EGG/N1/L peptides linked to the CTB template via disulfide bonds.

(viii) Finally an “e” is added when the TASP linker is a thioether bond and NOT a disulfide bond.

e.g.

Bowl/e/N2/L = Cavitand bowl linked to four N2/L peptides via a thioether bond.

3.2 Literature Methods for Linking Peptides to Templates

Some of the earliest reported TASPs were constructed by coupling protected peptide strands or amino acids to templates via backbone peptide bonds (Section 3.2.1). In the last decade a variety of chemoselective reactions have been used for the coupling of unprotected peptide strands and templates³²³ (Section 3.2.2). The field of chemoselective ligation, where peptide fragments (often containing non-natural groups) are spliced into, or added onto proteins is currently of great interest in the field of protein engineering.^{299,300,323,324}

3.2.1 TASPs Constructed using Peptides and Amino Acids with Protected Side Chains

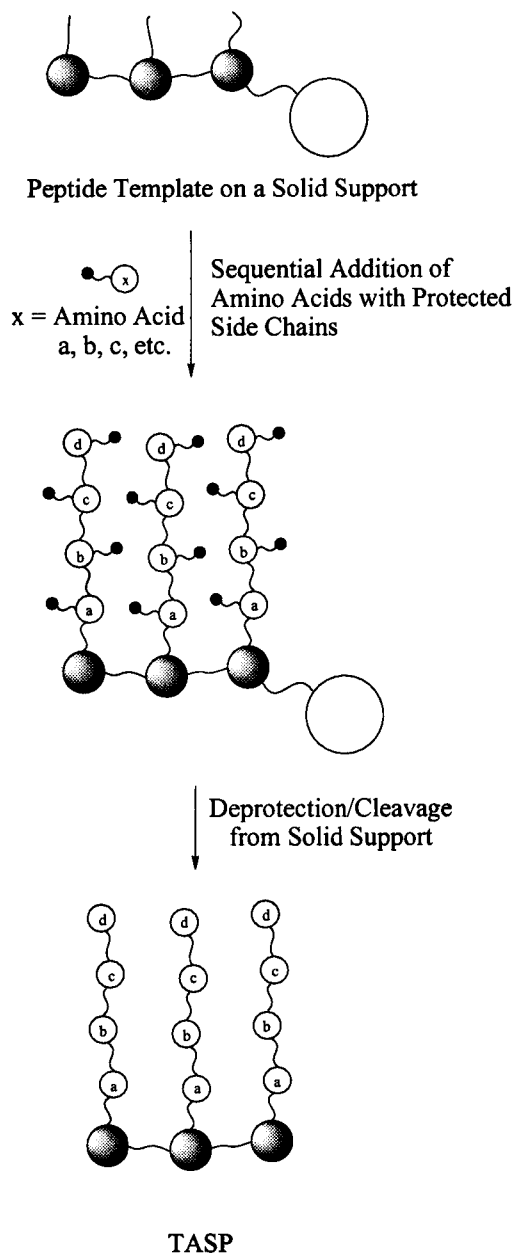
Mutter's original peptide strand TASP (Section 1.4.4.1) was synthesized using a method rooted in solid phase peptide synthesis.²⁵² The lysine rich peptide-template strand was attached to the solid support, and the protein was constructed by sequential addition of amino acids (Figure 3.3a). The TASP was isolated after side chain deprotection, and cleavage of the template from the solid support. Although unlimited in the amino acids it could employ, the major drawback of this method was that it is essentially impossible to purify the desired TASP from the multitude of closely related impurities (*i.e* due to the presence of multiple "deletion" peptides on the TASP).³²⁵

Sasaki and Kaiser used a protected peptide strand in the construction of their porphyrin based four-helix bundle TASP (introduced in Section 1.4.4.2).^{294,295} They reacted the free amino terminus of a protected peptide with activated carboxyl groups on the template, to give a peptide amide bond (Figure 3.3b). The protected TASP required fairly extreme solvent conditions (TFA, hot DMF, or hot DMSO) both to isolate it, and for the subsequent deprotection reaction. Both Degrado²⁶³ and Nishino²⁶⁴ used similar methodology in the construction of their porphyrin and pseudopeptide TASPs (introduced in Sections 1.4.4.2 and 1.4.4.3).

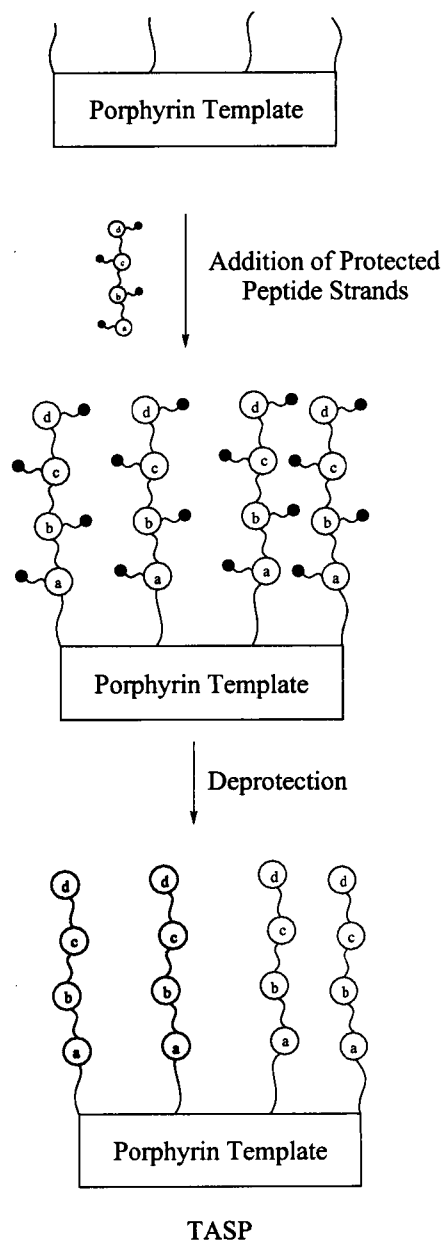
Finally, Inoue used amino acids with protected side chains, which he added onto his cyclotriphosphazene TASP²⁶⁸ (introduced in Section 1.4.4.3) using polymerization techniques. This method resulted in a TASP with peptides of an indeterminate length, made up of the same amino acid.

Figure 3.3: Schematic preparation of TASP's using protected amino acids and peptides. (a) Mutter's sequential addition of amino acids onto a resin-bound peptide template using standard solid state peptide synthesis methods. (b) Sasaki and Kaiser's addition of peptide strands with protected side chains and a free amino terminus, to a porphyrin template containing activated carboxylic acid groups on the template, to form an amide bond.

(a) Mutter



(b) Sasaki and Kaiser



3.2.2 Synthesis of TASP's using Unprotected Peptides

Many of the strategies for chemoselectively linking unprotected peptides to templates have used thioether bonds (Section 3.2.2.1). Thioester, oxime, amine, and amide bonds have also been used (Section 3.2.2.2), however, some of these methods are incompatible with certain reactive side chains (which have to be protected or omitted), or are restricted to coupling the terminus of the peptide backbone to the template.

3.2.2.1 The Thioether Bond

This template to peptide bond is usually formed from the reaction between a thiol nucleophile and an alkyl bromide or chloride. Often, the side chain of a cysteine residue in the peptide is the source of the thiol.

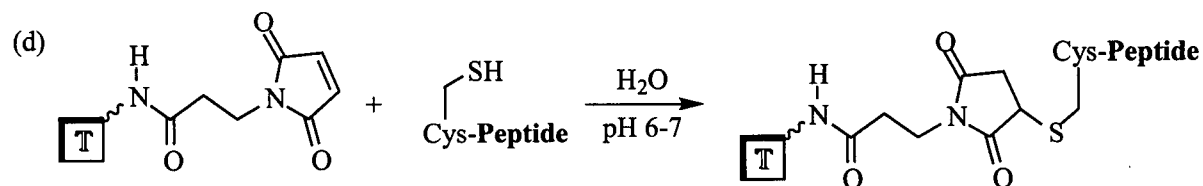
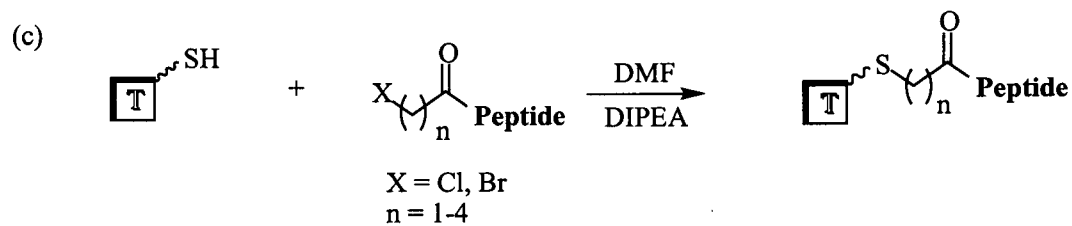
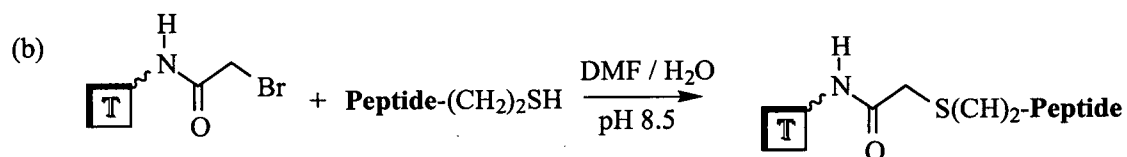
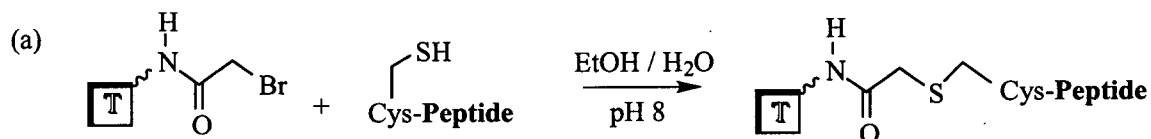
DeGrado porphyrin template³²⁶ contained bromoacetyl groups that were reacted with cysteine-containing peptides (Figure 3.4a).³²⁷ Fairlie also used bromoacetyl groups on all his aromatic templates (introduced in Section 1.4.4.3), but his peptides contained an thiol group on the C-terminus of the peptide backbone (Figure 3.4b).

Our research group has used aryl thiol moieties on the cavitand bowl template (introduced in Section 1.4.4.3), and peptides that contained alkyl bromide or chloride groups at the N-terminus of the peptide (Figure 3.4c).^{269,328} Similarly, both Futakai,³²⁹ and

Haehnel²⁵⁷ have also used bromoacetyl modified peptide strands in the construction of their TASPs, the template was a peptide strand that contained four cysteine residues.

A thioether linked cyclic template TASP was also made by Tuchscherer,³³⁰ who used maleimide groups on the template to form thioether bonds with cysteine-containing peptides (Figure 3.4d). This strategy was also employed by Haehnel and Willner to link a cysteine-containing TASP to a gold surface that had pendent maleimido groups.²⁵⁶

Figure 3.4: TASPs constructed using thioether bonds ("T" represents the template): (a) DeGrado's porphyrin TASP. (b) Fairlie's various aromatic TASPs. (c) Sherman's cavitand bowl TASP, Futakai's, and Haehnel's peptide strand TASPs. (d) Tuchscherer's peptide strand TASP.



3.2.2.2 Thioesters, Oximes, Amides, and Amines

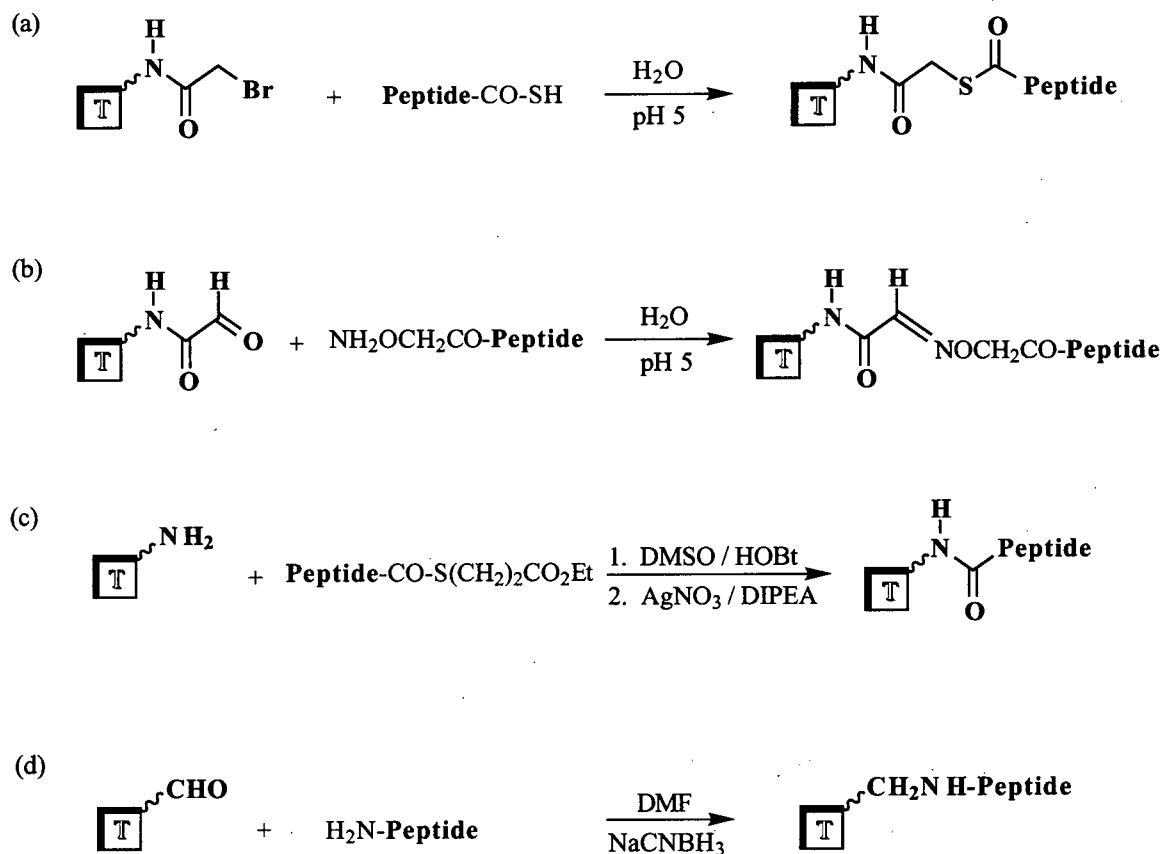
The thioester bond was used by Kent for attaching peptides which had a thioacid at their C-terminus, to a peptide template containing bromoacetylated lysine residues (Figure 3.5a).³³¹ Kent had originally developed this chemoselective ligation method for linking large peptide fragments together via a thioester bond.³³² Unfortunately, the thioester bond is somewhat unstable in basic conditions.

Both Rose,³³³ and Tuchscherer³³⁴ constructed TASPs via an oxime bond (Figure 3.5b). They both used a cyclic peptide strand as template, which contained lysine residues that had side chains that were derivatized with aldehyde moieties. Addition of peptide strands whose N-terminus contained an *O*-alkylhydroxylamine group resulted in an oxime bond. Unfortunately, this functional group is only stable below pH 7.

Futakai used peptides with their C-terminus activated with a thioester derivative, which acted as a leaving group.³³⁵ Addition of these activated peptides to a lysines rich peptide strand template (Figure 3.5c), resulted in an amide peptide bonded TASP.

Finally, Sasaki used the technique of reductive amination to make a TASP via an amine bond between the macrocycle and peptides (see Section 1.4.4.3).²⁶⁶ The N-terminal amino group of a peptide and one of the aldehyde groups on the template macrocycle are in equilibrium with an imine (plus water), a functional group which is reduced (irreversibly) to the amine in the presence of sodium cyanoborohydride (Figure 3.5d).

Figure 3.5: TASPs constructed using unprotected peptide strands with linkers other than thioethers (“T” represents the template): (a) Kent’s thioester linked peptide strand TASP. (b) Rose’s, and Tuchscherer’s oxime linked peptide strand TASPs. (c) Futakai’s amide linked peptide strand TASP. (d) Sasaki’s amine linked macrocyclic TASP.



3.3 Synthesis of CTB and Cavitand Bowl TASP

This section deals with the synthesis of all the TASP used in this thesis. The templates (see Chapter 2), and peptides (Section 3.3.1) were synthesized using published methodology. The remaining challenge was to synthesize our TASP using methods available in the literature. Initially we decided to make TASP via thioether bonds, using a cysteine residue as our thiol source, and an alkyl bromide on our template (Section 3.3.2). A better method for TASP synthesis was by using disulfide bonds, formed between the cysteine residue and an alkyl thiol on the template (Section 3.3.3); this was the method by which most of the TASP in this thesis were made.

3.3.1 Synthesis of Peptides

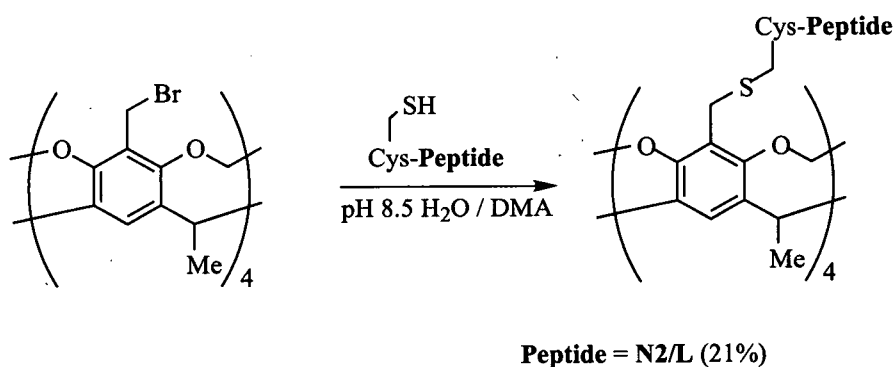
The peptides were synthesized using standard solid phase methods on an automatic peptide synthesizer using Fmoc chemistry (see experimental Section 3.4). Purification of the crude peptides was achieved using HPLC.

3.3.2 Cavitand Bowl TASP Synthesized via a Thioether Bond

This method employed the benzylbromide bowl **2a** and a cysteine containing peptide (Figure 3.6). The solvent system for this reaction was a 4:1 mixture of DMA:aqueous buffer; this solvent ratio was required to keep the cavitand bowl in solution. The optimum pH for this reaction was found to be 8.5. Below pH 8 the reaction proceeded very slowly, and above pH 9 formation of the peptide disulfide homo-dimer was a major side reaction.

Rigorous degassing of the reaction did reduce the rate of homo-dimer production. The template to peptide reaction did not work for peptides **EGG/N1/L**, **H-EGG/N1/L**, **EG/N1/L**, or **EEG/N1/L**. It only worked for peptide **N2/L**, which has a cysteine residue directly at the terminus of the peptide. If the cysteine was not at the terminus of the peptide strand no TASP product was observed, possibly because the reaction between the benzyl bromide group of the cavitand bowl **2a**, and the cysteine residue of the peptide is sterically unfavourable.

Figure 3.6: The reaction of benzyl bromide bowl **2a** and a cysteine containing peptide. The formation of peptide disulfide homo-dimer was a significant side reaction.

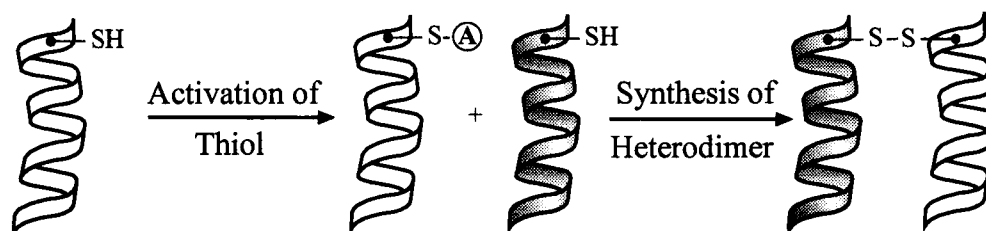


3.3.3 TASP's Synthesized via a Disulfide Bond

In a protein, a pair of cysteine residues can come together to form a disulfide bond.⁵ This oxidation reaction between two thiol moieties in a peptide can be driven by a variety of reagents (e.g. aqueous buffer alone,³³⁶ or DMSO^{337,338}), but this approach is only useful if a homo-dimer disulfide-linked peptide is desired. Normally when a mixture of cysteine-containing peptides are exposed to oxidizing conditions, then a mixture of disulfide linked

The most expedient way to form hetero-dimeric disulfide bonds is to first “activate” one thiol of the pair, and then expose this “reactive” species to its unactivated partner^{343,344} (Figure 3.8). This strategy for the synthesis of hetero-dimeric disulfide bonds has been used for linking prosthetic groups to proteins,^{320,345-349} in the construction of peptide heterodimers,²⁰³ and in the assembly of multiple peptide units.^{350,351} One reagent that has been used as a thiol activating agent is 2,2'-dipyridyl disulfide (DPDS, also known as 2,2'-dithiodipyridine or Aldrichthiol™). Therefore there were two approaches to forming the disulfide linked TASPs: The first was to activate the benzylthiol moieties on our template, and then expose these activated template thiols to the cysteine-containing peptide (Section 3.3.3.1); the second approach was to activate the cysteine residue on the peptide, and then expose this activated peptide thiol to the template (Section 3.3.3.2).

Figure 3.8: Stepwise synthesis of hetero-dimeric disulfide bonds by first activating one thiol of the pair.

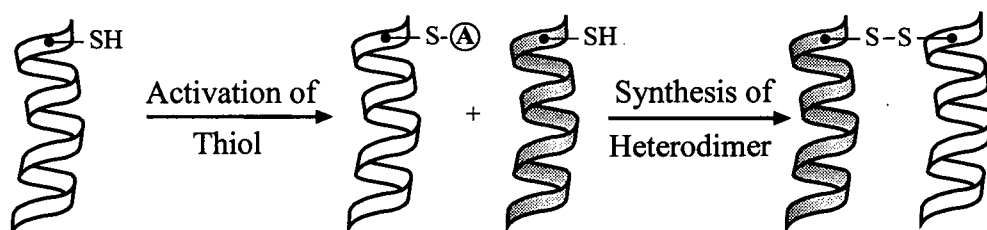


3.3.3.1 Activated Template Thiols plus a Cysteine-Containing Peptide

In this approach, we synthesized the bowl with activated thiol moieties, then added cysteine-containing peptides (Figure 3.9). The advantage of this approach is that once the

The most expedient way to form hetero-dimeric disulfide bonds is to first “activate” one thiol of the pair, and then expose this “reactive” species to its unactivated partner^{343,344} (Figure 3.8). This strategy for the synthesis of hetero-dimeric disulfide bonds has been used for linking prosthetic groups to proteins,^{320,345-349} in the construction of peptide heterodimers,²⁰³ and in the assembly of multiple peptide units.^{350,351} One reagent that has been used as a thiol activating agent is 2,2'-dipyridyl disulfide (DPDS, also known as 2,2'-dithiodipyridine or Aldrichthiol™). Therefore there were two approaches to forming the disulfide linked TASPs: The first was to activate the benzylthiol moieties on our template, and then expose these activated template thiols to the cysteine-containing peptide (Section 3.3.3.1); the second approach was to activate the cysteine residue on the peptide, and then expose this activated peptide thiol to the template (Section 3.3.3.2).

Figure 3.8: Stepwise synthesis of hetero-dimeric disulfide bonds by first activating one thiol of the pair.



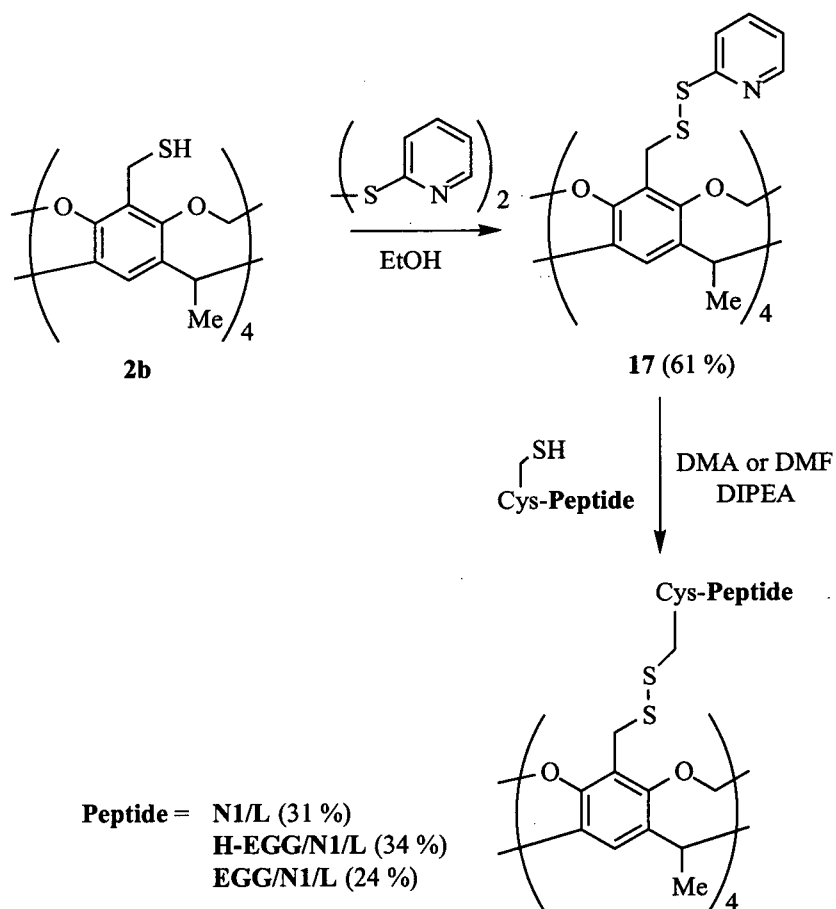
3.3.3.1 Activated Template Thiols plus a Cysteine-Containing Peptide

In this approach, we synthesized the bowl with activated thiol moieties, then added cysteine-containing peptides (Figure 3.9). The advantage of this approach is that once the

activated bowl is made, the synthesis of TASP_s would simply require the addition of peptides. Thus, we would have a method for the rapid synthesis of TASP_s.

The benzyl thiol cavitand bowl **2b** was activated with an S-pyridyl group, by treating it with DPDS (Figure 3.9). Addition of cysteine containing peptides to the activated cavitand bowl **17**, in the presence of the base *N,N*-diisopropylethylamine (DIPEA), resulted in the desired TASP (Figure 3.9). One of the major drawbacks of this reaction was the formation of the homo-dimeric disulfide-bonded peptides as a competing side reaction. Thorough degassing of the reaction system alleviated, but did not eliminate this problem.

Figure 3.9: Synthesis of TASPs via activated benzyl thiol groups on a template (cavitand bowl **2b** was used for these model reactions), and subsequent reaction with a cysteine containing peptide.

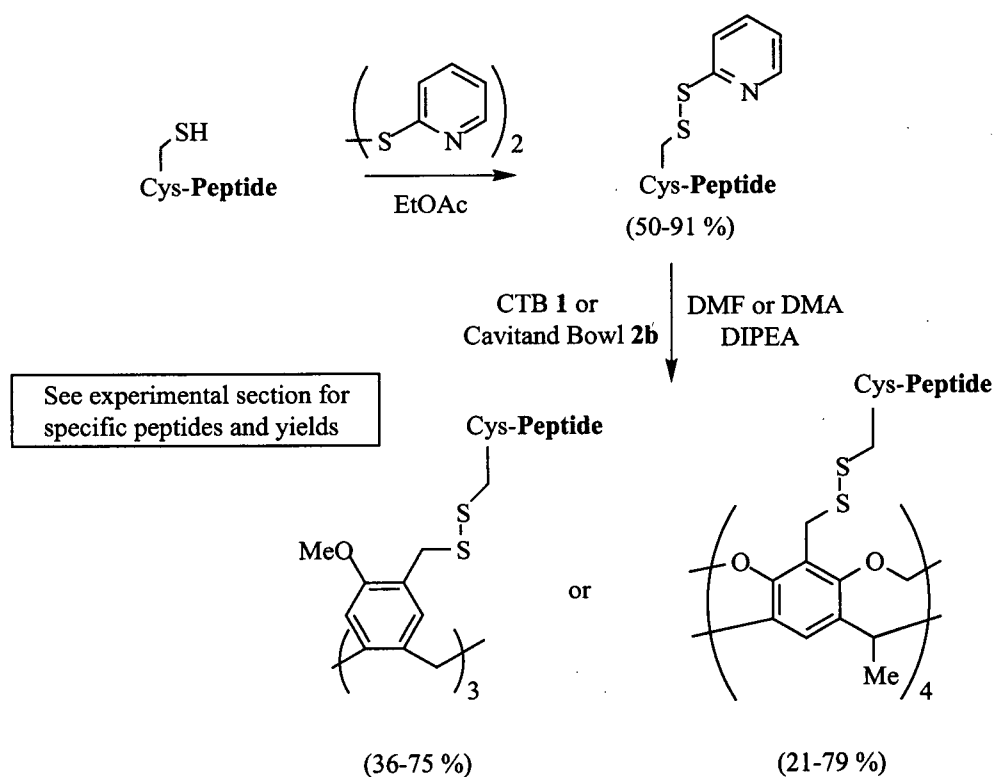


3.3.3.2 Activated Cysteine-Containing Peptide plus Benzyl Thiol Template

In this approach, the peptide's cysteine residue was first activated by an S-pyridyl group, and then coupled to the benzyl thiol template (Figure 3.10). This method prevents the formation of peptide disulfide homo-dimers. Homo-dimers of the template were never observed, even when the templates were exposed to each other in the absence of peptide.

The cysteine containing peptide was treated with DPDS, and this activated peptide was then reacted with the benzyl thiol template (CTB **1**, or cavitand bowl **2b**) in the presence of diisopropylethylamine base. As expected there was no peptide homo-dimer formed, and unreacted peptide starting material was recovered during the HPLC purification of the TASP. Peptides **H-EGG/N1(Spy)/L**, **EGG/N1(Spy)/L**, and all peptides with the cysteine at the terminus resulted in the desired cavitand bowl TASPs. However, peptides **EG/N1(Spy)/L** and **EEG/N1(Spy)/L** did not give any of the desired bowl TASP. This is possibly because their pre-helix "dangler" sequences were more sterically demanding with only one glycine "spacer" between the cysteine and the bulky glutamic acid of the "dangler" (as opposed to peptides **H-EGG/N1(Spy)/L** and **EGG/N1(Spy)/L**, which have two glycine residues as "spacers").

Figure 3.10: Synthesis of TASPs via disulfide bonds, made by first activating the peptide with an S-pyridyl group, and subsequent addition of this activated peptide to the benzyl thiol containing template (either CTB 1 or cavitand bowl 2b).



3.4 Conclusions for Chapter 3

Of the methods used for the synthesis of our TASP_s, the major problem associated with the use of a “free” cysteine on our peptides was the side reaction that resulted in disulfide linked peptide homo-dimers (Sections 3.3.2 and 3.3.3.1). Although performing the reactions under an inert atmosphere reduced the rate of this side reaction, it seemed more practical to avoid these “rigorous” experimental procedures. Therefore, all of the disulfide linked TASP_s in this thesis were made by first activating the thiol group of the peptides cysteine residues with an “S-pyridyl” group, and then reacting these activated peptides with the thiol-containing template (Section 3.3.3.2).

3.4 Experimental

General

Chemicals (Aldrich or BDH) used for the synthesis were reagent grade except dimethylformamide (DMF) and dimethylacetamide (DMA) which were dried over 4 Å molecular sieves and degassed by bubbling dry N₂ through them for 5 min before use. CTB **1**, and cavitand bowls **2a** and **2b** were prepared as reported in Chapter 2. Peptides were prepared on a 0.25 mmol scale using an Applied Biosystems 431A peptide synthesizer using FastMoc™ chemistry with peptide synthesis grade (Advanced Chemtech, Aldrich, or Richelieu Biotechnologies) reagents. The peptides and TASP s were purified using preparative scale C18 reverse phase HPLC (Perkin Elmer), using gradients of water (containing 0.1% TFA) and HPLC grade acetonitrile (containing 0.05% TFA); 229 nm was the wavelength monitored for product elution. Peptides were characterized by their correct mass using LSIMS performed on a Kratos Concept I IH32. TASP s were characterized using electrospray MS (ESMS) run on a Perkin Elmer SCIEX API 300 LC/MS/MS system. All reported molecular weights were within ± 3 Da. of the calculated values. ¹H NMR spectra were recorded on a Bruker AC-200E spectrometer at ambient temperature using residual ¹H signals from deuterated solvents as a reference (CDCl₃, 7.24 ppm). Microanalyses were performed on a Carlo-Erba CHN elemental analyzer, model 1106.

Tetra-pyridyl activated cavitand bowl:

7,11,15,28-Tetrakis(2-pyridyldisulfidomethyl)-1,21,23,25-tetramethyl-2,20:3,19-dimetheno-1*H*,21*H*,23*H*,25*H*-bis[1,3]dioxocino[5,4-*i*:5',4',*i*']benzo[1,2-*d*:5,4-*d'*]bis[1,3]benzodioxocin Stereoisomer:

The tetra-benzyl thiol bowl **2b** (50 mg, 64 μ mol) was dissolved in ethyl acetate (15 mL), and the solution degassed by bubbling nitrogen through it for 5 min. 2,2'-Dipyridyl disulfide (142 mg, 640 μ mol) was then added, and stirring was continued under a nitrogen atmosphere for 3 h. The solvent was removed *in vacuo*, and the residue was purified by silica gel column chromatography, using 1:1 ethyl acetate:hexanes as eluent. 44 mg (61 %) of product was isolated as a glassy, colourless solid.

¹H NMR (CDCl₃, 200 MHz): δ 8.38-7.09 (m, 20 H, Ar-H, Py-Hs); 5.94 (d, 4 H, H_{out}, J = 7 Hz); 4.99 (q, 4 H, Ar₂CH, J = 8 Hz); 4.50 (d, 4H, H_{in}, J = 7 Hz); 4.00 (s, 8 H, ArCH₂S); 1.76 (d, 12 H, CH₃, J = 8 Hz)

MS (LSIMS⁺, thioglycerol): 1212 (M+H)⁺

Analysis: Calculated for C₆₀H₅₂N₄O₈S₈.H₂O; C. 59.38, H. 4.48, N. 4.62: Found; C. 59.40, H. 4.70, N. 4.75.

Syntheses of Peptides:

Peptides were synthesized on Rink resin (which results in a C-terminal amide upon cleavage) using standard FastMoc™ procedures.³⁵² The N-termini of resin bound peptides (except **H-EGG/N1/L** – see text) were acetylated (after first deprotecting the N-terminus) using a standard DMF / acetic anhydride procedure.³⁵² The peptides were deprotected / cleaved from the resin using an 18:1:1 trifluoroacetic acid:water:ethane dithiol mixture and then isolated by filtration.³⁵² The crude peptides were then purified by HPLC and lyophilized. Peptides were characterized using LSIMS (see Table 3.1 below).

Table 3.1: LSIMS characterization of peptides.

Peptide	Mass (Da.)
N1/L	1814
H-EGG/N1/L	2015
EGG/N1/L	2057
EG/N1/L	2000
EEG/N1/L	2130
N2/L	1814
N2/AAAAA	1604
N2-GG/L	1928
N2-GGG/L	1985
N2-GGG/V	1971
N2-GGG/I	1985
N2-GGG/Nle	1985
C2-GGG/L	1985

Synthesis of Thioether Linked Cavitand Bowl TASP: Bowl/e/N2/L:

To a mixture of benzyl bromide cavitand bowl **2a** (2.0 mg, 2.1 μmol) and peptide **N2/L** (46 mg, 25 μmol) in DMA (10 mL, degassed) under a N_2 atmosphere at 0 $^\circ\text{C}$ (ice /water bath), was added pH 8.5 aqueous phosphate buffer (0.5 mL, 200 mM concentration). The reaction was stirred at 0 $^\circ\text{C}$ for 1 h. The solvent was removed *in vacuo*, and the residue dissolved in water (1.5 mL) containing 0.1% TFA, filtered (0.45 μm nylon filter), and purified by HPLC. Lyophilization resulted in 4 mg (21%) of **Bowl/e/N2/L**. NOTE: % Yields and characterization of products from these reactions are shown in Table 3.2. Reactions performed using peptides **H-EGG/N1/L**, **EGG/N1/L**, **EG/N1/L** and **EEG/N1/L** resulted in a multitude of products, none of which resembled the desired product.

Table 3.2: % Yields and ESMS characterization of thioether linked cavitand bowl TASP.

TASP	% Yield	Mass (Da.)
Bowl/e/H-EGG/N1/L	No Product	n/a
Bowl/e/EGG/N1/L	No Product	n/a
Bowl/e/EG/N1/L	No Product	n/a
Bowl/e/EEG/N1/L	No Product	n/a
Bowl/e/N2/L	21	7899

Preparation of TASPs using “Activated” Cavitand Bowl 17 and Cysteine-Containing Peptides: The preparation of **Bowl/H-EGG/N1/L** represents a typical example of this Procedure:

DIPEA (11 mg, 100 μ mol) was added to a mixture of activated cavitand bowl 17 (1.0 mg, 1.2 μ mol) and peptide **H-EGG/N1/L** (20 mg, 10 μ mol) in DMA (3 mL) under a N₂ atmosphere. The reaction was stirred at room temperature for 3 h. The solvent was removed *in vacuo*, keeping the temperature below 40 °C. The residue was re-dissolved in water (1.5 mL), filtered (0.45 μ m nylon filter), and injected directly onto a preparative HPLC column. Lyophilization resulted in 4 mg (34 %) of **Bowl/H-EGG/N1/L**. The % yield and ESMS characterization of the cavitand bowl TASPs prepared by this method are shown below (Table 3.3).

Table 3.3: % Yields and ESMS characterization of disulfide linked cavitand bowl TASPs, made via an “activated” cavitand bowl template.

TASP	% Yield	Mass (Da.)
Bowl/N1/L	31	8028
Bowl/H-EGG/N//L	34	8834
Bowl/EGG/N1/L	24	9002

Synthesis of Activated Peptides: The synthesis of **N1(SPy)/L** represents a typical example of this reaction:

Peptide **N1/L** (20 mg, 11 μmol) in 3 mL ethanol was added to a rapidly stirring solution of 2,2'-dipyridyl disulfide (12 mg, 55 μmol) in 2 mL ethanol. The reaction was stirred at room temperature for 1 h. The ethanol was reduced to 1 mL volume *in vacuo* (such that the reaction mixture is still in solution), and the solution pipetted onto ice-cold diethyl ether. This formed a precipitate that was filtered and washed with diethyl ether. This solid was re-dissolved in water (1.5 mL) containing 0.1% TFA, filtered (0.45 μm nylon filter), and purified by HPLC. Lyophilization resulted in 18 mg (85%) of **N1(Spy)/L**. The % yield and LSIMS characterization of all activated peptides is shown below (Table 3.4).

NOTE: Another method that was used to isolate the product from the reaction mixture is as follows: After the reaction was completed it was evaporated to dryness, and the residue partitioned between water (10 mL) and diethyl ether (5 mL). The layers were separated, and the aqueous phase was washed twice more with diethyl ether (2 x 5 mL). The aqueous phase was then reduced to 2 mL *in vacuo*, filtered (0.45 μm nylon filter), and purified by HPLC. Lyophilization produced yields similar to those seen for the "precipitation method" for isolating the product.

Table 3.4: % Yields and LSIMS characterization of S-pyridyl activated peptides.

Activated Peptide	% Yield	Mass (Da.)
N1(Spy)/L	85	1923
H-EGG/N1(SPy)/L	58	2124
EGG/N1(SPy)/L	65	2166
EG/N1(SPy)/L	54	2109
EEG/N1(SPy)/L	57	2239
N2(SPy)/L	65	1923
N2(SPy)/AAAAA	50	1713
N2(SPy)-GG/L	82	2037
N2(SPy)-GGG/L	75	2094
N2(SPy)-GGG/V	61	2080
N2(SPy)-GGG/I	68	2094
N2(SPy)-GGG/Nle	72	2094
C2(SPy)-GGG/L	91	2094

Synthesis of Disulfide Linked CTB TASPs using an “Activated” Peptide: The synthesis of CTB/N1/L represents a typical example of this procedure:

DIPEA (24 mg, 190 μ mol) was added to a mixture of activated peptide N1(SP_y)/L (36 mg, 19 μ mol) and benzyl thiol CTB 1 (1.6 mg, 3.2 μ mol) in DMF (7 mL) under a N₂ atmosphere. The reaction was stirred at room temperature for 1 h. The solvent was removed *in vacuo*, keeping the temperature of the water bath below 40 °C. The residue was re-dissolved in water (1.5 mL), filtered (0.45 μ m nylon filter), and injected directly onto a preparative HPLC column. Lyophilization resulted in 9 mg (45%) of CTB/N1/L. Also, 5 mg of N1(SP_y)/L was recovered. The % yield and ESMS characterization of all CTB TASPs are shown below (Table 3.5).

Table 3.5: % Yields and ESMS characterization of disulfide linked CTB TASP_s, made via “activated” peptides.

TASP	% Yield	Mass (Da.)
CTB/N1/L	45	5934
CTB-f/H-EGG/N1/L	31(54*)	6537
CTB-s/H-EGG/N1/L	23(54*)	6537
CTB/EGG/N1/L	36	6663
CTB/N2/L	65	5934
CTB/N2-GG/L	72	6278
CTB/N2-GGG/L	50	6447
CTB/N2-GGG/V	68	6408
CTB/N2-GGG/I	75	6447
CTB/N2-GGG/Nle	74	6447
CTB/C2-GGG/L	74	6447

***NOTE:** CTB/H-EGG/N1/L was prepared as described above, using activated peptide H-EGG/N1(SPy)/L. In this case it was possible to separate the two different diastereomers (see text) by HPLC. The faster eluting diastereomer CTB-f/H-EGG/N1/L, was recovered in 31% yield, and the slower moving diastereomer CTB-s/H-EGG/N1/L was recovered in 23% yield (therefore a 54% overall yield for this reaction).

Synthesis of Disulfide Linked Cavitand Bowl TASPs using an “Activated”

Peptide: The synthesis of **Bowl/N2-GGG/V** represents a typical example of this procedure:

DIPEA (13 mg, 100 μ mol) was added to a mixture of activated peptide **N2(SPy)-GGG/V** (36 mg, 12 μ mol) and benzyl thiol cavitand bowl **2b** (1.0 mg, 1.3 μ mol) in DMA (5 mL) under a N₂ atmosphere. The reaction was stirred at room temperature for 1 h. The solvent was removed *in vacuo*, keeping the temperature of the water bath below 40 °C. The residue was re-dissolved in water (1.5 mL), filtered (0.45 μ m nylon filter), and injected directly onto a preparative HPLC column. Lyophilization resulted in 9 mg (89%) of **Bowl/N2-GGG/V**. The % yield and ESMS characterization of all cavitand bowl TASPs are shown below (Table 3.6).

Table 3.6: % Yields and ESMS characterization of disulfide linked cavitand bowl TASP, made via “activated” peptides.

TASP	% Yield	Mass (Da.)
Bowl/N1/L	65	8028
Bowl/H-EGG/N1/L	21	8834
Bowl/EGG/N1/L	35	9002
Bowl/EG/N1/L	No Product	n/a
Bowl/EEG/N1/L	No Product	n/a
Bowl/N2/L	68	8027
Bowl/N2/AAAAA	70	7185
Bowl/N2-GG/L	47	8483
Bowl/N2-GGG/L	71	8713
Bowl/N2-GGG/V	89	8658
Bowl/N2-GGG/I	58	8712
Bowl/N2-GGG/Nle	79	8713
Bowl/C2-GGG/L	53	8712

Initial Three- and Four-Helix Bundle TASP

4.0 Introduction

The main aim when designing proteins from scratch (*i.e. de novo* design) is to produce a structure that has characteristics common with natural proteins: Namely, producing a polypeptide that undergoes controlled monomeric folding in solution, and results in a tertiary structure with high conformational specificity (a so called “native-like” structure). The global stability of the protein is also an important consideration, but can be sacrificed in the pursuit of “native-like” properties.^{210,353,354} This chapter first outlines some of the techniques used for the characterization of proteins in terms of their global structure/stability (Section 4.1), association state (Section 4.2), and conformational specificity (Section 4.3). Next, it introduces the design of the initial CTB and bowl TASP (Section 4.4), before reporting on their characterization (Section 4.5 to 4.10). The chapter concludes with a discussion of the results, and comparison to other TASP made in our group.

4.1 Methods for Characterizing the Global Structure and Stability of Proteins

Our CTB and bowl TASP s were designed to form three- and four-helix bundles in solution, respectively. One simple way to determine the secondary structure of a protein is using circular dichroism (CD) spectroscopy (Section 4.1.1).³⁵⁵ The effects of various reagents, (pH, salt and trifluoroethanol, see Section 4.1.2) on the secondary structure of a protein can be monitored by CD spectroscopy, and this can indirectly give information on the protein's tertiary structure. The stability of a protein is usually assayed by its resistance to unfolding, induced by either chaotropic reagents or temperature (Section 4.1.3). Thus, the global conformational stability of protein is the difference in free energy between its folded, and unfolded states.

4.1.1 Circular Dichroism (CD) Spectroscopy

This technique relies on the passage of beams of left- and right-circularly polarized light (in the far UV region) through a sample.³⁵⁶ As chiral molecules, L-amino acids interact differently with each component of the beam, altering their speeds, and thereby rotating the elliptically polarized light. The CD spectra is the difference between the absorbance values for the sample when using left and right circularly polarized light. The CD spectra for peptides are typically monitored between 300-190 nm. CD can be used to estimate the

conformation of polypeptides and proteins.³⁵⁷ The CD curve for an α -helix is characterized by minima at 222 and 208 nm, and a maximum at 295 nm.⁵ The CD signal at 222 nm is used to assess the amount of α -helix present during unfolding experiments (also it can be compared to a calculated theoretical maximum value).³⁵⁸

4.1.2 Effect of Varying pH. Addition of Salt, or Trifluoroethanol

Proteins can sometimes be denatured by, or undergo structural changes at extremes of pH.³⁵⁹ Extremes of pH can neutralize ionizable side chains (potentially removing stabilizing salt bridges). TASPs studied in pH 2-12 showed no noticeable difference in their structure (by CD).^{301,360} The presence of salt can screen any electrostatic interactions within the TASP.⁵ The addition of potassium chloride (KCl) to the TASPs, sometimes resulted in structural changes that were attributed to screening electrostatic interactions present in either the helix or the template/linker.

Addition of trifluoroethanol (TFE) to a polypeptide or protein in solution is known to induce α -helix structure.³⁶¹ The mechanism is not fully understood,³⁶² however with respect to helical bundles; TFE is thought to first disrupt inter-helix bundling and then stabilize the “isolated” α -helices.¹²⁶ The addition of TFE to the CTB and Bowl TASPs resulted in a very modest (14-16%) increase in their helicity, indicating that they are near to their maximum helicity.³⁶³

4.1.3 Stability to Chemical Denaturants and Temperature

We used guanidine hydrochloride (GnHCl) to test the stability of our CTB and Bowl TASP. GnHCl is thought to bind to protein surfaces, disrupting both hydrogen bonding and hydrophobic packing.^{5,364} Additionally GnHCl is a salt, and therefore screens any electrostatic interactions within the protein.³⁶⁵⁻³⁶⁷ The TASP. s all unfolded in a co-operative manner, and it was possible to assay their global conformational stability (also known as the free energy of unfolding) by analysis of the unfolding curves³⁶⁸ (see Section 4.12.2 for more details). TASP. s did not unfold in the presence of 10 M urea (a denaturant which does not screen electrostatic interactions).³⁶⁰ Thermally-induced unfolding of proteins is another method to assay the stability of proteins.⁵ However, the TASP. s did not unfold, even in the presence of GnHCl at 90°. ^{301,360}

4.2 Methods for Determining the Association State of Proteins

Ideally the TASP. s fold into monomers in solution. Self-association would affect the observed global stability and structure of the TASP. s. In order to detect any self-association of our TASP. s we tried several methods (Section 4.2.1-4.2.5) which included; assaying TASP. s for concentration dependent CD spectra or concentration dependent GnHCl-induced unfolding curves, and the techniques of size exclusion (gel permeation) chromatography,

electrospray mass spectrometry, and sedimentation equilibrium analytical ultracentrifugation.

4.2.1 Concentration Dependent Circular Dichroism

If a protein or peptide exists as a single species in solution, then its CD spectrum will not change with concentration. It is common for the helicity of self-associating helices to increase with concentration.²²⁴⁻²²⁸ However, the CTB and Bowl TASPs had concentration independent CD spectra but were undergoing self-association: CD measures secondary structure, and the helices in the TASPs were close to their maximum helicity (by CD), therefore self-association occurred with no detectable change.³⁶⁰

4.2.2 Concentration Dependent GnHCl-Induced Unfolding Curves

Another method used for detecting the association of proteins is by monitoring their unfolding curves (using temperature or chemical denaturation) at different concentrations.^{369,370} A higher concentration of protein (and therefore a higher degree of association) usually results in increased stability. Again, this method did not prove reliable for detecting TASP self-association: The TASPs were probably associated by weak forces,

and often low concentrations of denaturant were sufficient to disrupt this self-association before any detectable unfolding of the tertiary structure.

4.2.3 Size Exclusion Chromatography

Size exclusion (also known as gel permeation) chromatography separates molecules based on their size.⁵ The technique involves a column of molecular sieves (the solid phase), onto which a sample of protein is applied and eluted with aqueous buffer. The larger proteins tend to elute fastest. By calibrating the elution time with proteins of known size, it is possible to estimate the size of an unknown protein. This approach has been used to determine the oligomeric state of numerous *de novo* designed proteins; examples include coiled-coils and self-associating α -helices. All our TASPs eluted as a single species;³⁷¹ different solid phases (Sephadex, Biogel, ProteinPAK) were used. For reversibly self-associating proteins that elute as a single species, there are methods reported in the literature for assaying the degree of self-association using size exclusion chromatography. These methods involve multiple experiments using various concentrations of protein, and measurement of its elution profile.³⁷³ This labour intensive method was not attempted due to the availability of sedimentation equilibrium ultracentrifugation (Section 4.2.5).

4.2.4 Electrospray Mass Spectrometry

Electrospray mass spectrometry uses a soft ionization technique that has been used to study the non-covalent association of proteins.³⁷³ No self-association of the TASPs was observed. However, altering the experimental conditions to optimize the observation of non-covalent interactions, molecular weights consistent with up to 18 water molecules associated with our TASP were observed.³⁶⁰

4.2.5 Analytical Ultracentrifugation

Sedimentation equilibrium experiments using an analytical ultracentrifuge can be used to estimate the molecular weight of macromolecules in solution. This technique overcomes any uncertainties associated with the shape and hydration of the protein.³⁷⁴⁻³⁷⁷ Centrifugation of the solute is performed using a centrifugal force that is low enough for the tendency of a macromolecule to sediment, to be balanced by its tendency to diffuse: Thus, at equilibrium a concentration gradient is established throughout the solution. Analysis of the concentration gradient can give information on the molecular weight of the solute, and estimate the degree of self-association (see Section 4.12.3). This proved to be the most useful technique for determining the self-association of the TASPs in solution. Unfortunately, UBC is not equipped with an analytical ultracentrifuge, TASP samples were run at the Massachusetts Institute of Technology, and it was not possible to obtain exhaustive sedimentation equilibrium data.

4.3 Methods for Determining the Conformational Specificity of Proteins

The ultimate goal of this thesis is to produce TASP s with structures akin to those observed in natural proteins. Conformational specificity refers to the unique three-dimensional structure associated with proteins. Many *de novo* designed proteins have resulted in tertiary structures that resemble molten globules (see Section 1.3.3). A long term goal of this research is to investigate some of the features that are important to the formation of native-like structure. One technique for observing molten globules is binding of the hydrophobic dye, 1-anilinonaphthalene-8-sulfonate (ANS). A more popular technique is observation of the one-dimensional NMR spectra of the protein.^{378,379}

4.3.1 ANS Binding

The hydrophobic fluorescent dye 1-anilinonaphthalene-8-sulfonate (ANS), has frequently been used for detection of the molten globule state of proteins.³⁸⁰⁻³⁸² In the molten globule state, more of the hydrophobic surface of the protein core is loosely packed, therefore allowing the non-polar ANS dye to bind more readily. ANS has been found to bind preferentially to the molten globule, as opposed to the fully unfolded, or fully folded state of a protein.³⁸³ When the ANS is transferred from a polar (water) to a non-polar (interior of a protein) environment its wavelength of fluorescence is shifted (to a lower

wavelength), and increases in magnitude. A recent publication found that ANS did not bind to a *de novo* protein with a well packed core, a protein that lacked any conformational specificity according to a NMR study.²⁰⁶ The hydrophobic core of these TASPs is expected to be well packed and therefore this technique was not used.

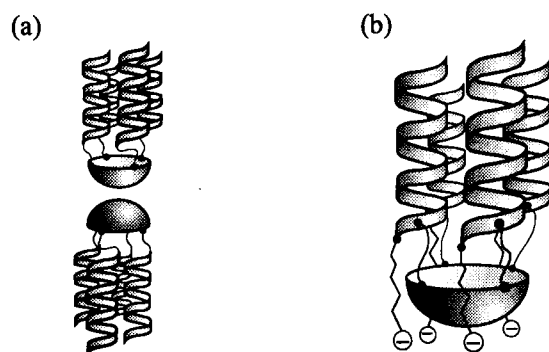
4.3.2 One Dimensional ^1H NMR Spectra

Numerous multi-dimensional NMR techniques exist for identifying the structure of proteins.³⁸⁴⁻³⁸⁶ These techniques would be difficult to employ due to the degenerate sequence/structure of our protein. Therefore we used one-dimensional NMR to qualitatively observe the conformational specificity of our TASPs. "Native-like" structure is manifested by large chemical shift dispersion and sharp lines associated with the protons in the TASP.³⁸⁷⁻³⁸⁹ Conversely a broad, poorly dispersed NMR spectrum is indicative of a molten globule structure.

4.4 Introduction to Initial CTB and Bowl TASP

The TASP described in this chapter were synthesized with the shortest template-to-helix link available (e.g. peptide Ac-CEKLLKELKELLEKG-NH₂, where the template-bound cysteine residue is adjacent to the “helix”). The code for these TASP is discussed in Section 3.1.1. The first CTB (Section 4.5) and cavitand bowl (Section 4.7) TASP were made using peptide N1/L (Ac-CEKLLKELKELLEKG-NH₂). This sequence is designed to fold into an α -helix that possessed both a hydrophilic and a hydrophobic face (see Section 3.1.1 for the design of amphiphilic α -helices that are used for making helical bundles). For the next CTB and bowl TASP in the series, peptide N1/L was lengthened at its N-terminus with a pre-cysteine three-residue sequence, to produce three- and four-helix bundle TASP incorporating either peptide EGG/N1/L (**Ac-EGG**CEKLLKELKELLEKG-NH₂), or peptide H-EGG/N1/L (**H-EGG**CEKLLKELKELLEKG-NH₂). At pH 7.0, EGG/N1/L has an overall negative charge on its pre-cysteine sequence, and H-EGG/N1/L (which had a free N-terminus) has both a positive and negative charge. The addition of these extra residues was an attempt to overcome any self-association between the hydrophobic organic macrocyclic templates in aqueous solution (Figure 4.1). This precaution was later found to be unnecessary, as TASP self-association was found to be primarily driven by the nature of the template-to-helix linker (see Chapter 5).

Figure 4.1: (a) Potential inter-macrocycle self-association in aqueous solution. (b) TASP containing charged, pre-cysteine residues to prevent inter-macrocycle self-association by charge-charge repulsion.



Following these initial investigations, we then altered the patterning of the hydrophilic residues in the peptide sequence, primarily to allow us to compare this series of TASPs to others made in our group (TASPs that differ in the template-to-helix linker - see Section 1.4.4). Thus, peptide N1/L (Ac-CEKLLKELKELLEKG-NH₂) was replaced by N2/L (Ac-CEELLKKLEELLKKG-NH₂). The resulting CTB (Section 4.6) and bowl (Section 4.8) TASPs could now be compared to other research done by our group,^{270,271} and the “EK/KE” patterning of hydrophilic residues in the helix, could be compared to the “EE/KK” pattern.

Next, we investigated the effect of changing all the leucine residues of N2/L to alanine (*i.e.* to peptide N2/AAAAA, Ac-CEEAAKKAAEEAAKKG-NH₂) on the bowl TASP. This change was made to investigate the stability of a TASP which lacked a hydrophobic core (Section 4.9). Finally (for this chapter), the effect of changing template-to-peptide linker from a disulfide to a thioether bond (Section 4.10) was investigated on the cavitand

bowl TASP, using the **N2/L** peptide. This chapter concludes with a discussion of the structural implications of these initial studies (Section 4.11), including a comparison between the three-helix bundle CTB, and the four-helix bundle bowl TASPs, as well as to earlier work done in our group using a different linker with the bowl TASP.

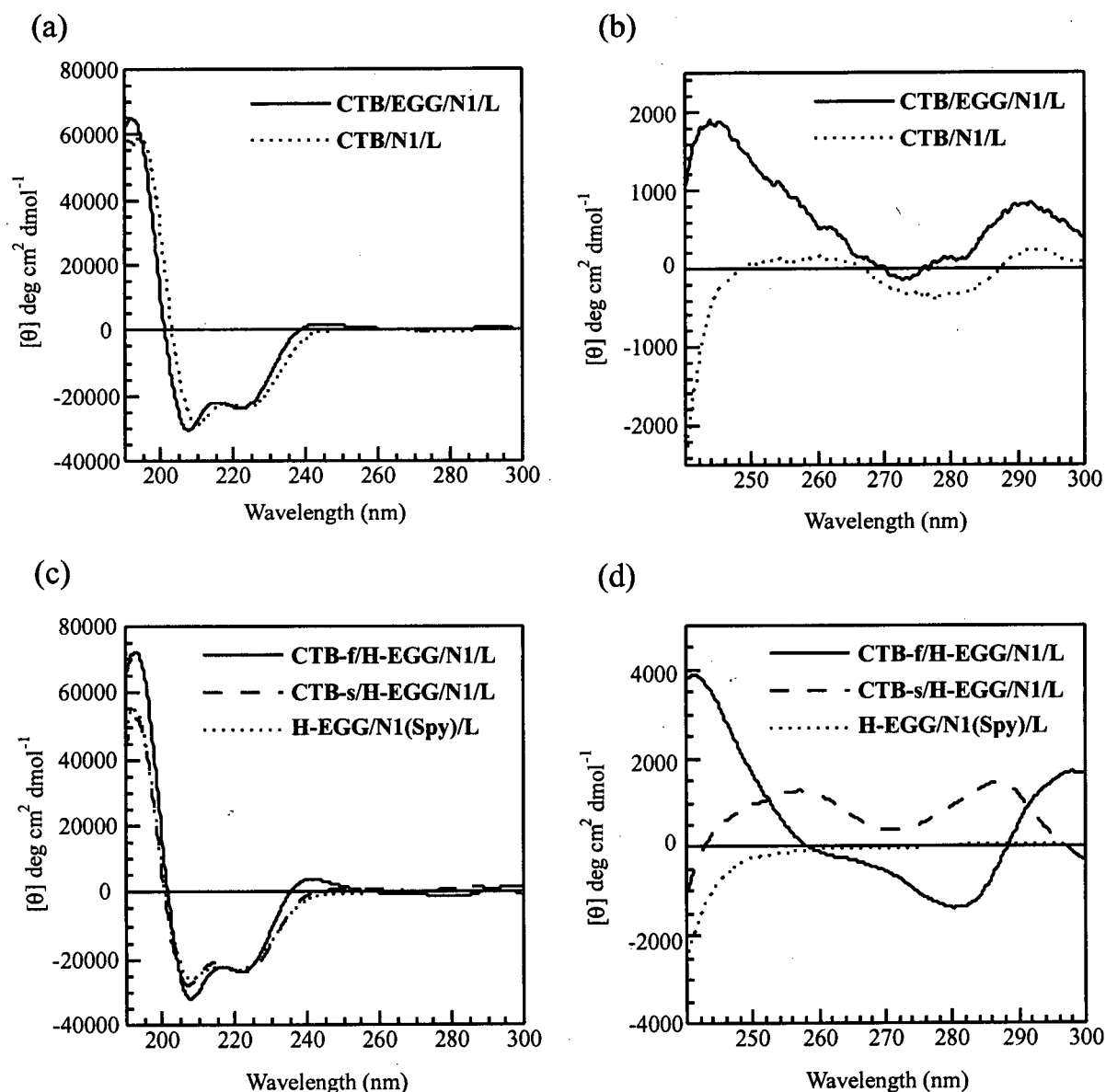
4.5 CTB/N1/L, CTB/EGG/N1/L, and CTB/H-EGG/N1/L

These three CTB TASPs contained disulfide-linked peptides **N1/L**, **EGG/N1/L** or **H-EGG/N1/L** (X-CEKLLKELKELLEKG-NH₂, where X = Ac-, Ac-EGG- or H-EGG-). One interesting feature of the **CTB/H-EGG/N1/L** TASP was the resolution of its two diastereomers (remember that the CTB template is a racemic mixture, and the addition of peptides made up of L-amino acids, results in a pair of diastereomers - see Section 2.1 for more details). These diastereomers were separated using reverse phase HPLC, and labeled as **CTB-f/H-EGG/N1/L** or **CTB-s/H-EGG/N1/L**, depending on whether it was the faster or slower eluting species. The resolution of these two diastereomers (which is due to their different physical characteristics), is an indication that **CTB/H-EGG/N1/L** exists as a monomer under the conditions used for the HPLC separation: If **CTB/H-EGG/N1/L** was associated during the HPLC separation, then one would not expect its elution profile to contain two peaks in a one-to-one ratio (unless self-association occurred only between the individual diastereomers). There was also some resolution between the two diastereomers of **CTB/EGG/N1/L** by HPLC, but there was no baseline separation between the two peaks, making it impractical to separate them.

The CD spectra of **CTB/N1/L**, **CTB/EGG/N1/L**, **CTB-f/H-EGG/N1/L**, **CTB-s/H-EGG/N1/L** and the peptide monomer **H-EGG/N1/L** (Figure 4.2a and 4.2c) are all characteristic of an α -helix, with minima at 208 and 222 nm, and a maximum at 195 nm. Also, these TASPs exhibit CD spectra above 240 nm (Figure 4.3b and 4.3d), whereas the peptide **H-EGG/N1/L** does not. If this near-UV CD signal for the resolved TASPs, **CTB-f/H-EGG/N1/L**, and **CTB-s/H-EGG/N1/L** resulted solely from the chiral CTB template, then they should be opposite in magnitude: They are not, and therefore this near-UV CD results from a pair of diastereomers, where the aromatic CTB chromophore (or the disulfide bond) is interacting with a chiral environment, *i.e.* the α -helices. The sum of the near-UV CD spectra of the resolved **CTB-f/H-EGG/N1/L** and **CTB-s/H-EGG/N1/L**, results in a similar shape to the near-UV CD signals seen for the unresolved CTB TASPs in this series.

In 2 M KCl (salt) the helicity of these TASPs increased slightly (**CTB/N1/L**, +5%; **CTB/EGG/N1/L**, +9%; **CTB-f/H-EGG/N1/L**, +3%; **CTB-s/H-EGG/N1/L**, +4%). Peptide helicity (*i.e.* secondary structure) is not necessarily related to a TASP's stability, nor its conformational specificity. A salt-induced increase in TASP helicity indicates that screening of electrostatic interactions changes the secondary structure (and therefore the tertiary structure). This change in TASP structure may result from the screening of electrostatic forces within the TASP molecule, or between TASP molecules that are associating in solution.

Figure 4.2: (a) Full (190-300 nm) and (b) expanded (240-300 nm) CD spectra of TASPs **CTB/N1L** (60 μ M, dotted line) and **CTB/EGG/N1/L** (64 μ M, solid line). (c) Full (190-300 nm) and (d) expanded (240-300 nm) CD spectra of TASPs **CTB-f/H-EGG/N1/L** (62 μ M, solid line), **CTB-s/H-EGG/N1/L** (64 μ M, dashed line) and peptide **H-EGG/N1/L** (180 μ M, dotted line). All CDs run in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



The CD spectra of these initial CTB TASPs were identical across a 50-fold range of TASP concentrations (3-150 μ M, data not shown). The peptide **H-EGG/N1/L** exhibited a concentration dependent CD spectrum (9-300 μ M, data not shown), with higher helicity at increasing concentrations. These observations are consistent with the CTB TASPs being monomers in solution, whereas the amphiphilic peptide strands associate readily. However, the helicity (judged by the magnitude of the signal at $[\theta]_{222}$) for these CTB TASPs is close to the theoretical maximum, so any self-association (usually accompanied by an increase in helicity at increasing concentration) may not be detected by CD, a technique which monitors only the secondary structure of a protein. Moreover, they may be associated over the entire concentration studied, where the aggregate remained intact even at low TASP concentration.

A more rigorous way to identify self-association for proteins is to compare the GnHCl-induced unfolding curves at different protein concentrations.^{369,370} Self-association is consistent with greater stability. The unfolding curves for chemical denaturation using GnHCl, are concentration independent (within experimental error, over a ten-fold concentration range) for TASPs **CTB/N1/L**, **CTB-f/H-EGG/N1/L**, and **CTB-s/H-EGG/N1/L** (Figure 4.3). However, TASP **CTB/EGG/N1/L** has concentration dependent unfolding curves (Figure 4.3b), consistent with a self-associating species. This evidence is consistent with TASPs **CTB/N1/L**, **CTB-f/H-EGG/N1/L**, and **CTB-s/H-EGG/N1/L** existing as monomers in solution, and the self-association of TASP **CTB/EGG/N1/L**. The GnHCl-induced unfolding of the peptide **H-EGG/N1/L** was not co-operative, and was concentration dependent; the bundling of amphiphilic peptide helices has been investigated by other researchers.²²⁴⁻²³⁰ Thus, pre-organization of the helical bundles by the CTB

templates contributes significantly to their overall conformational stability (Figure 4.4a compares the GnHCl-induced unfolding curves of peptide **N1(Spy)/L** to **CTB/N1/L**, and Figure 4.4b compares the unfolding curves of TASP **CTB/EGG/N1/L**, **CTB-f/H-EGG/N1/L** and **CTB-s/H-EGG/N1/L**).

Figure 4.3: Effect of GnHCl (at pH 7.0, 10 mM phosphate buffer) on the helicity of (a) CTB/N1/L (30 μ M, \circ , dashed line; 3.0 μ M, \times , solid line), (b) CTB/EGG/N1L (32 μ M, \circ , dashed line; 3.0 μ M, \times , solid line), (c) CTB-f/H-EGG/N1/L L (31 μ M, \circ , dashed line; 3.0 μ M, \times , solid line), and (d) CTB-s/H-EGG/N1/L (32 μ M, \circ , dashed line; 3.0 μ M, \times , solid line).

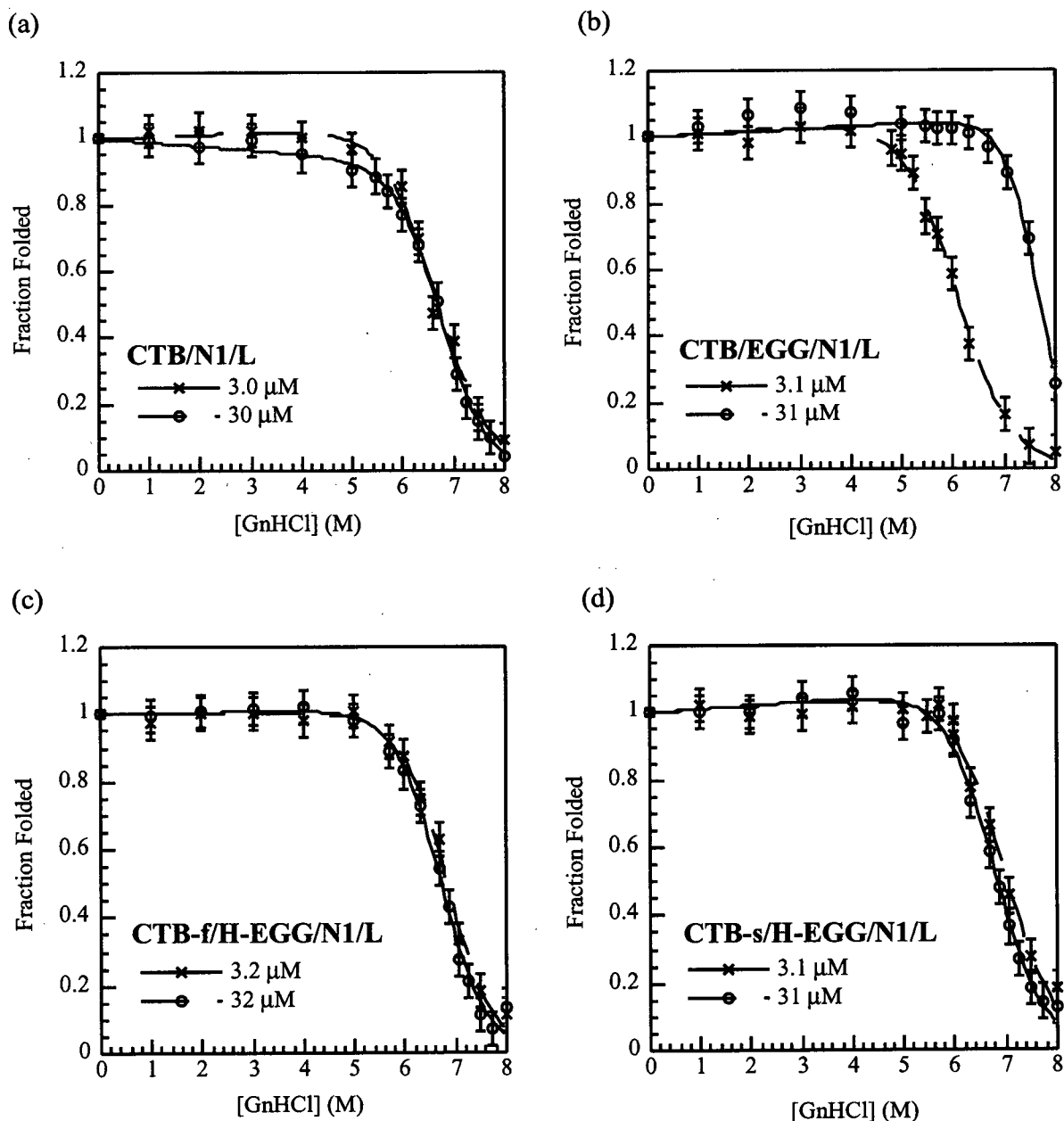
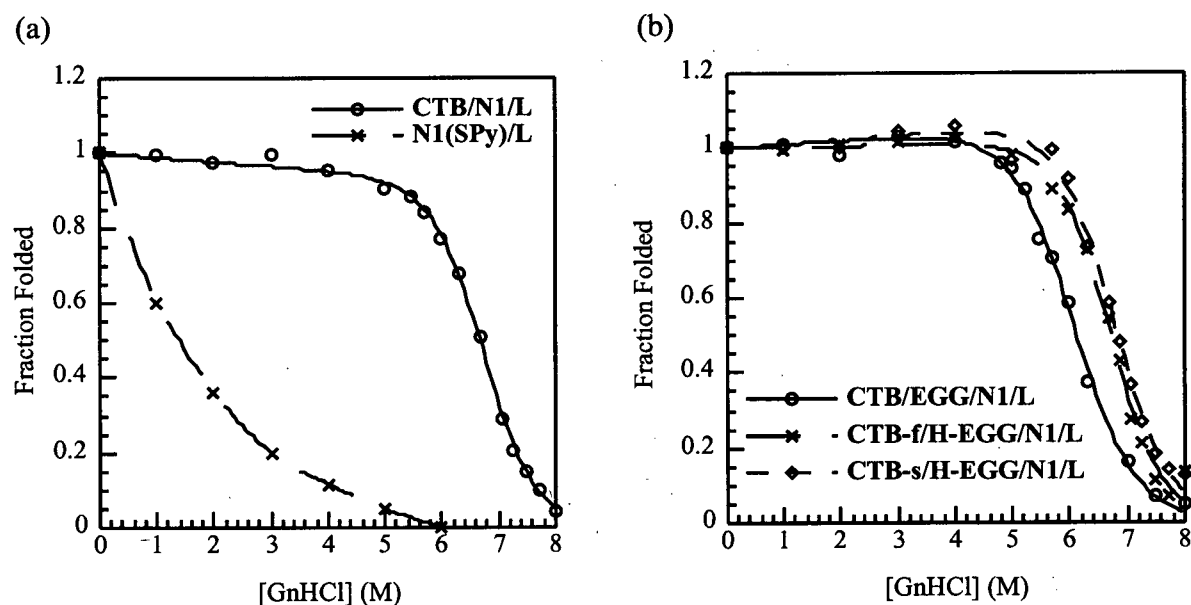


Figure 4.4: GnHCl-induced unfolding curves of (a) CTB/N1/L (3.0 μ M, O, solid line) versus peptide N1(Spy)/L (10.0 μ M, x, dashed line); and (b) CTB/EGG/N1/L (3.2 μ M, O, solid line) versus CTB-f/H-EGG/N1/L (3.2 μ M, x, long dashed line), and CTB-s/H-EGG/N1/L (3.1 μ M, \diamond , short dashed line);. All points measured in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



Unfolding of the TASPs, induced by the chemical denaturant GnHCl, allows us to assess their global conformational stability, in terms of the free energy change between the folded and unfolded state. A reversible, co-operative unfolding transition is consistent with a two-state unfolding process. These isothermal chemical denaturations give the mid-point of the unfolding curves, $[\text{GnHCl}]_{0.5}$, as a crude measure of global stability. More detailed analysis can be carried out using the “linear extrapolation method” (which assumes a two-

state unfolding transition - see Section 4.12.2 for further details): This method estimates the $\Delta G^\circ \text{H}_2\text{O}$ value (the conformational free energy of unfolding in the absence of denaturant), and also generates the molar co-solvent term m , (a reflection of the difference in solvent accessible surface area, between the folded and unfolded states). The CTB and bowl TASP presented in this chapter are not fully unfolded at 8 M GnHCl, therefore there is a larger error associated with the thermodynamic values obtained by estimating the post-transitional curve.

TASPs **CTB/N1/L**, **CTB-f/H-EGG/N1/L**, and **CTB-s/H-EGG/N1/L** all have similar values of m (a value that relates to the amount of buried hydrophobic surface that is exposed to the solvent on unfolding) and $[\text{GnHCl}]_{0.5}$ (Table 4.1). This indicates they have similar buried hydrophobic surface. Their predicted free energies of unfolding are also all similar within experimental error. At 3.1 μM (a concentration where its self-association is minimized), TASP **CTB/EGG/N1/L** has similar (but slightly lower) values for m and $\Delta G^\circ \text{H}_2\text{O}$, as the other TASPs in this series. However, at 31 μM , **CTB/EGG/N1/L** exhibits an increase in the value of m , consistent with a larger degree of buried surface in the folded state, *i.e.* self-association is apparent at 31 mM, and results in a greater degree of buried hydrophobic surface. This change in stability of the protein structure also indicates that the helices are responsible for self-association.

Table 4.1: Thermodynamic evaluation of **CTB/N1/L** variants, calculated from their GnHCl-induced denaturation. This data assumes that TASPs are monomers at the unfolding transition point.

TASP	Concentration (μ M)	[GnHCl] _{0.5} (M)	<i>m</i> kcal mol ⁻¹ M ⁻¹	$\Delta G^\circ_{H_2O}$ kcal mol ⁻¹
CTB/N1/L	3.0	6.7 \pm 0.1	1.3 \pm 0.1	8.9 \pm 0.3
CTB/N1/L	30	6.7 \pm 0.1	1.1 \pm 0.1	7.6 \pm 0.7
CTB/EGG/N1/L	3.1	6.2 \pm 0.1*	1.1 \pm 0.1*	6.8 \pm 0.4*
CTB/EGG/N1/L	31	7.7 \pm 0.1*	1.8 \pm 0.2*	13.5 \pm 1.6*
CTB-f/H-EGG/N1/L	3.2	6.7 \pm 0.1	1.3 \pm 0.1	8.9 \pm 0.7
CTB-f/H-EGG/N1/L	32	6.8 \pm 0.1	1.3 \pm 0.1	9.0 \pm 0.7
CTB-s/H-EGG/N1/L	3.1	6.8 \pm 0.1	1.2 \pm 0.1	8.3 \pm 0.6
CTB-s/H-EGG/N1/L	31	7.0 \pm 0.1	1.2 \pm 0.1	8.0 \pm 0.8

*This TASP has concentration dependent GnHCl-induced unfolding curves, therefore this data is likely to have errors greater than those shown.

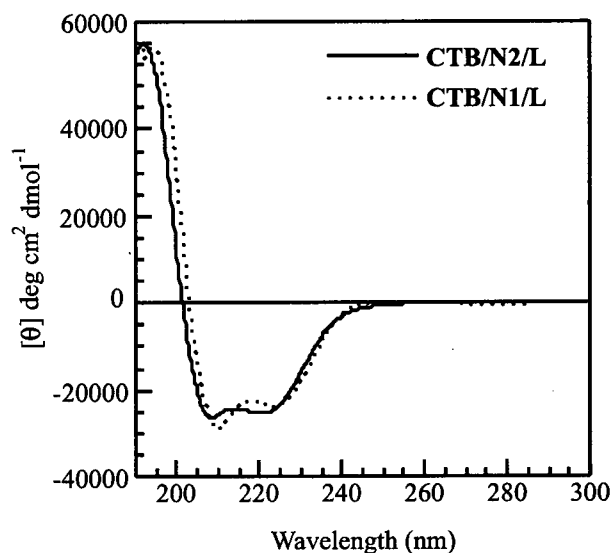
Thus, it appears that TASPs **CTB/N1/L**, **CTB-f/H-EGG/N1/L**, and **CTB-s/H-EGG/N1/L** exist as monomers in solution, and that **CTB/EGG/N1/L** undergoes a concentration dependent self-association. However, later in this thesis we will see that (for this series of TASPs) self-association is not always obvious when monitoring changes in secondary structure. Potential sources for TASP self-association will be speculated upon at the end of this chapter (Section 4.11).

4.6 CTB/N2/L

For this CTB TASP, the peptide sequence was changed from the original N1/L (Ac-**CEKLLKELKELLEK**G-NH₂) to N2/L (Ac-**CEELLKKLEELLKKG**-NH₂). The N2/L peptide contains sequentially “paired” hydrophilic lysine or glutamic acid residues. I was unable to separate the two CTB/N2/L diastereomers using HPLC.

The CD spectrum of CTB/N2/L (Figure 4.5) was insensitive to salt (2 M KCl) and was also concentration independent (data not shown). Unlike the analogous CTB/N1/L, CTB/N2/L did not exhibit any activity in its near-UV CD spectrum.

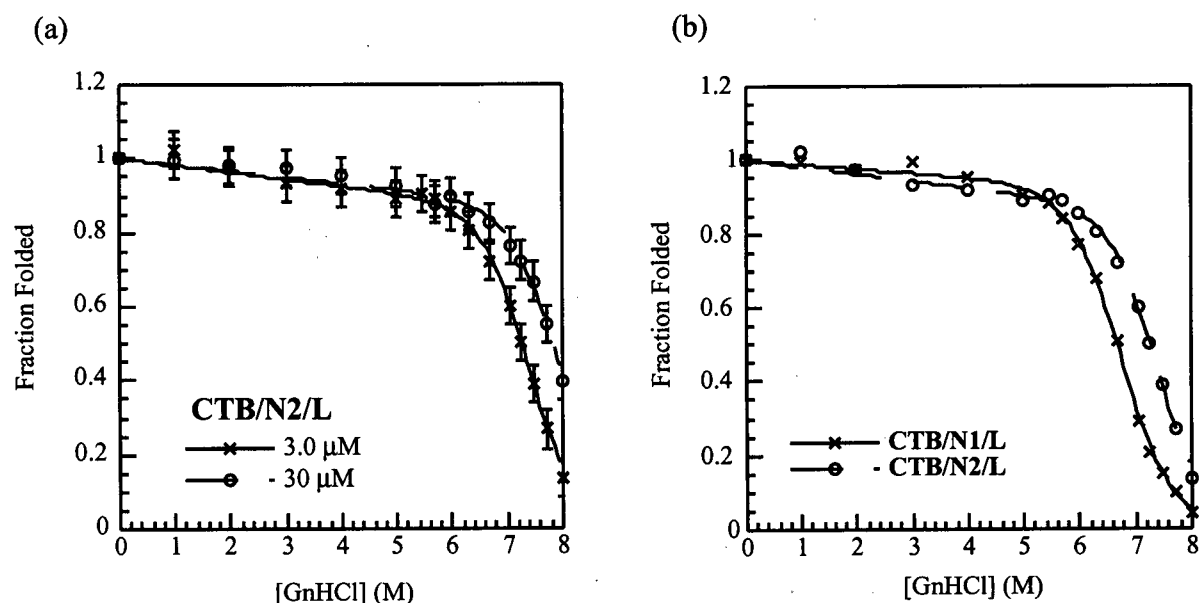
Figure 4.5: CD spectrum of CTB/N2/L (60 μ M, solid line), compared to CTB/N1/L (60 μ M, dotted line), in pH 7.0 phosphate buffer (10 mM) at 25 °C.



The oligomeric state of **CTB/N2/L** was assessed by its **GnHCl**-induced unfolding curves at different TASP concentrations (Figure 4.6). These curves appear to be dissimilar, consistent with a self-associating species in solution.

A sedimentation equilibrium ultracentrifugation study, estimated the “average” molecular weight to be 9300 ± 800 (fitting the data to a single ideal species, at $100 \mu\text{M}$ TASP concentration), which is greater than the calculated 5934. This confirms the self-association of **CTB/N2/L**, which exists as a monomer-dimer equilibrium in solution with a calculated association constant of $14100 \pm 600 \text{ M}^{-1}$ (see experimental section 4.12.3 for data).

Figure 4.6: **GnHCl**-induced denaturation curves of (a) **CTB/N2/L** ($30 \mu\text{M}$, \circ , dashed line; $3.0 \mu\text{M}$, \times , solid line) and (b) **CTB/N2/L** ($3.0 \mu\text{M}$, \circ , dashed line) versus **CTB/N1/L** ($3.0 \mu\text{M}$, \times , solid line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25°C .



Sedimentation equilibrium studies on other cavitand TASP s that describe a monomer-dimer equilibrium in aqueous buffer, have found the degree of association to be reduced in the presence of 6 M GnHCl, 2 M KCl, or 10% methanol.²⁷¹ Thus, there appears to be both an electrostatic and a hydrophobic component to the self-association of our TASP s. Is **CTB/N2/L** a monomer at the concentration of GnHCl required for it to unfold? The dissimilar GnHCl-induced unfolding curves do not support this. However, for simplicity, it is assumed that the TASP was not interacting with any other species at the unfolding transition point when calculating its global conformational stability (also, we have insufficient data to factor the degree of self-association into the analysis).

We see from the thermodynamic data calculated from the GnHCl-induced unfolding curves of **CTB/N2/L** (Figure 4.6 and Table 4.2), that this TASP is more stable than the analogous **CTB/N1/L**. This extra stability may be due to the greater degree of self-association of **CTB/N2L**. We're comparing our TASP s based on the assumption that they follow a two-state unfolding model; self-association (of partially unfolded TASP s, for example) would invalidate the data calculated from these unfolding curves, therefore the calculated thermodynamic values should be treated as an approximation.

The higher stability of **CTB/N2/L** versus **CTB/N1/L**, may result from the intrinsic stability of the **N2/L** peptide over the **N1/L** peptide (see Figure 4.7) rather than a higher degree of self-association. It is known that $i, i+1$ and $i, i+2$ oppositely charged interactions (E to K) are unfavourable to helix stability, and that $i, i+3$ and $i, i+4$ oppositely charged interactions (E to K) are favourable to helix stability (Section 1.4.1.4). **N2/L** has a total of

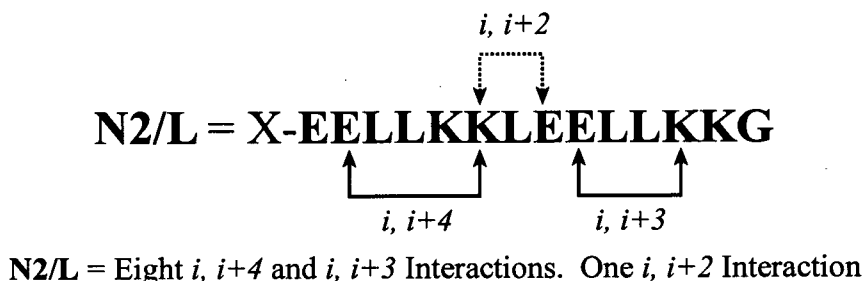
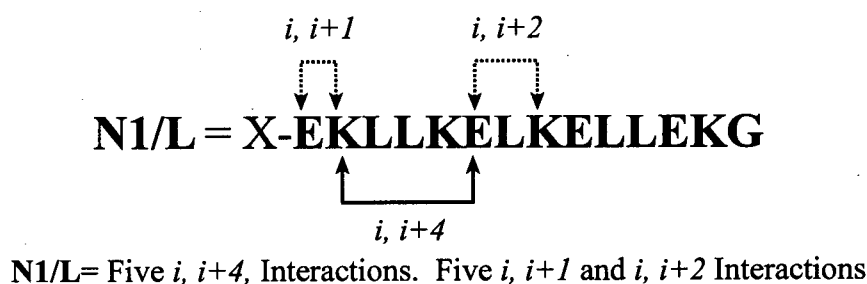
eight favourable and one unfavourable intra-helix electrostatic interactions, whereas N1/L has five favourable and five unfavourable electrostatic interactions.

Table 4.2: Thermodynamic evaluation of CTB/N2/L (compared to CTB/N1/L), calculated from its GnHCl-induced denaturation (this data assumes that each TASP is a monomer at the unfolding transition point).

TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m $\text{kcal mol}^{-1} \text{M}^{-1}$	$\Delta G^\circ \text{H}_2\text{O}$ kcal mol^{-1}
CTB/N1/L	3.0	6.7 ± 0.1	1.3 ± 0.1	8.9 ± 0.3
CTB/N1/L	30	6.7 ± 0.1	1.1 ± 0.1	7.6 ± 0.7
CTB/N2/L	3.0	$7.2 \pm 0.1^*$	$1.4 \pm 0.1^*$	$10.4 \pm 0.4^*$
CTB/N2/L	30	$7.8 \pm 0.1^*$	$1.3 \pm 0.1^*$	$10.4 \pm 0.9^*$

*This TASP has concentration dependent GnHCl-induced unfolding curves, therefore this data is likely to have errors greater than those shown.

Figure 4.7: The peptide sequences N1/L and N2/L. Examples of the favourable ($i, i+3$ and $i, i+4$) and the unfavourable ($i, i+1$ and $i, i+2$) interactions between oppositely charged residues in an α -helix are highlighted.



Natural proteins typically have NMR spectra with sharp, well dispersed peaks. The NMR spectrum of **CTB/N2/L** is broad, especially when compared to that of the peptide (Ac-EELLKKLEELLKKG-NH₂) alone (Figure 4.8). The most striking difference in the NMR spectra of **CTB/N2/L** versus the peptide, is in the N-H region. The peptide exhibits sharp N-H peaks, whereas **CTB/N2/L** shows a broad hump. The downfield (aliphatic-H) region is somewhat broadened for the TASP versus the peptide, probably due to the degenerate nature of the peptide sequence. The NMR evidence is consistent with **CTB/N2/L** existing in a multitude of slowly interchanging low energy conformers, *i.e.* a molten globule structure.

Another source of this line broadening in the NMR spectrum, could be the self-association of the TASP: However, the NMR spectra of other self-associating bowl TASPs (using the same "helix" sequence), are as sharp as that seen for the peptide.^{270,271} Therefore, this line broadening (which is characteristic of molten globule structure) is related to the nature of the template-to-helix linker.

Why does the peptide exhibit a NMR spectrum with sharp, well dispersed peaks? The N-H region of the peptide's NMR spectrum exhibits one sharp peak per residue: Is it forming a quaternary four-helix bundle protein with native-like structure, and near perfect symmetry? From the concentration dependent CD study of the peptide, we know that it is self-associating to form helical bundles. Unfortunately, we do not know the size and structure of these bundles, but if they were forming a native-like structure then one would expect them to have a greater resistance to denaturation, and exhibit a co-operative unfolding transition.

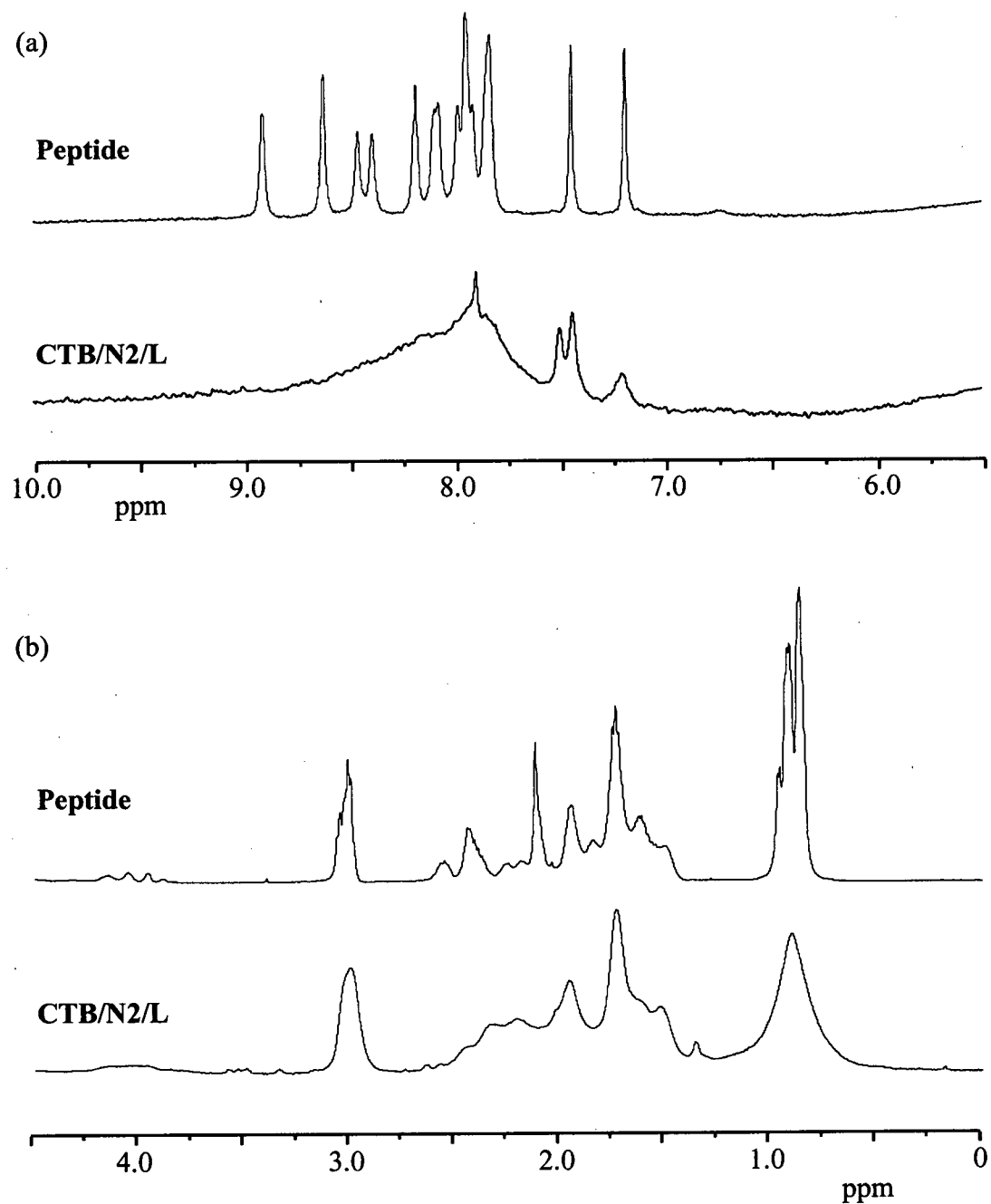
Another explanation to the sharp, well dispersed NMR spectrum of the peptide is that the helices form bundles that have a dynamic structure: This quaternary structure has rapidly inter-converting conformations that result in single peaks, *i.e.* from fast-exchange on the NMR time scale. A study on a 12 residue peptide (Ac-ELLKKLLEELKG-OH) similarly designed to associate into four-helix bundles was carried out using multidimensional, and concentration dependent NMR studies.³⁹⁰ The authors concluded that at higher concentrations this peptide formed helical bundles that associated reversibly, with rapidly inter-converting conformations. Another NMR study on a 16 residue peptide

designed to form four-helix bundles (Ac-GELEELLKKLKELLKG-NH₂) concluded that the peptide co-operatively forms tetramers, but that the symmetrical N-H region of the spectrum suggests the possibility of conformational averaging due to rapid side chain fluctuations, or rapidly equilibrating intermediates (e.g. dimers).²²⁶

The upfield, aliphatic region of **CTB/N2/L**'s NMR spectrum has similar chemical shift dispersion to the peptide, but with much broader peaks. The formation of a TASP with molten globule structure was not unexpected using this peptide sequence. *De novo* proteins that contain leucine residues occupying the hydrophobic core are known to form molten globules. Also, the peptide sequence was originally designed to form a four-, rather than a three-helix bundle. Thus, "forced" into a three-helix bundle topology, the core side chains of the helices may be unable to undergo unique packing, which is the hallmark of natural proteins.

Thus, the molten globule structure of **CTB/N2/L** was not unexpected; However there are four-helix bundle cavitand TASPs (that also differ in their template-to-helix linkers) made by our group that have resulted in what appears to be native-like structure (discussed further in Section 4.8). Would the **N2/L** peptide undergo specific packing of its side chains (to give a native like structure) when in a disulfide-linked four-helix bundle bowl TASP? Is the lack of conformational specificity observed in **CTB/N2/L** due to the disulfide template-to-helix linker, or is it due to the use of a peptide designed to form a four-helix bundle being "forced" into a three helix bundle? These questions will be addressed later in this chapter.

Figure 4.8: NMR spectra of **CTB/N2/L** (~0.3 mM) compared to peptide **Ac-EELLKKLEELLKKG-NH₂** (~2.2 mM), at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-Hs, aromatic-Hs) region, 10.0-5.5 ppm (magnified) and (b) Upfield (aliphatic-Hs) region, 4.5-0.0 ppm.

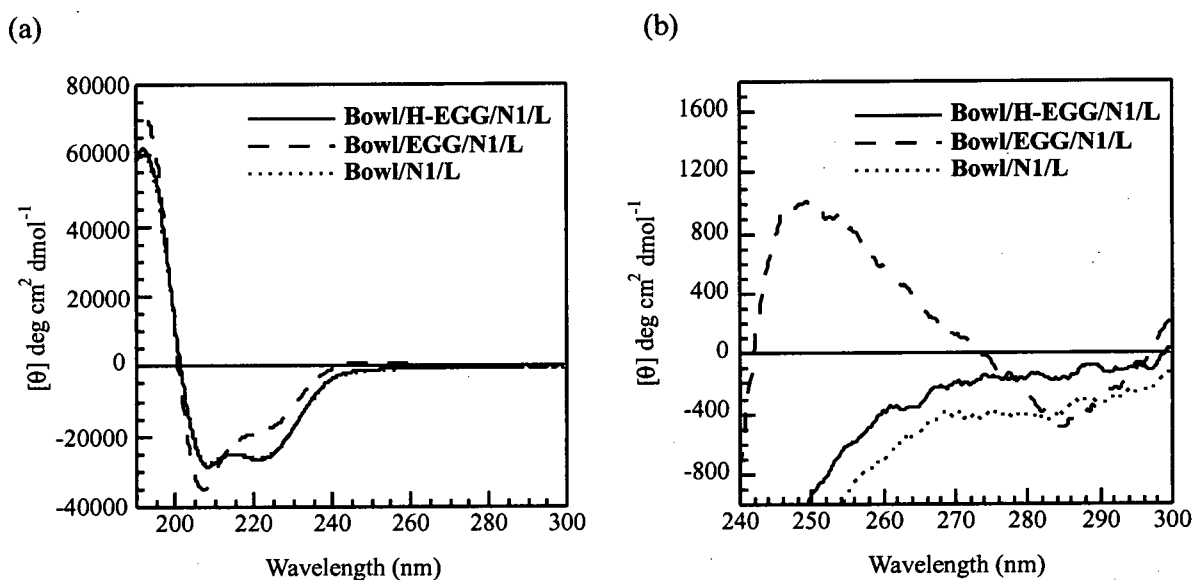


4.7 Bowl/N1/L, Bowl/EGG/N1/L, and Bowl/H-EGG/N1/L

These four-helix bundle TASP s were constructed (via disulfide bonds) using the cavitand template and either peptide N1/L (X-CEKLLKELKELLEKG-NH₂, X = Ac-), EGG/N1/L (X = Ac-EGG-), or H-EGG/N1/L (X = H-EGG-). These TASP s showed concentration independent CD spectra that were consistent with α -helical structure (Figure 4.9). Bowl/EGG/N1/L had a significantly lower amount of α -helical structure (as judged by CD) than either Bowl/N1/L or Bowl/H-EGG/N1/L. However, Bowl/EGG/N1/L does show a near-UV CD signal, induced as a result of the bowl (or disulfide) chromophore being in an asymmetric environment.

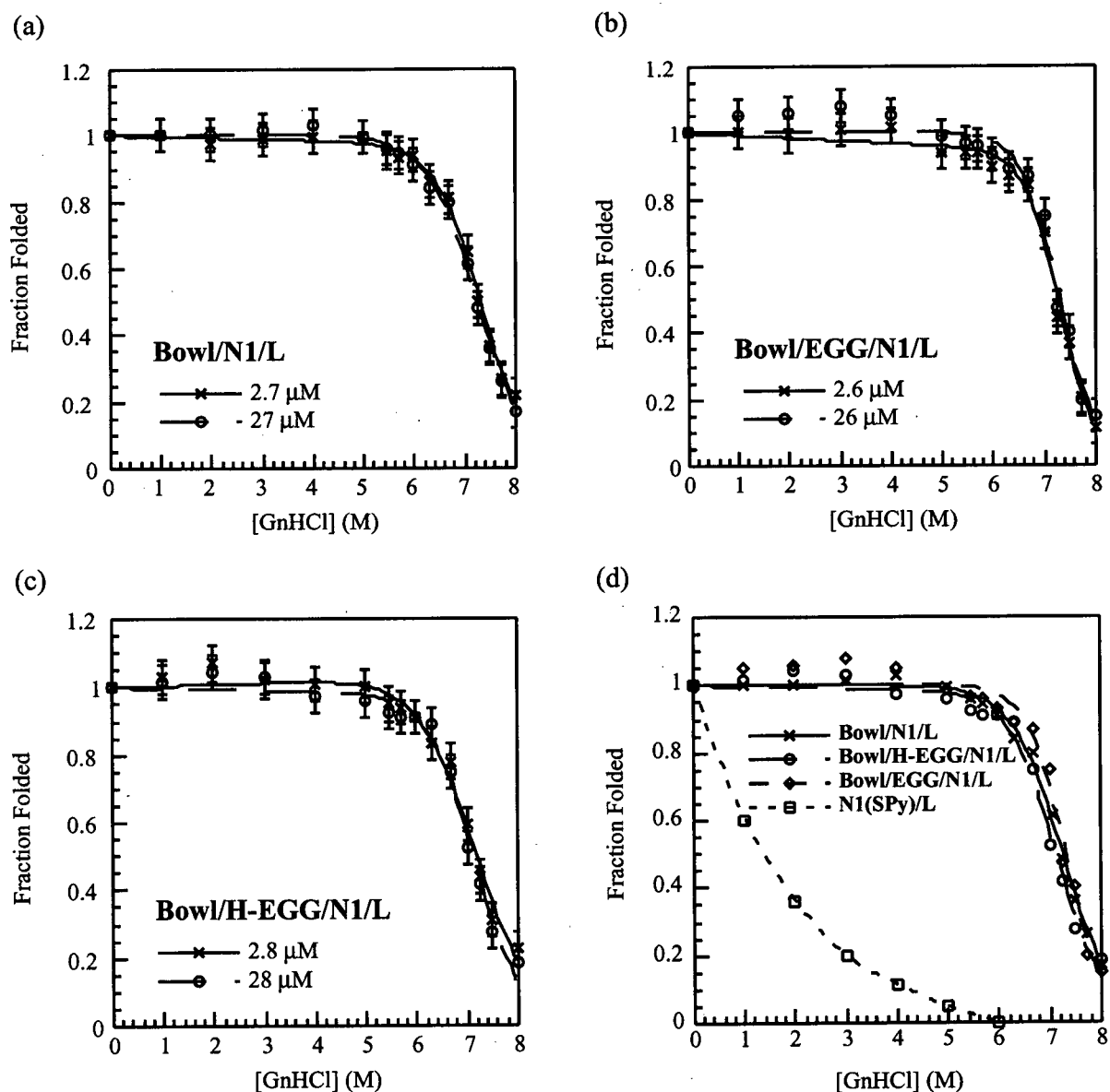
The presence of salt (2M KCl) increased the helicity of Bowl/N1/L (+ 21%), Bowl/EGG/N1/L (+ 51%), and Bowl/H-EGG/N1/L (+ 9%). These increases in helicity are significantly higher than those seen for the CTB three-helix bundle TASP s with the same peptide sequences; therefore, it is unlikely that these structural changes are intrinsic to the peptide alone, but are related either to the template/linker structure, or different tertiary interactions present in four-helix bundle structure. Salt screens electrostatic interactions, and also enhances hydrophobic bundling: Thus, screening of electrostatic forces associated with the bowl, linker, and/or the helices alters the structure of these TASP s. Potential sources of these electrostatic forces will be discussed later in this chapter (Section 4.11).

Figure 4.9: (a) Full and (b) Near-UV, CD spectra of **Bowl/N1/L** (54 μ M, dotted line), **Bowl/EGG/N1/L** (60 μ M, dashed line), and **Bowl/H-EGG/N1/L** (56 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C. NOTE: CD spectra of **Bowl/N1/L** and **Bowl/H-EGG/N1/L** are overlapping.



In the presence of low concentrations of GnHCl (a salt), the helicities of **Bowl/N1/L**, **Bowl/EGG/N1/L**, and **Bowl/H-EGG/N1/L** all increased by similar amounts to those described in the presence of 2 M KCl. Therefore in order for the TASPs to best describe a two-state unfolding transition, the GnHCl-induced unfolding experiments for these TASPs, were run in 100 mM KCl. Although this altered the pre-transition shape of the GnHCl-induced unfolding curves compared to those run without salt, the curves around the unfolding transition point were similar. The GnHCl-induced unfolding curves (in 100 mM KCl) of **Bowl/N1/L**, **Bowl/EGG/N1/L**, and **Bowl/H-EGG/N1/L** do not appear to be significantly dependent on TASP concentration, consistent with monomers near the transition point of unfolding (Figure 4.10).

Figure 4.10: GnHCl-induced unfolding curves for; (a) **Bowl/N1/L** (27 μ M, \circ , dashed line; 2.7 μ M, \times , solid line). (b) **Bowl/EGG/N1/L** (26 μ M, \circ , dashed line; 2.6 μ M, \times , solid line). (c) **Bowl/H-EGG/N1/L** (28 μ M, \circ , dashed line; 2.8 μ M, \times , solid line). (d) **Bowl/N1/L** (2.7 μ M, \times , solid line), **Bowl/EGG/N1/L** (3.6 μ M, \circ , long dashed line), **Bowl/H-EGG/N1/L** (2.8 μ M, \diamond , medium dashed line), compared to peptide **N1(Spy)/L** (10.0 μ M, \square , short dashed line). All spectra recorded in pH 7.0 phosphate buffer (10 mM) and 100 mM KCl at 25 $^{\circ}$ C.



Estimation of the global conformational stability of these TASP s (Table 4.3), has significant error as they were not fully denatured at 8 M GnHCl, therefore the post-transitional curve has been estimated (by the fitting program), also we make the assumption that all TASP s are monomers at the unfolding transition point. **Bowl/N1/L** and **Bowl/H-EGG/N1/L** appear to have similar stability (within error), **Bowl/EGG/N1/L** appears to be marginally more stable of these three TASP s (by about 2 kcal mol⁻¹). This extra stability may result from the stabilization of the helix macrodipole due to presence of negatively charged glutamic acid residues at the N-terminus of the helix.

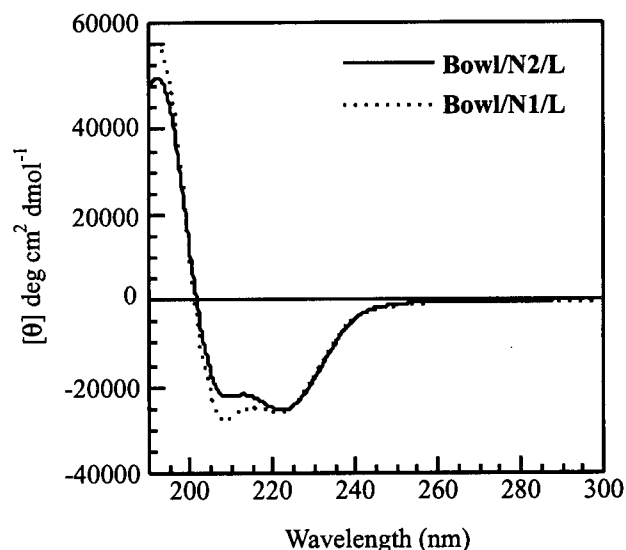
Table 4.3: Thermodynamic evaluation of **Bowl/N1/L** variants, calculated from their GnHCl-induced denaturation. This data assumes that the TASP s are monomers at the unfolding transition point.

TASP	Concentration (μ M)	[GnHCl] _{0.5} (M)	<i>m</i> kcal mol ⁻¹ M ⁻¹	ΔG° H ₂ O kcal mol ⁻¹
Bowl/N1/L	2.7	7.3 \pm 0.1	1.2 \pm 0.1	8.8 \pm 0.5
Bowl/N1/L	27	7.3 \pm 0.1	1.3 \pm 0.1	9.5 \pm 0.6
Bowl/EGG/N1/L	3.6	7.3 \pm 0.1	1.6 \pm 0.2	11.5 \pm 1.5
Bowl/EGG/N1/L	36	7.3 \pm 0.1	1.7 \pm 0.1	12.5 \pm 1.1
Bowl/H-EGG/N1/L	2.8	7.2 \pm 0.1	1.3 \pm 0.1	9.2 \pm 1.0
Bowl/H-EGG/N1/L	28	7.2 \pm 0.1	1.1 \pm 0.1	7.6 \pm 0.7

4.8 Bowl/N2/L

This TASP contained four of the N2/L peptides (Ac-CEELLKKLEELLKKG-NH₂), linked to the bowl via disulfide bonds. The concentration independent CD spectrum of **Bowl/N2/L** was consistent with an α -helical structure (Figure 4.11), and was unaltered in the presence of salt (as was the analogous **CTB/N2/L**).

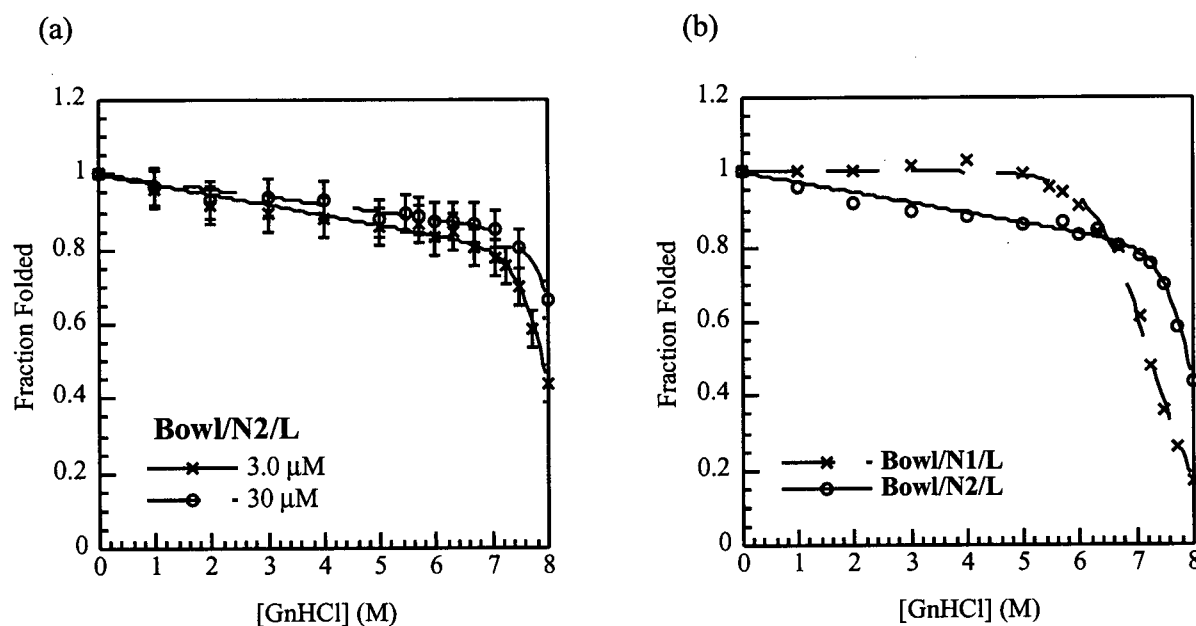
Figure 4.11: CD spectrum of **Bowl/N2/L** (60 μ M, solid line), compared to **Bowl/N1/L** (54 μ M, dotted line), in pH 7.0 phosphate buffer (10 mM) at 25 °C.



The GnHCl-induced unfolding curves of **Bowl/N2/L** (Figure 4.12) at varying TASP concentrations appear to be different, consistent with a self-associating species in solution.

Analysis of sedimentation equilibrium ultracentrifugation data of **Bowl/N2/L** (100 μM TASP concentration, see experimental Section 4.12.2) gives an average molecular weight of 8700 ± 1400 in solution, compared to the calculated molecular weight of 8027. This TASP can be described as a monomer within the uncertainty associated with the data analysis. However, a better fit to the data is obtained when **Bowl/N2/L** is described as a weakly associating monomer-dimer at equilibrium (with a calculated association constant of $1000 \pm 300 \text{ M}^{-1}$) in pH 7.0 phosphate buffer (50 mM).

Figure 4.12: GnHCl-induced denaturation curves of (a) **Bowl/N2/L** (30 μM , O, dashed line; 3.0 μM , x, solid line), and (b) **Bowl/N2/L** (3.0 μM , O, solid line) versus **Bowl/N1/L** (2.8 μM , x, dashed line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25 °C.



Analysis of the GnHCl-induced unfolding curve of **Bowl/N2/L** showed it to be more stable than **Bowl/N1/L** (Table 4.4). This is either due to the intrinsic stability of the **N2/L** peptide versus the **N1/L** peptide (already described in Section 4.2), or a higher association state of **Bowl/N2/L** versus **Bowl/N1/L** in pH 7.0 buffer solution.

The global conformational stability of **Bowl/N2/L** was also compared to the four-helix bundle bowl TASP **ArBowl/N2/L**, a TASP previously studied in our group. The structural differences between these two TASPs (Figure 4.13) are that **ArBowl/N2/L** contains an aryl- as opposed to a benzyl-thiol bowl template, also that its constituent peptides (which do not contain a cysteine residue) are linked via a thioether bond to the N-terminus of their peptide backbones (see section 3.3.2 for more details). **Bowl/N2/L** appears to have a greater conformational stability (about 5 kcal mol⁻¹) than **ArBowl/N2/L**. This difference in stability may be attributed to a higher degree of hydrophobic side chain burial in **Bowl/N2/L** versus **ArBowl/N2/L**, a structural facet of each TASP that is manifested by the different template-to-helix linkers. Both TASPs are not fully unfolded in 8 M GnHCl, therefore there are larger errors in the calculation of their conformational stability. Also this data assumes that they are monomers at the unfolding transition point, sedimentation equilibrium studies on **ArBowl/N2/L** show it to be in a monomer-dimer equilibrium in solution, similar to the self-association observed for **Bowl/N2/L**.

Figure 4.13: Structures of **ArBowl/N2/L** (peptide linked via N-terminus of polypeptide backbone), and **Bowl/N2/L** (peptide linked via N-terminal cysteine side chain).

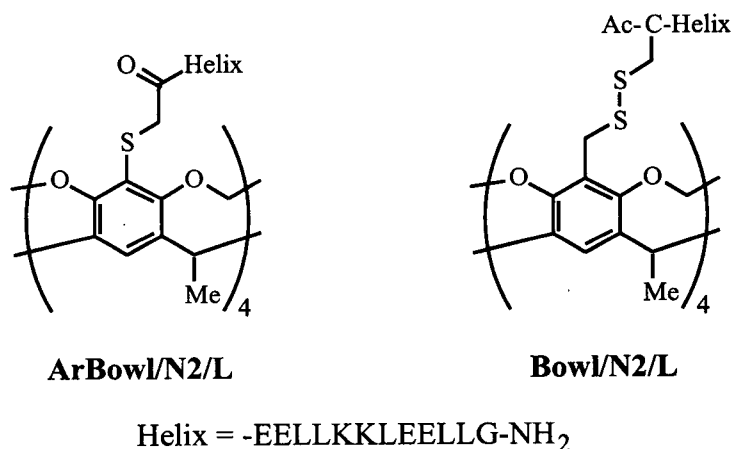
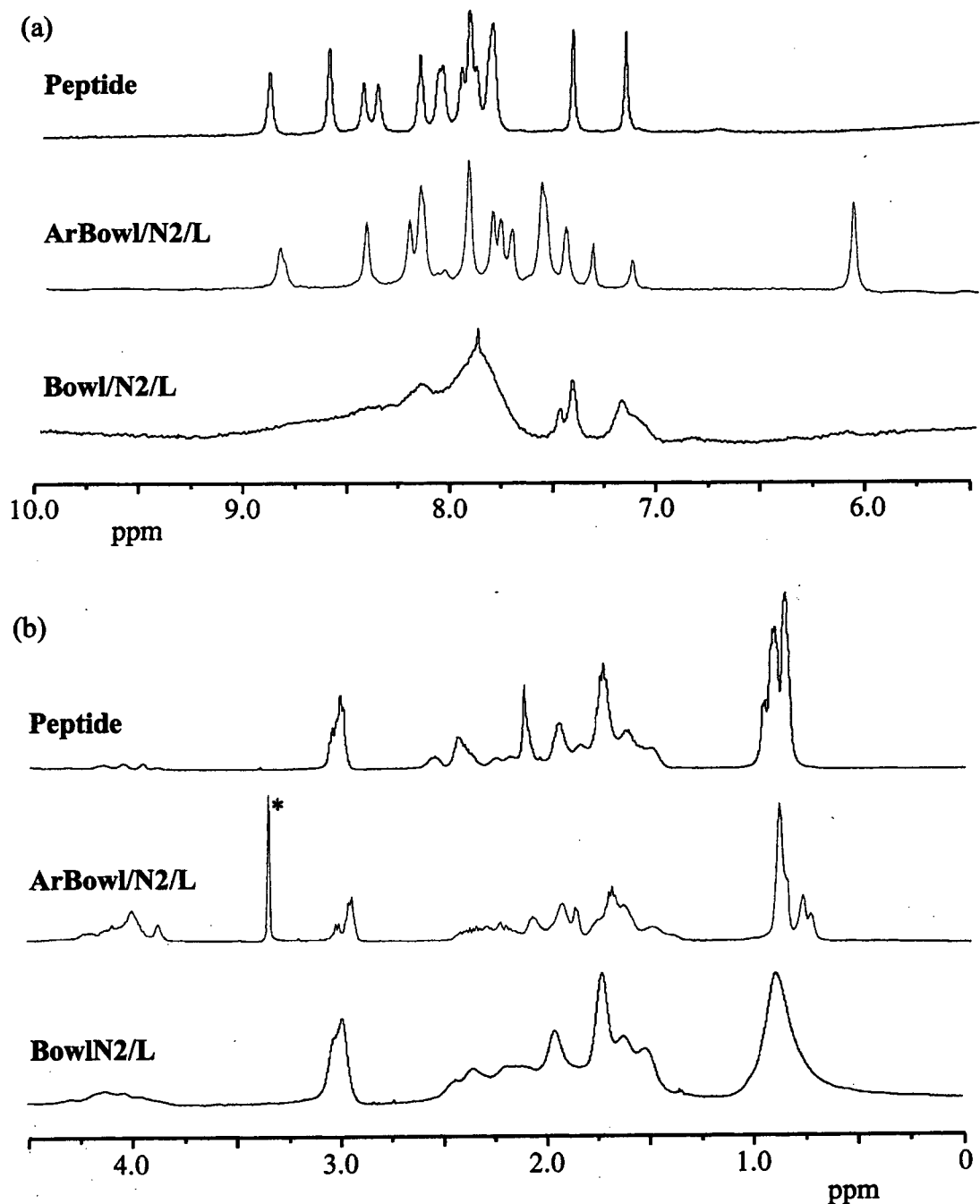


Table 4.4: Thermodynamic evaluation of **Bowl/N2/L** compared to **Bowl/N1/L** and **ArBowl/N2/L**, calculated from their GnHCl-induced denaturation. This data assumes that these TASP are monomers at the unfolding transition point.

TASP	Concentration (mM)	[GnHCl] _{0.5} (M)	<i>m</i> kcal mol ⁻¹ M ⁻¹	ΔG°H ₂ O kcal mol ⁻¹
Bowl/N1/L	2.7	7.3 ± 0.1	1.2 ± 0.1	8.8 ± 0.5
Bowl/N1/L	27	7.3 ± 0.1	1.3 ± 0.1	9.5 ± 0.6
Bowl/N2/L	3.0	7.9 ± 0.1	2.0 ± 0.2	16.2 ± 1.6
Bowl/N2/L	30	>8.0	1.9 ± 0.4	16.2 ± 3.3
ArBowl/N2/L	3.0	>8.0	1.3 ± 0.2	10.4 ± 1.1
ArBowl/N2/L	30	>8.0	1.4 ± 0.2	11.9 ± 1.5

The NMR spectrum of **Bowl/N2/L** exhibits broad peaks, when compared to both the peptide (Ac-EELLKKLEELLKKG-NH₂), and the analogous **ArBowl/N2/L** TASP (Figure 4.14). Therefore, **Bowl/N2/L** appears to have molten globule structure. As mentioned previously, there are many examples of *de novo* designed proteins with all leucine hydrophobic cores, that exist as molten globules, so this result was not unexpected. However, the NMR spectrum of the **ArBowl/N2/L** TASP exhibits sharp, well dispersed signals. **ArBowl/N2/L** was originally interpreted as having characteristics of “native-like” structure,^{270,271} and although this may be the correct interpretation one cannot exclude so-called “fast exchange” ($\gg 10^3 \text{ s}^{-1}$) on the NMR time-scale as the reason for its sharp NMR spectrum: The constituent helices of **ArBowl/N2/L** may be moving rapidly, giving rise to an “averaged” NMR spectrum. Unfortunately, unlike the self-associating peptide, it is not possible to do a NMR study on the concentration dependence of the chemical shifts of **ArBowl/N2/L** (for which the constituent peptides are already “self-assembled” by the template) to ascertain whether its structure is consistent with fast exchange. However, a hydrogen/deuterium amide exchange experiment (a technique that is used to assay the conformational stability of a protein³⁹¹) on **ArBowl/N2/L**, indicated it had characteristics of a molten globule. Therefore, it is possible that **Ar/Bowl/N2/L** exists as a molten globule with a rapidly ($\gg 10^3 \text{ s}^{-1}$) inter-converting tertiary structure. The increased stability of its helical structure (relative to the peptide) is due to non-specific hydrophobic interactions, enhanced by pre-organization of the helices on the template. This type of structure may be an artifact of the TASP approach, and may parallel the type of non-specific stabilizing interactions observed in early intermediates in the folding of proteins,⁶⁹ or in the quaternary structure of both *de novo* designed,²⁰⁶ and natural proteins⁷² (all studied by NMR).

Figure 4.14: NMR spectra of **Bowl/N2/L** (~0.3 mM) compared to peptide **Ac-EELLKKLEELLKKG-NH₂** (~2.2 mM), and **ArBowl/N2/L** (~0.3 mM), at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-Hs, aromatic-Hs) region, 10.0-5.5 ppm (magnified), and (b) Upfield (aliphatic-Hs) region, 4.5-0.0ppm.

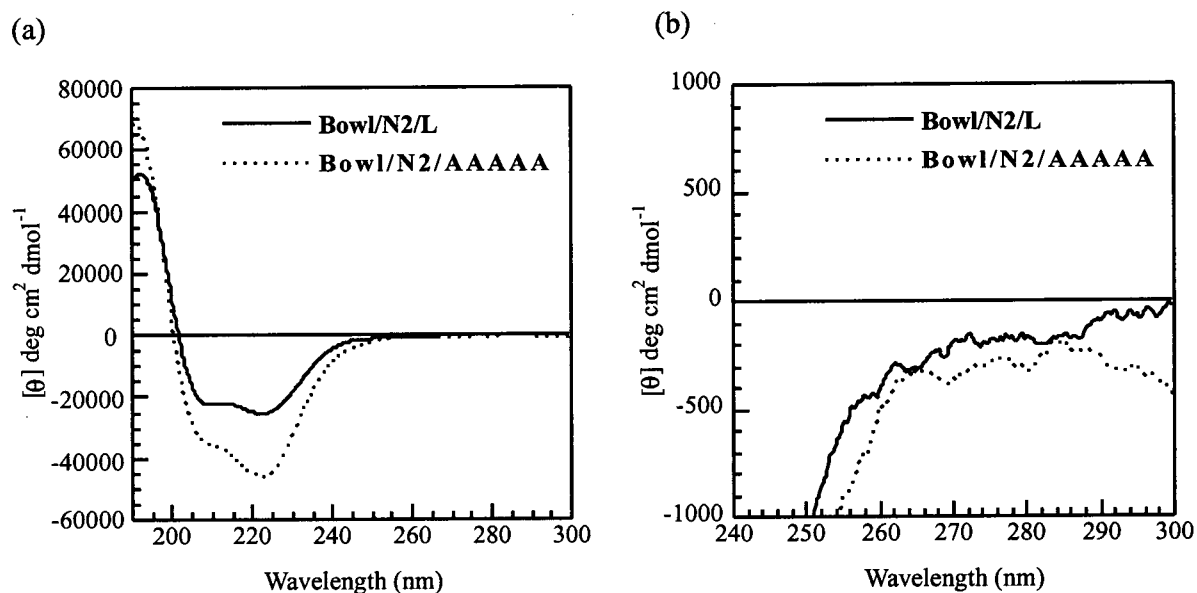


4.9 Bowl/N2/AAAAA

This TASP was similar to **Bowl/N2/L**, except all the leucine residues in the peptide were substituted by alanine (peptide **N2/AAAAA**, Ac-CEEAAKKAEEEAAKKG-NH₂). The CD spectrum of **Bowl/N2/AAAAA** (Figure 4.15) was highly helical. **Bowl/N2/AAAAA** did exhibit a concentration dependent CD spectrum, with higher helicity at increasing concentration of TASP. The cause of this self-association is unlikely to be solvent exposed hydrophobic leucine residues in the sequence; there are none! Therefore, self-association (in this case) appears to be driven by electrostatic interactions between the TASPs. Association of the hydrophobic templates in aqueous solution is still possible, but this would not explain the change in helix structure upon association.

In the presence of salt (2 M KCl), **Bowl/N2/AAAAA** loses about 8% of its helicity, but still displays a concentration dependent CD spectrum, indicating that non-electrostatic forces play a role in its self-association. The loss in helicity in the presence of salt may result from inhibiting the self-association of **Bowl/N2/AAAAA**, where self-association is driven by inter-TASP electrostatic interactions. The salt-induced loss in helicity may also result from screening the favourable (E to K) intra-helical salt bridges (thus, destabilizing the helix). The presence of these intrinsic intra-helix interactions would become more important in helices with weak tertiary interactions (*i.e.* in the absence of a stabilizing hydrophobic core).

Figure 4.15: CD spectrum of **Bowl/N2/AAAAA** (58 μ M, dotted line), compared to **Bowl/N2/L** (62 μ M, Solid line), in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.

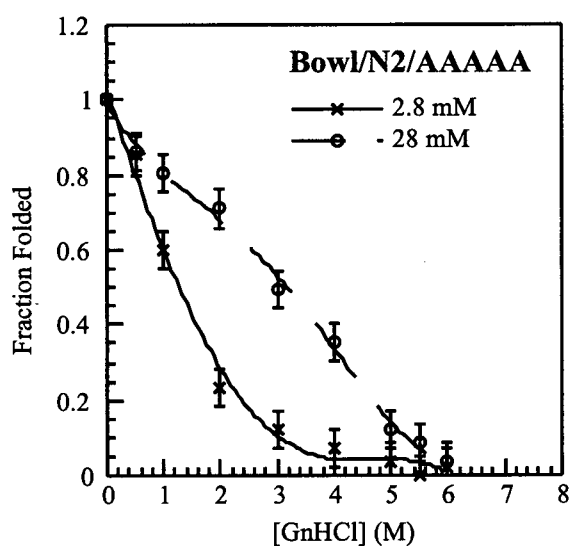


The GnHCl-induced unfolding curves of **Bowl/N2/AAAAA** were concentration dependent (Figure 4.16). The unfolding curves were also dissimilar when run in the presence of salt (up to 0.5 M KCl, data not shown).

Analysis of data from the sedimentation equilibrium study on **Bowl/N2/AAAAA** (100 μ M TASP concentration, see experimental Section 4.12.3), combined with information from the CD studies, indicate that it exists in a monomer-dimer equilibrium in pH 7.0 phosphate buffer (50 mM). The experimentally estimated average molecular weight from the sedimentation equilibrium study was 8600 ± 400 (the calculated molecular weight is 7185). This observation indicates that there was weak self-association that was driven by non-electrostatic forces. When the sedimentation equilibrium data for **Bowl/N2/AAAAA** is

fit to a monomer-dimer species in solution, the calculated association constant is $700 \pm 200 \text{ M}^{-1}$. Thus, the self-association of **Bowl/N2/AAAAA** occurs to a lesser degree than **Bowl/N2/L** (700 ± 200 versus $1000 \pm 300 \text{ M}^{-1}$), but unlike **Bowl/N2/L**, this self-association has a noticeable effect on the structure of the TASP's helices (assessed by CD spectra, a technique that monitors secondary structure).

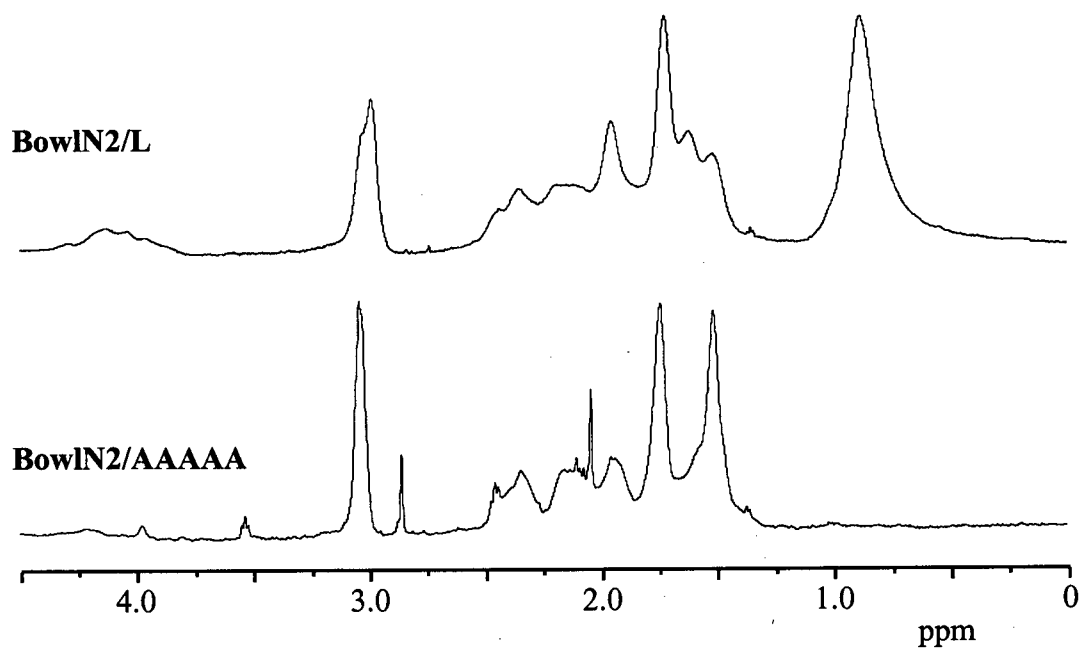
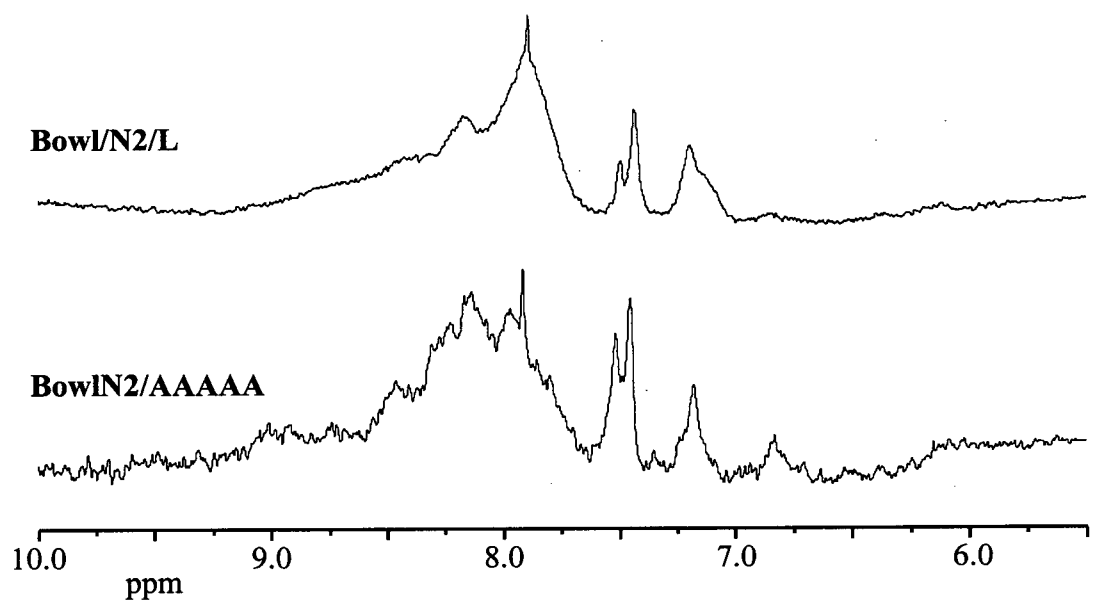
Figure 4.16: GnHCl-induced denaturation curves of **Bowl/N2/AAAAA** (28 μM , \circ , dashed line; 2.8 μM , \times , solid line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}\text{C}$.



The global conformational stability of **Bowl/N2/AAAAA** was not calculated from its GnHCl-induced unfolding curves, because it does not appear to undergo a co-operative unfolding transition. The unfolding curves for **Bowl/N2/AAAAA** at low concentration, resemble those seen for isolated peptides, thus, confirming the importance of a hydrophobic core for TASP stability.

The NMR spectrum of **Bowl/N2/AAAAA** was also broad (Figure 4.17). The source of the line broadening could result from interaction between the helices in a non-specific manner, resulting in a multitude of structures with similar, low energy conformations. This is akin to a molten globule, and the template is responsible for the high degree of interaction between the helices. It is interesting to note that the helices are not in “fast exchange” even though there should be little hydrophobic interaction between the alanine-containing faces of the helices. The motion of the helices may be slowed down to appear molten-globule like by inter-helical electrostatic interactions, or this may be a feature of the template-to-peptide disulfide linker. Another source of the line-broadening in the NMR could be TASP self-association, although this has not been seen for other self-associating TASPs made in our group.^{270,271}

Figure 4.17: NMR spectra of **Bowl/N2/AAAAA** (~0.3 mM) compared to **Bowl/N2/L** (~0.3 mM), at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-Hs, aromatic-Hs) region, 10.0-5.5 ppm (magnified) and (b) Upfield (aliphatic-Hs) region, 4.5-0.0 ppm.

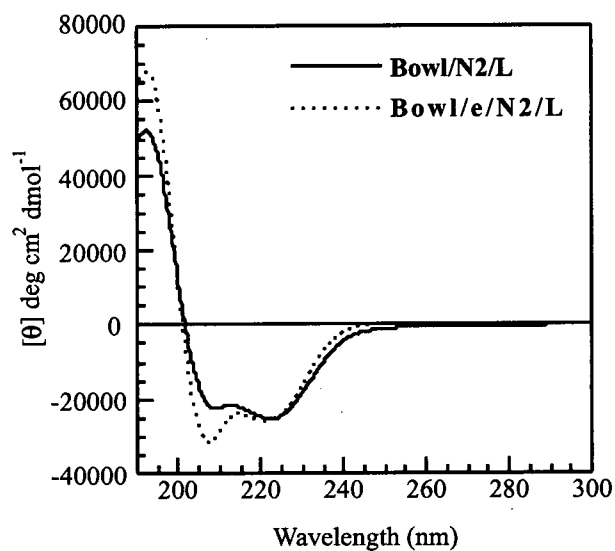


4.10 Bowl/e/N2/L

This TASP used the same peptide as **Bowl/N2/L** (*i.e.* peptide **N2/L**, Ac-CEELLKKLEELLKKG-NH₂), but the peptide was linked via a thioether bond, instead of the usual disulfide bond (see Section 3.3.2 for more details). Briefly, **Bowl/e/N2/L** was synthesized by coupling the tetra-benzylbromide cavitand **2a** with the cysteine side chain of peptide **N2/L**, in the presence of DIPEA base. The backbone of the attached peptides would be closer to the cavitand template, and therefore closer to each other: Thus, the helical bundle of **Bowl/e/N2/L** would possess a more tightly bound core than **Bowl/N2/L**, but the template-to-peptide linker would have different physical characteristics to the disulfide linkers employed so far.

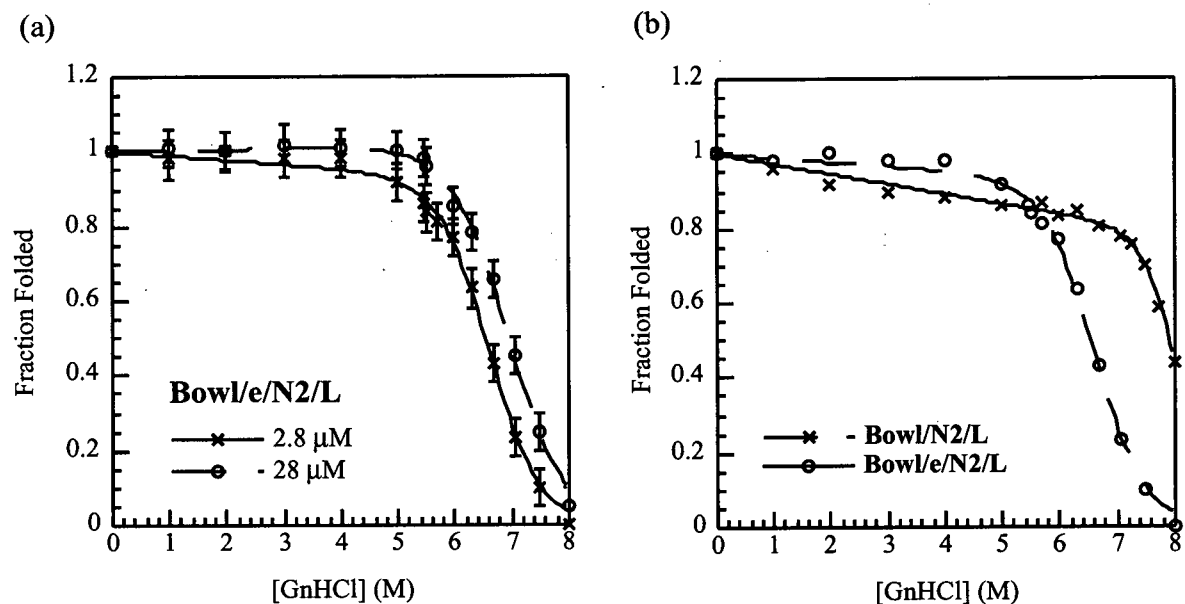
The CD spectrum of **Bowl/e/N2/L** indicated it was helical (Figure 4.18). The concentration independent helicity of this TASP was also unaffected in the presence of salt.

Figure 4.18: CD spectrum of **Bowl/e/N2/L** (58 μ M, dotted line), compared to **Bowl/N2/L** (60 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



The GnHCl unfolding curves at high and low concentrations of TASP **Bowl/e/N2/L** were not superimposable within experimental error (Figure 4.19), consistent with a self-associating species in solution. This is not unexpected, considering the observed self-association of **Bowl/N2/L**, and that **Bowl/e/N2/L** has a shorter template-to-helix linker.

Figure 4.19: GnHCl-induced denaturation curves of (a) **Bowl/e/N2/L** (28 μ M, O, dashed line; 2.8 μ M, x, solid line); and (b) **Bowl/e/N2/L** (2.8 μ M, O, dashed line) versus **Bowl/N2/L** (3.0 μ M, x, solid line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25 °C.



Analysis of the GnHCl-induced denaturation curves of **Bowl/e/N2/L** show it to have a lower conformational stability than **Bowl/N2/L** (Figure 4.19, and Table 4.5). This may be a result of overcrowding of the helices close to the bowl, resulting in less favourable inter-helix packing, and therefore fewer stabilizing hydrophobic interactions in the core of the TASP. A lower magnitude of the “*m*” value obtained for **Bowl/e/N2/L**, when compared to **Bowl/N2/L**, indicates that there is less buried hydrophobic surface in this TASP, consistent with the helices “splaying” outward from the template (see Section 4.11 for more details).

Table 4.5: Thermodynamic evaluation of **Bowl/e/N2/L** (compared to **Bowl/N2/L**), calculated from its GnHCl-induced denaturation. This data assumes that TAsPs are monomers at the unfolding transition point.

TASP	Concentration (mM)	[GnHCl] _{0.5} (M)	<i>m</i> kcal mol ⁻¹ M ⁻¹	ΔG°H ₂ O kcal mol ⁻¹
Bowl/N2/L	3.0	7.9 ± 0.1	2.0 ± 0.2	16.2 ± 1.6
Bowl/N2/L	30	>8	1.9 ± 0.4	16.2 ± 3.3
Bowl/e/N2/L	2.6	6.5 ± 0.1	1.4 ± 0.1	9.0 ± 0.6
Bowl/e/N2/L	26	7.0 ± 0.1	1.2 ± 0.1	8.5 ± 0.5

4.11 Conclusions for Chapter 4

These initial CTB and Bowl TASP's were found to have helical structure, consistent with the design. Furthermore, the TASP's resistance to GnHCl-induced unfolding was greatly increased relative to the individual peptide strands, indicating that template-promoted helical bundling is a major source of stability. The importance of this template-enhanced hydrophobic bundling to TASP stability was further highlighted by **Bowl/N2/AAAAA**, which without leucine residues in its core, did not unfold in a cooperative manner, and was far less stable than TASP's that contained a leucine core. This TASP also demonstrates that that template alone is insufficient to stabilize helix structure, the aforementioned hydrophobic bundling is essential to the overall stability of tertiary structure.

The self-association of the TASP's described in this chapter was sometimes evident from their concentration dependent GnHCl-induced unfolding curves. However, later in the thesis we will see that self-associating TASP's may manifest concentration independent GnHCl-induced denaturation curves (Chapter 5). For example, concentration dependent GnHCl-induced unfolding curves would not be observed in the case where a TASP is weakly associated, and the association is disrupted in low concentrations of GnHCl with no detectable change in helicity: This would lead to the conclusion that the TASP is monomeric in solution. The most definitive method for detecting self-association is using sedimentation equilibrium ultracentrifugation. Unfortunately UBC is not equipped with a

modern analytical ultracentrifuge, so not every TASP presented in this thesis was analyzed by this technique.

The self-association of the three- and four-helix bundle TASP described in this chapter seems to originate from the helices rather than from the macrocyclic templates, and appears to have both a hydrophobic and an electrostatic component. This self-association may be due to incorrect bundling of the helices, which would result from a template-to-helix linker that is too short, or from a poorly designed helix sequence. The importance of the linker flexibility between template and helix to allow optimal intra-helix bundling on a cavitand bowl TASP has been highlighted previously by our group,^{270,271} and will be described in the following chapter. For the TASP investigated in this chapter the template-bound cysteine residue was adjacent to the "helix" (e.g. peptide Ac-CEELLKKLEELLKKG-NH₂), and as a result the helices may be too tightly crowded around the base of the template and "splay out" from the closest point of contact (Figure 4.20a): Thus, some of the helix hydrophobic surface is exposed to the aqueous solvent, aiding in self-association of the TASP. Steric crowding at the template may be related to the linker, template and/or the peptide sequence (e.g. the leucine residues in the core are too bulky).

The templates, apart from their structural rigidity, also possess a dipole moment that may be unfavourable for the conformation of the template-to-helix linker (Figure 4.20b). The ether oxygens on the CTB template are free to rotate, thereby reducing the potential for unfavourable dipolar interactions with the sulfur atoms of the disulfide bond. In the case of the cavitand bowl template the lone pairs of electrons on the acetal oxygen bridging groups

can potentially interact with those of the sulfur atom(s) of the disulfide bond, but this electrostatic interaction was not found to be highly significant according to computer modelling.³⁶⁰

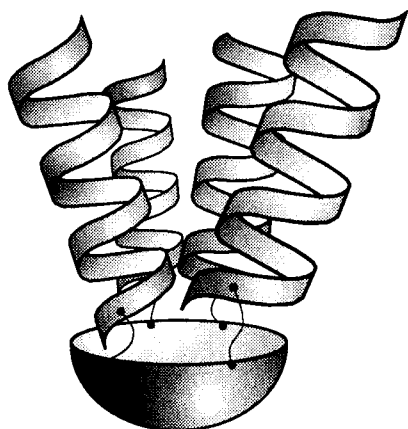
One other potential source of structural rigidity in the linker is the preference of disulfide bonds to have a dihedral angle of $\pm 90^\circ$ (Figure 4.20c). This dihedral angle is stable by about 10 kcal mol⁻¹ over other conformations, and is thought to be driven by electrostatic considerations (thus, screened by salt).³⁹²

The electrostatic component to self-association may also be related to the patterning of oppositely charged side chains (E and K) that are exposed to the solvent (Figure 4.20d) by each individual helix. This regular patterning of the oppositely charged side chains (E and K) may result in a pair of TASP s interacting in an anti-parallel manner, again this interaction would be screened in the presence of salt.

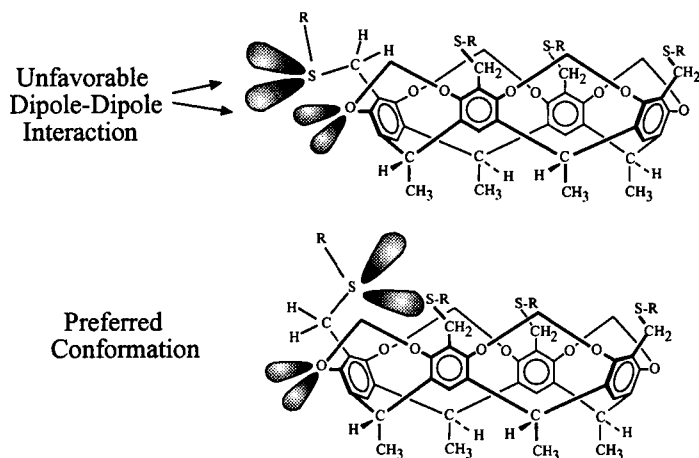
The structural features of the template-to-helix linker may, or may not have a significant effect on the structure/association of the helical bundle TASP s. However, the addition of salt seemed to have a greater effect on the structure of the bowl TASP s when compared to the analogous CTB TASP s, which suggests that the bowl template possesses more unsuitable electrostatic interactions, and/or that there were significant inter-helical electrostatic interactions (e.g. E to K) present in the four-, versus the three-helix bundle.

Figure 4.20: (a) Schematic diagram of a cavitand TASP with too short a linker that results in steric crowding of the helices around the template. (b) The dipolar interactions between the template and linker. (c) The preferred disulfide dihedral angle of $\pm 90^\circ$. (d) Self-association of two TASPs by complementary pairing of oppositely charged side chains.

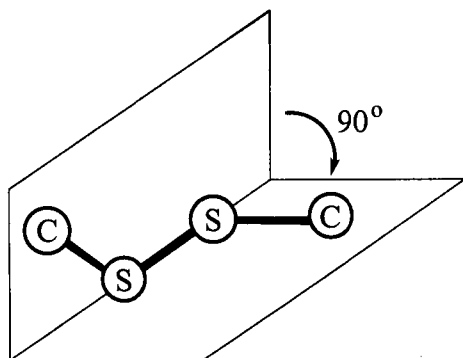
(a) "Splaying Out" of Helices



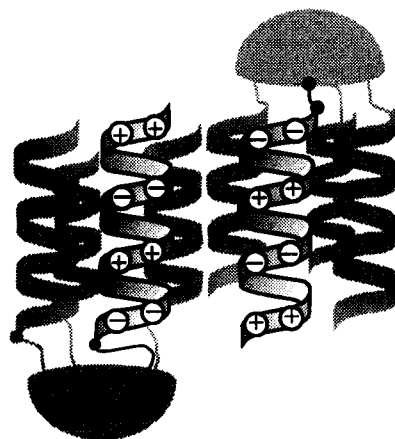
(b) Dipole-Dipole Interactions between Cavitand Bowl and Sulfur



(c) Preferred Dihedral Angle of a Disulfide Bond



(d) Self-Association of TASP via Oppositely Charged Side Chains



CTB and bowl TASP s containing the N1/L peptide increased in helicity to a greater degree than the N2/L TASP s upon the addition of salt. This may be due to the screening of the less favourable intra-helix electrostatic interactions present in the N1/L peptide (see Figure 4.8), or the screening of a less favourable helix macrodipole. The CTB and bowl TASP s containing the N2/L peptide were more stable to denaturation when compared to TASP s with the N1/L peptide, a feature that demonstrates the importance of hydrophilic residues in protein stability (this assumes that the extra stability is not due to a higher degree of self-association).

The EGG/N1/L sequence, which places three or four negative charges in the vicinity of the macrocycle, is especially disfavoured in both the CTB and bowl TASP s, probably due to intra-TASP charge-charge repulsion, resulting in a structural re-organization of the helices to minimize its effect: Thus, the constituent helices are even less favourably oriented for packing, and the degree of inter-TASP helix driven self-association increases.

The stability of the cavitand bowl TASP s were greater than those obtained for the analogous CTB TASP s. This observation is consistent with the burial of more hydrophobic surface in a four- versus three-helix bundle, increasing the global stability of a protein. Also, the CTB TASP s may have some hydrophobic surface exposed to the solvent in their folded state due to "over-packing" of their hydrophobic core: Thus, the free energy change on unfolding (which is a function of the change in solvent-exposed surface upon unfolding) would be lowered, because even in the folded state the CTB TASP s already have some of their hydrophobic surface exposed to the solvent.

NMR studies on **CTB/N2/L**, **Bowl/N2/L**, and **Bowl/N2/AAAAA** are consistent with dynamic, slowly inter-converting conformers, probably a molten globule-like structure. This is not entirely unexpected; *de novo* designed helix bundle proteins that have a degenerate sequence (and hydrophobic cores packed solely with leucine residues) have been found to exist as molten globules. The peptide on its own demonstrated a sharp, well resolved NMR spectrum which is consistent with either a “native-like” protein structure, but is most likely to be a dynamic structure with rapidly exchanging conformations on the NMR time-scale (see Section 4.8). The NMR of **Bowl/N2/L** was compared to that of another bowl TASP **ArBowl/N2/L** (which had a different template-to-helix linker) studied by our group (Section 4.8). Unlike **Bowl/N2/L**, **ArBowl/N2/L** demonstrated a sharp, well resolved NMR spectrum, similar to that observed for the peptide. This TASP was originally interpreted as having “native-like” structure, but we cannot formally rule out the possibility that **ArBowl/N2/L** has a rapidly inter-converting molten-globule-like structure, which may be an “artifact” of our TASP approach.

The major findings of this chapter are that these initial CTB and bowl TASPs fold into helical bundles with enhanced stability that results from template-promoted inter-helical hydrophobic interactions. Unfortunately, these initial TASPs self-associated in solution and lacked the conformational specificity associated with natural proteins, possibly a result of less than optimal inter-helix packing which arose from a template-to-helix linker that was too short: The effect of changing the length of this linker will be described in the next chapter.

4.12 Experimental

4.12.1 CD studies

All CD spectra were run on a Jasco J-710 spectropolarimeter (calibrated using *d*-10-camphorsulfonic acid),³⁹³ using quartz (SUPRASIL – Hellma®) cuvettes of 1 mm and 1 cm path length. The temperature was kept at 25 °C using a Haake FX 10 circulating bath. Samples were allowed 3 min to equilibrate to the desired temperature. Each CD spectrum was the average of three scans, from which the spectrum of the solvent blank was subtracted. Test samples for CD were prepared in duplicate or triplicate from stock TASP/peptide solutions. The concentrations of these stock solutions was obtained by amino acid analysis (Peptide/Protein Service, UBC). Guanidine Hydrochloride denaturation experiments were performed by dilution of an 8.0 M solution of (pH 7.0) GnHCl. The exact concentration of the 8 M GnHCl stock solutions were determined by refractometry.³⁶⁸ A 3 M KCl stock solution was diluted for use in the salt experiments. A 50 mM pH 7.0 phosphate buffer solution was used to adjust the sample solutions to 10 mM concentration of buffer. The pH was measured using a Fisher Scientific Accumet® pH meter model 915.

Raw CD spectra were normalized to mean residue ellipticity $[\theta]$, using the equation:

$$[\theta] = \theta_{\text{obs}} / 10 l c n$$

Where:

θ_{obs} = observed ellipticity (in millidegrees)

l = pathlength (cm)

c = peptide / TASP concentration (M)

n = number of residues in peptide/TASP

Table 4.6: Mean residual ellipticity at 222 nm, $[\theta]_{222}$, of CTB and bowl TASPs presented in Chapter 4.

TASP	Experimental $-[\theta]_{222}$	Theoretical ^a Maximum $-[\theta]_{222}$	Percent Helix
CTB/N1/L	23600	29300	81
CTB/EGG/N1/L	23800	31100	77
CTB-f/H-EGG/N1/L	23600	31100	76
CTB-s/H-EGG/N2/L	23900	31100	77
CTB/N2/L	25300	29300	86
Bowl/N1/L	25900	29300	88
Bowl/EGG/N1/L	18700	31100	60
Bowl/H-EGG/N1/L	26400	31100	85
Bowl/N2/L	25300	29300	86
Bowl/e/N2/L	25600	29300	87
Bowl/N2/AAAAA	46000	29300	157 ^b

a Theoretical helicity is calculated according to reference 394. This calculation includes all residues in the attached peptide, and not just those designed to take part in the “helix”.

b Greater than theoretical helicity! This may result from an error in determining the TASP concentration by amino acid analysis.

4.12.2 Global Conformational Stability of TASP: Calculated by Analysis of GnHCl-Induced Denaturation

This method for analysis assumes that the TASP undergoes a fully reversible, cooperative, two-state unfolding process. A non-linear least squares method³⁹⁵ was used for the analysis of the GnHCl-induced denaturation curves. This method estimates the pre- and post-transitional baselines, and assumes that the free energy of unfolding is a linear function of [GnHCl] according to the function:

$$\Delta G_{\text{obs}} = \Delta G^{\circ} \text{H}_2\text{O} - m[\text{GnHCl}]$$

Where:

ΔG_{obs} = the observed free energy of unfolding at a particular concentration of denaturant.

$\Delta G^{\circ} \text{H}_2\text{O}$ = the free energy of unfolding in the absence of denaturent

m = change in ΔG_{obs} with respect to [GnHCl]³⁹⁶

[GnHCl] = concentration of denaturant.

In order to account for pre- and post-transitional baselines, the data was fit to the following equation:

$$F_{\text{obs}} = F_N (f_n)(1 - a[\text{GnHCl}]) + F_U (1 - f_n)$$

Where:

F_{obs} = function of the fraction folded at a given concentration of GnHCl

F_N = function of the folded state in the absence of GnHCl

F_U = function of the unfolded state (in the cases where the TASP was not fully unfolded at 8 M GnHCl, the post-transitional baseline was estimated)

a = a constant

f_n = the fraction of the TASP in the folded state according to the following formula:

$$f_n = e^{((\Delta G^\circ \text{H}_2\text{O} - m [\text{GnHCl}])/RT)} / [1 + e^{((\Delta G^\circ \text{H}_2\text{O} - m [\text{GnHCl}])/RT)}]$$

Where:

$\Delta G^\circ \text{H}_2\text{O}$ = free energy of unfolding in the absence of denaturant

m = the change in free energy with respect to the concentration of GnHCl

R = universal gas constant

T = temperature

f_n = the fraction of the TASP in the folded state

The value of F_N was normalized to a value of 0.9999, and F_U is (nominally) set to 0.0001. The values of $\Delta G^\circ \text{H}_2\text{O}$, m , and a were determined by nonlinear least-squares analysis using KaleidaGraph 3.08 (Synergy Software). All TASPs were treated as

monomers when calculating their global stability. The presence of GnHCl, salt or methanol has been shown to disrupt (but not always overcome) the self-association of other members of our TASP family,²⁷¹ therefore we make the assumption that they are monomers at the GnHCl concentration required for unfolding. However, any self-association of the TASPs around the unfolding transition point may introduce errors into these calculations.

4.12.3 Sedimentation Equilibrium Ultracentrifugation

Sedimentation Equilibrium studies were performed by Dr. Adam Mezo, in the laboratory of Professor Barbara Imperiali at the Massachusetts Institute of Technology. The data was collected on a temperature-controlled Beckman XL-A analytical ultracentrifuge equipped with photoelectric scanner, at the noted UV wavelengths, 298 K, and a rotor speed of 33000 r.p.m. Samples were allowed to equilibrate for 21-24 h (equilibrium was judged to have been reached when scans taken at 3 h intervals were duplicated).

The raw data (sent as a Simpletext document from M.I.T.) was in the form of absorbance (A) versus radius (r). An initial treatment was to look for a single non-interacting species using (a variant) of the Lamm equation (which describes the movement of single, ideal, molecules in a centrifugal field).³⁹⁷ NOTE: Optical absorbance is used in place of solute concentration in this variant of the Lamm equation (this assumes that the solute obeys the Beer-Lambert law):

$$\delta \ln (A_r) / \delta (r^2) = M(1 - \bar{v} \rho) \omega^2 / 2RT$$

Where:

A_r = absorbance at radius r

r = radius (of measured absorbance)

M = molecular weight

\bar{v} = the partial specific volume of the peptide³⁹⁸

ρ = density of solvent³⁹⁹

ω = the angular velocity of the rotor (radians s^{-1}) = rotor speed (r.p.m.) $\times (2\pi/60)$

R = universal gas constant

T = absolute temperature

Thus, if a single ideal species exists in solution, the plot of $\ln (A)$ versus r^2 should yield a straight line with slope proportional to the molecular weight (M). Deviation from linearity represents either a non-ideal, or reversibly self-associating species (or the presence of impurities). If the species is reversibly self-associating in solution, then it is possible to estimate its oligomeric state by plotting tangents to the curve (for example, plotting a straight line using the first and last 20 data points in the curve), and comparing the ratio of their magnitude. Most of the TASPs in this thesis produced curves from this initial treatment, although some produced straight lines with molecular weights that were double that expected, indicating the major species present in solution to be a dimer.

After this first approximation of the molecular weight (and oligomeric state) of the TASPs, the data was analyzed using MacNONLIN-PPC,⁴⁰⁰ which performs a nonlinear

least-squares analysis.⁴⁰¹ Also, this program gives best estimates for unknown parameters. The major advantage of this method is that the data can be analyzed without further transformation. Also it can test for more complex models if there is an obvious difference between experimental and fitted data. This analysis calculates a series of curves to locate a best-fitting model of the data, successive iterations converge to stable values for the parameters being varied. For least-squared analysis, the differences between the fitted function and the experimental data are squared and summed, the parameters varied so as to minimize this sum, leading to a global minimum. In order to test the resultant parameters, additional statistical and graphical analysis is usually needed. Also, additional sample concentrations and/or rotor speeds should be assayed (this additional analysis was not carried out due to restricted instrument time).

The initial mathematical model for fitting of this data was chosen to describe an "ideal non-interacting species" in solution. For the TAsPs described in this thesis this analysis usually gave higher than expected molecular weights, consistent with a self-associating species in solution. The next mathematical model fitted the data to a monomer-dimer system that was discerned from the plots of $\ln(A)$ versus r^2 . Alternate models were also analyzed (from goodness of fit) to find the best description of the experimental data. The accuracy of the fit was evaluated on the basis of the randomness and magnitude of a plot of the deviation of the data, expressed as the difference between the theoretical curve and the experimental data. A good fit would demonstrate a random distribution of the deviation points on either side of the best fit line, with deviation values of less than ± 0.02 . The program also checks the fit parameters for physical reasonability.

In order to directly relate the data from the analytical ultracentrifuge (with A and r as the dependent and independent variables, respectively), an exponential solution to the Lamm equation can be derived (this fits the data to a single ideal species at equilibrium):

$$A_r = \text{Exp} \{ \ln (A_o) + M\omega^2[(1 - \bar{v} \rho)/(2RT)] (r^2 - r_o^2) \} + E$$

Where:

A_o = the absorbance at a reference radius r_o (usually the meniscus)

A_r = absorbance at radius r

r = radius (of measured absorbance)

r_o = reference radius

M = molecular weight

\bar{v} = the partial specific volume of the peptide

ρ = density of solvent

ω = the angular velocity of the rotor (radians s^{-1}) = rotor speed (r.p.m.)/21600

R = universal gas constant

T = absolute temperature

E = the baseline error correction factor

For self-associating models, the total absorbance results from the sum of the two (or more) species at equilibrium. The software could account for up to four ideal species in equilibrium according to the equation:

$$\begin{aligned}
 A_r = & A_{(\text{monomer}, r_o)} \exp \{ [(1 - \bar{v} \rho) \omega^2 / 2RT] [M(r^2 - r_o^2)] \} \\
 & + A_{(\text{monomer}, r_o)}^{n_2} K_{a,2} \exp \{ [(1 - \bar{v} \rho) \omega^2 / 2RT] n_2 [M(r^2 - r_o^2)] \} \\
 & + A_{(\text{monomer}, r_o)}^{n_3} K_{a,3} \exp \{ [(1 - \bar{v} \rho) \omega^2 / 2RT] n_3 [M(r^2 - r_o^2)] \} \\
 & + A_{(\text{monomer}, r_o)}^{n_4} K_{a,4} \exp \{ [(1 - \bar{v} \rho) \omega^2 / 2RT] n_4 [M(r^2 - r_o^2)] \} + E
 \end{aligned}$$

Where:

$A_{(\text{monomer}, r_o)}$ = the absorbance of the monomer at the reference radius r_o

n_2 = the stoichiometry for species 2

$K_{a,2}$ = the association constant for the monomer- n -mer equilibrium of species 2 (expressed in absorbance units)

n_3 = the stoichiometry for species 3

$K_{a,3}$ = the association constant for the monomer- n -mer equilibrium of species 3

n_4 = the stoichiometry for species 4

$K_{a,4}$ = the association constant for the monomer- n -mer equilibrium of species 4

This formula assumes that the extinction coefficient (absorbance) of the oligomer is the sum of its constituent parts. The monomer molecular weight is often calculated by first running the protein in denaturing conditions (this can be done for a more rigorous treatment of the molecule in question, but the lack of a readily available ultracentrifuge at UBC restricted the number of experiments performed).

The MacNonlin-PPC program, calculates association constant for the monomer-dimer self-association in absorbance units. In order to convert this association constant into the more generally accepted units of M^{-1} we applied the following formula:^{402,403}

$$K_{conc} = K_{abs} (\epsilon l / 2)$$

Where:

K_{conc} = the association constant in M^{-1}

K_{abs} = the association constant in terms of absorbance (estimated by MacNonlin-PPC directly from a best fit curve of a monomer-dimer self-associating system).

ϵ = the extinction coefficient in $cm^{-1} M^{-1}$ (obtained from an absorbance versus radius plot at time = 0)

l = path length in cm

CTB/N2/L (Figure 4.21), **Bowl/N2/L** (Figure 4.22), and **Bowl/N2/AAAAA** (Figure 4.23) are best described as being a monomer-dimer equilibrium in solution, with differing association constants (Table 4.7 includes all values used in the assessment of the sedimentation equilibrium data).⁴⁰⁴

Figure 4.21: Sedimentation equilibrium analysis of CTB/N2/L (100 μ M) at 25 $^{\circ}$ C in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

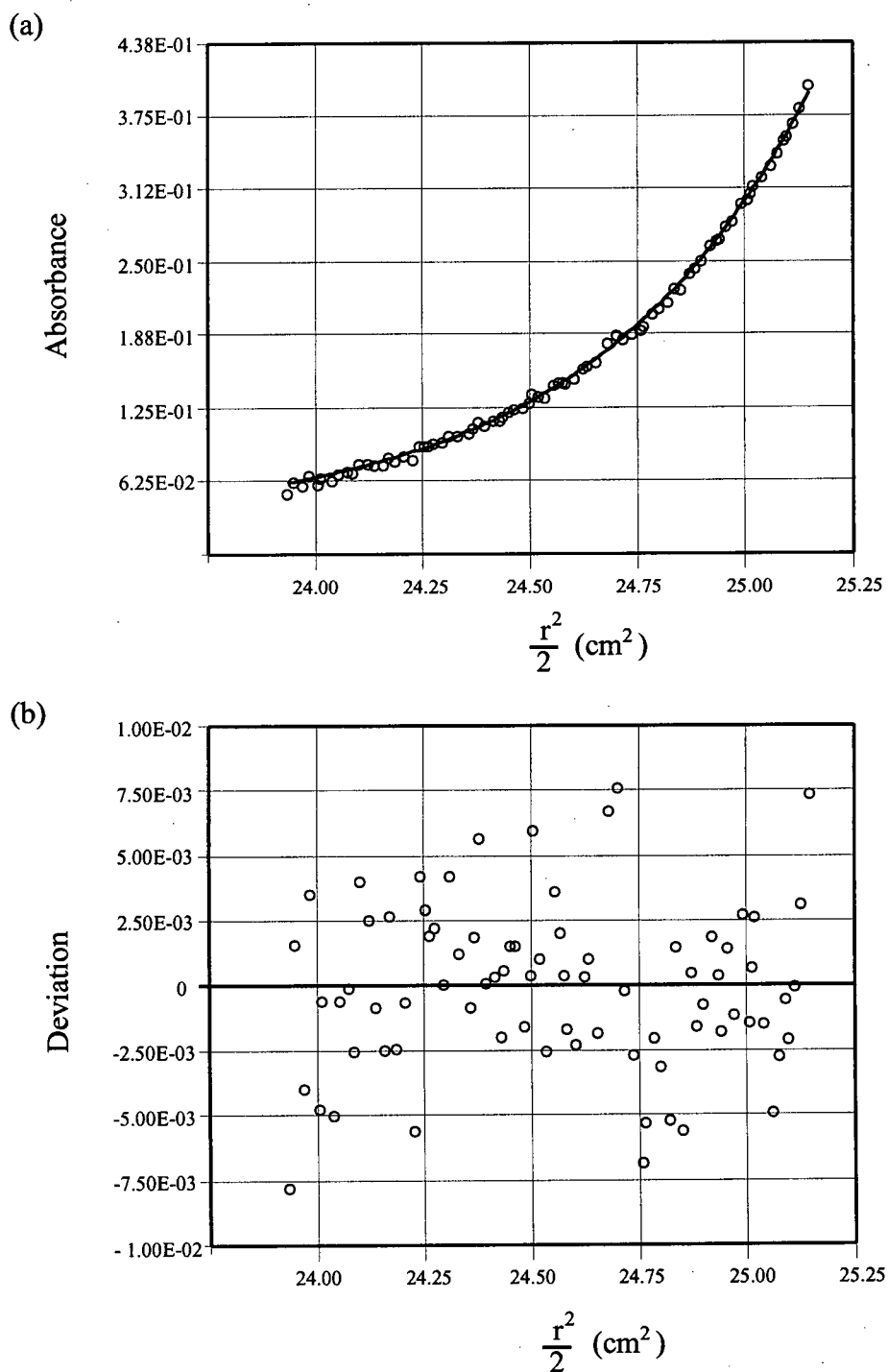


Figure 4.22: Sedimentation equilibrium analysis of **Bowl/N2/L** (100 μM) at 25 $^{\circ}\text{C}$ in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

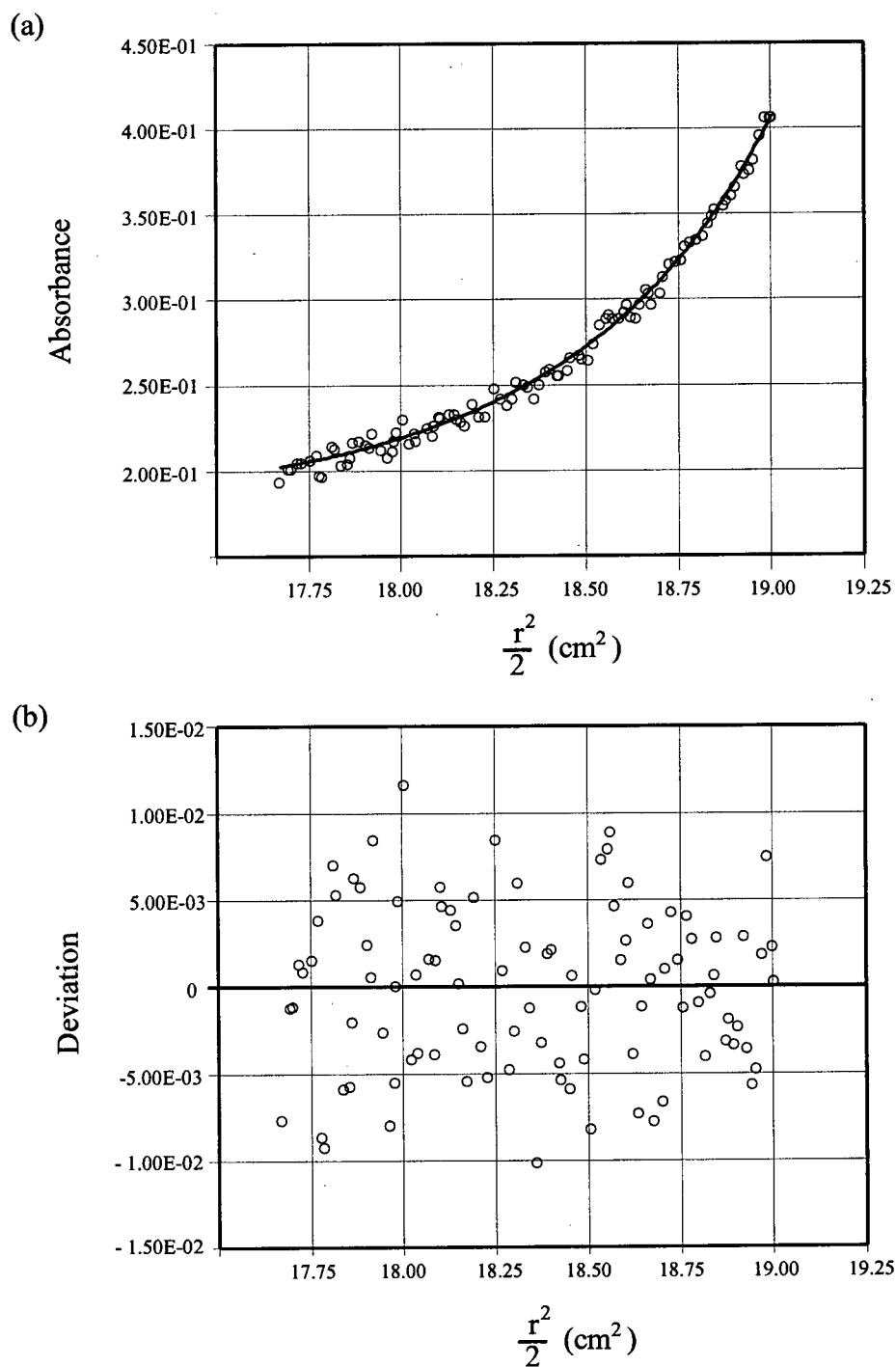


Figure 4.23: Sedimentation equilibrium analysis of **Bowl/N2/AAAAA** (100 μ M) at 25 $^{\circ}$ C in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

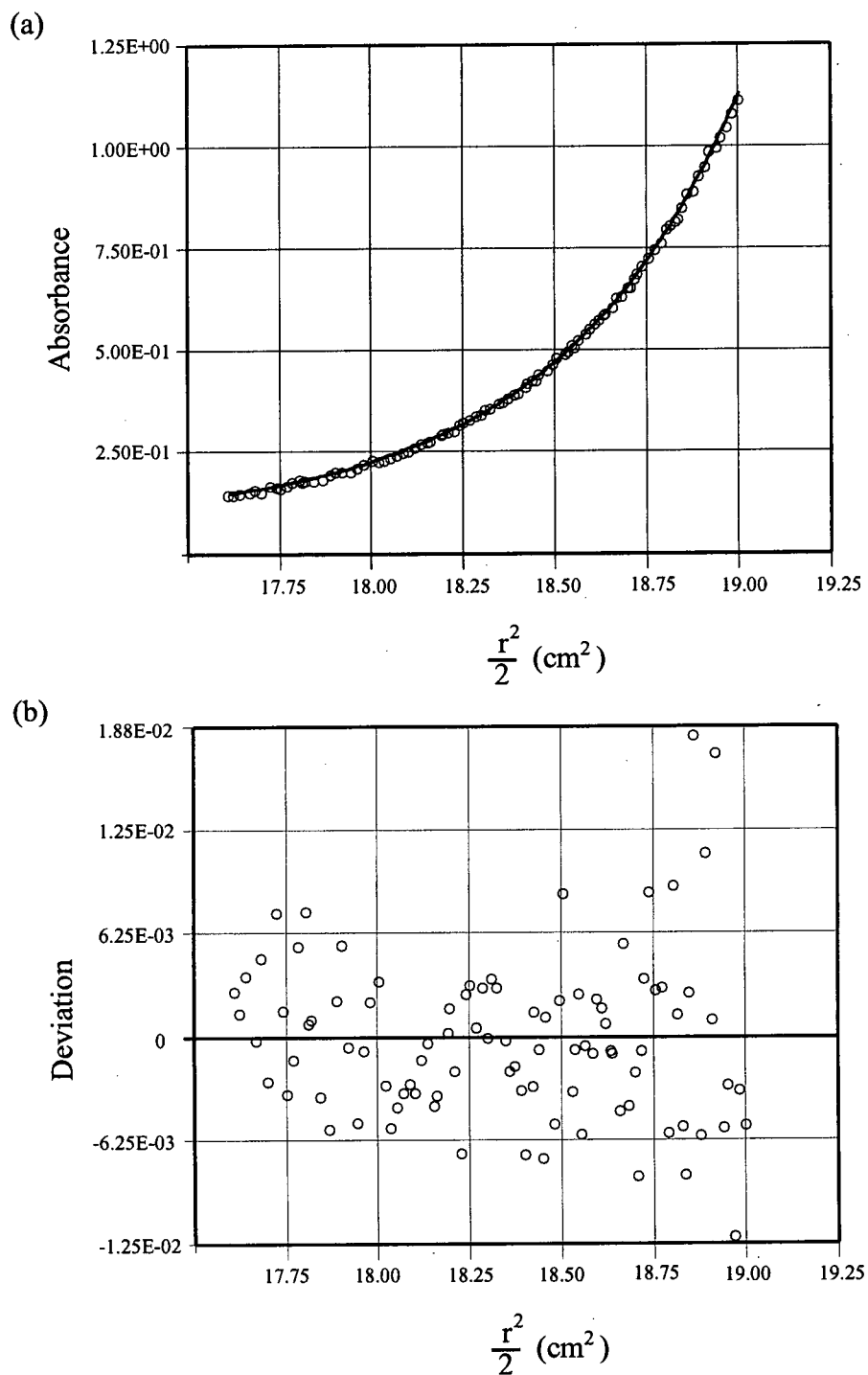


Table 4.7: Data used for, and obtained from, linear fit of sedimentation equilibrium data.

	CTB/N2/L	Bowl/N2/L	Bowl/N2/AAAAA
Mw [calculated] (Da)	5934	8027	7185
R (erg. mol ⁻¹ K ⁻¹)	8.314 x 10 ⁷	8.314 x 10 ⁷	8.314 x 10 ⁷
T (K)	298	298	298
\bar{v} (mL g ⁻¹)	0.7793	0.7793	0.7260
ρ (g mL ⁻¹)	1.001	1.001	1.001
ω (rad. s ⁻¹)	3456	3456	3456
Mw [average] (Da)	9300±800	8700±1400	8600±400
Mw [m↔d] (Da)	5900±400	7900±400	7200±100
K _{a(abs)} [m↔d]	7.7±1.5	0.73±0.05	0.17±0.04
$\epsilon_{[\lambda]}$ (cm ⁻¹ M ⁻¹ _[nm])	3580 _[262]	2490 _[295]	3590 _[309]
c (M)	1.00x10 ⁻⁴	1.00x10 ⁻⁴	1.00x10 ⁻⁴
l (cm)	1.10	1.10	1.10
K _{a(conc)} [m↔d] (M ⁻¹)	14100±600	1000±300	700±200

4.12.4 NMR Spectroscopy

1D- ^1H NMR spectra used to evaluate chemical shift dispersion were run at 500 MHz on a Bruker AMX500. Each sample (~ 0.3 mM) was run at 25 °C and dissolved in 45 mM phosphate buffer (90:10, $\text{H}_2\text{O}:\text{D}_2\text{O}$) at pH 7.0. A relaxation delay of 1.0 s was used, during which time the water signal was saturated using a frequency-selective low-power decoupling pulse. The NMR spectra were processed using SwaN-MR 3.4.9.⁴⁰⁵ A convolution function was applied to remove the residual water signal.

Effect of Linker Length on CTB and Bowl TASP

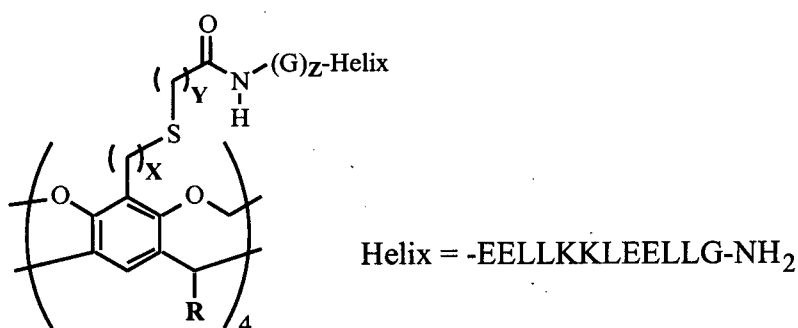
5.0 Introduction

Studies on our initial CTB and bowl TASP, which had the template-bound cysteine residue directly adjacent to the “helix” (Chapter 4, e.g. **CTB/N2/L**, or **Bowl/N2/L**), showed that they self-associate in aqueous solution, and that they lacked the conformational specificity associated with natural proteins. The self-association may indicate that the template-to-helix linker is too short/inflexible, resulting in incorrect bundling of the constituent helices, and exposure of the (non-packed) hydrophobic core to the solvent (which helps to drive inter-TASP association).

The importance of the template-to-helix linker to both the association state and “native-like” structure of a cavitand bowl TASP has been highlighted by previous work in our group.^{270,271} This research investigated the template-to-helix linker in terms of its functional group type and length between the template and helix (see Figure 5.1). Some of these bowl TASP appeared to be monomers according to sedimentation equilibrium ultracentrifugation. Fairlie, in the study of helical bundles that were templated by various organic macrocycles, found that a sufficiently long linker negates any effects implied by template size and shape (Section 1.4.4.3).²⁶⁵

When increasing the flexibility of the template-to-helix linker, one additional advantage was that the concentration of GnHCl required to unfold the TASPs was reduced. This results in TASPs that can be fully unfolded in less than 8 M GnHCl, which reduces the errors involved in calculating their global conformational stability from their unfolding curves (TASPs which are 50 % unfolded in 3-4 M GnHCl would be ideal).

Figure 5.1: Experimental “average” molecular weights (according to sedimentation equilibrium studies) of bowl TASPs made in our group that differ by their template-to-helix linker.



R	X	Y	Z	Experimentally Estimated MW	Calculated MW
-CH ₃	0	1	0	12850±50	7557
-CH ₃	0	1	1	18700±300	7788
-CH ₃	0	1	2	10000±600	8019
-CH ₃	0	1	3	8200±400	8250
-CH ₃	0	2	0	10800±100	7615
-CH ₃	0	4	0	22100±500	7725
-CH ₃	1	1	0	19300±100	7615
-(CH ₂) ₃ PO ₃ H ₂	1	1	0	11600±400	8112

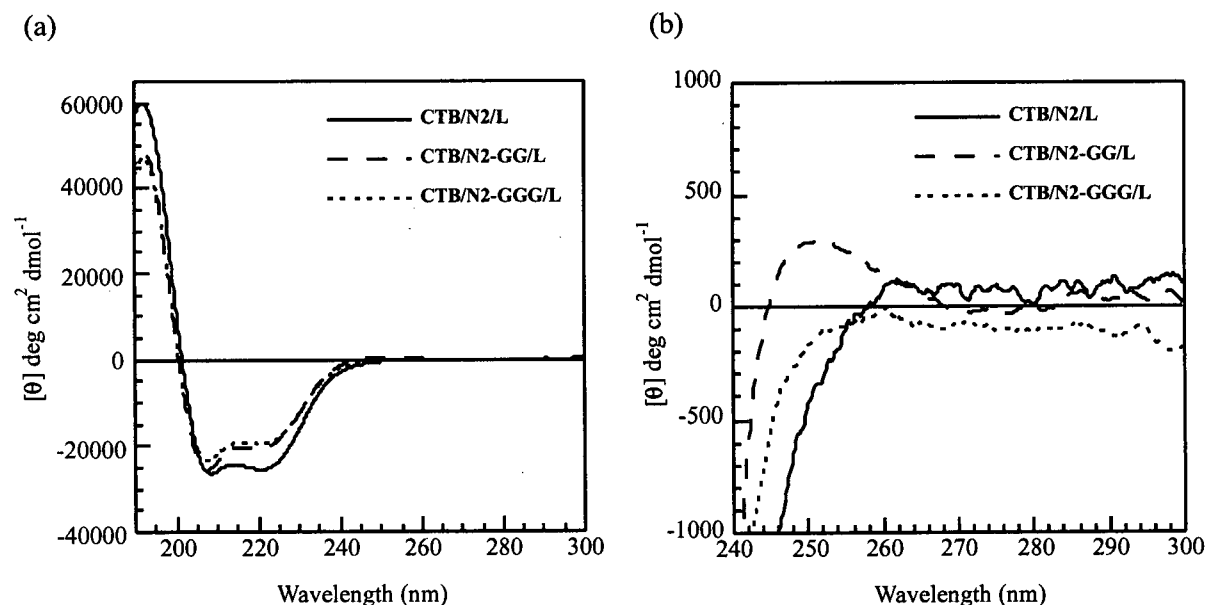
Thus, previous work done in our group has found that a simple solution to the problem of TASP self-association, is to insert glycine “spacer” residues between the template and helix.^{270,271} Therefore, we made CTB (Section 5.1) and bowl (Section 5.2) TASPs that contained peptides with either two (N2-GG/L, Ac-CGGEELLKKLEELLKKG-NH₂), or three (N2-GGG/L, Ac-CGGGEELLKKLEELLKKG-NH₂) glycine residues as flexible “spacers” between the template-bound cysteine residue and the “helix”.

We also investigated the effect of linking the peptide strands to the templates via their C-termini, as opposed to their N-termini. Thus, peptide C2-GGG/L (Ac-GEELLKKLEELLKKGGGC-NH₂) was linked (via disulfide bonds) to the CTB (Section 5.3) and bowl (Section 5.4) templates. This investigation will give information on whether the templates have any effect on the structure of the helical bundles, aside from pre-organization. If the template has no additional effect on the structure of the helical bundles, then there should be no difference between the stability and structure of the N- and C-terminally linked TASPs. However, it is possible that the template may interact (in a non-covalent manner) with the helical bundles; the template has both a dipole moment, and hydrogen bonding moieties.

5.1. CTB/N2-GG/L and CTB/N2-GGG/L

These three-helix bundle TASPs were synthesized by linking peptide N2-GG/L or N2-GGG/L (via disulfide bonds) to the CTB template. The CD spectra (Figure 5.2) of CTB/N2-GG/L and CTB/N2-GGG/L are both helical and are not dependent on TASP concentration. CTB/N2-GG/L has some activity in its near-UV CD spectrum (a small, positive band between 245-260 nm), but CTB/N2-GGG/L does not show any significant signal. The CD spectra were unchanged in the presence of 2 M KCl salt.

Figure 5.2: (a) Full CD, and (b) near-UV CD spectra of CTB/N2-GG/L (59 μ M, dashed line), CTB/N2-GGG/L (60 μ M, dotted line), and CTB/N2/L (54 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



Both **CTB/N2-GG/L** and **CTB/N2-GGG/L** were assessed to be monomers in solution from their concentration independent GnHCl-induced denaturation curves (Figure 5.3). However, sedimentation equilibrium experiments do not support this assumption (Table 5.1). In pH 7.0 phosphate buffer (50 mM), both **CTB/N2-GG/L** (100 μ M, average solution MW of 9000 ± 400 , versus calculated MW of 6278), and **CTB/N2-GGG/L** (100 μ M, average solution MW of 7800 ± 600 , versus calculated MW of 6450) can be best described as being in a monomer-dimer equilibrium. This inconsistency between the GnHCl-unfolding, and sedimentation equilibrium experiments, suggests that the weak association of these CTB TASPs in buffer solution is overcome at low concentrations of denaturant. Fitting the sedimentation equilibrium data to a monomer-dimer equilibrium gives association constants of $3700 \pm 1100 \text{ M}^{-1}$ for **CTB/N2-GG/L**, and $1500 \pm 400 \text{ M}^{-1}$ for **CTB/N2-GGG/L**. Repeating the sedimentation equilibrium experiments on **CTB/N2-GGG/L** in the presence of 2 M KCl reduces, but does not eliminate, the degree of self-association (average solution MW of 7100 ± 300 , versus calculated MW of 6450, with a calculated association constant for a monomer-dimer equilibrium of $550 \pm 200 \text{ M}^{-1}$). This observation implies that the self-association has both a hydrophobic and electrostatic component (*i.e.* screening of electrostatic forces reduces the degree of TASP self-association; thus, any residual self-association is driven by hydrophobic forces).

Figure 5.3: GnHCl-induced denaturation curves for (a) **CTB/N2-GG/L** (39 μ M, \circ , dashed line; 3.9 μ M, \times , solid line); (b) **CTB/N2-GGG/L** (40 μ M, \circ , dashed line; 4.0 μ M, \times , solid line); (c) **CTB/N2-GG/L** (59 μ M, \circ , long dashed line), **CTB/N2-GG/L** (60 μ M, \circ , short dashed line), compared to **CTB/N2/L** (40 μ M, \circ , solid line). All spectra recorded in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.

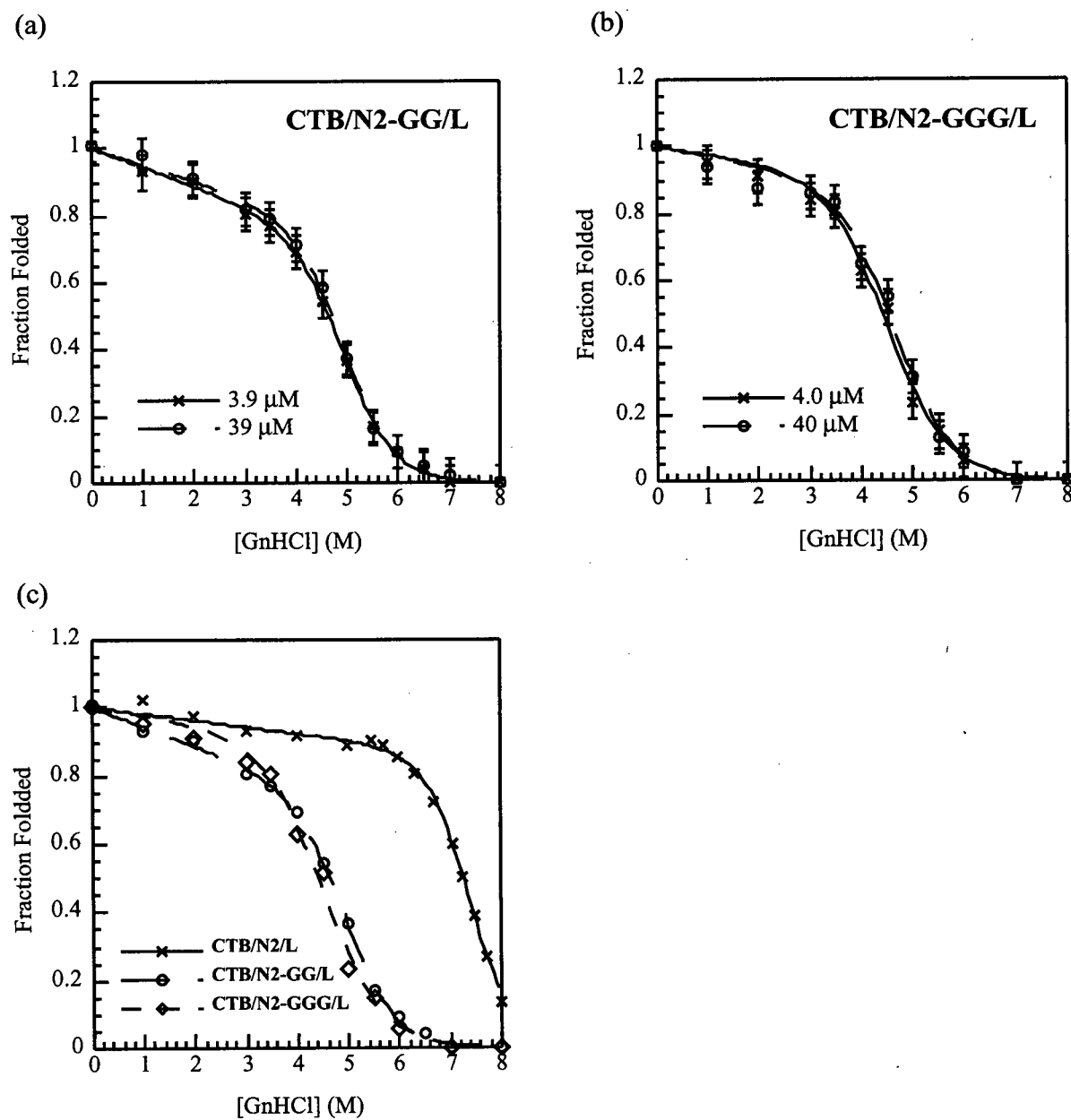


Table 5.1: Experimentally estimated molecular weights for **CTB/N2-GG/L**, and **CTB/N2-GGG/L** (which both describe a monomer-dimer in solution) determined by sedimentation equilibrium ultracentrifugation. Experiments were run at 100 μ M TASP concentration, 25 °C in 50 mM phosphate buffer, or 2 M KCl.

	CTB/N2-GG/L	CTB/N2-GGG/L	CTB/N2-GGG/L
MW [calculated] (Da)	6278	6450	6450
Solvent	Buffer	Buffer	2M KCl
MW [experimental] (Da)	9000 \pm 400	7800 \pm 200	7100 \pm 300

From the GnHCl-induced unfolding curves, the calculated global conformational stability data (Table 5.2) indicates that **CTB/N2-GG/L** is marginally (about 1 kcal mol⁻¹) more stable than **CTB/N2-GGG/L**. These weakly associated CTB TASPs are assumed to be monomers at the GnHCl concentration where the unfolding transition occurs. The higher stability of **CTB/N2-GG/L** versus **CTB/N2-GGG/L** may be attributed to a less flexible two versus three glycine linker between the template-bound cysteine residue and the “helix”. The GnHCl-induced unfolding curves for **CTB/N2-GG/L** is similar to **CTB/N2-GGG/L**, but both these TASPs are far less stable than **CTB/N2/L** (Figure 5.3c), which has no glycine spacers in its sequence, but exhibits a greater degree of self-association. This difference in stability may be solely related to the flexibility of the template-to-helix linker, but it may also be related to additional stabilizing interactions between the “helix” and the template for **CTB/N2/L** (which has no glycine “spacers”); for example, the template may help to

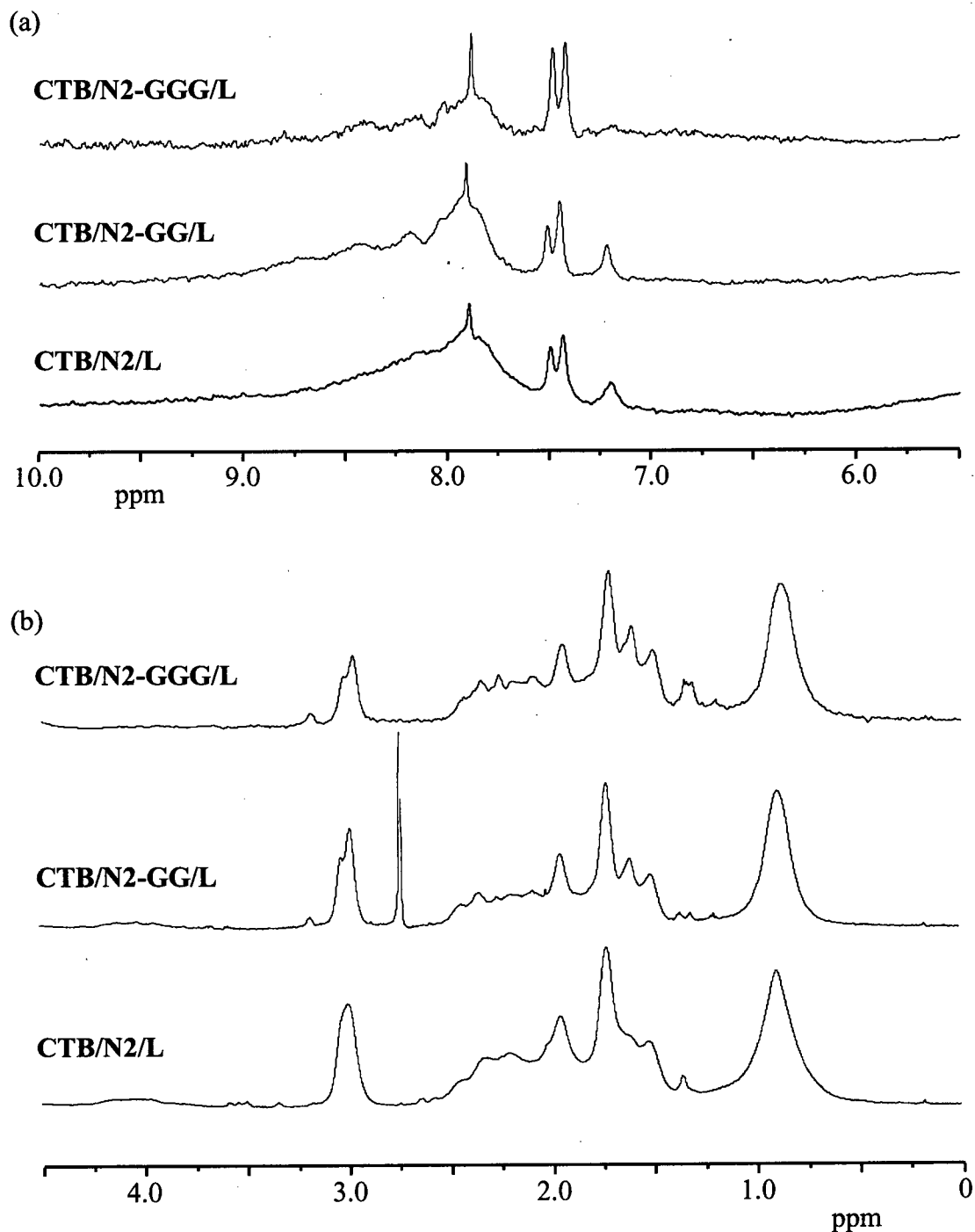
nucleate helix formation by hydrogen bonding with the “unsatisfied” backbone amide bonds, or via a hydrophobic interaction between the non-polar template and the first leucine residue in the core of the helix.

Table 5.2: Thermodynamic evaluation of **CTB/N2-GG/L**, **CTB/N2-GGG/L**, compared to **CTB/N2/L**, calculated from their GnHCl-induced denaturation.

TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m $\text{kcal mol}^{-1} \text{M}^{-1}$	$\Delta G^\circ \text{H}_2\text{O}$ kcal mol^{-1}
CTB/N2/L	2.8	7.2 ± 0.1	1.4 ± 0.1	10.4 ± 0.6
CTB/N2/L	28	7.8 ± 0.1	1.3 ± 0.1	10.4 ± 0.9
CTB/N2-GG/L	3.9	4.6 ± 0.1	1.2 ± 0.1	6.0 ± 0.3
CTB/N2-GG/L	39	4.7 ± 0.1	1.2 ± 0.1	6.2 ± 0.5
CTB/N2-GGG/L	4.0	4.4 ± 0.1	1.0 ± 0.1	4.7 ± 0.5
CTB/N2-GGG/L	40	4.5 ± 0.1	1.1 ± 0.1	5.1 ± 0.6

The broad NMR spectra of **CTB/N2-GG/L** and **CTB/N2-GGG/L** (Figure 5.4) indicate that they both exist in a multitude of slowly exchanging conformations (*i.e.* molten globules). Thus, a reduction in the degree of TASP self-association, which was probably a result of more optimal bundling of the TASP-helices, did not make any noticeable impression of the conformational specificity of these CTB TASPs. This lack of “native-like” structure for these three-helix bundles may result from using a helix which is designed to form four-helix bundles. Four-helix bundle TASPs will be investigated in section 5.3.

Figure 5.4: NMR spectra of **CTB/N2-GG/L** and **CTB/N2-GGG/L**, compared to **CTB/N2/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.



5.2 CTB/C2-GGG/L

For this CTB TASP, a peptide (with a three glycine “spacer”), was linked via its C-terminus (peptides have all been linked to the template by their N-termini so far in this thesis). Thus, peptide **C2-GGG/L** (Ac-GEELLKKLEELLKKGGGC-NH₂), was linked to the CTB template (via a disulfide bond).

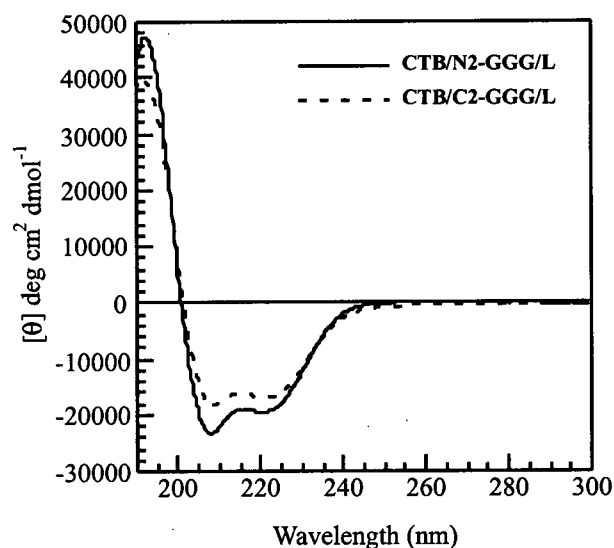
The CD spectrum of **CTB/C2-GGG/L** (Figure 5.5) was concentration independent, helical and insensitive to salt (2 M KCl). **CTB/C2-GGG/L** did not exhibit any significant signals in its near-UV CD spectrum.

CTB/C2-GGG/L exhibited concentration independent GnHCl-induced unfolding curves (Figure 5.6), implying that it is a monomer at the unfolding transition point. However, **CTB/C2-GGG/L** may be a weakly associating species in aqueous buffer (due to limited instrument time, we were unable to perform sedimentation equilibrium experiments on this TASP).

The global conformational stability of **CTB/C2-GGG/L** was calculated from its GnHCl-induced denaturation, and found to be lower (by about 1 kcal mol⁻¹) than for **CTB/N2-GGG/L** (Figure 5.5 and Table 5.3). The unfolding curves for **CTB/C2-GGG/L** and **CTB/N2-GGG/L** appear to be almost identical, the difference in stability is probably related to the steeper unfolding curve for **CTB/N2-GGG/L** at the unfolding transition point. This difference in stability indicates that the CTB template may be slightly more favoured as

a helix N-cap versus a C-cap. The difference in stability may also be related to the degree of fraying at the end of the helix not attached to the template, or related to the amount of hydrophobic surface exposed on unfolding. These factors could all influence the cooperativity of the unfolding transition.

Figure 5.5: CD spectra of **CTB/C2-GGG/L** (66 μ M, dotted line), compared to **CTB/N2-GGG/L** (60 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



The NMR spectrum of **CTB/C2-GGG/L** is broad, consistent with a molten globule in solution (Figure 5.7), but it appears to be slightly sharper than that seen for **CTB/N2-GGG/L**. This indicates that its conformation is either fluctuating faster on the NMR time-scale (and is therefore a more “averaged” spectrum), or that **CTB/C2-GGG/L** has more conformational specificity (*i.e.* a more “native-like” NMR).

Figure 5.6: GnHCl-induced denaturation curves for (a) **CTB/C2-GGG/L** (36 μM , \circ , dashed line; 3.6 μM , \times , solid line); (b) **CTB/C2-GGG/L** (3.6 μM , \circ , dashed line), compared to **CTB/N2-GGG/L** (4.0 μM , \times , solid line). All spectra recorded in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}\text{C}$.

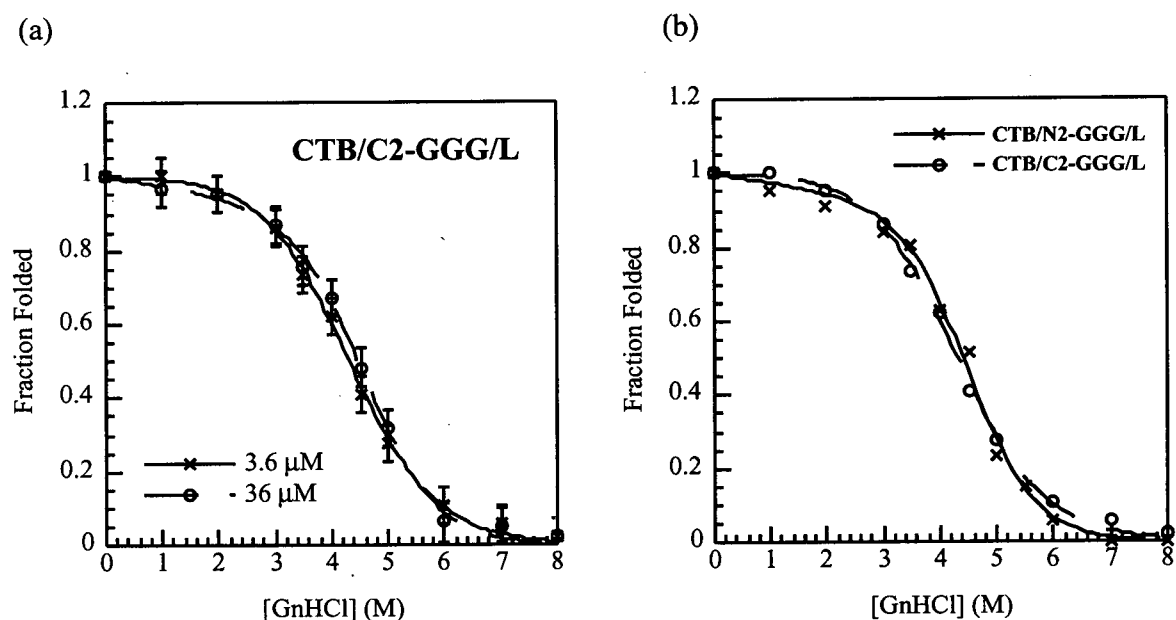
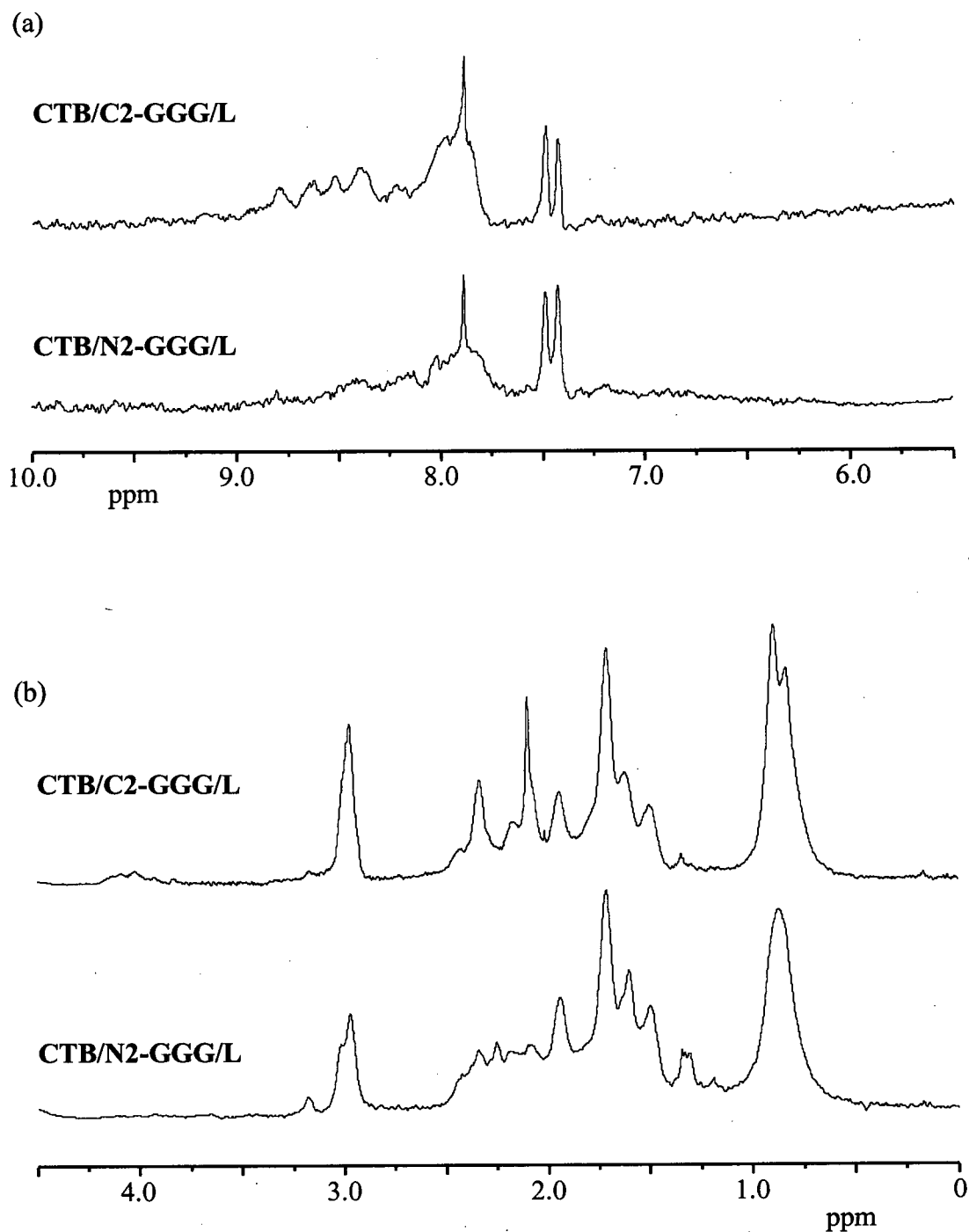


Table 5.3: Thermodynamic evaluation of **CTB/C2-GGG/L**, compared to **CTB/N2-GGG/L**, calculated from their GnHCl-induced denaturation.

TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m $\text{kcal mol}^{-1} \text{M}^{-1}$	$\Delta G^{\circ} \text{H}_2\text{O}$ kcal mol^{-1}
CTB/C2-GGG/L	3.6	4.3 ± 0.1	0.8 ± 0.1	3.7 ± 0.3
CTB/C2-GGG/L	36	4.4 ± 0.1	0.9 ± 0.1	4.3 ± 0.4
CTB/N2-GGG/L	4.0	4.4 ± 0.1	1.0 ± 0.1	4.7 ± 0.5
CTB/N2-GGG/L	40	4.5 ± 0.1	1.1 ± 0.1	5.1 ± 0.6

Figure 5.7: NMR spectra of **CTB/C2-GGG/L**, compared to **CTB/N2-GGG/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.

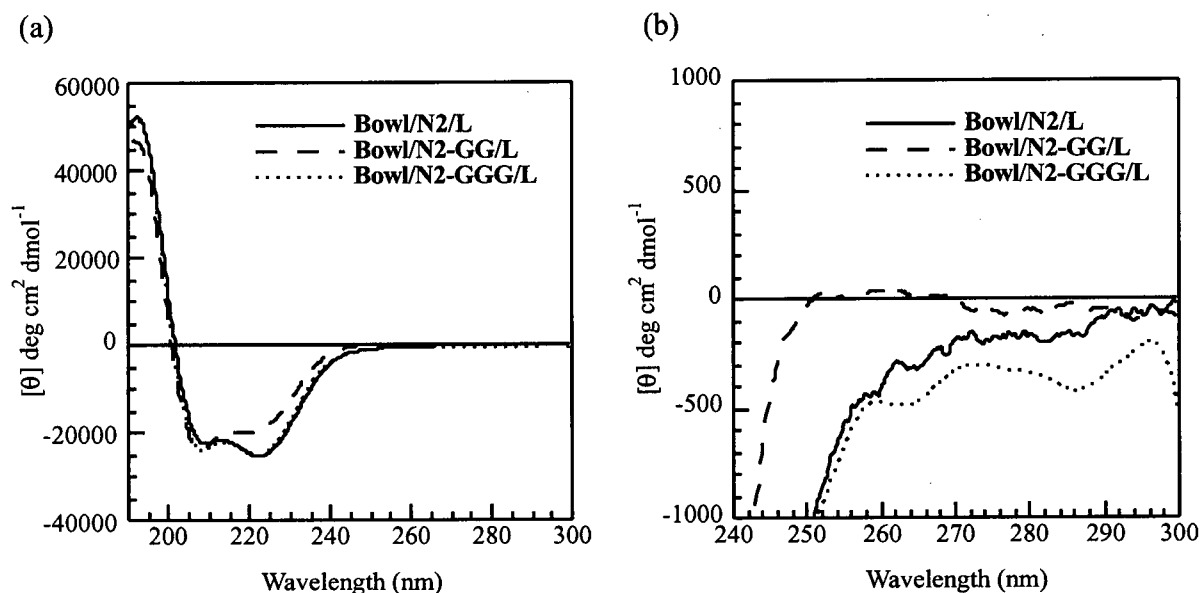


5.3 Bowl/N2-GG/L and Bowl/N2-GGG/L

These bowl TASP s contained peptides with either two or three glycine "spacers" between the template-bound cysteine residue and the "helix". Thus, these TASP s were synthesized using either peptide **N2-GG/L** (Ac-CGGEELLKKLEELLKKG-NH₂) or peptide **N2-GGG/L** (Ac-CGGGEELLKKLEELLKKG-NH₂), which were linked to the bowl via disulfide bonds. The concentration independent CD spectra indicated both **Bowl/N2-GG/L** and **Bowl/N2-GGG/L** were helical (Figure 5.8). **Bowl/N2-GGG/L** also exhibited a distinctive near-UV CD spectrum, indicating an interaction between the achiral bowl (or disulfide) chromophore, and the chiral environment surrounding the α -helix.

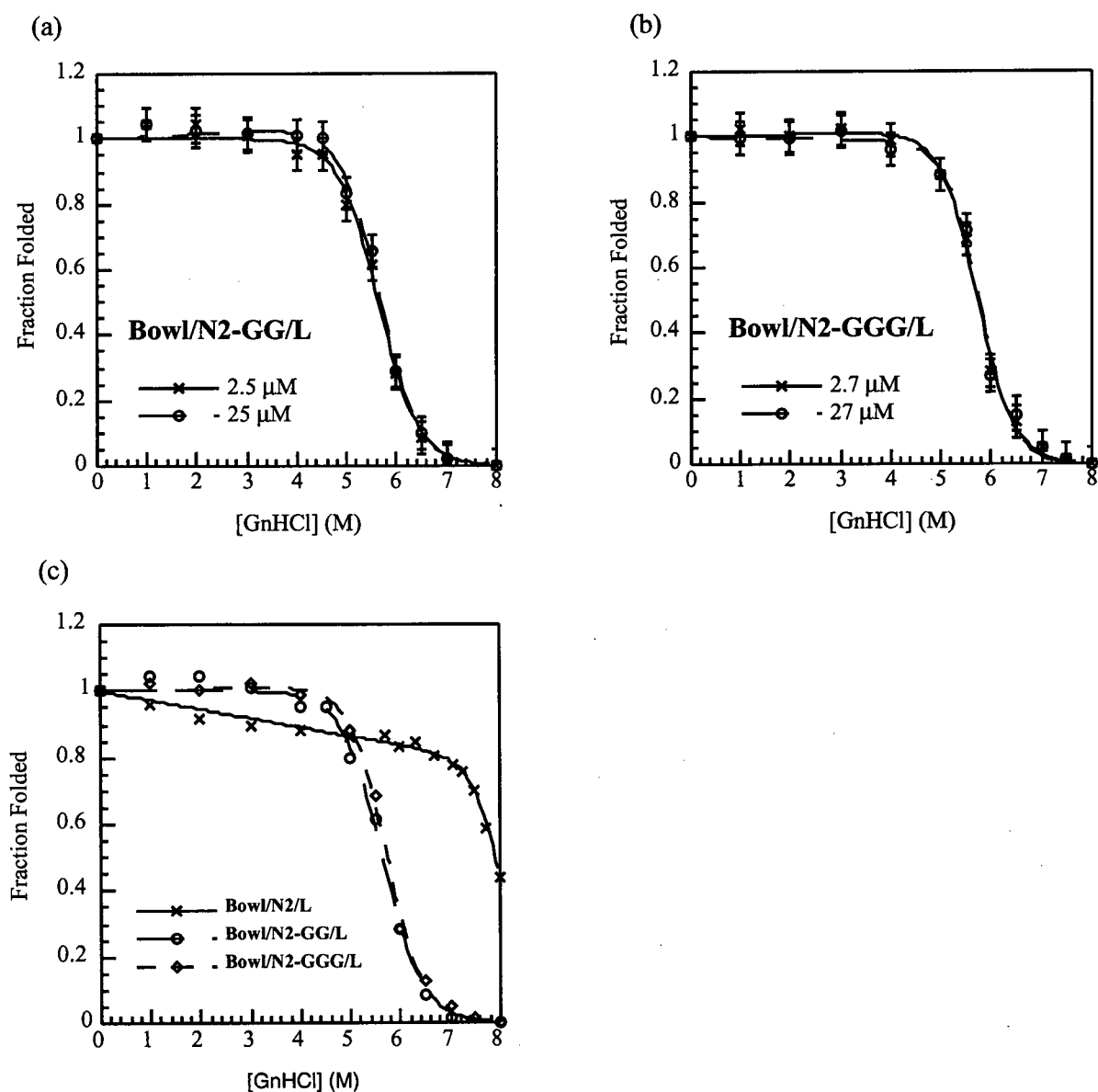
In 2 M KCl, the CD spectrum of **Bowl/N2-GGG/L** was unchanged; however, the helicity of **Bowl/N2-GG/L** increased (+ 10 %). The screening of unfavourable electrostatic forces present in the "helix" did not alter the secondary structure of the analogous **Bowl/N2/L** and **Bowl/N2-GGG/L**. Therefore, this change in the secondary structure of **Bowl/N2-GG/L** results from the screening of electrostatic forces present in the TASP s' tertiary structure/linker (the potential sources of these forces were discussed in Section 4.11).

Figure 5.8: (a) Full, and (b) Near-UV CD spectra of **Bowl/N2-GG/L** (50 μ M, dashed line), and **Bowl/N2-GGG/L** (54 μ M, dotted line), compared to **Bowl/N2/L** (60 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 °C. NOTE: The spectra of **Bowl/N2-GGG/L** and **Bowl/N2/L** overlap somewhat.



The GnHCl-induced unfolding curves of **Bowl/N2-GG/L** showed an increase in helicity (+ 10%) at low concentrations of denaturant. GnHCl is a salt, and screens electrostatic interactions in a similar manner to KCl. Therefore, the GnHCl unfolding experiments for **Bowl/N2-GG/L** were performed in the presence of 0.1 M KCl. This additional KCl did not affect the transition point (or shape) of the unfolding curve run in the absence of KCl. Both **Bowl/N2-GG/L** and **Bowl/N2-GGG/L** exhibited concentration independent GnHCl unfolding curves (Figure 5.9), consistent with monomers in solution.

Figure 5.9: GnHCl-induced denaturation curves of (a) **Bowl/N2-GG/L** (25 μ M, \circ , dashed line; 2.5 μ M, \times , solid line); (b) **Bowl/N2-GGG/L** (27 μ M, \circ , dashed line; 2.7 μ M, \times , solid line); (c) Comparison between **Bowl/N2-GG/L** (2.5 μ M, \circ , long dashed line), **Bowl/N2-GGG/L** (2.7 μ M, \diamond , short dashed line), and **Bowl/N2/L** (3.0 μ M, \times , solid line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C (**Bowl/N2-GG/L** contained 100 mM KCl).



Sedimentation equilibrium (Table 5.4) revealed that **Bowl/N2-GG/L** (100 μ M) is best described as a monomer-dimer in aqueous buffer (with a solution “average” MW of 10200 ± 500 , compared to a calculated MW of 8483, and an association constant calculated to be 700 ± 200). This contradicts the assumption made from the concentration independent GnHCl-induced unfolding curves that **Bowl/N2-GG/L** exists as a monomer in solution, however, it may still be a monomer at the GnHCl concentration required for unfolding. Sedimentation equilibrium studies on **Bowl/N2-GGG/L** confirm that it is indeed a monomer in both buffer and 2 M KCl (solution “average” MWs of 8000 ± 600 and 7800 ± 1500 , compared to a calculated MW of 8713).

Table 5.4: Experimentally estimated molecular weights for TASPs **Bowl/N2-GG/L** (monomer-dimer), and **Bowl/N2-GGG/L** (monomer) determined by sedimentation equilibrium ultracentrifugation. All experiments run at 100 μ M TASP concentration, 25 °C in 50 mM phosphate buffer.

	Bowl/N2-GG/L	Bowl/N2-GGG/L	Bowl/N2-GGG/L
MW [calculated] (Da)	8483	8713	8713
Solvent	Buffer	Buffer	2M KCl
MW [experimental] (Da)	10200 ± 500	8000 ± 600	7800 ± 1500

Analysis of the GnHCl-induced denaturation curves (Table 5.5) reveals that **Bowl/N2-GG/L** was less stable (by 2 kcal mol⁻¹) than **Bowl/N2-GGG/L**. This was unexpected. We assume that the longer (and therefore the more flexible) the linker, the less the effect the template has at directing the peptides, and therefore the easier it would be to unfold the TASP. This observation may indicate an unfavourable structural effect present in **Bowl/N2-GG/L** (this will be further speculated upon in Section 5.5). The data from the GnHCl-induced unfolding of **Bowl/N2-GGG/L** can be compared to another four-helix bundle bowl TASP **ArBowl/N2-GGG/L**, a TASP previously studied in our group that is also a monomer in solution.^{270,271} These TASP differ in their template-to-helix link (Figure 5.10), **ArBowl/N2-GGG/L** employs an aryl-, as opposed to a benzyl thiol bowl as template, and its constituent peptides are linked via a thioether bond to the polypeptide backbone (see Section 3.3.2 for more details), rather than via a disulfide bond to an amino acid's side chain. **Bowl/N2-GGG/L** and **ArBowl/N2-GGG/L** appear to have very similar stability (Table 5.5).

Figure 5.10: Structures of **ArBowl/GGG-N2/L** (peptide linked via the peptide backbone), and **Bowl/N2-GGG/L** (peptide linked via cysteine side chain).

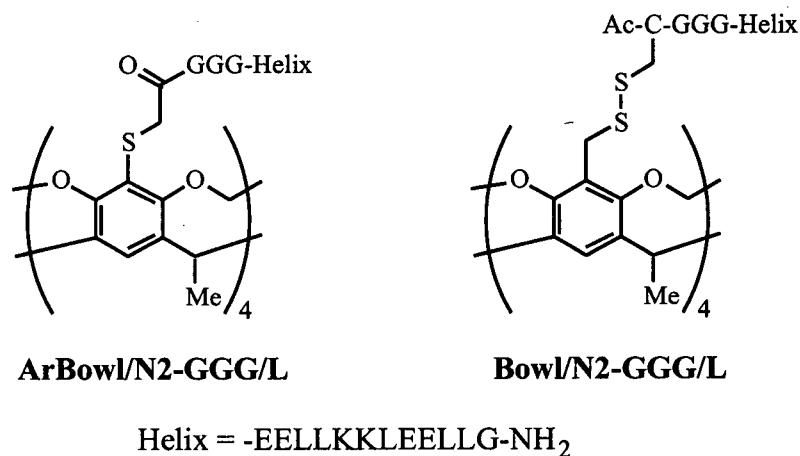


Table 5.5: Thermodynamic evaluation of **Bowl/N2-GG/L** and **Bowl/N2-GGG/L**, compared to **Bowl/N2/L** and **ArBowl/N2-GGG/L**, calculated from their GnHCl-induced denaturation curves. This data assumes that these TASP s are monomers at the unfolding transition point.

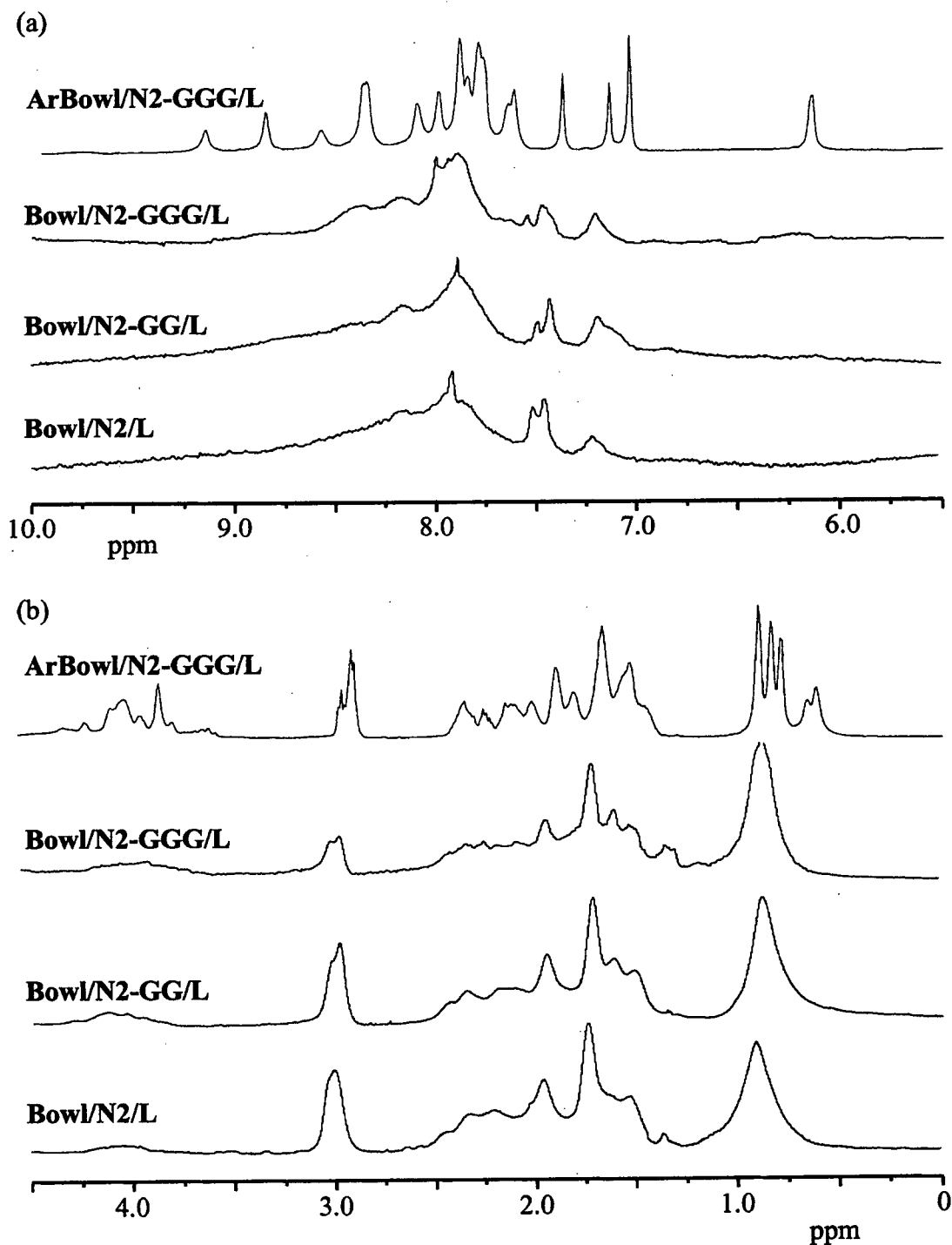
TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m $\text{kcal mol}^{-1} \text{M}^{-1}$	$\Delta G^{\circ} \text{H}_2\text{O}$ kcal mol^{-1}
Bowl/N2/L	3.0	7.9 ± 0.1	2.0 ± 0.2	16.2 ± 1.6
Bowl/N2/L	30	>8	1.9 ± 0.4	16.2 ± 3.3
Bowl/N2-GG/L	2.5	5.6 ± 0.1	1.4 ± 0.2	8.1 ± 1.0
Bowl/N2-GG/L	25	5.7 ± 0.1	1.4 ± 0.1	8.2 ± 0.8
Bowl/N2-GGG/L	2.7	5.7 ± 0.1	1.7 ± 0.1	9.4 ± 0.7
Bowl/N2-GGG/L	27	5.7 ± 0.1	1.8 ± 0.2	10.2 ± 1.1
ArBowl/N2-GGG/L	3.0	5.5 ± 0.1	1.8 ± 0.1	9.8 ± 0.6
ArBowl/N2-GGG/L	30	5.5 ± 0.1	1.9 ± 0.1	10.8 ± 0.5

The NMR spectra of **Bowl/N2-GG/L** and **Bowl/N2-GGG/L** (Figure 5.11) were both broad, consistent with molten globule structure. We now have strong evidence that the source of this broadening is not self-association: **Bowl/N2-GGG/L** is a monomer in buffer but exhibits a broad NMR spectrum.

By comparison, the thioether linked TASP **ArBowl/N2-GGG/L**, which has almost identical stability to **Bowl/N2-GGG/L**, demonstrates a significantly different NMR spectrum, with sharp, well resolved peaks (Figure 5.11). Thus, **ArBowl/N2-GGG/L** either demonstrates “native-like” conformational specificity, and forms a symmetrical four-helix bundle, or it consists of four helices that exhibit dynamic motion and are stabilized by non-specific hydrophobic interactions, akin to those seen in a molten globule or early folding intermediate (see also Section 4.8).

ArBowl/N2-GGG/L and **Bowl/N2-GGG/L** are both monomers in solution with similar global conformational stability but different conformational specificity. Assuming that the similar global stability of these TASPs is not just a coincidence, then it may indicate they have a similar buried (volume of, and interactions between) hydrophobic surface when folded, and therefore similar structures: This may indicate that **ArBowl/N2-GGG/L** lacks the specific side chain interactions present in “native-like” structure, and is a molten globule with rapidly inter-changing conformations on the NMR time-scale (*i.e.* a structure that is an artifact of the TASP approach). Unfortunately there is no easy way to answer these questions; because of the degenerate amino acid sequence of these TASPs, it is not possible to reliably interpret any data from a through space (e.g. nOe) NMR technique.

Figure 5.11: NMR spectra of **Bowl/N2-GG/L** and **Bowl/N2-GGG/L**, compared to **Bowl/N2/L**, and **ArBowl/N2-GGG/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.

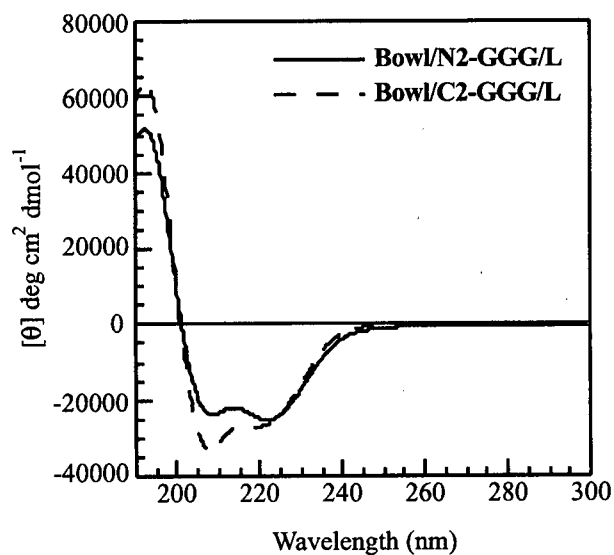


5.4 Bowl/C2-GGG/L

This TASP was made to investigate any effect the bowl template had on the helical bundles, specifically if it acted as a N- or a C-cap for the constituent helices. Thus, we linked peptide **C2-GGG/L** (Ac-GEELLKKLEELLKKGGGC-NH₂) to the bowl template via disulfide bonds.

The concentration independent CD spectrum (Figure 5.12) indicated that **Bowl/C2-GGG/L** was helical. It did not exhibit any significant near-UV CD spectrum. The CD spectrum was unchanged in the presence of salt.

Figure 5.12: CD spectrum of **Bowl/C2-GGG/L** (50 μ M, dashed line) compared to **Bowl/N2-GGG/L** (54 μ M, solid line), in pH 7.0 phosphate buffer (10 mM) at 25 °C.



The GnHCl-induced denaturation curves of **Bowl/C2-GGG/L** (Figure 5.13) were concentration independent, consistent with a monomer in solution (due to restricted instrument time we were unable to obtain sedimentation equilibrium data on this TASP).

The unfolding curve for **Bowl/C2-GGG/L** was less co-operative than that observed for **Bowl/N2-GGG/L**, and also unfolded at a lower concentration of GnHCl. The calculated global conformational stability of **Bowl/C2-GGG/L** (Table 5.6) is lower (by about 4 kcal mol⁻¹) than **Bowl/N2-GGG/L**. This is a significant difference in the stability of these two bowl TASPs, especially considering the relatively small difference in stability between the analogous CTB TASPs. This indicates that the bowl TASP stabilizes helices that are linked via their N-termini (see Section 4.11 for potential sources of the N-cap), or that it destabilizes helices that are attached via their C-termini, or both.

Once again the NMR of **Bowl/C2-GGG/L** was broad, consistent with a molten globule in solution (Figure 5.14).

Figure 5.13: GnHCl-induced denaturation curves of (a) **Bowl/C2-GGG/L** (25 μM , \circ , dashed line; 2.5 μM , \times , solid line), and (b) Comparison between **Bowl/C2-GGG/L** (2.5 μM , \circ , dashed line) and **Bowl/N2-GGG/L** (2.7 μM , \times , solid line). All points measured in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}\text{C}$.

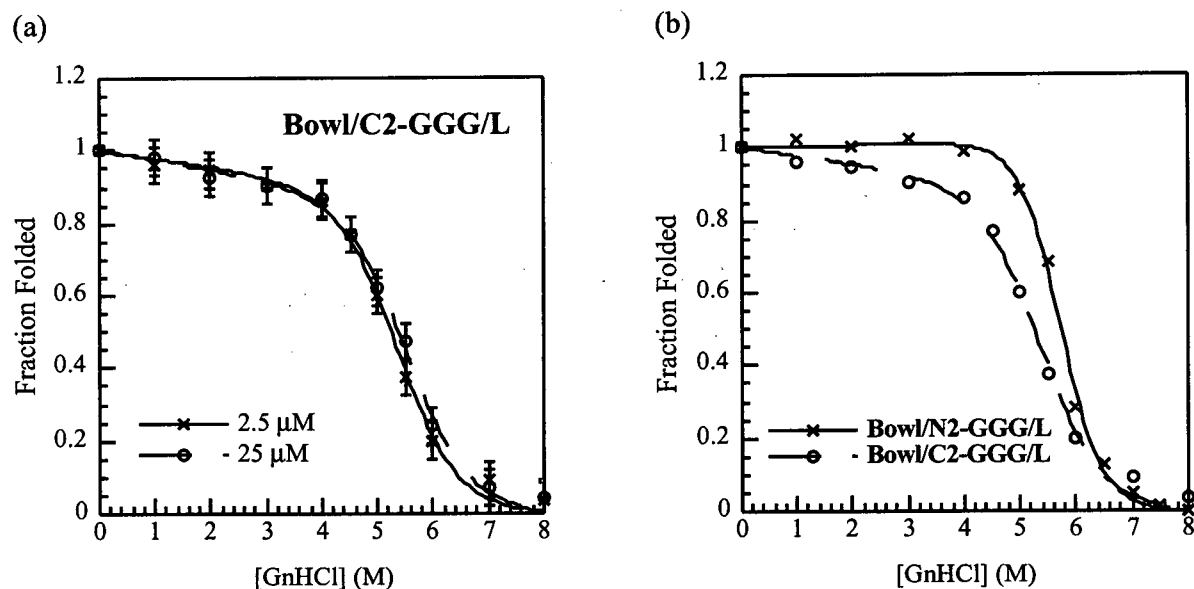
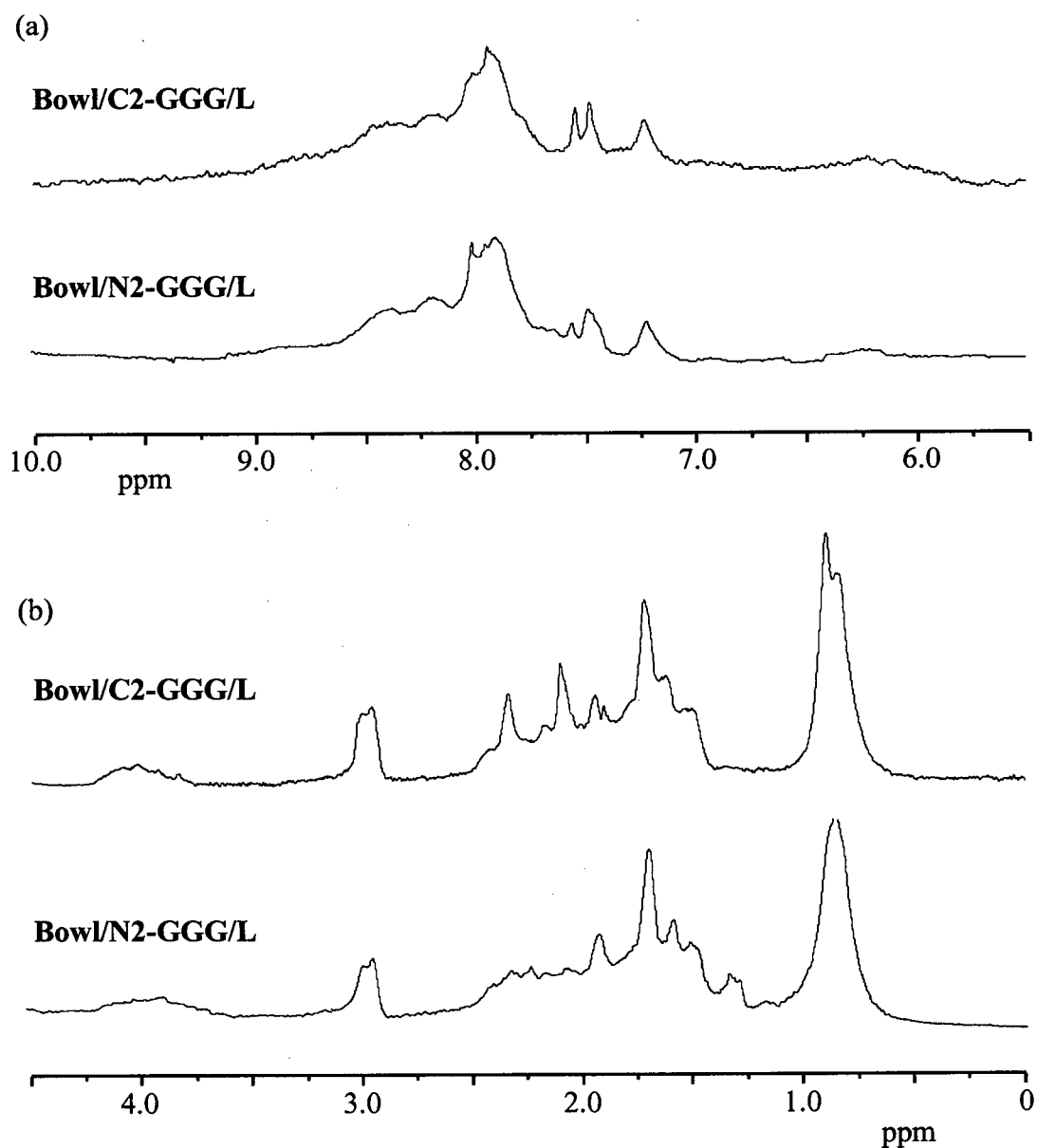


Table 5.6: Thermodynamic evaluation of **Bowl/C2-GGG/L** (compared to **Bowl/N2-GGG/L**), calculated from their GnHCl-induced denaturation curves.

TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m kcal mol $^{-1}$ M $^{-1}$	$\Delta G^{\circ}\text{H}_2\text{O}$ kcal mol $^{-1}$
Bowl/N2-GGG/L	2.7	5.7 ± 0.1	1.7 ± 0.1	9.4 ± 0.7
Bowl/N2-GGG/L	27	5.7 ± 0.1	1.8 ± 0.2	10.2 ± 1.1
Bowl/C2-GGG/L	2.5	5.2 ± 0.1	1.1 ± 0.1	6.0 ± 0.6
Bowl/C2-GGG/L	25	5.4 ± 0.1	1.1 ± 0.1	6.0 ± 0.5

Figure 5.14: NMR spectra of **Bowl/C2-GGG/L**, compared to **Bowl/N2-GGG/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.



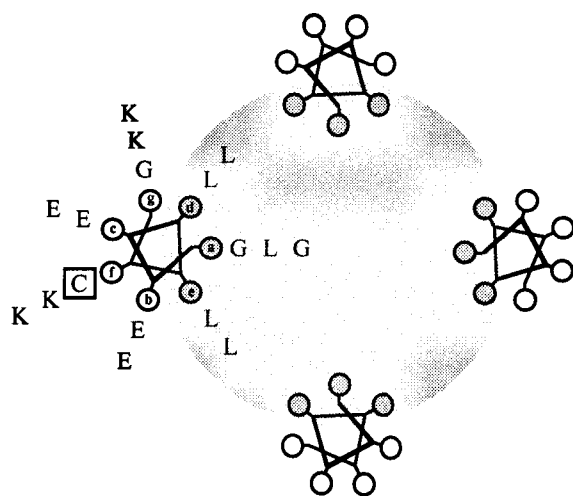
5.5 Conclusions for Chapter 5

The addition of three glycine residues between the template-bound cysteine and the “helix” decreases the self-association of both the bowl and CTB TASPs, probably because it enables optimal bundling of the helices. **Bowl/N2-GGG/L** proved to be a monomer in aqueous buffer by sedimentation equilibrium studies, whereas **Bowl/N2-GG/L** appeared to be a monomer-dimer in equilibrium. Indeed, **Bowl/N2-GG/L** was associated to a greater degree than **Bowl/N2/L**, a TASP with a shorter template-to-helix linker. Thus, **Bowl/N2-GG/L** may expose more of its hydrophobic core to the solvent, possibly a result of bad peptide/linker design (Figure 5.15). In this case the template-bound cysteine amino acid resides on the hydrophilic face of the helix (as opposed to the hydrophobic face in the cases of **Bowl/N2/L** and **Bowl/N2-GGG/L**). Although this assumes that the glycine “spacers” take some part in the helix, even if they are unstructured and act as flexible linkers, then the two glycine linker may not be long enough to reach from the template to the hydrophilic side of the helix to allow optimal bundling. Therefore the flexibility of the template-to-linker alone, is not the sole determinant factor in TASP intra-helix bundling versus inter-TASP self-association. This three glycine “spacer” is optimal for the helix sequence presented in this thesis, thus the design of other template-to-helix “spacers” would vary according to the context of the helix sequence.

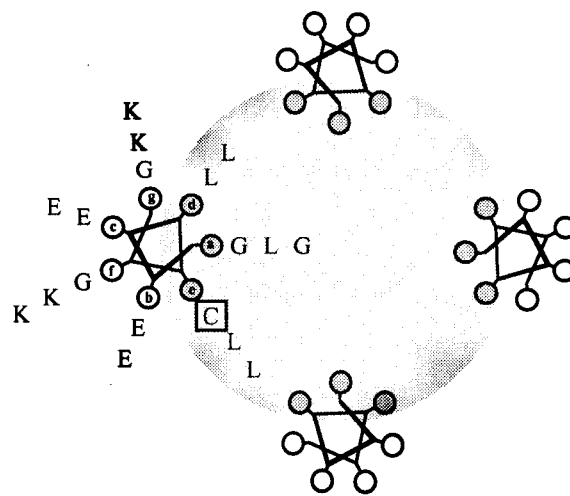
Figure 5.15: Diagram showing the positions of amino acid residues on a helix on (a) **Bowl/N2-GG/L**; and (b) **Bowl/N2-GGG/L**. Diagram shows helices in a coiled-coil heptad repeat (3.5 versus 3.6 residues per helical turn – see Section 1.4.2.2). Peptides are linked to the bowl template via a cysteine residue (which is the first residue in each sequence - shown on the diagram in a box – helices are read clockwise from the centre outward, for example **a→b→c→...** or **C→G→G→...**).

	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a
Bowl/N2-GG/L		C	G	G	E	E	L	L	K	K	L	E	E	L	L	K	K	G
Bowl/N2-GGG/L	C	G	G	G	E	E	L	L	K	K	L	E	E	L	L	K	K	G

(a) **Bowl/N2-GG/L**



(b) **Bowl/N2-GGG/L**



In aqueous solution, **CTB/N2-GG/L** and **CTB/N2-GGG/L** were both found to exist as monomer-dimers. The higher association state of **CTB/N2-GG/L** may occur for similar reasons as those given for **Bowl/N2-GG/L**, outlined above. The self-association of **CTB/N2-GGG/L** may result from the helix sequence being designed to form four-helix bundles, therefore, the hydrophobic core of the three-helix bundles is over-packed.

Exposure of the overpacked core leucine residues to the aqueous solvent would result in hydrophobically driven self-association. Thus, one result of forcing a helix designed to form four-helix bundles into a three-helix bundle is that the three helices are “seeking” the fourth helix to complete their preferred bundle structure; which leads to TASP self-association. The CTB and bowl TASPs present an opportunity to observe the effect of forcing the “wrong” three- or four-helix bundle sequence into an undesirable folding pattern. In this case, the peptide sequence that was designed to form four-helix bundles (*i.e.* –EELLKKLEELLKKG), cannot be “forced” into a three-helix bundle, highlighting the importance of the secondary structure design to the overall folding topology of a protein. One future goal is to design a helix that is designed to form three-helix bundles and observe the effect of “forcing” it into a four-helix bundle.

The global conformational stability (calculated from the chemical denaturation) of **CTB/N2-GGG/L** was about 1 kcal mol⁻¹ lower than **CTB/N2-GG/L**: TASPs with more flexible linkers are expected to have a lower stability, due to lower directing effect imparted by the template (assuming that the increase in stability does not come from the higher degree of self-association of **CTB/N2-GG/L**). However, when comparing the stability of the bowl TASPs, **Bowl/N2-GGG/L** was about 2 kcal mol⁻¹ more stable than **Bowl/N2-GG/L**! This observation also suggests a bad peptide/linker design for **Bowl/N2-GG/L** (see above, and Figure 5.14). **Bowl/N2-GGG/L** had a similar global conformational stability to the **ArBowl/N2-GGG/L** TASP; a monomeric TASP that had the same “helix” sequence, but differed in its template-to-helix linker (see Figure 5.10).

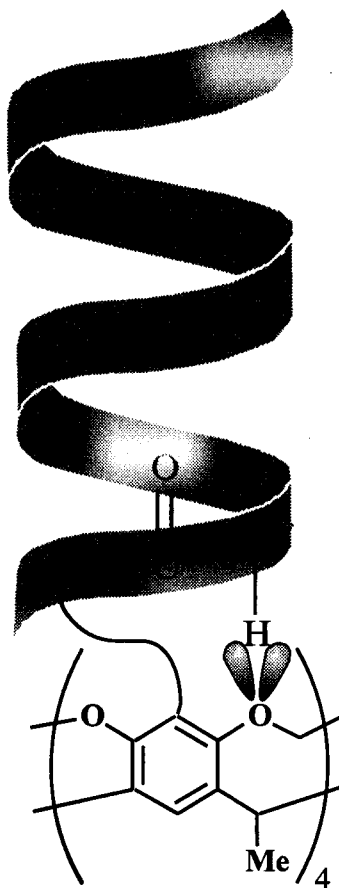
When comparing the global conformational stability of TASPs with peptides linked via their C- versus N-termini, **CTB/N2-GGG/L** was more stable than **CTB/C2-GGG/L** by about 1 kcal mol⁻¹, and **Bowl/N2-GGG/L** was more stable than **Bowl/C2-GGG/L** by about 4 kcal mol⁻¹. This indicates that both templates (especially the bowl) either stabilize the helices when they are attached via their N-terminus, or destabilize them when attached via their C-terminus, or both. If this difference in stability is solely dipolar in nature then the C-terminally linked helices should be more stable to denaturation; the dipole of the bowl points up from the bottom of the bowl, and the helix macrodipole points from its N- to C-terminus (*i.e.* both point in the same direction). The higher stability of the N-terminally linked helices may come from a hydrogen bond between one of the first four “unsatisfied” backbone N-Hs in the helix, and an oxygen atom on the template (Figure 5.16), thus helping to nucleate the helix formation. The bowl has eight (acetal) oxygen atoms on its upper rim, whereas the CTB (ether) has only three. The difference in helix end-capping ability of the templates complicates any comparisons made between the stability of a helical sequences situated on a CTB template, versus the same sequence on a bowl template.

All the TASPs described in this chapter were molten globules. **Bowl/N2-GGG/L**, which proved to be a monomer in solution, and therefore should have optimal packing of its constituent helices, did not possess conformational specificity. This observation is consistent with the structures of *de novo* designed helical bundles with all-leucine residues occupying their hydrophobic core. **ArBowl/N2-GGG/L** although of similar global conformational stability to **Bowl/N2-GGG/L** (TASPs that differed by their template-to-helix linker, see Section 5.3) demonstrated a different conformational specificity by NMR.

Unfortunately it would be difficult to prove whether **ArBowl/N2-GGG/L** (which is not part of this thesis) has native-like structure, or if it has four highly mobile, mutually stabilizing helices. One indirect method to probe this question may be to replace all the leucine residues with norleucine, a residue that is known to induce molten globule structure in proteins, if the NMR spectrum of this TASP remains sharp then **ArBowl/N2-GGG/L** may also have a dynamic conformation. If **ArBowl/N2-GGG/L** does prove to induce native-like structure on TASPs with an all-leucine core, when there is precedent for molten globule structure, then it may not be a suitable model for the investigation of protein structure, however it may be useful in the construction of molecular machines.

The major finding of this chapter are that a three glycine “spacer” between the template-bound cysteine residue and the “helix” is optimal for bundling to occur. However, the best spacer was unable to overcome the self-association of the CTB three-helix bundle TASP, probably a result of the helix design. It was also apparent that both templates imparted greater stability to helices that were linked via their N- versus their C-termini. All the TASPs presented in this chapter proved to be molten globules, probably due to their all-leucine hydrophobic core. In the next chapter the central amino acid of the helix will be “mutated”, and any effect of these changes on the global conformational stability, and conformational specificity of the CTB and bowl TASPs will be investigated.

Figure 5.16: Proposed helix nucleating hydrogen bond between an “unsatisfied” helix backbone N-H, and lone pair of electrons of an acetal oxygen on bowl template.



5.6 Experimental

TASPs were synthesized as described in Chapter 3. The CD (see Table 5.7), sedimentation equilibrium, and NMR experiments were performed as outlined in Chapter 4.

The sedimentation equilibrium data used for calculating the oligomeric state of **CTB/N2-GG/L** (Figure 5.17), **CTB/N2-GGG/L** in buffer alone (Figure 5.18), and **CTB/N2-GGG/L** in 2 M KCl (Figure 5.19), are shown in Table 5.8. The sedimentation equilibrium data for **Bowl/N2-GG/L** (Figure 5.20), **Bowl/N2-GGG/L** in buffer alone (Figure 5.21), and **Bowl/N2-GGG/L** in 2 M KCl (Figure 5.22), are shown in Table 5.9. As mentioned in the previous chapter there are large errors in calculating the association constants obtained from the sedimentation equilibrium experiments.

Table 5.7: Mean residual ellipticity at 222 nm, $[\theta]_{222}$, of CTB and bowl TASP_s presented in Chapter 5.

TASP	Experimental $-\theta_{222}$	Theoretical* $-\theta_{222}$	Percent Helix
CTB/N2/L	25300	29300	86
CTB/N2-GG/L	20100	30200	67
CTB/N2-GGG/L	19300	31100	62
CTB/C2-GGG/L	16700	31100	54
Bowl/N2/L	25300	29300	86
Bowl/N2-GG/L	19500	30200	65
Bowl/N2-GGG/L	25300	31100	81
Bowl/C2-GGG/L	26400	31100	85

* Theoretical helicity is calculated according to reference 394. This calculation includes all residues in the attached peptide, and not just those designed to take part in the "helix".

Table 5.8: Data and numerical constants used for, and obtained from, linear least-squares fit of sedimentation equilibrium data.

	CTB/N2-GG/L	CTB/N2-GGG/L	CTB/N2-GGG/L
Mw [calculated] (Da)	6278	6450	6450
R (erg. mol ⁻¹ K ⁻¹)	8.314 x 10 ⁷	8.314 x 10 ⁷	8.314 x 10 ⁷
T (K)	298	298	298
\bar{v} (mL g ⁻¹)	0.7626	0.7561	0.7561
ρ (g mL ⁻¹)	1.001	1.001	1.077
ω (rad. S ⁻¹)	3456	3456	3456
Mw [average] (Da)	9000±400	7800±600	7100±300
Mw [m↔d] (Da)	6200±300	6500±200	6200±300
K _{a(abs)} [m↔d]	2.1±0.6	0.3±0.2	0.4±0.2
$\epsilon_{[\lambda]}$ (cm ⁻¹ M ⁻¹ _[nm])	3200 _[306]	4300 _[295]	2300 _[306]
c (M)	1.00x10 ⁻⁴	1.00x10 ⁻⁴	1.00x10 ⁻⁴
l (cm)	1.10	1.10	1.10
K _{a(conc.)} [m↔d] (M ⁻¹)	3700±1100	1500±400	550±200

Table 5.9: Data and numerical constants used for, and obtained from, linear least-squares fit of sedimentation equilibrium data.

	Bowl/N2-GG/L	Bowl/N2-GGG/L	Bowl/N2-GGG/L
Conditions	Buffer Only	Buffer Only	2M KCl
Mw [calculated] (Da)	8483	8713	8713
R (erg. mol ⁻¹ K ⁻¹)	8.314 x 10 ⁷	8.314 x 10 ⁷	8.314 x 10 ⁷
T (K)	298	298	298
\bar{v} (mL g ⁻¹)	0.7626	0.7561	0.7561
ρ (g mL ⁻¹)	1.001	1.001	1.077
ω (rad. S ⁻¹)	3456	3456	3456
Mw [average] (Da)	10200±500	8000±600	7800±1600
Mw [m↔d] (Da)	8500±300	-	-
K _{a(abs)} [m↔d]	0.23±0.08	-	-
$\epsilon_{[\lambda]}$ (cm ⁻¹ M ⁻¹ _[nm])	5300 _[309]	1770 _[295]	1050 _[306]
c (M)	1.00x10 ⁻⁴	1.00x10 ⁻⁴	1.00x10 ⁻⁴
l (cm)	1.10	1.10	1.10
K _{a(conc.)} [m↔d] (M ⁻¹)	700±200	-	-

Figure 5.17: Sedimentation equilibrium analysis of **CTB/N2-GG/L** (100 μM) at 25 $^{\circ}\text{C}$ in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

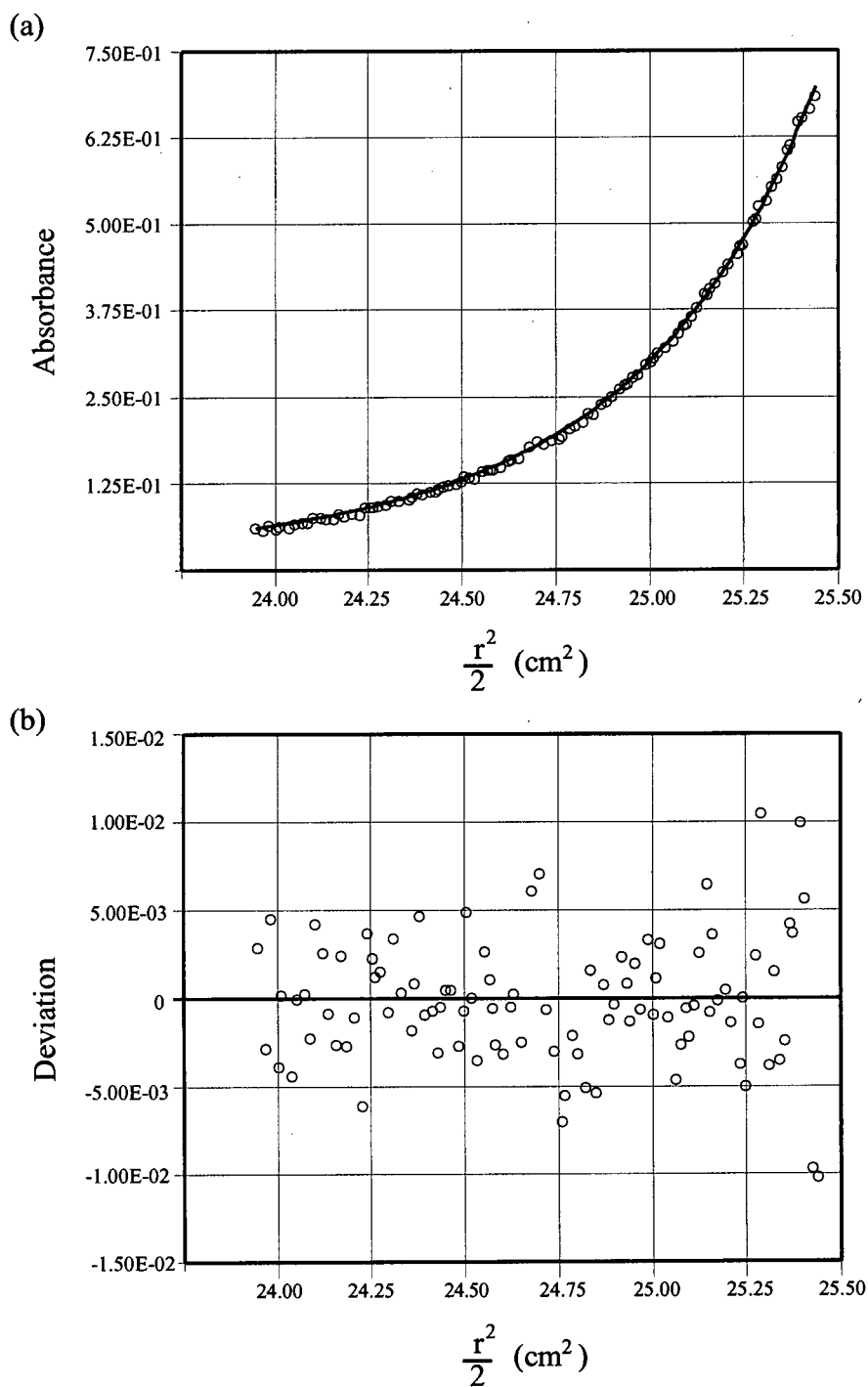


Figure 5.18: Sedimentation equilibrium analysis of **CTB/N2-GGG/L** (100 μM) at 25 $^{\circ}\text{C}$ in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

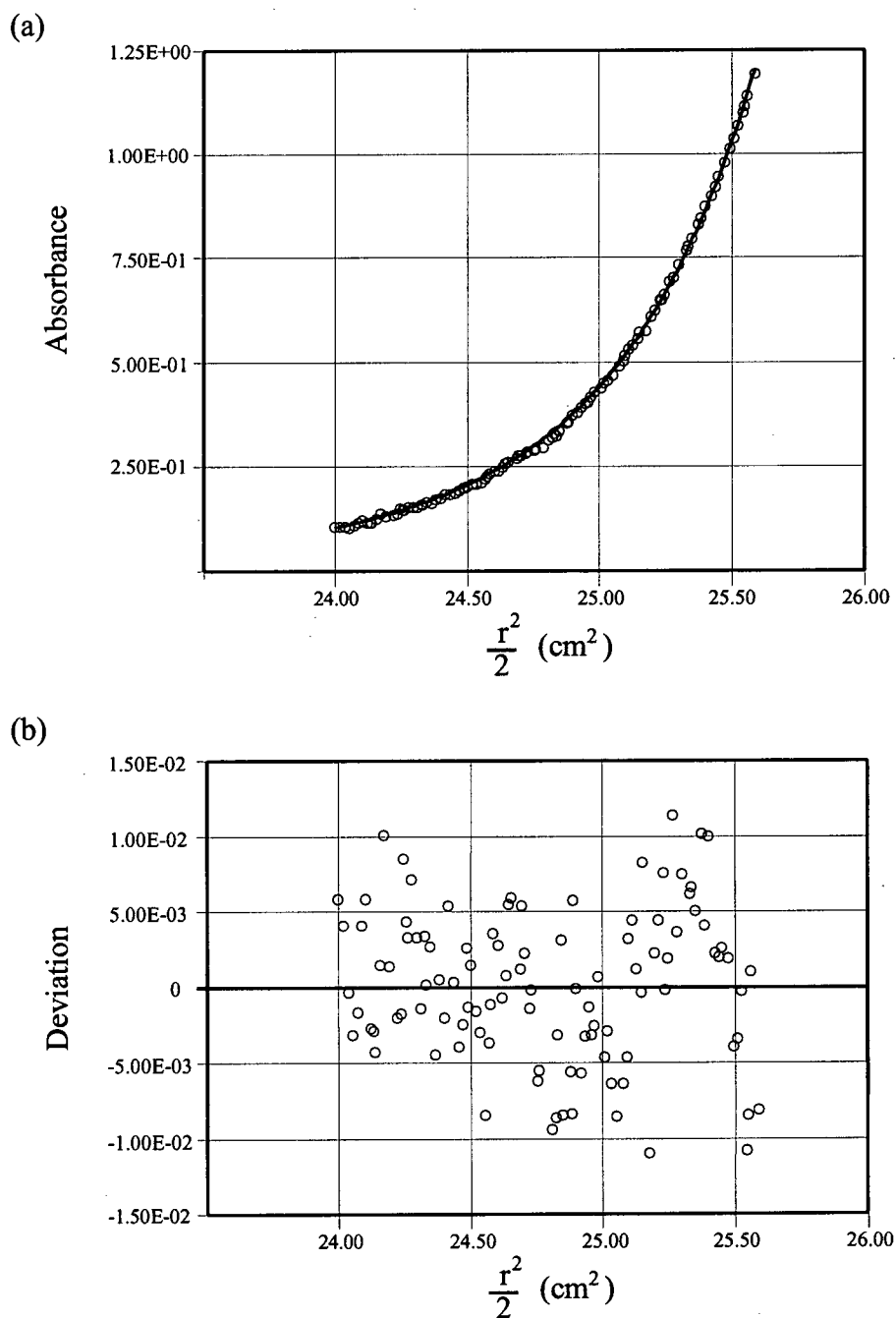


Figure 5.19: Sedimentation equilibrium analysis of **CTB/N2-GGG/L** (100 μM) at 25 $^{\circ}\text{C}$ in 2 M KCl. (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

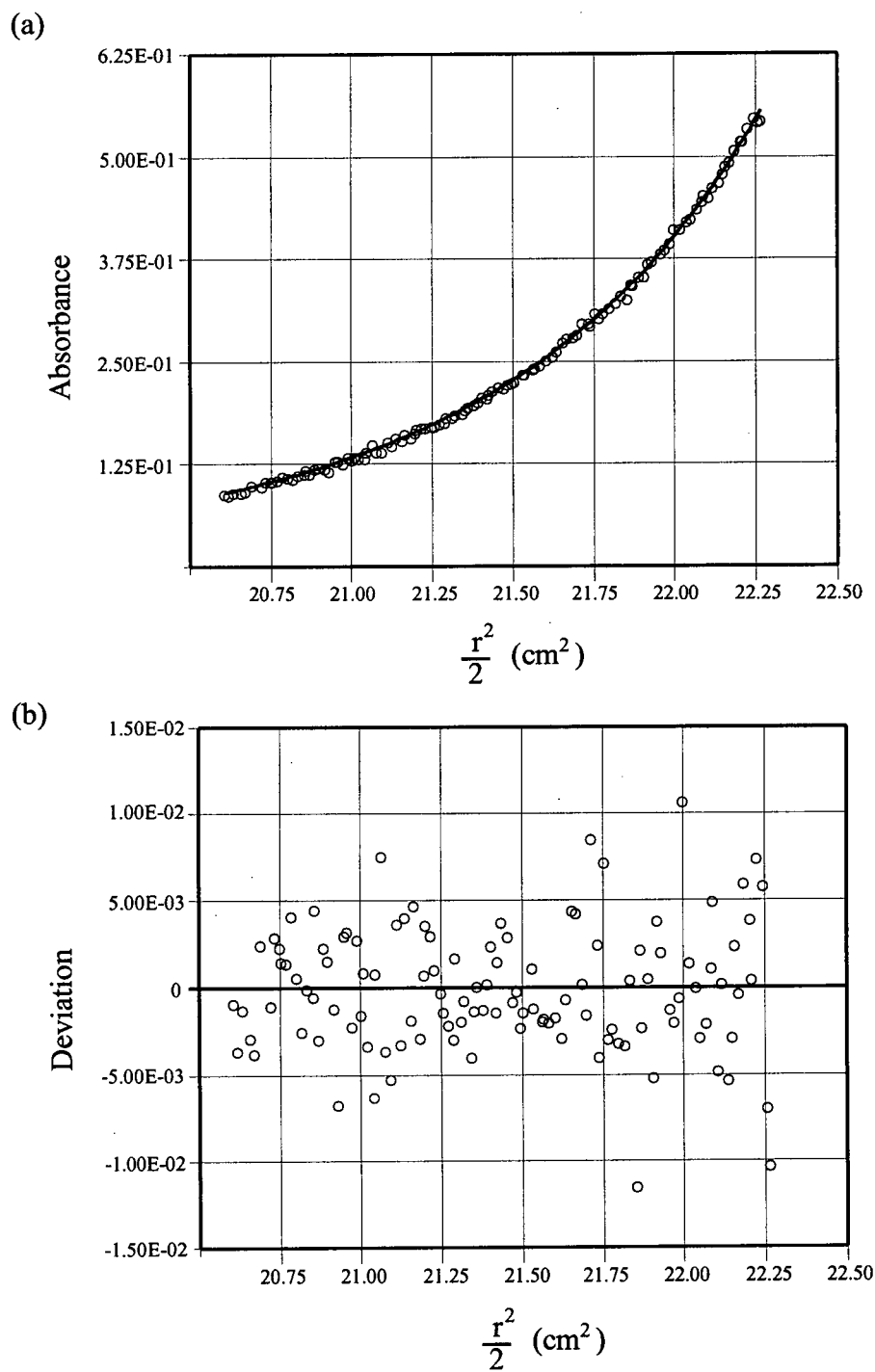


Figure 5.20: Sedimentation equilibrium analysis of **Bowl/N2-GG/L** (100 μ M) at 25 $^{\circ}$ C in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer-dimer in equilibrium. (b) Scatter of data points around best-fit line.

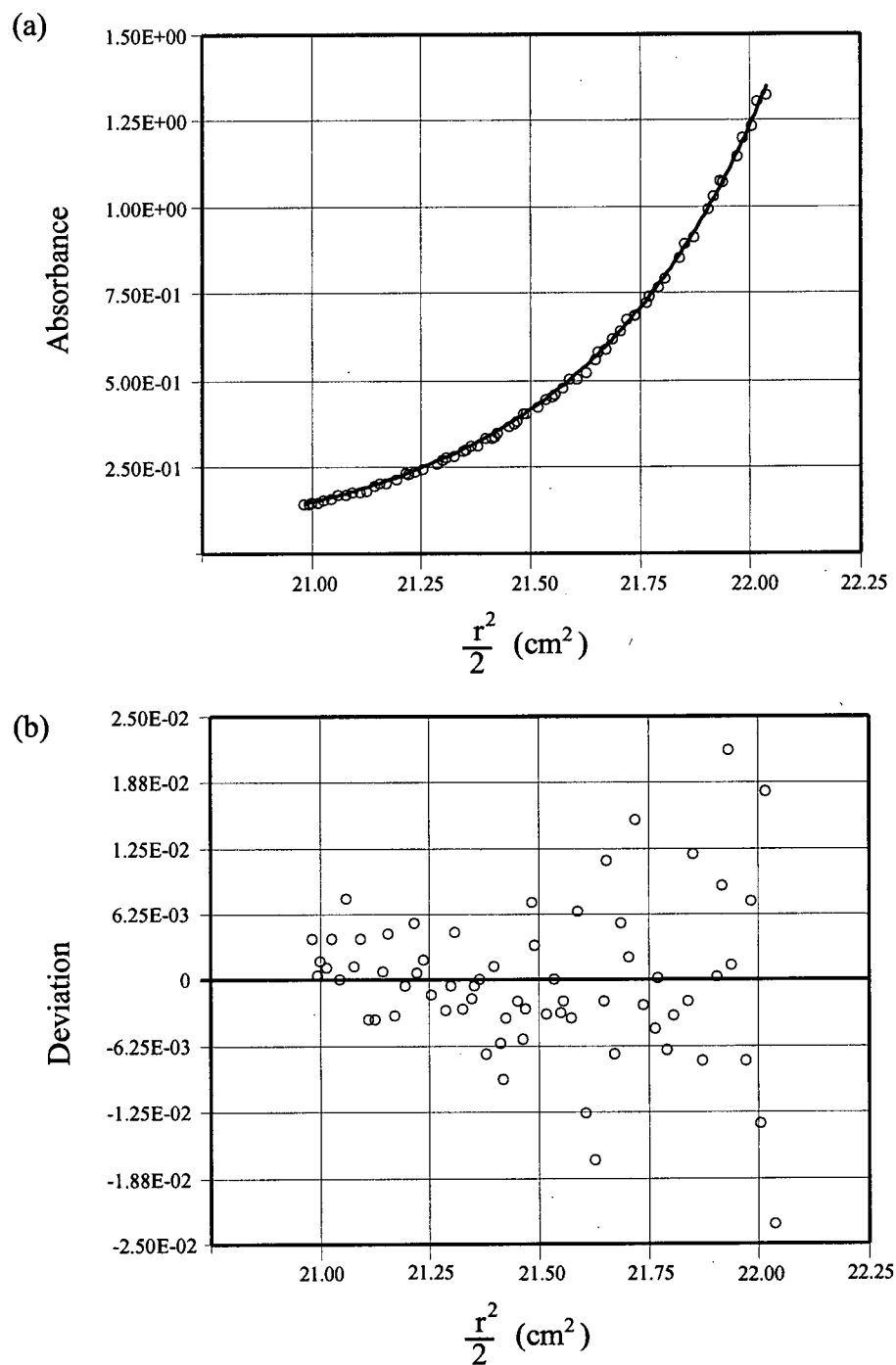


Figure 5.21: Sedimentation equilibrium analysis of **Bowl/N2-GGG/L** (100 μM) at 25 $^{\circ}\text{C}$ in pH 7.0 phosphate buffer (50 mM). (a) Raw data, the solid line represents the theoretical fit to a monomer. (b) Scatter of data points around best-fit line.

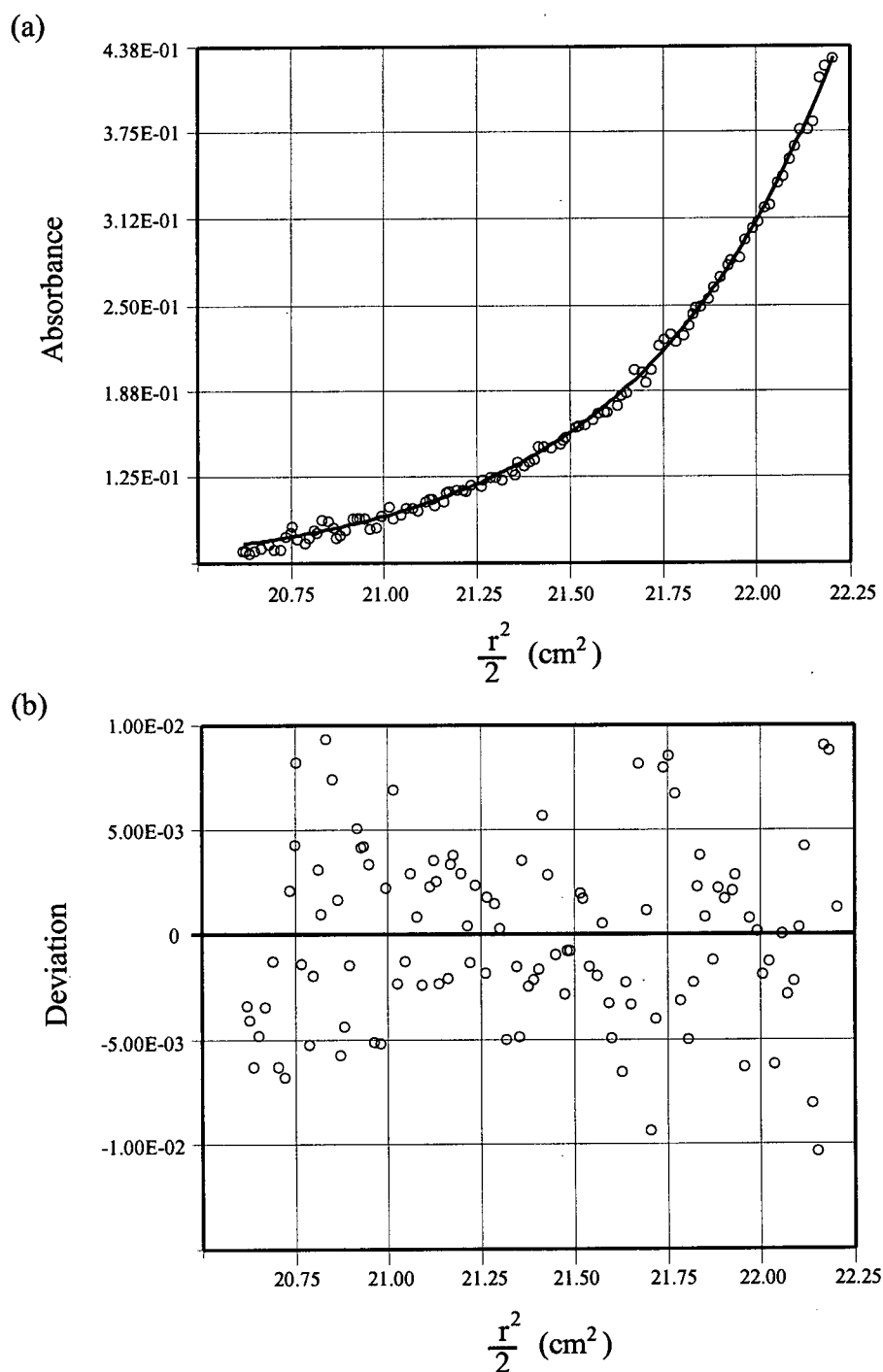
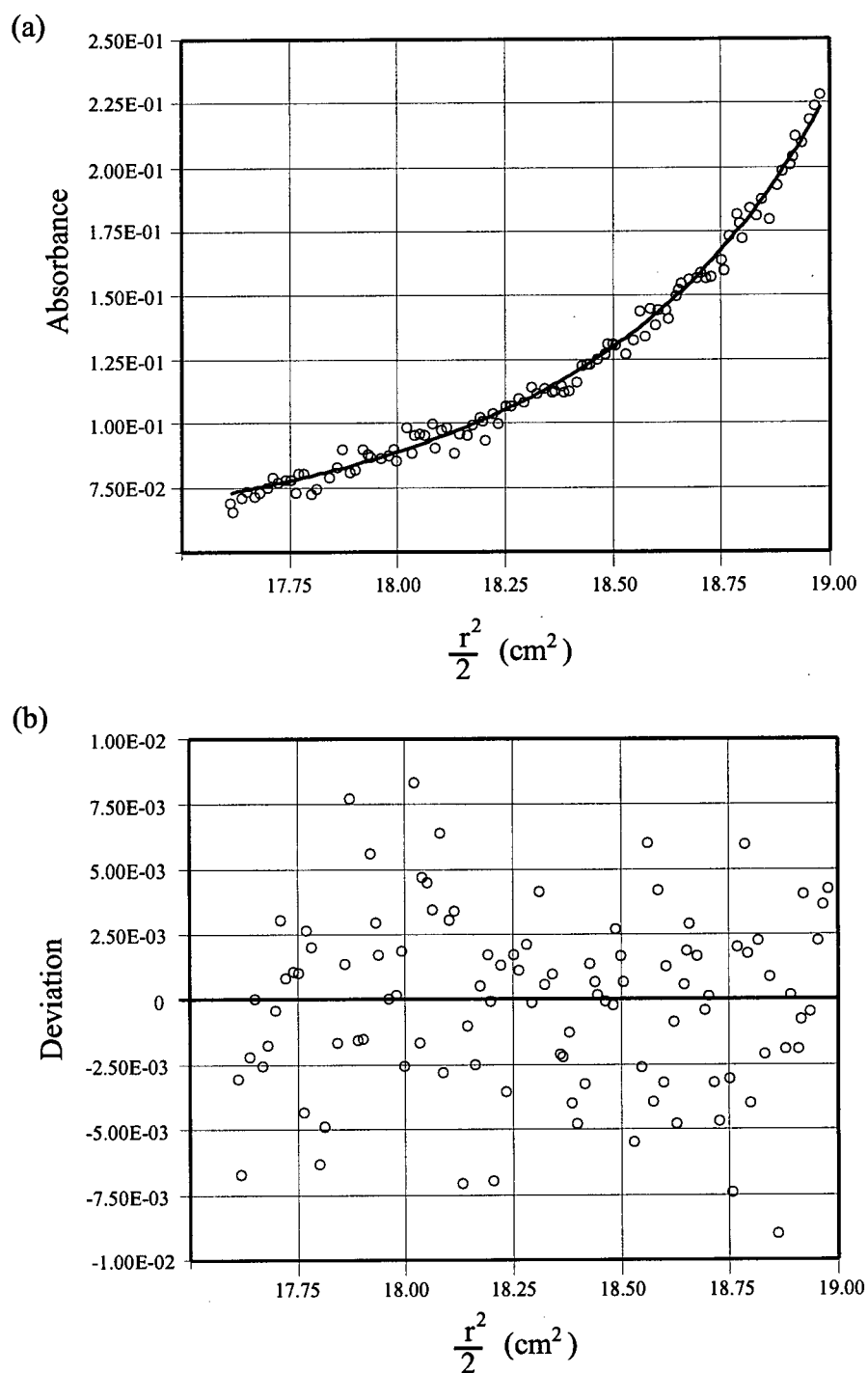


Figure 5.22: Sedimentation equilibrium analysis of **Bowl/N2-GGG/L** (100 μM) at 25 $^{\circ}\text{C}$ in 2M KCl. (a) Raw data, the solid line represents the theoretical fit to a monomer. (b) Scatter of data points around best-fit line.



Effect of Changing the Central Leucine Amino Acid in the Helices of CTB and Bowl TASP

6.0 Introduction

In the previous chapter we determined that a three glycine spacer between the template-bound cysteine residue and the “helix” is optimal for the intra-molecular bundling of helices on a CTB or bowl TASP. However, all the TASP s produced in Chapters 4 and 5 lacked the conformational specificity usually associated with natural proteins. Therefore, in this chapter we are working towards the production of TASP s with improved conformational specificity. We used peptides with a three glycine spacer between the template-bound cysteine residue and the “helix” for this investigation into the effect of changing the central leucine residue on TASP stability and structure. The peptides were all attached via their N-termini.

The amino acids in the central portion of an α -helix are the most important in determining its overall stability and structure.⁵ This is especially true for the amino acids in the core of α -helical bundles.⁵ Thus, we investigated the effect of changing the central leucine residue of the “helix” on the overall structure of CTB and Bowl TASP s. The central leucine residue of peptide N2-GGG/L (Ac-CGGGEELLKK~~X~~EELLKKG-NH₂ , where X =

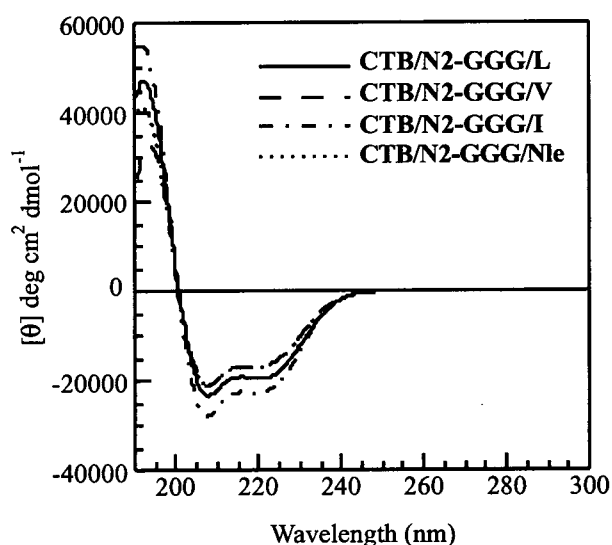
L), was “mutated” to either isoleucine (peptide **N2-GGG/I**, **X** = I), valine (peptide **N2-GGG/V**, **X** = V), or the unnatural amino acid norleucine (peptide **N2-GGG/Nle**, **X** = Nle). Leucine, isoleucine, and norleucine are structural isomers of each other (iso-butyl, sec-butyl, or n-butyl side chain, respectively): Thus, we could investigate any effect on the TASP structure imparted by subtle differences in the size and shape of the central, core amino acid. Changes in the size and shape of core amino acids have been shown to affect both the structural specificity and the global stability of proteins,^{210,353} as well as altering the quaternary structure (association state) of coiled-coils.¹⁸⁴ The β -branched amino acids, isoleucine and valine are known to induce conformational specificity into helical bundles, a feature that is generally attributed to the fact that the side chain of β -branched amino acids can adopt only one low energy conformation when in an α -helix.⁴⁰⁶⁻⁴⁰⁹ Additionally, valine has a smaller side chain than isoleucine, so any effects of the size of the side chain on the stability of the TASP could also be probed. The unnatural amino acid norleucine has been shown to induce molten globule structure into natural proteins, and was therefore introduced into the CTB and bowl TASPs in an attempt to produce a “fast-exchange” molten globule (see section 5.5) with a sharper NMR spectrum.

6.1 CTB/N2-GGG/I, CTB/N2-GGG/V, and CTB/N2-GGG/Nle

These TASP s were synthesized by linking the CTB macrocycle (via disulfide bonds) to either peptide **N2-GGG/I** (Ac-CGGGEELLKK~~X~~EELLKKG-NH₂, X = I), **N2-GGG/V** (X = V), or **N2-GGG/Nle** (X = Nle). The concentration independent CD spectra (Figure 6.1) of **CTB/N2-GGG/I**, **CTB/N2-GGG/V**, and **CTB/N2-GGG/Nle**, indicated they were all helical. The CD spectra were unaffected by salt and did not exhibit any significant near-UV CD activity.

The GnHCl-induced unfolding curves for of **CTB/N2-GGG/I**, **CTB/N2-GGG/V**, and **CTB/N2-GGG/Nle** were concentration independent (Figure 6.2). As we saw in Chapter 4, this observation is consistent with either a monomer, or a species that is a monomer at the unfolding transition point (sedimentation equilibrium ultracentrifugation experiments were not run on these TASP s due to limited instrument time).

Figure 6.1: CD spectra of **CTB/N2-GGG/I** (59 μ M, dotted and dashed line), **CTB/N2-GGG/V** (60 μ M, dashed line), and **CTB/N2-GGG/Nle** (54 μ M, dotted line), compared to **CTB/N2-GGG/L** (60 μ M, dotted and dashed line). All spectra were recorded in pH 7.0 phosphate buffer (10 mM) at 25 °C. NOTE: The spectra of **CTB/N2-GGG/L** and **CTB/N2-GGG/Nle** are overlapping.



The global conformational stability of these CTB TASP variants was assessed from their GnHCl-induced unfolding curves (Table 6.1). **CTB/N2-GGG/L** and **CTB/N2-GGG/I** have similar stability, **CTB/N2-GGG/Nle** is marginally less stable (by about 1 kcal mol⁻¹), and **CTB/N2-GGG/V** is the least stable (by about 2.5 kcal/mol⁻¹ when compared to **CTB/N2-GGG/L**). A protein's global conformational stability is related to the amount of buried hydrophobic surface (and other factors) in the folded structure. Thus, it is not surprising that **CTB/N2-GGG/V**, which has three less methylene units buried in its core is

the least stable. Remember that these changes in the TASP's conformational stability are context dependent, *i.e.* the central amino acid under investigation is surrounded by four leucine amino acids in the core of the helical bundle. The differences in the stability of these CTB TASP will be discussed further in Section 6.3.

The NMR spectra of **CTB/N2-GGG/I**, **CTB/N2-GGG/V**, and **CTB/N2-GGG/Nle** (Figure 6.3 – compared to **CTB/N2-GGG/L**) all resembled molten globules. Thus, there was no noticeable difference in the conformation specificity in these CTB TASP's: This may be because the hydrophobic patterning in the helix is more suited to forming a four-helix bundle, a structure that will be investigated in the next section.

Figure 6.2: GnHCl-induced denaturation curves for (a) **CTB/N2-GGG/I** (32 μ M, \circ , dashed line; 3.2 μ M, \times , solid line); (b) **CTB/N2-GGG/V** (31 μ M, \circ , dashed line; 3.1 μ M, \times , solid line); (c) **CTB/N2-GG/Nle** (39 μ M, \circ , dashed line; 3.9 μ M, \times , solid line); (d) **CTB/N2-GGG/I** (3.2 μ M, \circ , long dashed line), **CTB/N2-GGG/V** (3.1 μ M, \diamond , medium dashed line), **CTB/N2-GGG/Nle** (3.9 μ M, \square , short dashed line), compared to **CTB/N2-GGG/L** (4.0 μ M, \times , solid line). All spectra recorded in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.

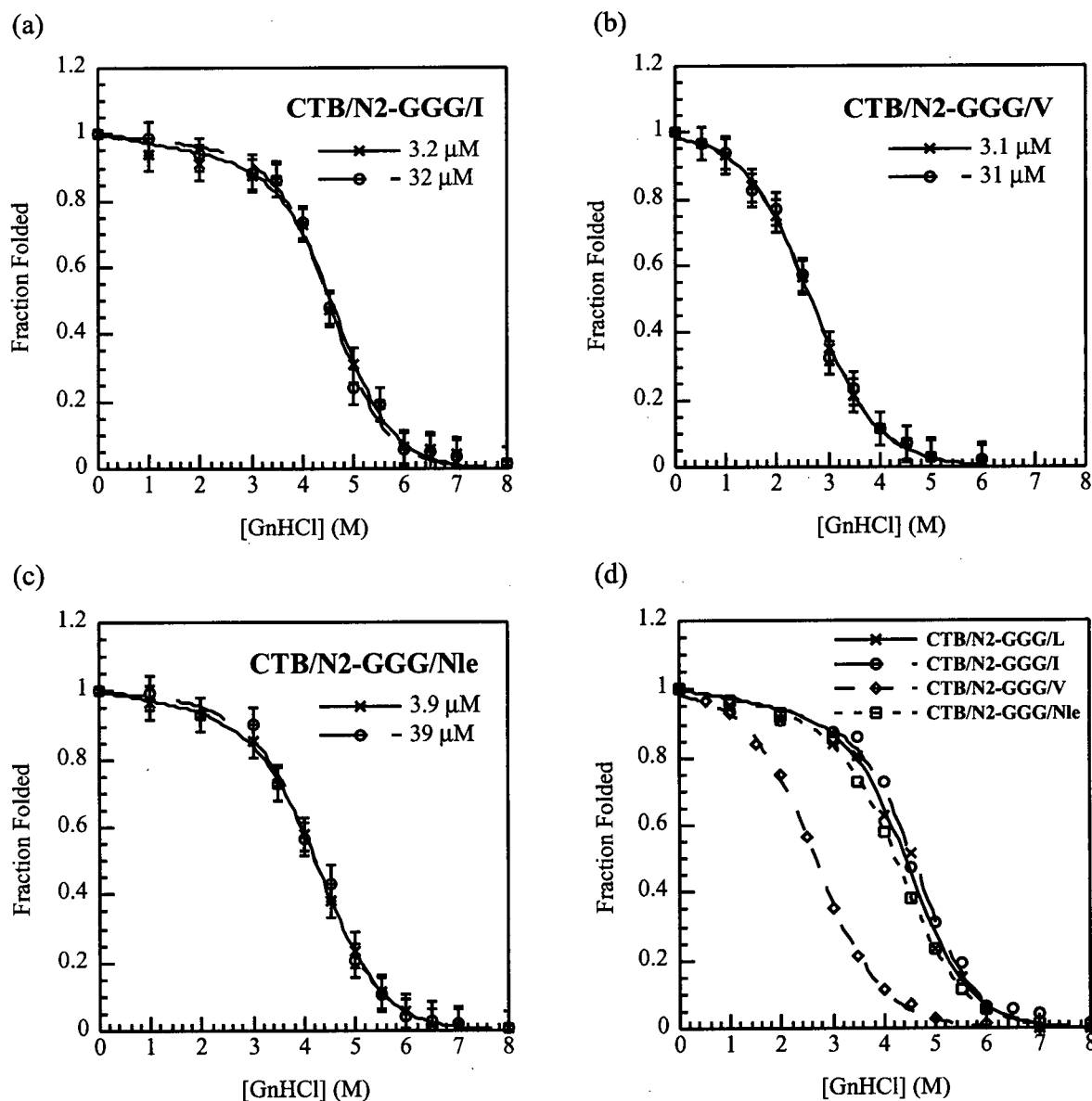
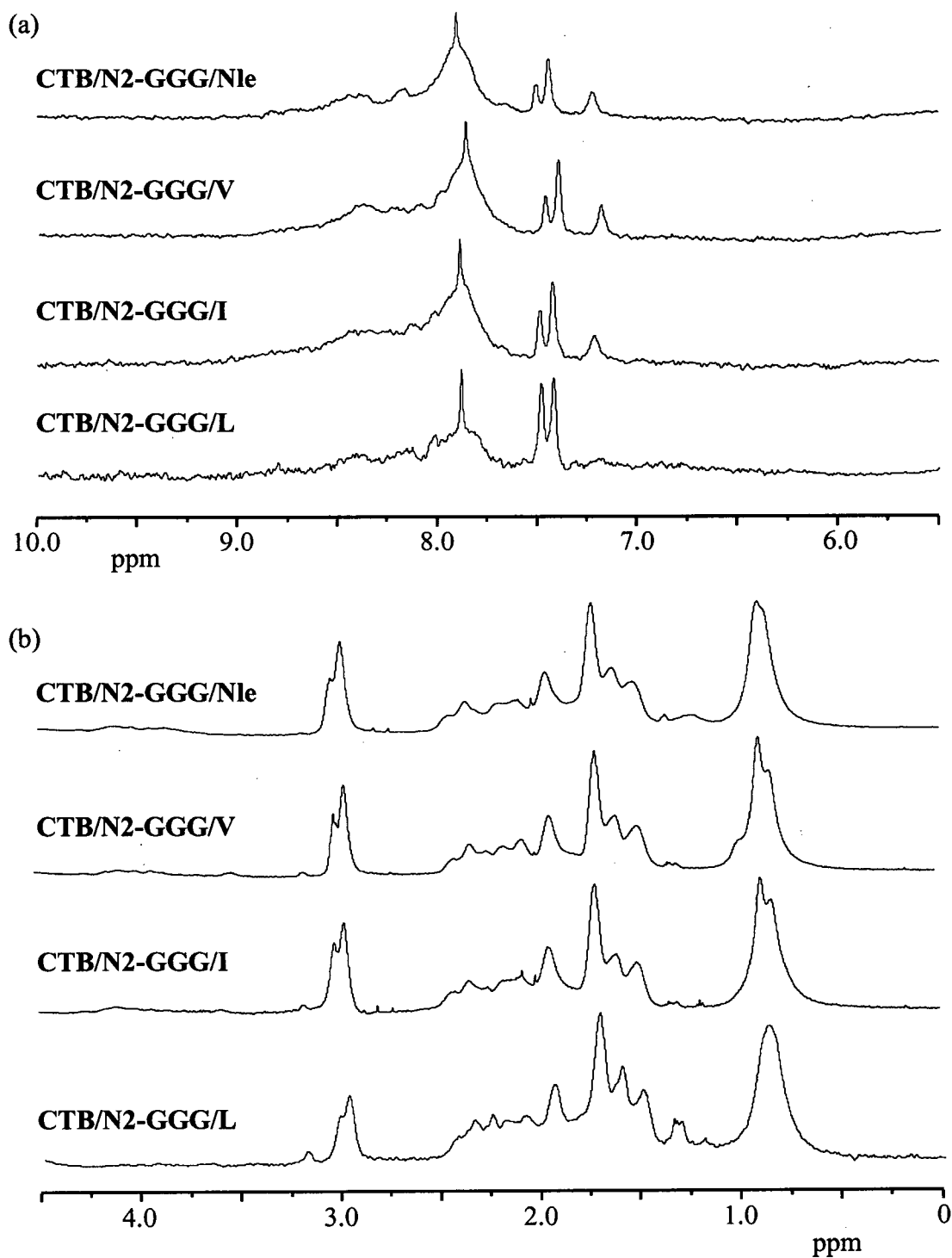


Table 6.1: Thermodynamic evaluation of **CTB/N2-GGG/I**, **CTB/N2-GGG/V**, **CTB/N2-GGG/Nle**, compared to **CTB/N2-GGG/L**, calculated from their **GnHCl**-induced denaturation.

TASP	Concentration (μM)	$[\text{GnHCl}]_{0.5}$ (M)	m $\text{kcal mol}^{-1} \text{M}^{-1}$	$\Delta G^\circ \text{H}_2\text{O}$ kcal mol^{-1}
CTB/N2-GGG/L	4.0	4.4 ± 0.1	1.0 ± 0.1	4.7 ± 0.5
CTB/N2-GGG/L	40	4.5 ± 0.1	1.1 ± 0.1	5.1 ± 0.6
CTB/N2-GGG/I	3.2	4.8 ± 0.1	1.1 ± 0.1	5.0 ± 0.6
CTB/N2-GGG/I	32	4.7 ± 0.1	1.1 ± 0.1	5.1 ± 0.6
CTB/N2-GGG/V	31	2.6 ± 0.1	0.9 ± 0.1	2.4 ± 0.1
CTB/N2-GGG/V	3.1	2.6 ± 0.1	0.9 ± 0.1	2.4 ± 0.1
CTB/N2-GGG/Nle	3.9	4.2 ± 0.1	1.0 ± 0.1	4.1 ± 0.1
CTB/N2-GGG/Nle	39	4.2 ± 0.1	1.1 ± 0.1	4.1 ± 0.1

Figure 6.3: NMR spectra of **CTB/N2-GGG/L**, **CTB/N2-GGG/V**, and **CTB/N2-GGG/Nle**, compared to **CTB/N2-GGG/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.

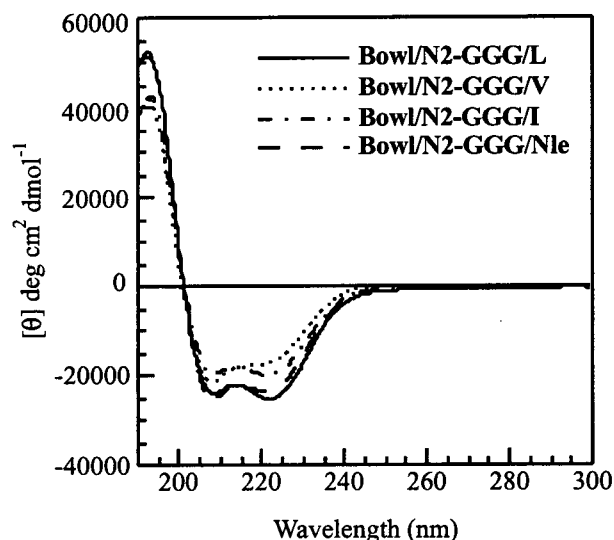


6.2 Bowl/N2-GGG/I, Bowl/N2-GGG/V, and Bowl/N2-GGG/Nle

These TASP s were synthesized (via disulfide bonds) using the bowl macrocycle and peptides that differed in their central hydrophobic amino acid (namely, **N2-GGG/I**, **N2-GGG/V**, and **N2-GGG/Nle**). The concentration independent CD spectra (Figure 6.4) of **Bowl/N2-GGG/I**, **Bowl/N2-GGG/V**, and **Bowl/N2-GGG/Nle** indicated the presence of helical structure. None of these TASP s had any significant signal in its near-UV CD region (consistent with a flexible linker). The CD spectrum **Bowl/N2-GGG/Nle** was unaffected in the presence of salt (2 M KCl), whereas **Bowl/N2-GGG/I** lost 10 % of its helicity, and **Bowl/N2-GGG/V** lost 24 % of its helicity. This suggests that electrostatic interactions (between the helices – see Section 6.4 for more details) are more important to the structure of **Bowl/N2-GGG/I** and **Bowl/N2-GGG/V**.

The GnHCl-induced unfolding curves of **Bowl/N2-GGG/I**, **Bowl/N2-GGG/V**, and **Bowl/N2-GGG/Nle** were concentration independent (Figure 6.5), indicating that these TASP s are all monomers near the unfolding transition point. The presence of very low concentrations of GnHCl, resulted in a 25 % loss of helicity for **Bowl/N2-GGG/V**, before its unfolding transition: This “step down” before the unfolding point would not yield a two-state unfolding curve, therefore the GnHCl-induced denaturation curves for this TASP were run in the presence of 100 mM KCl (this gave a similar transition point for the unfolding curve of **Bowl/N2-GGG/V** to that obtained in the absence of salt).

Figure 6.4: CD spectra of **Bowl/N2-GGG/I** (48 μ M, dotted and dashed line), **Bowl/N2-GGG/V** (47 μ M, dotted line), and **Bowl/N2-GGG/Nle** (59 μ M, dashed line), compared to **Bowl/N2-GGG/L** (40 μ M, solid line). All CD spectra were run in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C.



The global conformational stability of these bowl variants was assessed from their GnHCl-induced denaturation curves (Table 6.2). **Bowl/N2-GGG/I** and **Bowl/N2-GGG/Nle** appeared to have equal global stability, but both were less stable than their isomer **Bowl/N2-GGG/L** by about 2 kcal mol⁻¹. **Bowl/N2-GGG/V** was significantly less stable than the other bowl TASPs, about 5.5 kcal mol⁻¹ less than **Bowl/N2-GGG/L**. This difference can be primarily attributed to **Bowl/N2-GGG/V** containing four fewer methylene units in its hydrophobic core.

Figure 6.5: GnHCl-induced denaturation curves for (a) **Bowl/N2-GGG/I** (28 μ M, \circ , dashed line; 2.8 μ M, \times , solid line); (b) **Bowl/N2-GGG/V** (32 μ M, \circ , dashed line; 3.2 μ M, \times , solid line); (c) **Bowl/N2-GGG/Nle** (27 μ M, \circ , dashed line; 2.7 μ M, \times , solid line); (d) **Bowl/N2-GGG/I** (2.8 μ M, \circ , long dashed line), **Bowl/N2-GGG/V** (3.2 μ M, \diamond , medium dashed line), **Bowl/N2-GGG/Nle** (2.7 μ M, \square , short dashed line), compared to **Bowl/N2-GGG/L** (3.0 μ M, \times , solid line). All spectra were recorded in pH 7.0 phosphate buffer (10 mM) at 25 $^{\circ}$ C, **Bowl/N2-GGG/V** was recorded in the additional presence of 100 mM KCl.

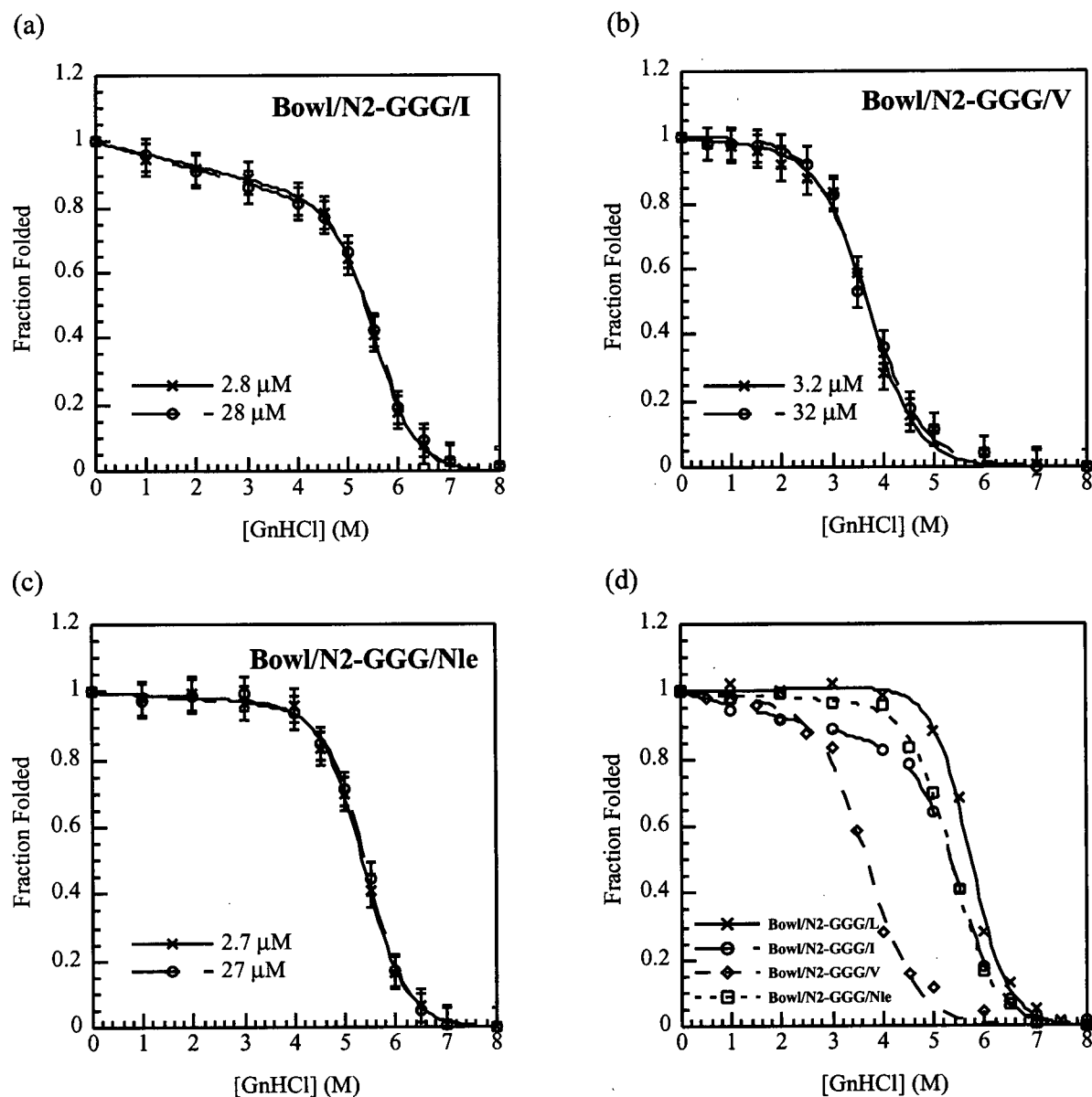
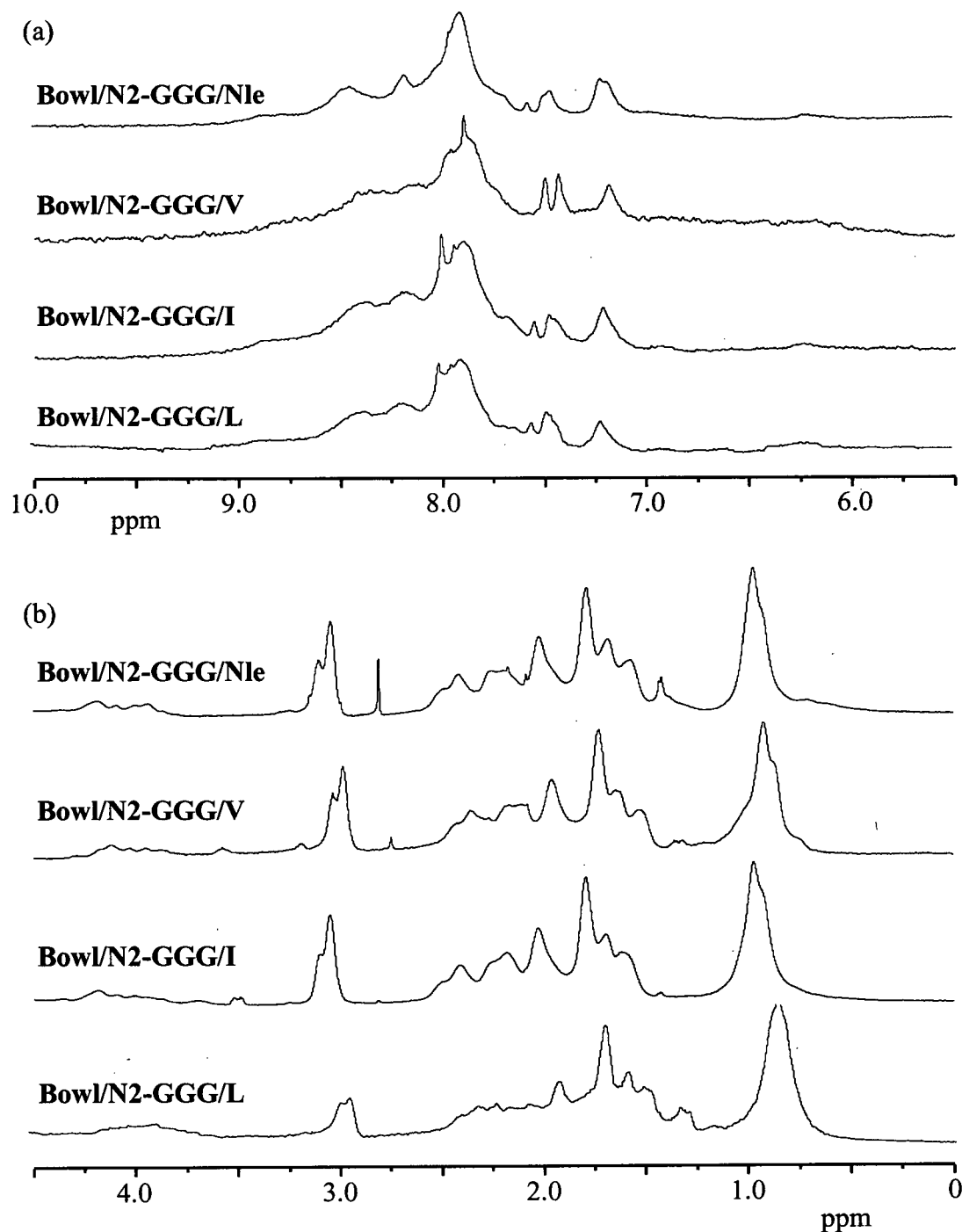


Table 6.2: Thermodynamic evaluation of **Bowl/N2-GGG/I**, **Bowl/N2-GGG/V**, **Bowl/N2-GGG/Nle**, compared to **Bowl/N2-GGG/L**, calculated from their GnHCl-induced denaturation.

TASP	Concentration (μ M)	[GnHCl] _{0.5} (M)	<i>m</i> kcal mol ⁻¹ M ⁻¹	$\Delta G^\circ_{H_2O}$ kcal mol ⁻¹
Bowl/N2-GGG/L	3.0	5.7 \pm 0.1	1.7 \pm 0.1	9.4 \pm 0.7
Bowl/N2-GGG/L	30	5.7 \pm 0.1	1.8 \pm 0.2	10.2 \pm 1.1
Bowl/N2-GGG/I	2.8	5.3 \pm 0.1	1.5 \pm 0.2	8.2 \pm 0.3
Bowl/N2-GGG/I	28	5.4 \pm 0.1	1.5 \pm 0.2	8.4 \pm 0.5
Bowl/N2-GGG/V	3.2	3.6 \pm 0.1	1.2 \pm 0.1	4.6 \pm 0.4
Bowl/N2-GGG/V	32	3.6 \pm 0.1	1.1 \pm 0.1	4.0 \pm 0.4
Bowl/N2-GGG/Nle	2.7	5.4 \pm 0.1	1.4 \pm 0.2	7.8 \pm 0.4
Bowl/N2-GGG/Nle	27	5.4 \pm 0.1	1.5 \pm 0.2	8.1 \pm 0.5

The NMR spectra of **Bowl/N2-GGG/I**, **Bowl/N2-GGG/V**, and **Bowl/N2-GGG/Nle** were broad (Figure 6.6), therefore there was no noticeable change in their conformational specificity relative to **Bowl/N2-GGG/L**.

Figure 6.6: NMR spectra of **Bowl/N2-GGG/L**, **Bowl/N2-GGG/V**, and **Bowl/N2-GGG/Nle**, compared to **Bowl/N2-GGG/L**. TASP concentration was ~ 0.3 mM, spectra were recorded at 298 K in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-5.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.



6.3 Further Investigation of TASP Structure

The global conformation of a representative TASP, namely **Bowl/N2-GGG/I**, was investigated by a temperature dependent NMR study (Figure 6.7). The effect of raising the temperature on a molten globule structure is to increase the intra-molecular motion, resulting in a dynamic averaging for each N-H resonance and a sharper NMR spectrum (but with a loss in chemical shift dispersion). Decreasing the temperature may “freeze out” the lowest energy conformation of the molten globule, possibly leading to a tertiary structure which exhibits conformational specificity, resulting in a sharp NMR spectrum with greater chemical shift dispersion in its N-H region. We also investigated the effect of trifluoroethanol (TFE) on the NMR spectrum of this TASP (Figure 6.8). TFE is thought to disrupt the tertiary interactions within a protein, at the same time promoting α -helix structure by promoting intra-helix backbone hydrogen bonds. Therefore we can look at the NMR spectrum of what are essentially isolated α -helices on the TASP.

For *de novo* proteins with degenerate amino acid sequences, the upfield aliphatic-H region of the NMR spectrum often contains multiple, overlapping peaks, therefore a sharp, well dispersed downfield N-H region is considered necessary for the presence of native-like protein structure.

At 358 K (high temperature) **Bowl/N2-GGG/I** is still helical by CD, and its NMR spectrum appears to be sharper (Figure 6.7). Some of the N-Hs were not visible at this temperature, presumably due to rapid exchange with the solvent. This is consistent with the

faster motion of the helices in the context of the tertiary structure. At 358 K (high temperature), the increased motion (and therefore “averaging”) of the TASP structure also results in the appearance of peaks associated with the cavitand template (at 4.7 and 5.9 ppm). Cooling to 277 K leads to broadening of the N-H region of the NMR spectrum. This is either due to the slower motion of the tertiary (molten globule-like) structure, or a reduction in the rate of tumbling of the TASP (therefore, longer relaxation times for the protons, however the aliphatic region of the NMR spectrum does not broaden by an equivalent amount). These observations are consistent with a temperature dependent NMR study of melittin, a tetrameric α -helical protein that consisted of a dimer of dimers.⁴¹⁰ As temperature increased, the protein sub-units remained associated and helical, the NMR spectrum became sharper and less well dispersed, but not consistent with a random coil. The authors concluded that the inter-helical contact (which is primarily via hydrophobic side chains) between the monomer units lost its conformational specificity (*i.e.* fast motion of helices which are slipping over each other, also the inter-helical nOe’s between the side chains disappeared).

Addition of TFE results in a sharper, and more dispersed N-H region of the NMR spectrum of **Bowl/N2-GGG/I** (Figure 6.8). Thus, a reduction of the tertiary interactions between the helices (induced by TFE), also results in a better defined N-H region of the NMR spectrum: This result, coupled with the high temperature NMR data, is consistent with the inter-helical interactions between the side chains being responsible for the molten globule structure of the TASP.

In other experiments, we attempted to titrate small, non-polar molecules into TASP solutions, to see if there would be any binding in the hydrophobic core. Small non-polar molecules have been shown to bind into cavities in a protein's core, leading to a change in conformation and stability.^{411,412} Also, the binding of small, non-polar molecules has led to allosteric control of the association state of coiled-coil proteins.⁴¹³ **Bowl/N2-GGG/V** was chosen for this study because it is most likely to have a "hole" in its hydrophobic core due to the contextual leucine-lined core. However, the addition of benzene (forming a saturated solution in aqueous buffer) resulted in no change in the GnHCl-induced unfolding curve, and no change in its NMR spectrum. It is possible that conformational specificity is required before any significant binding of a non-polar molecule into the hydrophobic core can occur; or there may be no cavity since the structure may rearrange to minimize the presence of any thermodynamically unfavourable cavity; or benzene may be a poor guest.

Figure 6.7: NMR spectra of **Bowl/N2-GGG/I** at 277, 298 and 358 K. TASP concentration was ~ 0.3 mM, spectra were recorded in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-4.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.

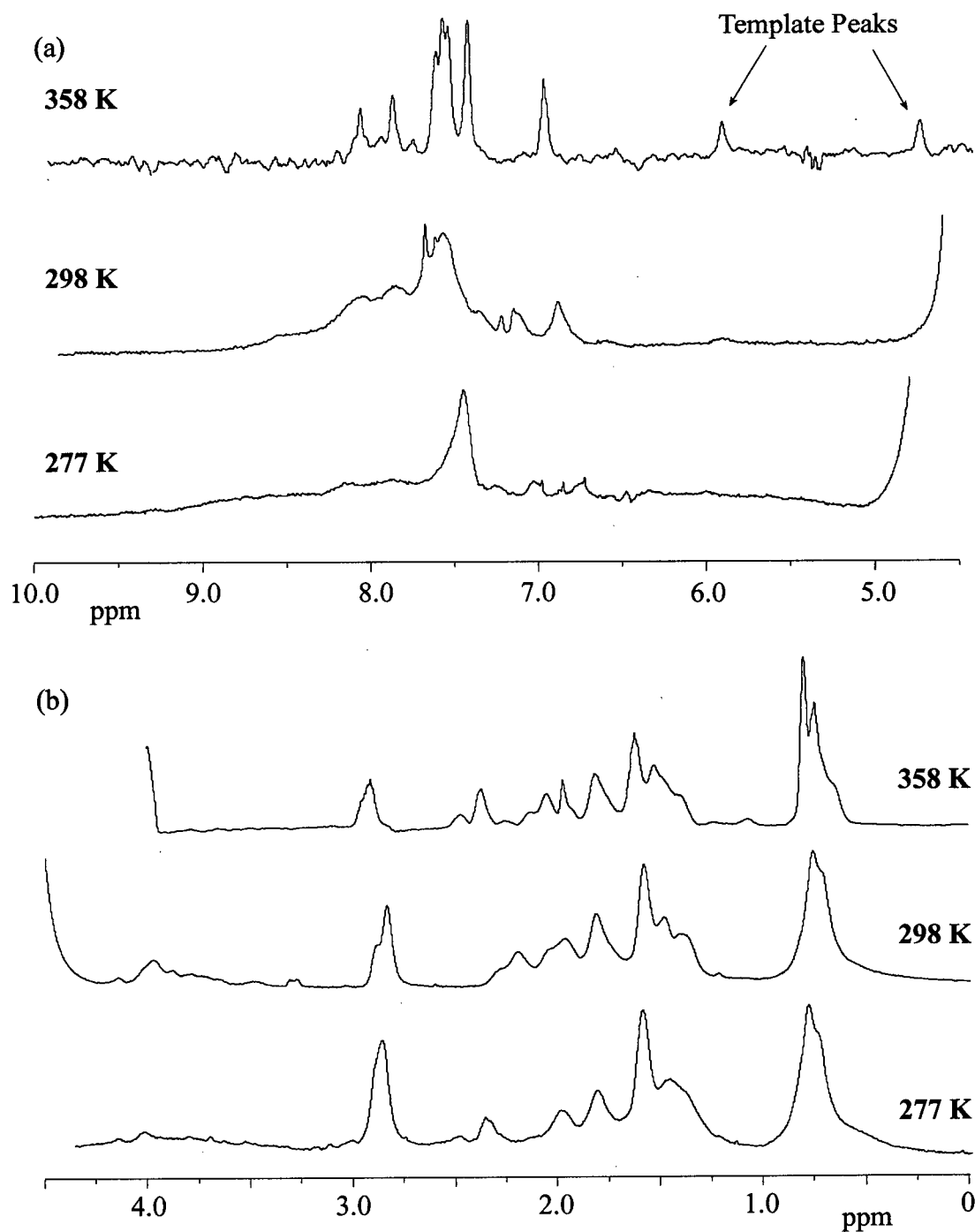
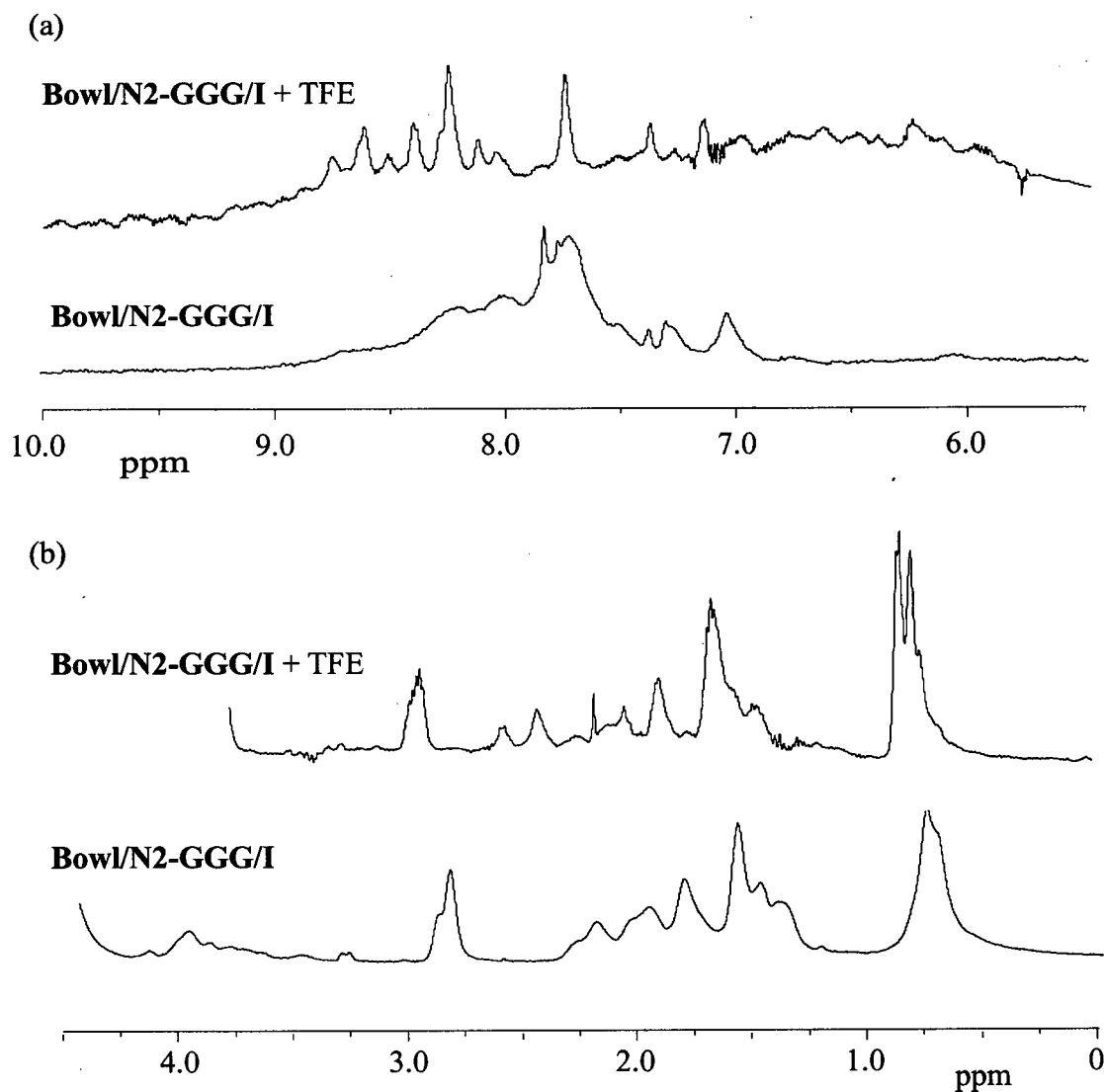


Figure 6.8: NMR spectra of **Bowl/N2-GGG/I** in pH 7.0 phosphate buffer and in 1:9 TFE:buffer. TASP concentration was ~ 0.3 mM, spectra were recorded in 45 mM phosphate buffer (pH 7.0). (a) Downfield (N-H, aromatic-H) region, 10.0-4.5 ppm (magnified); and (b) Upfield (aliphatic-H) region, 4.5-0.0 ppm.



6.4 Conclusions for Chapter 6

The effect of changing the central hydrophobic amino acid of the TASP-helices was investigated in this chapter, primarily in an attempt to increase the conformational specificity of the CTB and bowl TASPs. However, these changes still resulted in CTB and bowl TASPs with molten globule structures.

The global conformational stability (calculated from the GnHCl-induced unfolding curves) of each TASP is a reflection of the energy difference between the folded and unfolded state of the protein. The bowl TASPs were more stable than the related CTB TASPs (Table 6.3). This is due to the greater hydrophobic volume that is buried in the four- versus the three-helix bundles, and also in part to the greater N-capping ability of the cavitand bowl versus CTB template (see Chapter 5). As mentioned in the previous chapter, the CTB TASPs may have an over-packed hydrophobic core that further reduces the magnitude of the energy change upon unfolding.

Any change in the global conformational stability of these TASPs is related to the properties of the “mutated” amino acid residue in the helix, specifically its hydrophobicity, helix propensity, and any size/shape effect it has on the bundle structure. The hydrophobicity of the leucine, isoleucine, and norleucine side chains are roughly equivalent, and greater than valine. The preference of each amino acid to take part in an α -helix (*i.e.* its helix propensity, see Section 1.4.1.1) is in the order; norleucine \approx leucine>isoleucine>valine. Any size/shape effects that the side chain of the “mutated” amino acid has on the overall

tertiary structure of the TASP may also be manifested by a change in global conformational stability.

For the CTB TASP, the global stability to GnHCl denaturation did not appear significantly different between **CTB/N2-GGG/L** and **CTB/N2-GGG/I**; **CTB/N2-GGG/Nle** was marginally less stable than the other two TASP that contained the same number of carbon atoms in their core. The lack of conformational specificity may negate any effect the shape of the side chain has on the conformational stability. However, **CTB/N2-GGG/Nle** is less stable than **CTB/N2-GGG/I**, a TASP which contained isoleucine residues that are energetically less favoured to take part in an α -helix: Thus, the size/shape of the isoleucine side chain in the core of this three-helix bundle TASP more than compensates for its lower intrinsic helix propensity. **CTB/N2-GGG/V** was less stable than **CTB/N2-GGG/L** by about $2.5 \text{ kcal mol}^{-1}$, probably due to its smaller size, this equates to about $0.8 \text{ kcal mol}^{-1}$ per methylene group removed from the core of these CTB TASP.

Table 6.3: Comparison of the global conformational stability of bowl and CTB TASPs with different central amino acids. NOTE: These values represent the average of four experiments at different TASP concentrations (for clarity errors have been omitted).

Central Amino Acid	$\Delta G^\circ \text{H}_2\text{O}$ (kcal mol ⁻¹)			
	Bowl		CTB	
	TASP	per helix	TASP	per helix
L	9.8	2.5	4.9	1.6
I	8.3	2.1	5.1	1.7
V	4.3	1.1	2.4	0.8
Nle	8.0	2.0	4.1	1.4

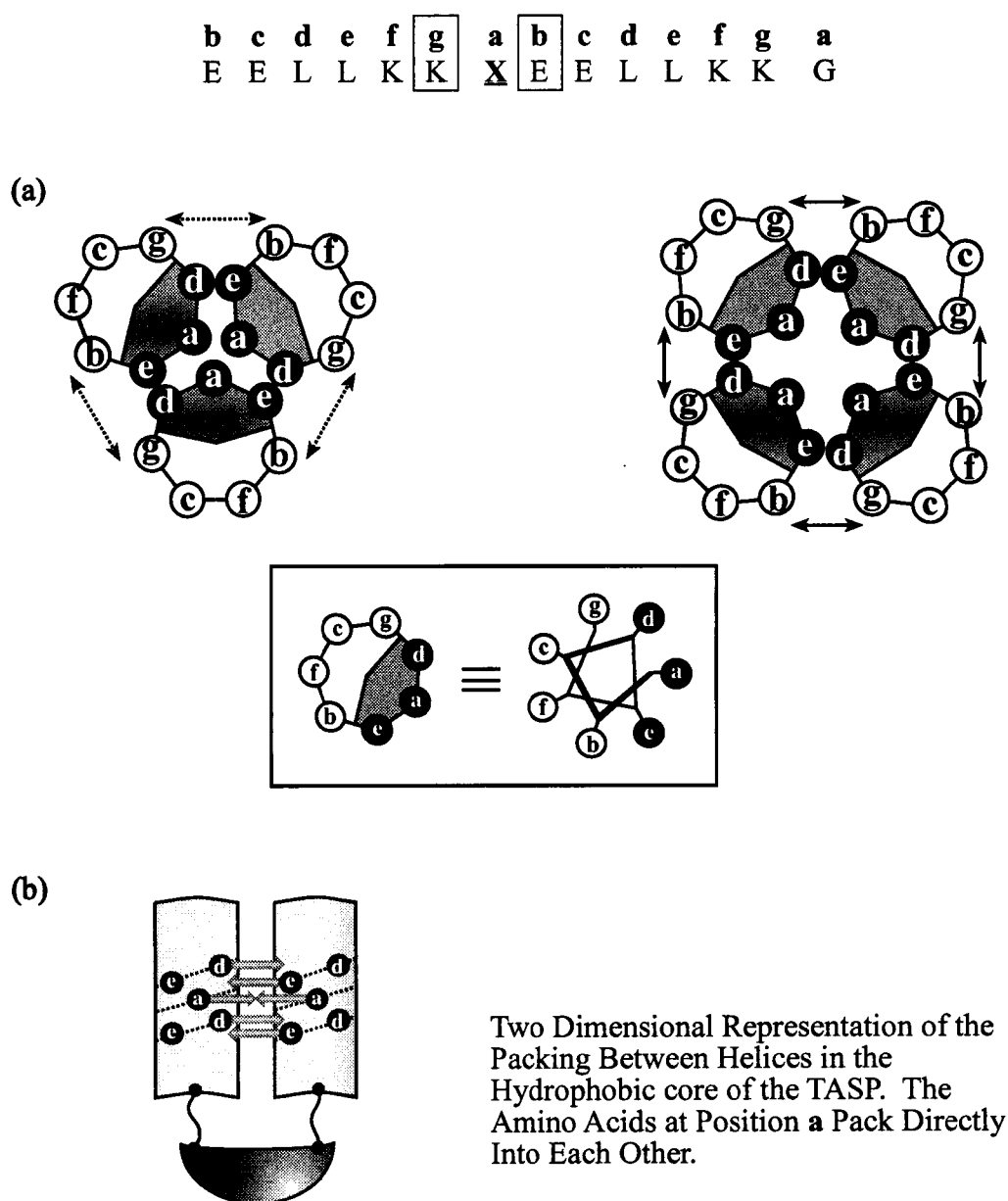
Bowl/N2-GGG/L was more stable than **Bowl/N2-GGG/I** and **Bowl/N2-GGG/Nle** by about 1.5-1.8 kcal mol⁻¹. This may be due to more optimal filling of the core by the leucine side chain in a four-helix bundle, versus isoleucine (which also has a lower helix propensity) and norleucine, within the context of leucine-lined hydrophobic core (Figure 6.9). Again, these TASPs lacked conformational specificity, so this is only a crude assessment of their structure. **Bowl/N2-GGG/V** was less stable than **Bowl/N2-GGG/L** by about 5.5 kcal mol⁻¹ (about 1.4 kcal mol⁻¹ per methylene group removed), and also it was less stable than **Bowl/N2-GGG/I** and **Bowl/N2-GGG/Nle** by 4.0-3.7 kcal mol⁻¹ (about 1.0-0.9 kcal mol⁻¹ per methylene group removed from the core, a similar value to that seen for the analogous CTB TASPs).

The structures of **Bowl/N2-GGG/V** and **Bowl/N2-GGG/I** were affected by the addition of salt, which suggests that the hydrophobic forces were not totally dominant in these TASP, and that there was a stabilizing electrostatic interaction present: Inter-helical (E to K) salt bridges on the bowl TASP are one source of this favourable electrostatic interaction, which will become more significant with a smaller volume filling the core (*i.e* isoleucine or valine, Figure 6.9). Salt had no effect on the analogous CTB TASP, suggesting that any inter-helical salt bridges are not present, or that they are not highly significant to the three-helix bundle structure of these TASP.

The difference in the effect of the individual amino acid side chain in a four-, compared to a three-helix bundle is shown in Table 6.3. Two general assumptions can be made from this data; that leucine seems to be more greatly favoured in the centre of the leucine-lined four-helix bundle, and that valine is strongly disfavoured in the four- versus the three-helix bundle (with leucine-lined cores). The reasons for these observations are probably the size and shape of the side chains, and the role they play in the bundle (see Figure 6.9 for an explanation).

Our failure to increase the conformational specificity of our TASP in this chapter may be due to a lack in the variety of, and complementary packing of the core amino acids. Future structures that are aimed at improving the conformational specificity of these three- and four-helix bundle TASP will be outlined in the next chapter.

Figure 6.9: Simplified diagram showing the proposed packing, and interactions in the three- and four-helix bundle TASPs (positions **a**, **d**, and **e**, of the helix, shown in a heptad repeat representation, are the hydrophobic core residues): (a) The potential inter-helix salt bridges between position **b** and **g** (in boxes) of adjacent helices in a four-helix bundle, these interactions may not occur in a three-helix bundle because **b** and **g** are farther apart. (b) In a TASP with parallel helices (shown as “flat” cylinders), the **a** positions will be at the same “level”, whereas positions **d** and **e** will occur at different pitch of the α -helix.



6.5 Experimental

TASPs were synthesized as described in Chapter 3. The CD (see Table 6.4) and NMR experiments were performed as outlined in Chapter 4.

Table 6.4: Mean residual ellipticity at 222 nm, $[\theta]_{222}$, of CTB and bowl TASPs presented in Chapter 6.

TASP	Experimental $-\theta_{222}$	Theoretical $-\theta_{222}$	Percent Helix
CTB/N2-GGG/L	19300	31100	62
CTB/N2-GGG/I	22300	31100	72
CTB/N2-GGG/V	16500	31100	53
CTB/N2-GGG/Nle	16600	31100	53
Bowl/N2-GGG/L	25300	31100	81
Bowl/N2-GGG/I	19900	31100	64
Bowl/N2-GGG/V	17300	31100	56
Bowl/N2-GGG/Nle	23400	31100	75

* Theoretical helicity is calculated according to reference 394. This calculation includes all residues in the attached peptide, and not just those designed to take part in the “helix”.

Summary, Conclusions, and Future Work

7.0 Introduction

This chapter will review the progress towards the goals of this thesis that were outlined at the end of Chapter 1: Namely, to find an efficient method for the synthesis of CTB and bowl TASP (Section 7.1), produce TASP that were monomeric in solution (Section 7.2), investigate their structure and stability (Section 7.3), and to produce TASP with conformational specificity (Section 7.4). Each section will review the major findings of each investigation, and propose future studies that could answer some of the questions raised by this research. The ultimate aim of the work presented here is to add to the understanding of protein folding and structure.

7.1 Synthesis of CTB and Bowl TASP

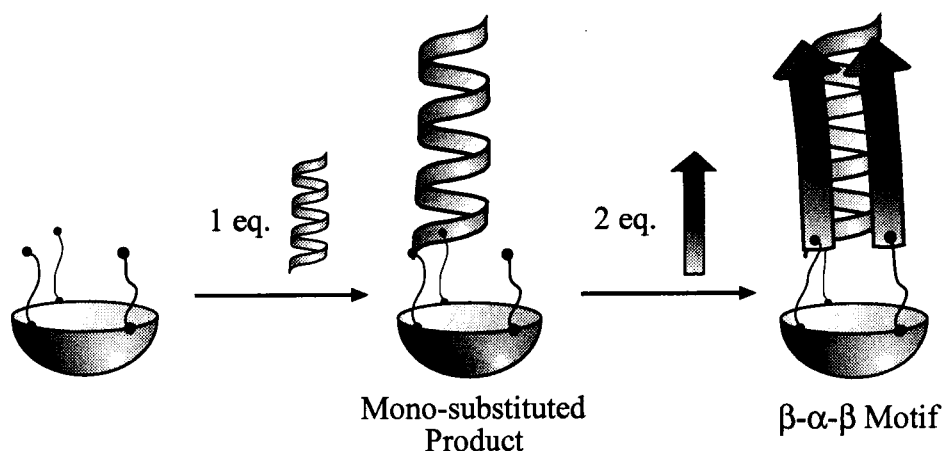
The first goal of this thesis was to find an efficient method of linking the peptides to the templates. The most effective method for achieving this was using a disulfide bond between the template and peptide (Chapter 3). This method employed the activated thiol

moiety of a cysteine residue in the peptide sequence, which was coupled to the CTB or Bowl template (both synthesized in Chapter 2) that contained three- or four- benzyl thiol moieties. The additional advantages of this method were that it allowed attachment of the peptide at almost any position in its sequence, and that the activated peptide starting material could be recovered easily.

Sequential addition of peptides that are linked via disulfide bonds is one approach that could be employed in the production of TASPs with non-identical peptide strands. Important targets that require different peptide strands on the template include anti-parallel three- and four-helix bundles, as well as another common super-secondary structure, the β - α - β motif (see Figure 7.1 for suggested synthesis).

The study of a β - α - β structure has been shown to be a useful tool in the investigation of β -sheets.⁴¹⁴⁻⁴¹⁶ β -sheets are a less well understood secondary structure when compared to α -helices, primarily because isolated β -strands are more difficult to study, therefore a β - α - β TASP is an attractive target. A hydrophobic face on the α -helix of a β - α - β TASP, could act as a scaffold for the β -sheets, which would also contain a hydrophobic and hydrophilic face. The peptide sequences for this motif could be based on *de novo* designed,^{108,121-124,417} or natural protein β - α - β structure.⁴¹⁸⁻⁴²¹ A TASP that contains β -strand peptides only, may not result in a stable β -sheet, due to a lack of conformational rigidity in the peptide backbone.

Figure 7.1: Suggested method for the synthesis of TASPs containing different peptide sequences. For example, the β - α - β structure outlined here would be synthesized on a CTB template by first linking an α -helical peptide (via its C-terminus), and subsequent addition of two β -strand peptides (via their N-termini).



7.2 Oligomeric State of TASPs

After synthesis of the three- and four-helix bundle TASPs, the second goal of this thesis was to produce TASPs that were monomeric in solution. The TASPs made with no glycine spacers between the template-bound cysteine residue and the “helix” (for example, peptide CEELLKKLEELLKKG, Chapter 3) resulted in TASPs that all reversibly self-associated in solution. The macrocyclic template was initially suspected as the cause of this self-association, but the use of a peptide with a charged, “pre-helix” amino acid sequence which resided in the vicinity of the template did nothing to alleviate this aggregation, in

some cases making it more noticeable (Chapter 4). This led to the conclusion that the source of the TASP self-association appeared to come from helices, possible due to the inflexibility of the template-to-helix linker: Incorrect intra-molecular helix bundling led to inter-TASP self-association. Thus, flexible glycine residues were added between the template-bound cysteine residue and “helix” (for example, peptide CGGEELLKKLEELLKKG, Chapter 5), a three glycine spacer was found to be optimal for intra-helix bundling. The four-helix bundle bowl TASP appeared to be a monomer in solution with this three glycine spacer, but the analogous three-helix bundle CTB TASP still exhibited some degree of self-association. The three glycine “spacer” is optimal for the specific sequence of hydrophobic and hydrophilic residues in the “helix”; if the “helix” sequence were “scrambled”, then the “spacer” may have to be redesigned.

The self-association of the three-helix bundle CTB TASP is almost certainly due to the “helix” sequence, a design that is known to form a four-helix bundle structure in solution. Thus, when this “helix” was added to the CTB template, it was “forced” into a three-helix bundle topology, which probably results in an over-packed hydrophobic core that is exposed to the solvent, and thus results in inter-TASP association. A peptide sequence with the same pattern of hydrophobic and hydrophilic amino acids, and an overall charge (for example, Ac-CGGGEQLLKKLEQLLKKG-NH₂, see Figure 7.2a) may discourage self-association of the CTB TASP via charge-charge repulsion (at pH 7.0). The resulting three-helix bundle would have an over-packed core, and probably lack conformational specificity, but it would be interesting to learn if we could control the self-association of the resultant CTB TASP through the use of charge-charge repulsion. The use of a “helix” with an overall

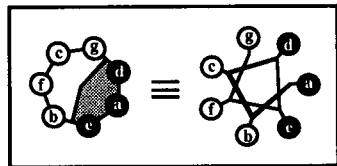
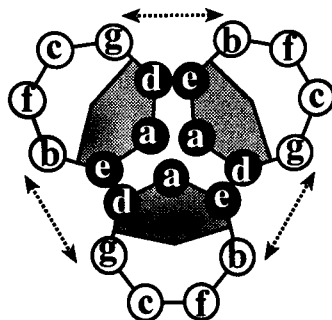
charge could also be used to further discourage self-association of the four-helix bundle bowl TASP. Currently, all our TASP have an overall charge of zero in the "helix".⁴²²

Conquering self-association of the CTB three-helix bundle TASP may be a little more complicated than giving it an overall charge. The "helix" (i.e. -EELLKKLEELLKKG) used for making our CTB and bowl TASP was originally designed to form four-helix bundles, which may have led to the higher degree of self-association seen for the CTB TASP versus the bowl TASP. Therefore, we could design a peptide that is better suited to the formation of three-helix bundles. The hydrophobic and hydrophilic amino acids in this sequence (Ac-CGGGIAAIESKIAAIESKG-NH₂, for example), would be patterned in a similar manner to the many three-helix bundle coiled-coils reported in the literature (Figure 7.2b).^{135,167-171} Furthermore, it may be interesting to measure the stability of this sequence when it is on a bowl template, and "forced" into a four-helix bundle. This work may allow us to discern some of the important design features required for a helix when it is required to take part in a three- or a four-helix bundle: Design features that have not been investigated because there are few systems that can be used to embark on such a study.

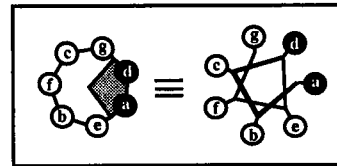
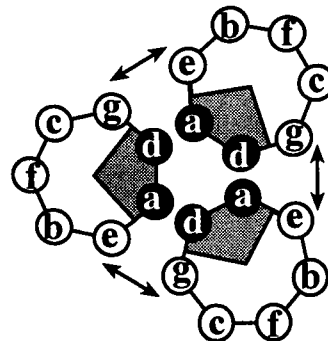
Figure 7.2: (a) A heptad repeat representation showing a “helix”(-EQLLKKLEQLLKKG) with an overall charge, which may discourage TASP self-association by charge-charge repulsion. (b) A heptad repeat representation of a “helix” (-IAAIESKIAAIESKG) that is more suited to forming a three-helix bundle; there is also the possibility of favourable (E to K) inter-helix electrostatic interactions (this sequence is currently “charge-neutral”, the serine residues may also be “mutated” to lysine to give the helix an overall charge – see sequence 8.2a).

	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	
(a)		C	G	G	G	E	Q	L	L	K	K		L	E	Q	L	L	K	K	G
(b)	C	G	G	G	I	A	A	I	E	S	K		I	A	A	I	E	S	K	G

(a)



(b)



7.3 Structure and Stability of TASPs

All TASPs reported in this thesis had helical structure, and TASP self-association did not appear to affect this helical (secondary) structure (except for **Bowl/N2/AAAAA**, which lacked a hydrophobic core). CTB TASPs tended to have a lower helicity than the analogous bowl TASPs: The design of the helix to take part in a four-helix bundle may result in the CTB TASP trading helicity (secondary structure) in order to maximize the packing of its hydrophobic core (tertiary structure).

The TASPs demonstrated enhanced stability when compared to the peptides. This is primarily due to template-enhanced hydrophobic bundling of the amphiphilic helices. There seems to be a relationship between the amount of buried hydrophobic surface and increased stability. Evidence for this is seen by comparing bowl and CTB TASPs (bowl TASPs have a greater degree of buried hydrophobic surface); also in the lack of stability of **Bowl/N2/AAAAA** (which lacked a hydrophobic core); and finally in the lower stability of the valine-containing TASPs, when comparing them to TASPs containing the structural isomers of leucine. There were some noticeable size/shape effects between the leucine structural isomers on the stability of the CTB and bowl TASPs: Isoleucine seemed to be favoured as the central core amino acid in the three-helix bundles, and leucine seemed to be strongly favoured in the four-helix bundles. The stability of the CTB and bowl TASPs was also affected by the pattern of hydrophilic residues (*i.e.* **N1/L** peptide versus **N2/L** peptide), although this difference in stability was not as significant as for the changes made in the hydrophobic core (and may be due to a difference in the association state of these TASPs).

In terms of the “spacer” between the template-bound cysteine residue, and the template, it appears that having the cysteine directly adjacent to the “helix” (for example, peptide CEELLKKLEELLKKG) leads to TASPs with the greatest stability. However, this increase in stability may be a function of increased TASP self-association in solution. Two and three glycine “spacers” between the template-bound cysteine residue, resulted in TASPs that were less stable to GnHCl-induced denaturation. This loss in stability may be due to an increase in the flexibility of the linker, which reduces the directing effect of the template. However, the bowl TASP with a two glycine “spacer” was less stable than that with three glycines, indicating that there are other structural considerations to take into account when designing a template-to-helix linker. Another explanation for the increased stability of the TASPs with no glycine “spacers”, is that the first leucine side chains of the helix can be stabilized by the hydrophobic templates, thus nucleating (and therefore stabilizing) helix formation.

When linking the helix to the template via its N- versus its C-terminus, it appeared that the bowl template acted as a good N-cap for the helices. The CTB TASP was marginally more stable when the peptides were attached via their N-termini. Therefore, to compare the stability of the three- and four-helix bundle TASPs, the different N-capping effects of the bowl and CTB templates must be taken into account. This is also true when comparing peptides attached to the templates via their N- versus C-termini.

7.4 Conformational Specificity

After a suitable template-to-helix linker had been established, the final, and perhaps the major goal of this thesis was to work towards TASP s that exhibit “native-like” structure. TASP s containing the helices with all leucine core residues appeared to be molten globules by NMR. This was not surprising, there are several examples of *de novo* designed proteins with all-leucine cores that are molten globules. However, the conformational specificity of these disulfide-linked TASP s differed from earlier work done by our group on four-helix bundle TASP s that had the peptide strands attached via their backbone N-termini (for example, comparing **Bowl/N2-GGG/L** and **ArBowl/N2-GGG/L**, see Chapter 5). When studied by NMR, the TASP s with the “backbone-linked” peptides appeared to have either a “native-like” structure, or a dynamic, rapidly inter-converting conformation. It may prove difficult to discern the correct conformation of TASP s with degenerate amino acid sequences, but one method may involve the use of N-H exchange rates, a method that has been used to determine the conformational mobility of symmetrical coiled-coil proteins.⁴²³⁻⁴²⁹

Mutating the central amino acid of the “helix” of the CTB and bowl TASP s to either isoleucine, norleucine, or valine did not alter the conformational specificity of the three- and four-helix bundle TASP s. Achieving conformational specificity may be as simple as changing one amino acid in the core of the TASP, but a better approach may be to design complementary packing interactions in the core of the TASP.^{233-236,354,430-432} Dutton, for example, has investigated the effects of changing multiple amino acids on a four-helix bundle.³⁵³ Results from this work could serve as a starting point for our own studies.

Reducing the number of identical side chains in the core of the protein may also allow them to be characterized using multi-dimensional NMR techniques. Thus, the major challenge for this work (and for other researchers) is to investigate the features that are required to form a “native-like” structure: Our CTB and bowl TASP s provide an excellent method to investigate these features.

References

1. Stryer, L. *Biochemistry*, 3rd ed.; W.H. Freeman and Co., New York, **1988**.
2. Sanger, F. *Adv. Protein Chem.* **1952**, *7*, 1-67.
3. Anfinsen, C.B. *Science* **1973**, *181*, 223-230.
4. Creighton, T.E. *Protein Folding*; W.H. Freeman and Co., New York, **1992**.
5. Creighton, T.E. *Proteins: Structures and Molecular Properties*, 2nd ed.; W.H. Freeman and Co., New York, **1993**.
6. Richards, F.M. *Scientific American* **1991**, January, 54-63.
7. Chan, H.S.; Dill, K.A. *Physics Today* **1993**, February, 24-32.
8. Dill, K.A. *Biochemistry* **1990**, *29*, 7133-7155.
9. Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland, New York, **1991**.
10. Collins, F.S.; Jegalian, K.G. *Scientific American* **1999**, December, 86-91.
11. Benner, S.A.; Gerloff, D.L.; Jenny, T.F. *Science* **1994**, *265*, 1642-1644.
12. Bajorath, J.; Aruffo, A. *Bioconjugate Chem.* **1994**, *5*, 173-181
13. Sander et. al. *Proteins SFG* **1992**, *12*, 105-110.
14. Dahiyat, B.I.; Mayo, S.L. *Science* **1997**, *278*, 82-87.
15. Frishman, D.; Argros, P. *Proteins SFG* **1997**, *27*, 329-335.
16. Banavar, J.R.; Cieplak, M.; Maritan, A.; Nadig, G.; Seno, F.; Vishveshwara, S. *Proteins SFG* **1998**, *31*, 10-20.
17. Srinivasan, R.; Rose, G.D. *Proteins SFG* **1995**, *22*, 81-99.
18. Butler, D. *Nature* **1999**, *402*, 705-706.
19. Wilson, E.K. *C&EN* **2000**, September 25, 41-44.
20. Pace, C.N. *Trends. Biochem. Sci.* **1990**, *15*, 14-17.

21. Lesk, A.M.; Chothia, C. *J. Mol. Biol.* **1980**, *136*, 225-270.
22. Doolittle, R.F. *Science* **1980**, *214*, 149-159.
23. DeGrado, W.F.; Wasserman, Z.R.; Lear, J.D. *Science* **1989**, *243*, 622-628.
24. Finkelstein, A.V.; Ptitsyn, O.B. *Prog. Biophys. Molec. Biol.* **1987**, *50*, 171-190.
25. MacArthur, M.W.; Thornton, J.M. *J. Mol. Biol.* **1996**, *264*, 1180-1195.
26. Pauling, L.; Corey, R.B.; Branson, H.R. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205-210.
27. For a review see: Kumar, S.; Bansal, M. *Biophysical J.* **1998**, *75*, 1935-1944.
28. Barlow, D.J.; Thornton, J.M. *J. Mol. Biol.* **1988**, *201*, 601-619.
29. Wada, A. *Adv. Biophys.* **1976**, *9*, 1-63.
30. Hol, W.G.J.; van Duijnen, P.T.; Berendsen, H.J.C. *Nature* **1978**, *273*, 443-446.
31. Aurora, R.; Creamer, T.P.; Srinivasan, R.; Rose, G.D. *J. Biol. Chem.* **1997**, *272*, 1413-1416.
32. Kamtekar, S.; Hecht, M.H. *FASEB J.* **1995**, *9*, 1013-1022.
33. Calvert, R.; Kahana, E.; Gratzer, W.B. *Biophysical J.* **1996**, *71*, 1605-1610.
34. Johansson, M.U. et. al. *FEBS Lett.* **1995**, *374*, 257-261.
35. Fairman, R.; Beran-Steed, R.K.; Handel, T.M. *Protein Sci.* **1997**, *6*, 175-184.
36. Olszewski, K.A.; Kolinski, A.; Skolnick, J. *Proteins SFG* **1996**, *25*, 286-299.
37. Kraulis, P.J. et. al. *FEBS Lett.* **1996**, *378*, 190-194.
38. Pascual, J.; Pfuhl, M.; Rivas, G.; Pastore, A.; Saraste, M. *FEBS Lett.* **1996**, *383*, 201-207.
39. Kline, A.D. et. al. *FEBS Lett.* **1997**, *407*, 239-242.
40. Katz, L. Hanson, P.I.; Heuser, J.E.; Brennwald, P. *EMBO J.* **1998**, *17*, 6200-6209.
41. Babcock, G.T.; Wikstrom, M. *Nature* **1992**, *356*, 301-309.
42. Cohen, C.; Parry, D.A.D. *Trends Biochem. Sci.* **1986**, *11*, 245-248.

43. Kohn, W.D.; Hodges, R.S. *Trends. Biotechnol.* **1998**, *16*, 379-389.
44. For example see: Tsukada, M.; Will, E.; Gallwitz, D. *Mol. Biol. Cell* **1999**, *10*, 63-75.
45. For example see: Lindblom, A. et. al. *J. Biol. Chem.* **1999**, *274*, 6374-6380.
46. For example see: Kjer-Nielsen, L.; Teasdale, R.D.; van Vliet, C.; Gleeson, P.A. *Curr. Biol.* **1999**, *9*, 385-388.
47. For example see: Hakansson, K.; Lim, N.K.; Hoppe, H.-J.; Reid, K.B.M. *Structure* **1998**, *7*, 255-264.
48. For example see: Ginsburg, A. et. al. *Biochemistry* **1997**, *36*, 7876-7883.
49. For example see: Steven, A.C. et. al. *J. Mol. Biol.* **1988**, *200*, 351-365.
50. Crick, F.H.C. *Acta. Crystallog.* **1953**, *6*, 689-698.
51. Muller, N. *Acc. Chem. Res.* **1990**, *23*, 23-28.
52. Kellis, J.T.; Nyberg, K.; Sali, D.; Fersht, A.R. *Nature* **1988**, *333*, 784-786.
53. Budisa, N.; Pifat, G. *Croatica Chimica Acta* **1998**, *71*, 179-187.
54. Privalov, P.L.; Gill, S.J. *Pure and Appl. Chem.* **1989**, *61*, 1097-1104.
55. Lins, L.; Brasseur, R. *FASEB J.* **1995**, *9*, 535-540.
56. Privalov, P.L.; Gill, S.J. *Adv. Prot. Chem.* **1988**, *39*, 191-234.
57. Chothia, C.; Finkelstein, A.V. *Annu. Rev. Biochem.* **1990**, *59*, 1007-1039.
58. Doig, A.J.; Sternberg, M.J.E. *Protein Sci.* **1995**, *4*, 2247-2251.
59. Morii, H.; Uedaira, H.; Ogata, K.; Ishii, S.; Sarai, A. *J. Mol. Biol.* **1999**, *292*, 909-920.
60. Karpusas, M.; Baase, W.A.; Matsumura, M.; Matthews, B.W. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8237-8241.
61. Schwehm, J.M.; Kristyanne, E.S.; Biggers, C.C.; Stites, W.E. *Biochemistry* **1998**, *37*, 6939-6948.
62. Kono, H.; Nishiyama, M.; Tanokura, M.; Doi, J. *Protein Eng.* **1998**, *11*, 47-52.
63. Pace, C.N.; Shirley, B.A.; McNutt, M.; Gajiwala, K. *FASEB J.* **1996**, *10*, 75-83.

64. Koh, J.T.; Cornish, V.W.; Schultz, P.G. *Biochemistry* **1997**, *36*, 11314-11322.
65. Allewell, N.M.; Oberoi, H. *Methods Enzymol.* **1991**, *202*, 3-19.
66. Matsumura, M.; Signor, G.; Matthews, B.W. *Nature* **1989**, *342*, 291-293.
67. Yon, J. M. *Cell. Mol. Life Sci.* **1997**, *53*, 557-567.
68. Baker, D.; Agard, D.A. *Biochemistry* **1994**, *33*, 7505-7509.
69. For an investigation into *flavodoxin* see: van Mierlo, C.P.; Steensma, E. *J. Biotechnology* **2000**, *79*, 281-298.
70. Baldwin, R.L. *Nature* **1994**, *369*, 183-184.
71. Laurents, D.V.; Baldwin, R.L. *Biophysical J.* **1998**, *75*, 428-434.
72. Roder, H. et. al. *Biochemistry* **1993**, *32*, 10271-10276.
73. Neri, D.; Billeter, M.; Wider, G.; Wuthrich, K. *Science* **1992**, *257*, 1559-1563.
74. Akiyama, S. Takahashi, S.; Ishimori, K.; Morishima, I. *Nature Structural Biology* **2000**, *7*, 514-520.
75. Wong, K.-B.; Clarke, J.; Bond, C.J.; Neira, J.L.; Freund, S.M.V.; Fersht, A.R.; Daggett, V.J. *Mol. Biol.* **2000**, *296*, 1257-1282.
76. Fisher, K.F.; Marqusee, S. *J. Mol. Biol.* **2000**, *302*, 701-712.
77. Baum J.; Dobson, C.M.; Evans, P.A.; Hanley, C. *Biochemistry* **1989**, *28*, 7-13.
78. For an investigation into the folding of *bovine pancreatic trypsin inhibitor* see: van Mierlo, C.P.; Kemmink, J.; Neuhaus, D.; Darby, N.J.; Creighton, T.E. *J. Mol. Biol.* **1994**, *235*, 1044-1061.
79. Dill, K.A.; Chan, H.S. *Nature Structural Biology* **1997**, *4*, 10-19.
80. Service, R.F. *Science* **1996**, *373*, 29-30.
81. Baldwin, R.L.; Rose, G.D. *Trends. Biochem. Sci.* **1999**, *24*, 26-33.
82. Baldwin, R.L.; Rose, G.D. *Trends. Biochem. Sci.* **1999**, *24*, 77-83.
83. Roder, H.; Colon, W. *Curr. Opin. Struct. Biol.* **1997**, *7*, 15-28.

84. For an example see: De Filippis, V.; de Laureto, P.P.; Toniutti, N.; Fontana, A. *Biochemistry* **1996**, 35, 11503-11511.
85. For an example see: Nitta, K. et. al. *Biochemistry* **2000**, 39, 3248-3257.
86. For an example see: Demarest, S.J.; Boice, J.A.; Fairman, R.; Raleigh, D.P. *J. Mol. Biol.* **1999**, 294, 213-221.
87. Ruddon, R. W.; Bedows, E. *J. Biol. Chem.* **1997**, 272, 3125-3128.
88. Wittung-Stafshede, P.; Grey, H.B.; Winkler, J.R. *J. Am. Chem. Soc.* **1997**, 119, 9562-9563.
89. Helenius, A. *Mol. Biol. Cell* **1994**, 5, 253-265.
90. Kang, X.; Carey, J. *J. Mol. Biol.* **1999**, 285, 463-468.
91. Jaenicke, R. *Prog. Biophys. Mol. Biol.* **1987**, 49, 117-237.
92. Buchner, J. *FASEB J.* **1996**, 10, 10-19.
93. Martin, J.; Hartl, F.U. *Structure* **1993**, 1, 161-164.
94. Frydman, J.; Hartl, F.U. *Science* **1996**, 272, 1497-1502.
95. Yang, J.; Spek, E.J.; Gong, Y.; Zhou, H.; Kallenbach, N.R. *Protein Sci.* **1997**, 6, 1264-1272.
96. Altmann, K.-H.; Florsheimer, A.; Mutter, M. *Int. J. Peptide Protein Res.* **1986**, 314-319.
97. Yuan, S.-M.; Clark, N.D. *Proteins SFG* **1998**, 30, 136-143.
98. Kaiser, E.T.; Kezdy, F.J. *Science* **1984**, 223, 249-255.
99. Dalal, S.; Regan, L. *Protein Sci.* **2000**, 9, 1651-1659.
100. DeGrado, W.F.; Lear, J.D. *J. Am. Chem. Soc.* **1985**, 107, 7684-7689.
101. Matsumura, M. et. al. *J. Biol. Chem.* **1989**, 264, 16059-16066.
102. Myers, J.K.; Oas, T.G. *J. Mol. Biol.* **1999**, 289, 205-209.
103. For a recent review on forming stable helical peptides using natural and artificial amino acids see: Andrew, M.J.I.; Tabor, A.B. *Tetrahedron* **1999**, 55, 11711-11743.
104. Chou, P.Y.; Fasman, G. *Annu. Rev. Biochem.* **1978**, 47, 251-276.

105. Chou, P.Y.; Fasman, G. *Biochemistry* **1974**, *13*, 211-221.
106. Chou, P.Y.; Fasman, G. *Biochemistry* **1974**, *13*, 222-245.
107. Levitt, M. *Biochemistry* **1978**, *17*, 4277-4285.
108. Chakrabartty, A.; Kortemme, T.; Baldwin, R.L. *Protein Sci.* **1994**, *3*, 843-852.
109. Rohl, C.A.; Chakrabartty, A.; Baldwin, R.L. *Protein Sci* **1996**, *5*, 2623-2637.
110. Lyu, P.C.; Liff, M.I.; Marky, L.A.; Kallenbach, N.R. *Science* **1990**, *250*, 669-673.
111. Gans, P.J.; Lyu, P.C.; Manning, M.C.; Woody, R.W.; Kallenbach, N.R. *Biopolymers* **1991**, *31*, 1605-1614.
112. Blaber, M.; Zhang, X. J.; Mathews, B. W. *Science* **1993**, *260*, 1637-1640.
113. Blaber, M.; Zhang, X.J.; Lindstrom, J.D.; Pepiot, S.D.; Baase, W.; Mathews, B.W. *J. Mol. Biol.* **1994**, *235*, 600-624.
114. Horovitz, A.; Mathews, J.M.; Fersht, A.R. *J. Mol. Biol.* **1992**, *227*, 560-568.
115. O'Neil, K.T.; DeGrado, W.F. *Science* **1990**, *250*, 646-651.
116. Zhou, et. al. *Protein and Peptide Letters* **1994**, *1*, 114-119.
117. Monera, et. al. *J. Peptide Sci.* **1995**, *1*, 319-329.
118. Chakrabartty, A.; Baldwin, R.L. *Adv. Protein Chem.* **1995**, *46*, 141-176.
119. Myers, J.K.; Pace, C.N.; Scholtz, J.M. *Biochemistry* **1997**, *36*, 10923-10929.
120. Reddy, K.S.; Datta, S.; Tiwari, S. *Protein Eng.* **1998**, *11*, 1137-1145.
121. Pace, C.N.; Scholtz, J.M. *Biophysical J.* **1998**, *75*, 422-427.
122. Creamer, T.P.; Rose, G.D. *Proteins SFG* **1994**, *19*, 85-97.
123. Creamer, T.P.; Rose, G.D. *Proc. Natl. Acad. Sci.* **1992**, *89*, 5937-5941.
124. Luo, P.; Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4930-4935.
125. Lau, S.Y.M.; Taneja, A.K.; Hodges, R.S. *J. Biol. Chem.* **1984**, *259*, 13253-13261.
126. Su, J.Y.; Hodges, R.S.; Kay, C.M. *Biochemistry* **1994**, *33*, 15501-15510.

127. Ihara, S.; Ooi, T.; Takahashi, S. *Biopolymers* **1982**, *21*, 131-145.
128. Shoemaker, K.R.; Kim, P.S.; York, E.J.; Stewart, J.M.; Baldwin, R.L. *Nature* **1987**, *326*, 563-567.
129. Huyghues-Despointes, B.M.P.; Scholtz, J.M.; Baldwin, R.L. *Protein Sci.* **1993**, *2*, 1604-1611.
130. Lockhart, D.S.; Kim, P.S. *Science* **1993**, *260*, 198-202.
131. Kohn, et. al. *J. Peptide Science* **1997**, *3*, 209-233.
132. Sholtz, J.M.; Qian, H.; Robbins, V.H.; Baldwin, R.L. *Biochemistry* **1993**, *32*, 9668-9678.
133. Smith, J.S.; Sholtz, J.M. *Biochemistry* **1998**, *37*, 33-40.
134. Padmanabhan, S.; Baldwin, R.L. *J. Mol. Biol.* **1994**, *241*, 706-713.
135. Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D.K.; DeGrado, W.F.; Eisenberg, D. *Science* **1993**, *259*, 1288-1293.
136. Padmanabhan, S.; Baldwin, R.L. *Protein Sci.* **1994**, *3*, 1992-1997.
137. Armstrong, K.M.; Fairman, R.; Baldwin, R.L. *J. Mol. Biol.* **1993**, *230*, 284-291.
138. For a review on helix capping see: Aurora, R.; Rose, G.D. *Protein Sci.* **1998**, *7*, 21-38.
139. Presta, L.G.; Rose, G.D. *Science* **1988**, *240*, 1632-1641.
140. Richardson, J.S.; Richardson, D.C. *Science* **1988**, *240*, 1648-1652.
141. Parker, M.H.; Hefford, M.A. *Protein Eng.* **1997**, *10*, 487-496.
142. Lyu, P.C.; Wemmer, D.E. *Biochemistry* **1993**, *32*, 421-425.
143. Joshi, H.V.; Meier, M.S. *J. Am. Chem. Soc.* **1996**, *118*, 12038-12044.
144. Bruch, M.D.; Dhingra, M.M.; Gierasch, L.M. *Proteins SFG* **1991**, *10*, 130-139.
145. Lyu, P.C.; Wemmer, D.E.; Zhou, H.Z.; Pinker, R.J.; Kallenbach, N. R. *Biochemistry* **1993**, *32*, 421-425.
146. Forood, B.; Feliciano, E.J.; Nambiar, K.P. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 838-842.

147. Forood, B.; Reddy, H.K.; Nambiar, K.P. *J. Am. Chem. Soc.* **1994**, *116*, 6935-6936.
148. Chakrabartty, A.; Doig, A.J.; Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11332-11336.
149. Doig, A.J.; Baldwin, R.L. *Protein Sci.* **1995**, *4*, 1325-1336.
150. Petukhov, M.; Yumoto, N.; Murase, S.; Onmura, R.; Yoshikawa, S. *Biochemistry* **1996**, *35*, 387-397.
151. Serrano, L.; Fersht, A.R. *Nature* **1989**, *342*, 296-299.
152. Serrano, L.; Neira, J.L.; Sancho, J.; Fersht, A.R. *Nature* **1992**, *356*, 453-455.
153. Serrano, L.; Sancho, J.; Hirshberg, M.; Fersht, A.R. *J. Mol. Biol.* **1992**, *227*, 544-599.
154. Bell, J.A.; Becktel, W.J.; Sauer, U.; Baase, W.A.; Mathews, B.W. *Biochemistry* **1992**, *31*, 3590-3596.
155. Aurora, R.; Srinivasan, R.; Rose, G.D. *Science* **1994**, *264*, 1126-1130.
156. Walther, D.; Eisenhaber, F.; Argos, P. *J. Mol. Biol.* **1996**, *255*, 536-553.
157. Chothia, C.; Levitt, M.; Richardson, D. *J. Mol. Biol.* **1981**, *145*, 215-250.
158. Chothia, C. *Ann. Rev. Biochem.* **1984**, *53*, 537-572.
159. Taylor, K.S.; Lou, M.-Z.; Chin, T.-M.; Yang, N.C.; Gravito, R.M. *Protein Sci.* **1996**, *5*, 414-421.
160. Bowie, J. U. *Nature Structural Biology* **1997**, *4*, 915-919.
161. Walther, D.; Springer, C.; Cohen, F.E. *Proteins SFG* **1998**, *33*, 457-459.
162. Preissner, R.; Goede, A.; Frommel, C. *Protein Eng.* **1999**, *12*, 825-831.
163. Bowie, J.U. *Protein Science* **1999**, *8*, 2711-2719.
164. Bowie, J.U. *Nature Structural Biology* **2000**, *7*, 91-94.
165. For a "mini-review" see: Efimov, A.V. *FEBS Lett.* **1999**, *463*, 3-6.
166. Conway, J.F.; Parry, D.A.D. *Int. J. Biol. Macromol.*, **1990**, *12*, 328-335.
167. Sorger, P.K.; Nelson, H.C. *Cell*, **1989**, *59*, 807-813.

168. Peteranderl, R.; Nelson, H.C. *Biochemistry*, **1992**, *31*, 12272-12276.
169. Harbury, P.B.; Kim, P.S.; Alber, T. *Nature* **1994**, *371*, 80-83.
170. Ogihara, N.L.; Weiss, M.S.; DeGrado, W.F.; Eisenberg, D. *Protein Sci.* **1997**, *6*, 80-88.
171. Vazquez, M.-I.; Rivas, G.; Cregut, D.; Serrano, L.; Esteban, M. *J. Virology* **1998**, *72*, 10126-10137.
172. Betz, S.F.; DeGrado, W.F. *Biochemistry* **1996**, *35*, 6955-6962.
173. Johansson, J.S.; Gibney, B.R.; Rabanal, F.; Reddy, K. S.; Dutton, P.L. *Biochemistry* **1998**, *37*, 1421-1429.
174. Gibney, B.R.; Rabanal, F.; Reddy, K.S.; Dutton, P.L. *Biochemistry* **1998**, *37*, 4635-4643.
175. Sharp, R.E.; Diers, J.R.; Bocian, D.F.; Dutton, P.L. *J. Am. Chem. Soc.* **1998**, *120*, 7103-7104.
176. Robertson, D.E.; Farid, R.S.; Moser, C.C.; Mullholland, S.E.; Pidikiti, R.; Lear, J.D.; Wand, A.J.; DeGrado, W.F.; Dutton, P.L. *Nature* **1994**, *368*, 425-432.
177. Schnappinger, D.; Schubert, P.; Pfeiderer, K.; Hillen, W. *EMBO J.* **1998**, *17*, 535-543.
178. Sutton, R.B.; Fasshauer, D.; Jahn, R.; Brunger, A.T. *Nature* **1998**, *395*, 347-353.
179. Nooren, I.M.A.; Kaptein, R.; Sauer, R.T.; Boelens, R. *Nature Structural Biology* **1999**, *6*, 755-759.
180. Malashkevich, V.N.; Kammerer, R.A.; Efimov, V.P.; Schulthess, T.; Engel, J. *Science* **1996**, *274*, 761-765.
181. Broo, K.S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 11362-11372.
182. Broo, K.S.; Nilsson, H.; Nilsson, J.; Flodberg, A.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 4063-4068.
183. Broo, K.S.; Nilsson, H.; Nilsson, J.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 10287-10295.
184. Harbury, P.B.; Zhang, T.; Kim, P.S.; Alber, T. *Science* **1993**, *262*, 1401-1407.

185. Johansson, J.S.; Gibney, B.R.; Skalicky, J.J.; Wand, A.J.; Dutton, P.L. *J. Am. Chem. Soc.* **1998**, *120*, 3881-3886
186. O'Shea, E.K.; Rutkowski, R.; Kim, P.S. *Science* **1989**, *243*, 538-542.
187. Quinn, T.P.; Tweedy, N.B.; Williams, R.W.; Richardson, J.S.; Richardson, D.C. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8747-8751.
188. Kohn, W.D.; Kay, C.M.; Hodges, R.S. *J. Peptide Sci.* **1997**, *3*, 209-211.
189. Durr, E.; Jelesarov, I.; Bosshard, H.R. *Biochemistry* **1999**, *38*, 870-880.
190. Wagschal, K.; Triplet, B.; Hodges, R.S. *J. Mol. Biol.* **1999**, *285*, 785-803.
191. Moitra, J.; Szilak, L.; Krylov, D.; Vinson, C. *Biochemistry* **1997**, *36*, 12567-12573.
192. Yan, Y.; Erickson, B. W. *Protein Sci.* **1994**, *3*, 1069-1073.
193. Lu, M.; Shu, W.; Ji, H.; Spek, E.; Wang, L.; Kallenbach, N.R. *J. Mol. Biol.* **1999**, *288*, 743-752.
194. Blondelle, S. E.; Forood, B.; Houghten, R. A.; Perez-Paya, E. *Biochemistry* **1997**, *36*, 8393-8400.
195. Schenck, H.L.; Gellman, S.H. *J. Am. Chem. Soc.* **1998**, *120*, 4869-4870.
196. Hodges, R.S. et. al. *Biopolymers* **1998**, *47*, 101-123.
197. Mondera, O.D.; Sereda, T.J.; Zhou, N.E.; Kay, C.M.; Hodges, R.S. *J. Peptide Sci.* **1995**, *1*, 319-329.
198. Zhou, N.E.; Kay, C.M.; Hodges, R.S. *Biochemistry* **1993**, *32*, 3178-3187.
199. Monera, O.D.; Kay, C.M.; Hodges, R.S. *Protein Sci.* **1994**, *3*, 1984-1991..
200. Sharman, G.J.; Searle, M.S. *J. Am. Chem. Soc.* **1998**, *120*, 5291-5300.
201. Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253-256.
202. Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, *30*, 153-161.
203. Lumb, K.J.; Kim, P.S. *Biochemistry* **1995**, *34*, 8642-8648.
204. Bosshard, H.R. et. al. *Biochemistry* **1997**, *36*, 204-213.
205. Lombardi, A.; Bryson, J.W.; DeGrado, W.F. *Biopolymers* **1997**, *40*, 495-504.

206. Wand, A.J. et. al. *J. Am. Chem. Soc.* **1999**, *121*, 4941-4951.
207. For a review on *de novo* peptide and protein design see: DeGrado, W.F. *Adv. Protein Chem.* **1988**, *39*, 51-124.
208. For a review on *de novo* protein design see: DeGrado, W.F. et. al. *Science* **1995**, *270*, 935-941.
209. For a review on *de novo* protein design see: DeGrado, W.F. *Annu. Rev. Biochem.* **1999**, *68*, 779-819.
210. Hill, R.B.; Raleigh, D.P.; Lombardi, A.; DeGrado, W.F. *Acc. Chem. Res.* **2000**, *33*, 745-754.
211. Cohen, C.; Parry, D.A.D. *Proteins SFG* **1990**, *7*, 1-15.
212. Munson, M.; O'Brien, R.; Sturtvant, J.M.; Regan, L. *Protein Sci.* **1994**, *3*, 2015-2022.
213. Osterman, D.G.; Kaiser, E.T. *J. Cell. Biol.* **1985**, *29*, 1-16.
214. Osterman, D.G.; Mora, R.; Kezdy, F.J.; Kaiser, E.T.; Meredith, S.C. *J. Am. Chem. Soc.* **1984**, *106*, 6845-6847.
215. Struthers, M.D.; Cheng, R.P.; Imperiali, B. *Science* **1996**, *271*, 342-345.
216. Struthers, M.D. Cheng, R.P.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3073-3081.
217. Dahiyat, B.I.; Sarisky, C.A.; Mayo, S.L. *J. Mol. Biol.* **1997**, *273*, 789-796.
218. Malakauskas, S.M.; Mayo, S.L. *Nature Structural Biology* **1998**, *5*, 470-475.
219. Kohn, W.D.; Kay, C.M.; Sykes, B.D.; Hodges, R.S. *J. Am. Chem. Soc.* **1998**, *120*, 1124-1132.
220. Fezoui, Y.; Weaver, D.L; Osterhout, J.J. *Protein Sci.* **1995**, *4*, 286-295.
221. Fezoui, Y.; Braswell, E.H.; Xian, W.; Osterhout, J.J. *Biochemistry* **1999**, *38*, 2796-2804.
222. Hecht, M.H.; Richardson, J.S.; Richardson, D.C.; Ogden, R.C. *Science* **1990**, *249*, 884-891.
223. Schneider, J.P.; Lombardi, A.; DeGrado, W.F. *Folding & Design* **1998**, *3*, 29-40.
224. Ho, S.P.; DeGrado, W.F. *J. Am. Chem. Soc.* **1987**, *109*, 6751-6758.

225. Regan, L.; DeGrado, W.F. *Science*, **1988**, *241*, 976-978.
226. DeGrado, W.F. et. al. *J. Am. Chem. Soc.* **1992**, *114*, 331-337.
227. Handel, T.M.; Williams, S.A.; DeGrado, W.F. *Science* **1993**, *261*, 879-885.
228. Raleigh, D.P.; DeGrado, W.F. *J. Am. Chem. Soc.* **1992**, *114*, 10079-10081.
229. Raleigh, D.P.; Betz, S.F.; DeGrado, W.F. *J. Am. Chem. Soc.* **1995**, *117*, 7558-7559.
230. Handel, T.; DeGrado, W.F. *J. Am. Chem. Soc.* **1990**, *112*, 6710-6711.
231. Bryson, J.W.; Desjarlais, J.R.; Handel, T.M.; DeGrado, W.F. *Protein Sci.* **1998**, *7*, 1404-1414.
232. Walsh, S.T.R.; Cheng, H.; Bryson, J.W.; Roder, H.; DeGrado, W.F. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5486-5491.
233. Brive, L.; Dolphin, G.T.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 8598-8607.
234. Dolphin, G. T.; Brive, L.; Johansson, G.; Baltzer, L. *J. Am. Chem. Soc.* **1996**, *118*, 11297-11298.
235. Gibney, B.R.; Rabanal, F.; Skalicky, J.J.; Wand, A.J.; Dutton, P.L. *J. Am. Chem. Soc.* **1997**, *119*, 2323-2324.
236. Jiang, X.; Bishop, E.J.; Farid, R.S. *J. Am. Chem. Soc.* **1997**, *119*, 838-839.
237. Kamtekar, S.; Schiffer, J.M.; Xiong, H.; Babik, J.M.; Hecht, M. H. *Science* **1993**, *262*, 1680-1685.
238. Roy, S.; Ratnaswamy, G.; Boice, J.A.; Fairman, R.; McLendon, G.; Hecht, M.H. *J. Am. Chem. Soc.* **1997**, *119*, 5302-5306.
239. Efimov, A.V. *Protein Eng.* **1991**, *4*, 245-250.
240. Nagi, A.D.; Anderson, K.S.; Regan, L. *J. Mol. Biol.* **1999**, *286*, 257-265.
241. Efimov, A.V. *Curr. Opin. Struct. Biol.* **1993**, *3*, 379-384.
242. Mutter, M.; Vuillemeir, S. *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 535-554.
243. Vuilleumier, S.; Mutter, M. *Biopolymers* **1993**, *33*, 389-400.

244. Suich, D.J.; Mousa, S.A.; Singh, G.; Liapakis, G.; Reisine, T. DeGrado, W.F. *Bioorganic and Medicinal Chemistry* **2000**, *8*, 2229-2241.
245. Kemp, D.S.; Curran, T.P.; Davis, W.M.; Boyd, J.G.; Muendel, C. *J. Org. Chem.* **1991**, *56*, 6672-6682.
246. Kemp, D.S.; Curran, T.P.; Boyd, J.G.; Allen T.J. *J. Org. Chem.* **1991**, *56*, 6683-6697.
247. Kemp, D.S.; Boyd, J.G.; Muendel, C. *Nature* **1991**, *352*, 451-454.
248. Diaz, H.; Tsang, K.Y.; Kelly, J.W. *Tetrahedron* **1993**, *49*, 3533-3540.
249. Mutter, M. et. al. *Helv. Chim. Acta.* **1988**, *71*, 835-847.
250. Kemp, D.S. *Trend. Biotechnol.* **1990**, *8*, 249-255.
251. Reddy, B.V.B.; Blundell, T.L. *J. Mol. Biol.* **1993**, *233*, 464-479.
252. Mutter, M. et. al. *J. Am. Chem. Soc.* **1992**, *114*, 1463-1470.
253. Mutter, M. et. al. *Protein Sci.* **1992**, *1*, 1377-1386.
254. Mutter, M. et. al. *Angew. Chem. Intl. Ed. Engl.* **1996**, *35*, 1482-1486.
255. Mutter, M. et. al. *Helv. Chim. Acta.* **1993**, *76*, 1539-1563.
256. Katz, E.; Heleg-Shabatai, V.; Willner, I.; Rau, H.R.; Haehnel, W. *Angew. Chem. Int. Ed.* **1998**, *37*, 3253-3256.
257. Rau, H.K.; DeJonge, N.; Haehnel, W. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11526-11532.
258. Rau, H.K.; Haehnel, W. *J. Am. Chem. Soc.* **1998**, *120*, 468-476.
259. Montal, M.; Montal, M.S.; Tomich, J.M. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6929-6933.
260. Futaki, S. *Biopolymers* **1998**, *47*, 75-81.
261. Sasaki, T.; Kaiser, E.T. *J. Am. Chem. Soc.* **1989**, *111*, 380-381.
262. Sasaki, T.; Kaiser, E.T. *Biopolymers* **1990**, *29*, 79-88.
263. Akerfeld, K.S.; Kim, R.M.; Camac, D.; Groves, J.T.; Lear, J.D.; DeGrado, W.F. *J. Am. Chem. Soc.* **1992**, *114*, 9656-9657.

264. Mihara, H.; Nishino, N.; Hasegawa, R.; Fujimoto, T. *Chem. Lett.* **1992**, 1805-1808.
265. Wong, A.K.; Jacobsen, S.; Winzor, D.J.; Fairlie, D.P. *J. Am. Chem. Soc.* **1998**, *120*, 3836-3841.
266. Tahmassebi, D.C.; Sasaki, T. *J. Org. Chem.* **1998**, *63*, 728-731.
267. Nishino, N. et. al. *Chem. Lett.* **1996**, 49-50.
268. Inoue, K.; Miyahara, A.; Itaya, T. *J. Am. Chem. Soc.* **1997**, *119*, 6191-6192.
269. Gibb, B.C.; Mezo, A.R.; Sherman, J.C. *Tetrahedron Lett.* **1995**, *36*, 7587-7590.
270. Mezo, A.R.; Sherman, J.C. *J. Am. Chem. Soc.* **1999**, *121*, 8983-8994.
271. Mezo, A.R. *Ph.D. Thesis, University of British Columbia, Vancouver*, **1999**.
272. Ghadiri, M.R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, 4000-4002.
273. Lieberman, M.; Sasaki, T. *J. Am. Chem. Soc.* **1991**, 1470-1471.
274. Lieberman, M.; Sasaki, T. *Tetrahedron* **1993**, *49*, 3677-3689.
275. Lieberman, M.; Tabet, M.; Sasaki, T. *J. Am. Chem. Soc.* **1994**, *116*, 5035-5044.
276. Ghadiri, M.R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, 825-831.
277. Ghadiri, M.R.; Case, M.A. *Angew. Chem. Intl. Ed. Engl.* **1993**, *32*, 1594-1597.
278. Cuenoud, B.; Schepartz, A. *Science* **1993**, *259*, 510-513.
279. Minoura, N. *J. Chem. Soc. Chem. Commun.* **1993**, 196-197.
280. Ueno, M.; Murakami, K.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575-12576.
281. For a review on the use of peptidic frameworks for the construction of molecular receptors and devices see: Voyer, N.; Lamothe, J. *Tetrahedron* **1995**, *51*, 9241-9284.
282. Tirrell, J.G.; Fournier, M.J.; Mason, T.L.; Tirrell, D.A. *C&EN* **1994**, December 19, 40-51.
283. Tuchscherer, G.; Scheibler, L.; Dumy, P.; Mutter, M. *Biopolymers* **1998**, *47*, 63-73.
284. Benson, D.R.; Hart, B.R.; Zhu, X.; Doughty, M.B. *J. Am. Chem. Soc.* **1995**, *117*, 8502-8510.

285. DeGrado, W.F. *Nature* **1993**, 365, 488-489.
286. Szilak, L.; Moitra, J.; Vinson, C. *Protein Sci.* **1997**, 6, 1273-1283.
287. Dutton, P.L. et. al. *Biochemistry* **2000**, 39, 11041-11049.
288. Shifman, J.M.; Gibney, B.R.; Sharp, R.E.; Dutton, P.L. *Biochemistry* **2000**, 39, 14813-14821.
289. DeGrado, W.F. et. al. *J. Am. Chem. Soc.* **1994**, 116, 856-865.
290. Lombardi, A.; Bryson, J.W.; Ghirlanda, G.; DeGrado, W.F. *J. Am. Chem. Soc.* **1997**, 119, 12378-12379.
291. Pecoraro, V.L. et. al. *J. Am. Chem. Soc.* **1997**, 119, 6195-6196.
292. Cheng, R.P.; Fisher, S.L.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, 118, 11349-11356.
293. Gibney, B.R.; Dutton, P.L. *Protein Sci.* **1999**, 8, 1888-1898.
294. Baltzer, L.; Broo, K.S.; Nilsson, H.; Nilsson, J. *Bioorganic & Medicinal Chemistry* **1999**, 7, 83-91.
295. Hahn, K.W.; Klis, W.A.; Stewart, J.M. *Science* **1990**, 248, 1544-1547.
296. Ghirlanda, G.; Lear, J.D.; Lombardi, A.; DeGrado, W.F. *J. Mol. Biol.* **1998**, 281, 379-391.
297. Lear, J.D.; Wasserman, Z.R.; DeGrado, W.F. *Science* **1988**, 240, 1177-1181.
298. Montal, M.S.; Blewitt, R.; Tomich, J.M.; Montal, M. *FEBS* **1992**, 313, 12-18.
299. Tam, J.P. (Editorial comment) *Biopolymers* **1999**, 51, 309-310.
300. Ayers, B. et. al. *Biopolymers* **1999**, 51, 343-354.
301. Causton, A.S.; Sherman, J.C. *Bioorganic & Medicinal Chemistry* **1999**, 7, 23-27.
302. Menger, F.M.; Takeshita, M.; Chow, J.F. *J. Am. Chem. Soc.* **1981**, 103, 5938-5939.
303. Cram, D.J.; Cram, J.M. *Container Compounds and Their Guests*; The Royal Society of Chemistry: Cambridge, **1994**.
304. Collet, A.; Gabard, J.; Jacques, J.; Cesario, M.; Guilhem, J.; Pascard, C. *J. Chem. Soc., Perkin Trans. I* **1981**, 1630-1638.

305. Robinson, G.M. *J. Chem. Soc.* **1915**, 107, 267-276.
306. Lindsey, A.S. *Chem. Ind.* **1963**, 823-824.
307. For a review see: Collet, A. *Tetrahedron* **1987**, 43, 5725-5759.
308. Keipert, S.J. *Ph.D. Thesis, University of California, Los Angeles*, **1985**.
309. For a recent review of resorcinarenes see: Timmerman, P.; Verboom, W.; Reinhoudt, D.N. *Tetrahedron* **1996**, 52, 2663-2704.
310. Erdtman, H.; Hogberg, S.; Adrahamsson, S.; Nilsson, B. *Tetrahedron Lett.* **1968**, 9, 1679-1682.
311. Fraser, J.R.; Borecka, B.; Trotter, J.; Sherman, J.C. *J. Org. Chem.* **1995**, 60, 1207-1213.
312. Fraser, J.R.; Sherman, J.C. *unpublished results*.
313. Sorrell, T.N.; Pigge, F.C. *J. Org. Chem.* **1993**, 58, 784-785.
314. Cram, D.J.; Karbach, S.; Kim, H.-E.; Knobler, C.B.; Maverick, E.F.; Ericson, J.L.; Helgeson, R.C. *J. Am. Chem. Soc.* **1988**, 110, 2229-2237.
315. Cram, D.J.; Karbach, S.; Kim, Y.H.; Baczynskyj, L.; Marti, K.; Sampson, R.M.; Kalleymeyn, W. *J. Am. Chem. Soc.* **1988**, 110, 2554-2560.
316. Buehler, C.A.; Harris, J.O.; Shacklett, C.; Block, B.P. *J. Am. Chem. Soc.* **1946**, 68, 574-577.
317. Elliot, M.; Janes, N.F.; Pearson, B.P. *J. Sci. Fd. Agric.* **1967**, 18, 325-331.
318. Cram, D.J.; Tanner, M.E.; Keipert, S.J.; Knobler, C.B. *J. Am. Chem. Soc.* **1991**, 113, 8909-8916.
319. Frisch, B.; Boeckler, C.; Schuber, F. *Bioconjugate Chem.* **1996**, 7, 180-186.
320. Haselgrübler, T.; Amerstorfer, A.; Schindler, H.; Gruber, H.J. *Bioconjugate Chem.* **1995**, 6, 242-248.
321. Carlsson, J.; Drevin, H.; Axen, R. *Biochem. J.* **1978**, 173, 723-737.
322. Shoemaker, K.R.; Kim, P.S.; Brems, D.N.; Marqusee, S.; York, E.J.; Chaiken, I.M.; Stewart, J.M.; Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **1985**, 82, 2349-2353.

323. Brinkley, M. *Bioconjugate Chem.* **1992**, *3*, 2-13.
324. Means, G.E.; Feeney, R.E. *Bioconjugate Chem.* **1990**, *1*, 2-12.
325. This method has recently been improved: Peluso, S.; Dumy, P.; Nkubana, C.; Yokokawa, Y.; Mutter, M. *J. Org. Chem.* **1999**, *64*, 7114-7120.
326. Choma, C.T.; Kaestle, K.; Akerfeldt, K.S.; Kim, R.M.; Groves, J.T.; DeGrado, W.F. *Tetrahedron Lett.* **1994**, *35*, 6191-6194.
327. Sasaki also used a similar method in the construction of his porphyrin that was "strapped" with an α -helix: Grier, G.R.; Sasaki, T. *Tetrahedron Lett.* **1997**, *38*, 3821-3824.
328. Sherman, J.C. et. al. *Tetrahedron* **1995**, *32*, 8719-8732.
329. Futaki, S.; Ishikawa, T.; Niwa, M.; Yagami, T.; Kitagawa, K. *Tetrahedron Lett.* **1995**, *36*, 201-204.
330. Mutter, M. et. al. *Tetrahedron* **1997**, *53*, 7231-7236.
331. Dawson, P.E.; Kent, S.B.H. *J. Am. Chem. Soc.* **1993**, *115*, 7263-7266.
332. Schnolzer, M.; Kent, S.B.H. *Science* **1992**, *256*, 221-225.
333. Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30-33.
334. Tuchscherer, G. *Tetrahedron Lett.* **1993**, *34*, 8419-8422.
335. Futaki, S.; Kitagawa, K. *Tetrahedron Lett.* **1994**, *35*, 1267-1270
336. Ahmed, A.K.; Schaffer, S.W.; Wetlaufer, D.B. *J. Biol. Chem.* **1975**, *250*, 8477-8482.
337. Tam, J.P.; Wu, C.-R.; Liu, W.; Zhang, J.-W. *J. Am. Chem. Soc.* **1991**, *113*, 6657-6662.
338. Fristad, W.E.; Peterson, J.R. *Synthetic Communications* **1985**, *15*, 1-5.
339. Pons, M. et. al. *J. Org. Chem.* **1993**, *58*, 6319-6328.
340. Ramage, R.; Stewart, A.S.J. *J. Chem. Soc. Perkin Trans. 1* **1993**, 1947-1952.
341. Pons, M. et. al. *J. Am. Chem. Soc.* **1998**, *120*, 6639-6650.
342. Hober, S.; Forsberg, G.; Palm, G.; Hartmanis, M.; Nilsson, B. *Biochemistry* **1992**, *31*, 1749-1756.

343. See also: Shih, H. *J. Org. Chem.* **1993**, *58*, 3003-3008.
344. See also: Regan, L.; Rockwell, A.; Wasserman, Z.; DeGrado, W.F. *Protein Sci.* **1994**, *3*, 2419-2427.
345. Webb, R.R.; Kaneko, T. *Bioconjugate Chem.* **1990**, *1*, 96-99.
346. Harpp, D.N.; Back, T.G. *J. Org. Chem.* **1971**, *36*, 3828-3829.
347. VanHorn, W.F. et. al. *Tetrahedron Lett.* **1970**, *11*, 3551-3554.
348. Mukaiyama, T.; Takahashi, K. *Tetrahedron Lett.* **1968**, *9*, 5907-5908.
349. Liu, L.; Tanke, R.S.; Miller, M.J. *J. Org. Chem.* **1986**, *51*, 5332-5337.
350. Futaki, S. et. al. *Tetrahedron Lett.* **1995**, *36*, 5203-5206.
351. Thioether bonds were also used see: Futaki, S et. al. *Bioorganic & Medicinal Chemistry* **1997**, *5*, 1883-1891.
352. Users Manual for *Applied Biosystems 431A Peptide Synthesizer* **1992**.
353. Gibney, B.R.; Rabanal, F.; Skalicky, J.J.; Wand, A.J.; Dutton, P.L. *J. Am. Chem. Soc.* **1999**, *121*, 4952-4960.
354. Hill, R.B.; Hong, J.-K.; DeGrado, W.F. *J. Am. Chem. Soc.* **2000**, *122*, 746-747.
355. For a review of this and other methods for spectroscopic analysis of protein structure see: Pelton, J.T.; McLean, L.R. *Anal. Biochem.* **2000**, *277*, 167-176.
356. Fasman, G.D. (Editor) *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press, New York, **1996**.
357. Greenfield, N.J. *Anal. Biochem.* **1996**, *235*, 1-10.
358. Johnson, W.C. *Proteins SFG* **1990**, *7*, 205-214.
359. For example see: Anderson, D.E. et. al. *Biochemistry* **1990**, *29*, 2403-2408.
360. Causton, A.S.; Sherman, J.C. *unpublished results*.
361. Nelson, J.W.; Kallenbach, N.R. *Proteins SFG* **1986**, *1*, 211-217.
362. Walgers, R.; Lee, T.C.; Cammers-Goodwin, A. *J. Am. Chem. Soc.* **1998**, *120*, 5073-5079.

363. Typically the helicity TASPs increased from about 75% to about 90% of the maximum theoretical value, upon addition of TFE.
364. Ahmad, F.; Bigelow, C.C. *Biopolymers* **1986**, *25*, 1623-1633.
365. Smith, J.S.; Scholtz, J.M. *Biochemistry* **1996**, *35*, 7292-7297.
366. Gupta, R.; Yadav, S.; Ahmed, F. *Biochemistry* **1996**, *35*, 11925-11930.
367. Gonnelli, M.; Strambini, G.B. *Biochemistry* **1997**, *36*, 16212-16220.
368. Pace, C.N. *Methods Enzymol.* **1986**, *131*, 266-280.
369. Sancho, J. et. al. *Protein Sci.* **1996**, *5*, 1376-1388.
370. Jana, R.; Hazbun, T.R.; Mollah, A.K.M.M.; Mossing, M.C. *J. Mol. Biol.* **1997**, *273*, 402-416.
371. Reverse phase HPLC also resulted in monomeric proteins using MeCN/H₂O as mobile phase (see Section 4.4).
372. Stellwagon, E. *Methods Enzymol.* **1990**, *182*, 317-328.
373. Loo, J.A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.
374. For a review on analytical ultracentrifugation see: Cole, J.L.; Hansen, J.C. *J. Biomol. Tech.* **1999**, *10*, 163-176.
375. For a review on analytical ultracentrifugation see: Minton, A.P. *Anal. Biochem.* **1989**, *176*, 209-216.
376. Pollet, R.J.; Haase, B.A.; Standaert, M.L. *J. Biol. Chem.* **1979**, *254*, 30-33.
377. For a review on analytical ultracentrifugation of complex macromolecular systems see: Hansen, J.C.; Lebowitz, J.; Demeler, B. *Biochemistry* **1994**, *33*, 13155-13163.
378. Structural Dynamics have also been probed using aromatic residues in the core, see:
(a) Handel, T.M.; Williams, S.A.; Menyhard, D.; DeGrado, W.F. *J. Am. Chem. Soc.* **1993**, *115*, 4457-4460.
379. Structural Dynamics have also been probed using aromatic residues in the core, see:
Mihara, H.; Tanaka, Y.; Fujimoto, T.; Nishino, N. *J. Chem. Soc. Perkin Trans. 2* **1995**, 1915-1921.
380. Kuwajima, K. *Proteins SFG* **1989**, *6*, 339-377.

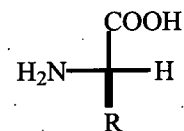
381. For a review on probing molten globule structure with ANS see: Gilmanshin, R.I. et. al. *Biopolymers* **1991**, *31*, 119-128.
382. For a review on probing molten globule structure with ANS see: Poklar, N.; Lah, J.; Salobir, M.; Macek, P.; Vesnaver, G. *Biochemistry* **1997**, *36*, 14345-14352.
383. Semisotnov, G.V. et. al. *Biopolymers* **1991**, *31*, 119-128.
384. For a review on protein NMR see: Gronenborn, A.M.; Clore, G.M. *Proteins SFG* **1994**, *11*, 273-276.
385. Hill, R.B.; Bracken, C.; DeGrado, W.F.; Palmer, A.G. *J. Am. Chem. Soc.* **2000**, *122*, 11610-11619.
386. Walsh, S.T.R.; Sukharev, V.I.; Betz, S.F.; Vekshin, N.L.; DeGrado, W.F. *J. Mol. Biol.* **2001**, *305*, 361-373.
387. Wishart, D.S.; Sykes, B.D.; Richards, F.M. *Biochemistry* **1992**, *31*, 1647-1651.
388. Case, D.A. *Curr. Opin. Struct. Biol.* **2000**, *10*, 197-203.
389. Bystrov, V.F. *Prog. NMR Spect.* **1976**, *10*, 41-81.
390. Ciesla, D.J.; Gilbert, D.E.; Feigon, J. *J. Am. Chem. Soc.* **1991**, *113*, 3957-3961.
391. Nitta, K. et. al. *Proteins SFG* **2000**, *40*, 579-589.
392. Jocelyn, P.C. *Biochemistry of the SH Group*; Academic Press, London/New York, **1972**.
393. Chen, G.C.; Yang, J.T. *Anal. Lett.* **1977**, *10*, 1195-1207.
394. Lyu, P.C.; Sherman, J.C.; Chen, A.; Kallenbach, N.R. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5317-5320.
395. Santoro, M.M.; Bolen, D.W. *Biochemistry* **1988**, *27*, 8063-8068.
396. Myers, J.K.; Pace, C.N.; Scholtz, J.M. *Protein Sci.* **1995**, *4*, 2138-2148.
397. McRorie, D.K.; Voelker, P.J. *Self-Associating Systems in the Analytical Ultracentrifuge*; Beckman Instruments, Inc., California, **1993**.
398. The value of \bar{v} (the partial specific volume of the peptide) was calculated from the amino acid composition according to: Harding, S.E.; Rowe, A.J.; Horton, J.C. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; The Royal Society of Chemistry, Oxford, **1992**.

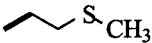

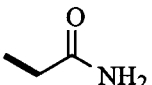
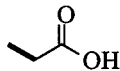
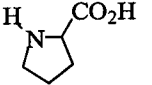
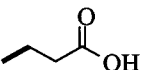
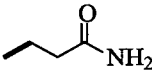
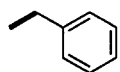
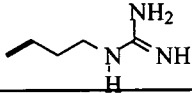

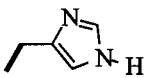

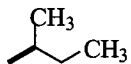

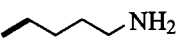
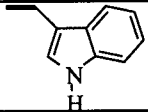
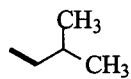
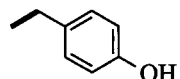


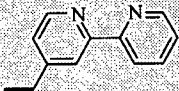

399. The densities of solvents (ρ) were calculated using the buffer concentrations: Hansen, J.C.; Lebowitz, J.; Demler, B. *Biochemistry* **1994**, *33*, 13155-13163.
400. URL: {<http://www.beckman.com/Beckman/biorsrch/BioLit/BioLitList.asp>}
401. Johnson, M.L.; Correia, J.J.; Yphantis, D.A.; Halvorson, H.R. *Biophys. J.* **1981**, *36*, 575-588.
402. Technical information article T-1782A: "α-Chymotrypsin: Characterization of a Self-Associating System in the Analytical Ultracentrifuge". Obtained from the Beckman web site: URL: {<http://www.beckman.com/Beckman/biorsrch/BioLit/BioLitList.asp>}.
403. Wilson, S.H. et. al. *Biochemistry* **1991**, *30*, 11707-11719.
404. This conclusion comes from the computer analysis of the data. Analysis of the data "by hand" (which does not account for any baseline absorbance) indicates that many of the TASP's are predominantly dimers at 100 μ M concentration.
405. URL: {<http://qobruue.usc.es/jsgroup/swan/index.html>}
406. Schrauber, H.; Eisenhaber, F.; Argos, P. *J. Mol. Biol.* **1993**, *230*, 592-612.
407. Isogai, Y.; Ishii, A.; Fujisawa, T.; Ota, M.; Nishikawa, K. *Biochemistry* **2000**, *39*, 5683-5690.
408. Lee, K. H.; Xie, D.; Freire, E.; Amzel, L.M. *Proteins SFG* **1994**, *20*, 68-84.
409. McGregor, M.J.; Islam, S.A.; Sternberg, M.J.E. *J. Mol. Biol.* **1987**, *198*, 295-310.
410. Iwadate, M.; Asakura, T.; Williamson, M.P. *Eur. J. Biochem.* **1998**, *257*, 479-487.
411. Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R.S. *Protein Sci.* **1999**, *8*, 2312-2329.
412. Feher, V.A.; Baldwin, E.P.; Dahlquist, F.W. *Nature Structural Biology* **1996**, *3*, 516-521.
413. Gonzales, L.; Plecs, J.J.; Alber, T. *Nature Structural Biology* **1996**, *3*, 510-515.
414. Stone, M.J.; Gupta, S.; Snyder, N.; Regan, L. *J. Am. Chem. Soc.* **2001**, *123*, 185-186.
415. Smith, C.K.; Withka, J.M.; Regan, L. *Biochemistry* **1994**, *33*, 5510-5517.
416. Minor, D.L.; Kim, P.S. *Nature*, **1994**, *367*, 660-663.

417. Yang, X.-M.; Yu, W.-F.; Fuchs, J.; Tasayco, M.L. *J. Am. Chem. Soc.* **1998**, *120*, 7985-7986.
418. Luger, K.; Hommel, U.; Herold, M.; Hofsteenge, J.; Kirschner, K. *Science* **1989**, *243*, 206-210.
419. Orengo, C.A.; Thornton, J.M. *Current Biology* **1993**, *1*, 105-120.
420. Lesk, A.M.; Branden, C.-I.; Chothia, C. *Proteins SFG* **1989**, *5*, 139-148.
421. Taylor, W.R.; Thornton, J.M. *Nature* **1983**, *301*, 540-542.
422. TASP s containing peptide EGG/N1/L have an overall charge that resides around the template.
423. Oas, T.G.; McIntosh, L.P.; O'Shea, E.K.; Dahlquist, F.W.; Kim, P.S. *Biochemistry* **1990**, *29*, 2891-2894.
424. Saudek, V. et. al. *Protein Eng.* **1990**, *4*, 3-10.
425. Rohl, C.A.; Baldwin, R.L. *Biochemistry* **1997**, *36*, 8435-8442.
426. Zhang, M.; Wu, B.; Baum, J.; Taylor, J.W. *J. Peptide Res.* **2000**, *55*, 398-408.
427. Greenfield, N.J.; Montelione, G.T.; Farid, R.S.; Hitchcock-DeGregori, S.E. *Biochemistry* **1998**, *37*, 7834-7843.
428. Goodman, E.M.; Kim, P.S. *Biochemistry* **1991**, *30*, 11615-11620.
429. Junis, F.K.; Mackay, J.P.; Bubb, W.A.; Jensen, S.A.; Weiss, A.S.; King, G.F. *Biochemistry* **1995**, *34*, 6164-6174.
430. Kay, M.S.; Ramos, C.H.I.; Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2007-2012.
431. Desjarlais, J.D.; Handel, T.M. *Protein Sci.* **1995**, *4*, 2006-2018.
432. Regan, L. et. al. *Protein Sci.* **1996**, *5*, 1584-1593.

Appendix A

The one- and three-letter codes for the amino acids mentioned in this thesis, and their side chains:



A	Ala	Alanine	$-\text{CH}_3$	M	Met	Methionine	
C	Cys	Cysteine		N	Asn	Asparagine	
D	Asp	Aspartic Acid		P	Pro	Proline*	
E	Glu	Glutamic Acid		Q	Gln	Glutamine	
F	Phe	Phenylalanine		R	Arg	Arginine	
G	Gly	Glycine	$-\text{H}$	S	Ser	Serine	
H	His	Histidine		T	Thr	Threonine	
I	Ile	Isoleucine		V	Val	Valine	
K	Lys	Lysine		W	Trp	Tryptophan	
L	Leu	Leucine		Y	Tyr	Tyrosine	
-	Cha	Cyclohexyl-alanine		-	Nle	Norleucine	
-	Bpa	Bipyridyl-alanine		-	Aib	α -aminoiso-butyrlic acid	

*The side chain of proline is covalently bonded to the nitrogen atom.