

**Template-Assembled Peptide Models of the N-Peptide
Helix Bundle from HIV-1 Gp41**

by

WEIMING XU

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Chemistry and Chemical Biology

written under the direction of

Dr. John W. Taylor

and approved by

New Brunswick, New Jersey

October, 2007

ABSTRACT OF THE DISSERTATION

Template-Assembled Peptide Models of the N-Peptide Helix Bundle from HIV-1 Gp41

By Weiming Xu

Dissertation Director:

Dr. John W. Taylor

The HIV-1 pandemic is one of the most serious health problems facing the world today. Infection of target cells by HIV-1 is initiated by fusion of viral and cell membranes, which is mediated by the viral glycoproteins, gp120 and gp41. After initial cell binding by gp120, the folding of gp41 to form a stable six-helix bundle structure is directly associated with membrane fusion. This helix bundle is composed of an α -helical trimer of gp41 N-helices, with three copies of the α -helical gp41 C-peptides folded onto it in an antiparallel orientation. Peptides that interfere with the formation of this six-helix bundle structure by targeting either the N-helix or the C-helix are believed to be able to block the cell fusion process and therefore prevent HIV-1 infection.

In this study, we successfully synthesized three-helix peptide structures of gp41 as models of the internal N-helix bundle, by assembling three copies of N-peptides onto three-fold symmetric templates. The templates were derived from

cis,cis-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid (Kemp's Triacid, KTA) or tris(2-aminoethyl)amine (TREN), which represent choices of rigid and flexible molecules, respectively, for organization of the folding of the N-peptide three-helix bundle.

Biophysical analysis of the four synthetic model structures, KTA-3N29, TREN-3N29, KTA-3N29b and TREN-3K-N29b, demonstrated that, at neutral pH, they all exist as monomers with high helix contents. Binding isotherms, measured by circular dichroism spectropolarimetry in the presence of physiological salt, indicated that KTA-3N29b binds three copies of the C-peptide native sequence, with a K_D of about 260 nM, while TREN-3K-N29b binds to C-peptide with a K_D in the low-micromolar range. Within the scope of this investigation, KTA-Br, the relatively more rigid template derived from KTA, was the best design for a template for the N-peptide bundle, since it generated the most helical, and most stable three-helix bundle structure with the highest binding affinity for C-peptides. These features of KTA-3N29b demonstrate that this templated three-helix bundle serves as a functional model for the native N-peptide structure that will allow detailed studies of the folding and thermodynamic stability of the gp41 six-helix bundle, and may aid the future development of potent HIV-1 fusion inhibitors and immunogens.

Dedications

To my parents, my husband, my son and all my family.

Acknowledgement

I would like to thank Dr. Taylor for his constant support and help during my entire PhD research.

I would like to thank Dr. Larry Romsted, Dr. Ralf Warmuth and Dr. Peter Lobel for serving as my committee members.

I would like to thank Dr. Norma Greenfield for her training in using CD instrument.

I would like to thank my parents, my parents-in-laws, my husband, my son and my whole family for their support, especially during the past one year when my entire life changed to a completely new stage due to the birth of my lovely son. Without them, I could not make it this far.

Table of Contents

Abstract.....	ii
Dedications.....	iv
Acknowledgement.....	v
Table of Contents.....	vi
List of Tables.....	ix
List of Illustrations.....	xi
List of Schemes.....	xv
List of Abbreviations.....	xvi
I. Introduction.....	1
II. Literature Survey.....	23
A. HIV-1 Membrane fusion.....	23
B. Current Developments of HIV-1 Fusion Inhibitors Targeting Gp41.....	31
C. Current Approaches to Construct HIV-1 N-peptide Three-helix Bundle Analog.....	47
D. Protein Design of Helix Bundle Structures.....	52
III. Experimental.....	91
A. Materials and Instruments.....	91
B. Synthesis Methods.....	94
B.1. Synthesis of template molecules.....	94
B.2. Peptide synthesis.....	108
B.3. Synthesis of N-peptide three-helix bundle model structures.....	113
B.4. Synthesis of thiol-blocked N-peptide sequences.....	115

C. Methods for CD Characterization.....	116
C.1. Sample preparation.....	116
C.2. Collecting CD spectra.....	117
C.3. Melting curves.....	117
C.4. Urea denaturation experiments.....	117
C.5. C-peptide binding studies.....	118
IV. Characterization of Three-helix Bundle Model Structures.....	129
A. Initial design: helix-bundle model structures built using Cys-N29 (KTA-3N29 and TREN-3N29)	129
A.1. CD spectrum and curve fitting results under neutral conditions.....	130
A.2. Urea denaturation and curve-fitting under neutral conditions.....	132
A.3. Thermal denaturation studies and curve-fitting under acidic conditions.....	135
A.4. Concentration dependence of melting temperature under acidic conditions.....	140
B. Modified design: helix bundle model structures built using Cys-N29b (KTA-3N29b, and TREN-3K-N29b).....	141
B.1. CD spectrum and curve fitting results under neutral conditions.....	143
B.2. Urea denaturation and curve-fitting under neutral conditions.....	146
B.3. Thermal denaturation studies and curve-fitting under acidic conditions.....	151
B.4. Concentration dependence of melting temperature under acidic conditions.....	155

C. Studies on binding of C-peptide to KTA-3N29b and TREN-3K-N29b under neutral conditions.....	156
C.1. Preliminary experiments.....	157
C.2. Derivation of binding model for binding isotherms.....	158
C.3. Results of binding isotherms.....	160
V. Discussions.....	188
A. Conformational Studies.....	188
B. Stability Studies.....	192
C. Aggregation State Studies.....	201
D. Binding Studies.....	202
VI. Conclusions.....	212
Curriculum Vitae	216

Lists of Tables

Table 1.1 Summary of N-peptide three-helix bundle model structures prepared and studied within the scope of this thesis.	13
Table 3.1 Comparison of Experimental and Calculated Mass of the Model Structures	115
Table 4.1 Curve fitting results for CD spectra in Figure 4.2.	131
Table 4.2 Curve fitting results of urea denaturation studies in 20mM Tris.HCl, pH=7.40 and 25 °C.	134
Table 4.3 Curve-fitting results for thermal denaturation of 16.4 µM KTA-3N29 at pH 2.50 analyzed using different data ranges.	138
Table 4.4 Curve-fitting results for thermal denaturation of 16.3µM TREN-3N29 at pH 2.50 analyzed using different data range.	138
Table 4.5 Melting temperatures (t_m) calculated from curve-fitting (0-85 °C) of melting curves obtained for samples with a wide range of concentrations.	141
Table 4.6 Curve fitting results for CD spectra in Figure 4.10 using LINCOMB software and the standard data set of Brahms and Brahms.	145
Table 4.7 Curve-fitting results for urea denaturation of 6.7 µM three-helix bundle model structure at 25 °C in 20 mM Tris.HCl., pH 7.40.	149
Table 4.8 Curve-fitting results for thermal denaturation of 16.4 µM KTA-3N29b at pH 2.50 using different temperature ranges.	152
Table 4.9 Curve-fitting results for thermal denaturation of 16.9 µM TREN-3K-N29b at pH 2.50 using different temperature ranges.	153

Table 4.10 Complete comparison of thermodynamic parameters of folding process of three-helix bundle model structures at pH 2.50 (using 0-85°C data).	154
Table 4.11 List of t_m 's calculated from curve-fitting (0-85°C) of melting curves obtained from samples with a wide range of concentration.	156
Table 4.12 Curve-fitting results for binding isotherm of KTA-3N29b.	161
Table 4.13 Curve-fitting results for binding isotherm of TREN-3K-N29b.	161
Table 4.14 Comparison of the conformation change of peptide C28 during binding to KTA-3N29b or TREN-3K-N29b (assume $n = 3$)	163

List of Illustrations

Figure 1.1.a Schematic graph of the glycoproteins (spikes) on the surface of HIV-1.	17
Figure 1.1.b Proposed mechanism of HIV entry.	18
Figure 1.2 a. Structural regions of gp41. b. Peptide sequences in the current study.	19
Figure 1.3.a Schematic presentation of core structure of gp41 in its fusion-active conformation.	20
Figure 1.3.b Helical wheel presentation of N36 and C34 complex crystal.	21
Figure 1.4 Overall design of this project.	22
Figure 2.1 A proposed mechanism for the conformational changes of gp41 in pore formation and stabilization.	75
Figure 2.2 Model of intermediates during HIV-1 Env-mediated fusion.	76
Figure 2.3 The HIV fusion model proposed by Y. Shai.	77
Figure 2.4 Original (in boxes) and modified peptides derived from the gp41 N-helix and C-helix.	78
Figure 2.5 Modified short peptides with anti-HIV activities.	79
Figure 2.6 Several examples of anti-HIV compounds targeting gp41.	80
Figure 2.7 Schematic representation of 5-helix and N36(L6)C34 mutants.	81
Figure 2.8 Schematic representations of current approaches to construct N-peptide three helix bundle mimics.	82
Figure 2.9 Hypothetical free energy diagram for a protein. The native state is the lowest energy state.	83
Figure 2.10 Schematic of the heptad repeat of an antiparallel coiled coil.	84
Figure 2.11 Acute packing in the three-stranded coiled coil GCN4-pII.	85

Figure 2.12 Schematic representation of template-assembled synthetic proteins (TASP).	86
Figure 2.13 Examples of TASP using linear or cyclic peptide template molecules.	87
Figure 2.14 Examples of branched organic molecules served as templates in TASP.	88
Figure 2.15 Examples of TASP formed via metal-assisted self-assembly process.	89
Figure 2.16 Other templates used in construction of TASP.	90
Figure 3.1.a LC/MS characterization of KTA-Br template (reaction mixture of the step c in Scheme 3.1).	120
Figure 3.1.b The ^1H NMR spectrum of the purified KTA-Br.	121
Figure 3.2.a Reverse-phase HPLC of the crude reaction mixture used to prepare TREN-Br template (Scheme 3.2).	122
Figure 3.2.b ESI-MS of HPLC-purified TREN-Br template (peak 11 from HPLC shown in Figure 3.2.a).	123
Figure 3.2.c The ^1H NMR spectrum of the purified TREN-Br template.	124
Figure 3.3 The ^1H NMR spectrum of purified template TREN-3[K(Br)-NH $_3^+$].	125
Figure 3.4 The ^1H NMR spectrum of purified template TREN-3[K(NH $_3^+$)-Br]	126
Figure 3.5.a. HPLC (reverse phase) UV trace of the reaction mixture used to prepare KTA-3N29.	127
Figure 3.5.b ESI-MS of purified KTA-3N29.	128
Figure 4.1 CD spectra collected at 25 °C of a) KTA-3N29, b) TREN-3N29 and c) Cys-N29 alone in 20 mM Tris.HCl, pH 7.40.	166
Figure 4.2 Comparison of CD spectra of model structures, KTA-3N29, TREN-3N29 and Cys-N29	167

Figure 4.3 Comparison of the CD spectra collected at 0 °C before and after heating up to 80 °C under neutral conditions.	168
Figure 4.4 Reversibility of urea denaturation of KTA-3N29 under neutral pH.	169
Figure 4.5 Urea denaturation studies carried out as described in section III.C.4.	170
Figure 4.6 Comparison of CD spectra collected at 0 °C before and after heating up to 80 °C under acidic conditions.	171
Figure 4.7 Thermal denaturation of KTA-3N29, TREN-3N29 and Cys-N29 under acidic conditions.	172
Figure 4.8 Curve fitting results in narrow temperature ranges.	173
Figure 4.9 Melting curves of compounds with a wide range of concentrations.	174
Figure 4.10 CD spectra collected at 25 °C of a) KTA-3N29b, b) TREN-3K-N29b and c) Cys-N29b alone in 20 mM Tris.HCl, 100 mM KF, pH 7.40.	175
Figure 4.11 Comparison of CD spectra collected at 25 °C of KTA-3N29b, TREN-3K-N29b and Cys-N29b.	176
Figure 4.12 Reversibility of urea denaturation of KTA-3N29b in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer.	177
Figure 4.13 Urea denaturation studies carried out as described in section III.C.	178
Figure 4.14 Comparison of urea denaturation curves. a) salt effect; b) KTA-3N29 versus KTA-3N29b	179
Figure 4.15 Comparison of CD spectra collected at 0 °C before and after heating up to 80 °C under acidic conditions.	180
Figure 4.16 Thermal denaturation of KTA-3N29b, TREN-3K-N29b and Cys-N29b under acidic conditions (phosphate buffer, pH = 2.50).	181

Figure 4.17 Melting curves of compounds with a wide range of concentrations.	182
Figure 4.18 CD spectra showing C28 binding to KTA-3N29b.	183
Figure 4.19 CD spectra showing C28 binding to TREN-3K-N29b.	184
Figure 4.20 Concentration dependence of the CD signal at 222 nm of C28.	185
Figure 4.21 Assay of C28 binding to KTA-3N29b.	186
Figure 4.22 Assay of C28 binding to TREN-3K-N29b.	187
Figure 5.1 The Hofmeister series.	211

List of Schemes

Scheme 3.1 Synthesis of the KTA-Br template.	96
Scheme 3.2 Synthesis of TREN-Br template.	98
Scheme 3.3 Synthesis of the TREN-3[K(Br)-NH ₃ ⁺] template.	102
Scheme 3.4 Synthesis of the TREN-3[K(NH ₃ ⁺)-Br] template.	107
Scheme 3.5 Coupling of the template molecule to three copies of N-peptide.	113

List of Abbreviations

AIDS.....	Acquired immunodeficiency syndrome
HIV	Human immunodeficiency virus
FDA.....	Food and drug administration
SIV	Simian immunodeficiency viruses
Env	Envelop glycoprotein
SHB.....	Six-stranded α -helix bundle
NHR	N-terminal heptad repeat
CHR.....	C-terminal heptad repeat
PHI.....	Pre-hairpin intermediate
KTA	<i>cis,cis</i> -1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid
TREN	<i>tris</i> -(2-aminoethyl)-amine
DMF	<i>N,N'</i> -Dimethylformamide
DCM.....	Dichloromethane
i-PrOH.....	Isopropanol
DCC.....	<i>N,N'</i> -Dicyclohexylcarbodiimide
EDC	1-(3-Dimethylaminopropyl)-3-ethyl carbodiimide hydrochloride
Fmoc.....	9-Fluorenylmethoxycarbonyl
Boc.....	<i>t</i> -Butoxycarbonyl
RINK amide MBHA resin	
..4-(2', 4'-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetamido-norleucyl-MBHA resin	
DIEA.....	<i>N,N</i> -Diisopropylcarbodiimide

PyBop.. (benzotriazol-1-yloxy)-tris(pyrrolidino)phosphonium hexafluorophosphate
HOBt.....1-hydroxybenzotriazole
TFA..... Trifluoroacetic acid
TIS.....Triisopropylsilane
TRIZMA BASE..... *tris*-[hydroxymethyl]-aminomethane
DTT..... Dithiothreitol
DCU.....*N,N'*-Dicyclohexylurea
eq.....equavelent

HPLC.....High performance liquid chromatography
CD..... Circular dichroism
IC₅₀.....Concentration for 50% inhibition
T_m.....Apparent melting temperature
Gdn.HCl.....Guanidine hydrochloride
LC/MS.....Liquid chromatography coupled mass spectrometry
TIC.....Total ion current chromatography

I. Introduction

HIV-1 infection leading to AIDS is one of the most serious health threats facing the world today, with the population of HIV-1 infected individuals worldwide at 39.5 million by the end of year 2006. HIV-1 was discovered as the etiological agent for AIDS by Robert C. Gallo and coworkers in early 1980's (1). Depletion of CD4⁺ cells was recognized as the major pathogenic mechanism of AIDS progression, while other profound immune suppressions were also associated with HIV-1 infection (2).

HIV-1 is a member of a group of retroviruses that possess complex genomes and exhibit cone-shaped capsid core particles. Like all retroviruses, the HIV-1 genome is encoded by RNA, which is reverse-transcribed to viral DNA by viral reverse transcriptase (RT) upon entering a new host cell. The replication cycle of HIV can be divided into entry and post entry steps. The entry steps feature the recognition and binding of the target cell by the mature virion, and fusion of target cell membrane with viral cell membrane leading to the release of viral genetic materials into the target cell; while the post-entry steps include (a) reverse transcription of viral RNA to viral DNA, and integration of the genomic DNA into the chromosome of the host cell; (b) regulated expression of the integrated proviral genome, (c) synthesis and assembly of viral protein and (d) virus budding and maturation. Structural biology studies of the viral proteins involved in this entire process were reviewed in (3).

While a vaccine against HIV-1 is considered by many to be required for eradication of this disease, the development of candidate vaccines has been

problematic and recent phase-III clinical trials (VaxGen) of the first candidate have been unsuccessful. Meanwhile, efforts to control the AIDS epidemic have also focused heavily on the development of potent and affordable anti-HIV drugs.

First generation of HIV therapy targeted the post entry steps of HIV infection. Drugs currently approved by the FDA include nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs). Clinical application of these drugs in combination, known as highly active antiretroviral therapy (HAART) (reviewed in (4)) consists of two or three reverse-transcriptase inhibitors combined with at least one inhibitor of viral protease. In many people, this combination antiretroviral therapy results in a dramatic reduction in viral load, decreasing the rate of decline in CD4 cell and delaying the progression to AIDS. However, the high cost and high toxic effects of this therapy have greatly limited its wider application, and the recent emergence of drug-resistant virus strains is beginning to compromise its effectiveness. Therefore it is essential and an international priority to develop more effective and less toxic anti-HIV drugs targeting the HIV-1 entry steps.

New classes of anti-HIV drugs target the entry steps, which inhibit the fusion of viral and target cell membrane and are usually called “entry inhibitors” or “fusion inhibitors”. This direction is very exciting because: first, viral fusion represents a new target for drug therapy, so that resistant viral strains have not yet emerged; second, for some peptide-based therapeutics, the design phase of the development of effective inhibitory drugs is minimal because they are based

directly on the native viral sequence; and third, side effects from the treatment by such peptides are likely to be minimal because they are readily degraded to natural amino acids, and they do not affect any known human biochemical process. In phase-III clinical trials performed in 2002, the first peptide of this type, T-20, administered twice daily at a dose of 100mg/day, was demonstrated effective at reducing viral loads in the blood stream by about one order of magnitude, even in patients who had developed a resistance to the traditional therapeutic drugs. Side-effects were limited to blistering and swelling at the injection sites.

Even though the production cost has been greatly cut, T-20 is still one of the most expensive drugs. Designing smaller and more potent inhibitors is still of great interest to many scientists.

The HIV-1 virion surface contains about 70 envelope spikes (5, 6). Each spike is composed of two non-covalently associated subunits, the surface subunit gp120, which determines direct target-cell recognition, and the transmembrane subunit gp41, which directly mediates fusion of the viral and cell membranes (7) (Figure 1.1.a). As illustrated in Figure 1.1, complexes of gp120/gp41 exist as trimers on the surface of viruses. In their native state, most of the exposed surface of this complex is that of gp120. During the fusion process (Figure 1.1.b), gp120 binds the CD4 receptor on the target cell (8), and subsequently a chemokine co-receptor (9), and gp41 is proposed to undergo conformational changes from a non-fusogenic state to a fusion-active state (10). This leads to the fusion of the viral envelope and the cellular membrane and the delivery of the

HIV-1 genome into the target cells (for detailed description, please refer to Chapter II). Each of these steps could serve as a target for the HIV entry inhibitors, which, therefore, consist of three categories: gp120-CD4 binding inhibitors, gp120-co-receptor binding inhibitors and fusion inhibitors targeting gp41. For complete review on development of all type of entry inhibitors, please refer to (11) - (13) (will be discussed in Chapter II). In this study, we will only focus on the design targeting the fusion steps mediated by gp41.

HIV-1 gp41 and gp120 is derived from the precursor gp160 by proteolytic cleavage. Gp41 is composed of 345 amino acid residues, corresponding to the sequence of 512-856 (numbering corresponds to HIV-1 gp160 for virus strain HXB2), with a molecular weight of 41 kD (Figure 1.2). It consists of three domains: an extracellular domain (ectodomain, residues 512-683), a transmembrane domain (TM, residues 684-704), and an intracellular domain (endodomain, residues 705-856). The gp41 ectodomain is composed of three important functional regions which are closely related to its fusion activity. From the N-terminus, these regions are (a) the fusion peptide (FP) region consisting of a cationic and hydrophobic, glycine-rich sequence, (b) the 4-3 heptad repeats designated NHR, and (c) another region of 4-3 heptad repeats designated CHR. Both NHR and CHR regions contain hydrophobic residues predicted to form α -helices, denoted N- and C-helix, respectively. Synthetic peptides derived from the NHR and CHR regions of gp41 are termed N-peptides and C-peptides (Figure 1.2).

In 1997, three crystal structures were constructed representing the structure of ectodomain of gp41, namely, P. Kim's structure (10), N36/C34, prepared by standard Fmoc peptide synthesis; D. C. Wiley's structure (14), prepared as a GCN4/gp41 chimera and M. Lu's structure (15), N34(L6)C28, prepared by standard recombinant DNA techniques. Although the three groups applied different strategies, their crystal structures share many essential features.

As all of the authors described, the overall core structure of gp41 is a six-stranded helical bundle (SHB). At the center, three N-terminal helices form a typical coiled-coil structure, and three C-terminal helices wrap in the reverse direction into the three hydrophobic grooves on the surface of the N-terminal trimer (Figure 1.3). To achieve the optimal packing, highly conserved hydrophobic residues at the **a** and **d** positions (referring to the heptad repeat positions on a helical wheel representation) of the N-helices pack together to form the central coiled coil and residues at the **a** and **d** positions of the C-helices pack against the **e** and **g** positions of the N-helices to form the outside antiparallel layer (Figure 1.3.b).

Further biophysical studies showed that the six-helix bundle structure folds in the absence of gp120. The isolated structure is exceedingly stable to thermal denaturation and mutations in gp41 that abolish infectivity and membrane fusion often map to residues that are expected to stabilize the gp41 core structure. Meanwhile, given the similarity of the gp41 core structure to the low-pH-induced conformation of influenza HA₂ and to the structure of Mo-MLV TM, each of which has been proposed to represent a fusion-active conformation, it is believed that

this six-helix bundle structure is the fusogenic (fusion-active) end-point state of gp41 (10).

Base on the above X-ray crystal structure studies and biophysical characterizations, it is proposed that: following the binding of gp120 to CD4 and the coreceptor, a series of conformational changes of gp41 occurs to free its fusion peptide, resulting in the formation of a transient prehairpin intermediate, which spans two membranes, with its transmembrane region in the viral membrane and the fusion peptide inserted into the host-cell membrane. At this stage, the N-peptide coiled coil and likely the C-peptide region are exposed, which make them vulnerable to inhibitory molecules. In the absence of fusion inhibitors, the prehairpin intermediate folds into the trimer-of-hairpins formed in the core gp41 structure and membrane fusion occurs. The three crystal structures described above represent the hairpin like structure presented in the final fusogenic stage (7).

The 3D structure of the prehairpin intermediate (PHI) has not been characterized to date. However, the proposal of a prehairpin intermediate not only helps explain some experimental facts, but also gives scientists an ideal target for the design of fusion inhibitors, because those structures (either polypeptide or synthetic molecules) that bind the exposed regions (either N- or C-helical regions) of this intermediate are considered to be able to either block or trigger premature folding of this structure into the six-helix bundle formation, either of which would lead to the inhibition of fusion. T-20, which is one of the best known examples of such a fusion inhibitor, corresponds to gp41 residues

638-673 and overlaps with the C-peptide sequence (628-655). T-20 is assumed (not fully proved) to be involved in a competition with the viral C-peptides for binding sites on the helical trimeric N-peptide core. Many other designs of inhibitors have also been published, including “D-peptide” (16) and a short constrained C-peptide (17) targeting the hydrophobic pocket region of N-peptide triple helix; C-peptide (10) and derivatives (18) targeting the N-helix; 5-helix (19) and IQN17 (20) targeting the C-terminal helix and others (11).

The objective of this thesis is to construct three-helix bundle peptide model structures that mimic the gp41 NHR core. By studying the stabilities of these model structures and their binding affinities for CHR, we will obtain insight into the details of the molecular mechanism of the formation of gp41 SHB and these results should provide the framework needed for the development of more potent fusion inhibitors for clinical use.

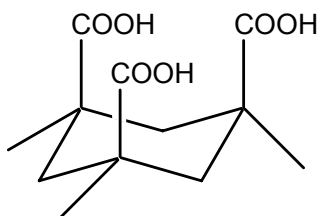
Characterization of the six-helix bundle structure formed by gp41, together with its proposed role in the viral fusion mechanism (7), has generated considerable interest in the design of analogs of the helical N-peptide trimer at the core of this structure. Appropriately designed analogs that are functional mimics of the native structure have potential applications both as inhibitors to the fusion process and as immunogens in the development of an HIV-1 vaccine. To date, several approaches have been applied to the construction of models of the N-peptide trimer. These include the chimeric variants prepared by Kim and coworkers, in which soluble trimeric coiled coils are fused to portions of the gp41

N peptide (20); the three-helix bundle structure prepared by Clore and coworkers, in which three intermolecular disulfide bridges were formed by cysteines (21,22) inserted into the native sequence of gp41; the three-helix bundle structure prepared by Gochin *et al.*, in which a metal-ion induced complex was formed by tris-bipyridyl groups added onto the N-terminal end of the N-peptide sequence (23); and, most recently, the coiled-coil mimic prepared by Bianchi and coworkers, in which both a fused coiled coil and intermolecular disulfide bonds were introduced in to direct the assembly of the N-peptide three-helix bundle (24).

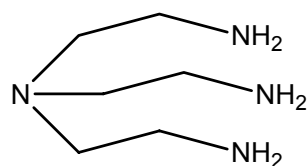
Our strategy to construct N-peptide three-helix bundle model structures is assembling three copies of N-peptide sequences onto a three-fold symmetric synthetic scaffold (Figure 1.4). Assembling linear peptide sequences onto a molecular template to achieve desired folded structures, particularly helix-bundles, has long been proposed by Mutter and co-workers (25). A wide variety of templating structures have been explored for the purposes of various investigations, in the groups of Sasaki and Kaiser (26), DeGrado (27), Ghadiri (28), Goodman (29, 30), Jensen (31) and others (for detailed review please refer to Chapter II). In each case, the covalent connection of multiple peptide chains to a directing template structure has the purpose of reducing the unfavorable entropy loss associated with the folding process, through pre-organization of the peptide chains in the correct orientation relative to one another for intramolecular assembly. In most cases, this assembly of

relatively short peptides onto a template also has the considerable advantage of bringing globular protein-like structures into the range of synthetic accessibility.

Goodman's group have successfully incorporated three collagen peptide chains onto two kinds of template (scaffold) molecules, forming mimetics of the triple-helical structure of collagen (29, 30). The scaffolds include *cis,cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid (Kemp's Triacid, KTA) and tris(2-aminoethyl)amine (TREN). Detailed studies showed that the TREN scaffold induces triple helical formation more effectively than the KTA scaffold for longer peptide chain ensembles, while the KTA scaffold is better able to stabilize shorter triple helices than the TREN scaffold, as a result of the energetics of folding. The KTA scaffold was better minimizing entropy since it is rigid, while the TREN was better at maximizing enthalpy because of its flexibility. The optimal folding is a balance between minimizing entropy and maximizing enthalpy.



KTA



TREN

It is worth mentioning that, although they are both three-helix bundle structures, the collagen structure is quite different from that of the N-peptide trimer that we propose to investigate. First, each single α -chain of collagen has Gly-Xaa-Yaa triplets as a repeat unit and adopts a left-handed polyproline-II helix, which makes its secondary structure more extended and slimmer than a

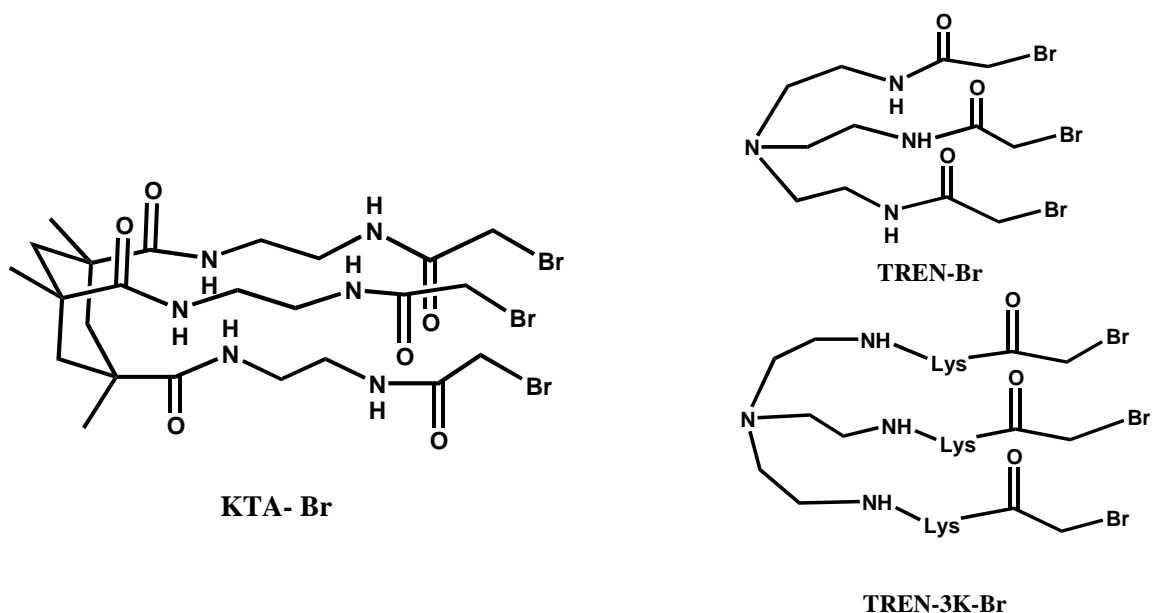
normal right-handed α -helix. Second, to form a collagen molecule, three copies of single chains are intertwined with each other into a right-handed triple-helical structure, but out of register from one strand to the next with a one-residue shift. In contrast, the gp41N-peptide trimer from HIV-1 is a well packed α -helical coiled-coil structure with a three fold axis of symmetry.

In our approach, analogs of the trimeric N-peptide α -helical core of the six-helix bundle were synthesized by assembling three copies of N-peptide sequences onto a three-fold symmetric synthetic scaffold. Conformation, stability and binding affinity of these resultant structures to a C-peptide analog were then characterized to evaluate the effectiveness of our designed template molecules on stabilizing three-helix bundle conformation and therefore generating mimics of the native N-peptide core of gp41 (Figure 1.4). This template-assembly approach was designed to provide a straightforward way to achieve the rapid assembly of multiple structural variants, allowing further optimization of the conformational and C-peptide binding properties of the model N-peptide helix-bundle.

The template molecules we explored included “KTA-Br”, derived from KTA, “TREN-Br” and “TREN-3K-Br” derived from TREN, representing choices of rigid or flexible templates (as shown below), respectively.

KTA itself was shown to favor the chair conformation in which three carboxylic acid groups in the cyclohexane ring occupy the axial positions, while the three methyl groups are equatorial (32). Therefore, derivatization of each of these axial carboxylic acids by connection to the HIV-1 N-peptide appeared to us

to provide an ideal orienting effect favoring the formation of a three-fold symmetric three-helix bundle structure that would mimic that formed by the N-peptide in the gp41 six-helix bundle. In the KTA-Br template we chose to explore, the three axial carboxylic acid groups have each been extended out from the cyclohexane ring by insertion of 1,2-diaminoethane as a spacing group to allow proper folding. These linking arms were then capped with bromoacetyl groups at their ends to provide the functional groups needed for peptide assembly onto the template.



TREN itself is a three-fold symmetric, flexible molecule. In template “TREN-Br”, each arm was capped with bromoacetyl ends to provide the bromides needed for final peptide assembly. In template “TREN-3K-Br”, a Lys residue with a free amine group on its side chain was inserted into each arm. This modification was carried out with the major intention being to increase the solubility of the final construct by introducing extra cationic groups into the

molecules under neutral pH conditions. The difference in length of each arm also allows us to examine the effect of length upon stabilization of the helix bundle by these template molecules.

End-capped Cys-N29 was the first N-peptide sequence we used to construct an N-peptide helix bundle mimic (Figure 1.2). It corresponds to the native sequence of gp41 (553-581). The N-terminal Cys was added to provide the thiol functional group used for the final S_N2 coupling reaction with the template molecules.

Cys-N29: ac-**C**NNLLRA IEAQQHL LQLTVWG IKQLQAR IL-nh₂

The modified N-peptide sequence Cys-N29b, which we used in later studies, comprises the Gln⁵⁶⁰-substituted gp41 (553-581) peptide amide. The Gln⁵⁶⁰ substitution for Glu⁵⁶⁰ in the native sequence was made after preliminary studies reported earlier (33) suggested that this substitution might enhance the solubility of the peptide at neutral pH.

Cys-N29b: ac-**C**NNLLRA IQAQQHL LQLTVWG IKQLQAR IL-nh₂

After the templates and the N-peptides were both ready, these two segments were connected together by rapid S_N2 reactions between the thiols on the side chains of the N-peptides and each of the three bromides on the arms of the template molecules. The resultant model structures that were built and studied within the scope of this thesis are listed in Table 1.1.

The properties of these constructed model structures were then analyzed and compared with untemplated N-peptide alone in terms of their conformations, stabilities, aggregation states and binding affinities for C-peptide analog. All

experimental methods are reported in Chapter III; results are reported in Chapter IV and discussed in Chapter V.

Table 1.1 Summary of N-peptide three-helix bundle model structures prepared and studied within the scope of this thesis. Structures were named after their component parts.

<i>TEMPLATE</i> <i>N-PEPTIDE</i>	<i>KTA-Br</i>	<i>TREN-Br</i>	<i>TREN-3K-Br</i>
<i>Cys-N29</i>	KTA-3N29	TREN-3N29	
<i>Cys-N29b</i>	KTA-3N29b		TREN-3K-N29b

References:

1. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, Palker TJ, Redfield R, Oleske J, Safai B, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*. 1984, **224**:500-3.
2. Fauci AS. The human immunodeficiency virus: infectivity and mechanisms of pathogenesis. *Science*. 1988, **239**:617-22.
3. Turner BG, Summers MF. Structural biology of HIV. *J Mol Biol*. 1999, **285**(1):1-32.
4. Matsushita, S., Current status and future issues in the treatment of HIV-1 infection. *Int. J. Hematol*. 2000, **72**: 20-27.
5. Gelderblom HR. Assembly and morphology of HIV: potential effect of structure on viral function. *AIDS*. 1991, **5**:617-37.
6. Zhu P., Liu J., Bess J., Chertova E., Lifson J. D., Grise H., Ofek G. A., Taylor K. A., Roux K. H.. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature*. 2006, **441**: 847-852
7. Eckert, D. M., Kim, P. S. Mechanism of viral membrane fusion and its inhibition. *Annu. Rev. Biochem*. 2001, **70**: 777-810.
8. Dalglish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*. 1984-1985, **312**:763-7.
9. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*. 1996, **272**:872-7.
10. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*. 1997, **89**:263-73.
11. Jiang S., Zhao Q., Debnath, A. K. Peptide and non-peptide HIV fusion inhibitors. *Curr. Pharm. Des*. 2002, **8**: 563-580
12. Liu S., Wu S., Jiang S., HIV entry inhibitors targeting gp41: from polypeptides to small molecule compounds. *Curr. Pharm, Des*. 2007, **13** (2): 143-62.
13. LaBranche C. C., Galasso G., Moore J. P., Bolognesi D. P., Hirsch M. S., Hammer S. M., *Antiviral Research*, 2001, **50**: 95–115.
14. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., Wiley, D. C., Atomic structure of the ectodomain from HIV-1 gp41. *Nature*, 1997, **387**: 426-430
15. Tan, K., Liu, J., Wang, J., Shen, S., Lu, M., Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. USA*, 1997, **94**: 12303-12308

16. Echert, D.M., Malashkevich, V. N., Hong, L. H., Carr, P. A., Kim, P. S., Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. *Cell*, 1999, **99**: 103-115
17. Sia, S. K., Carr, P. A., Cochran, A. G., Malashkevich, V. N., Kim, P. S., Short constrained peptides that inhibit HIV-1 entry. *Proc. Natl. Acad. Sci. USA*, 2002, **99**: 14664-14669
18. Otaka, A., Nakamura, M., Nameki, D., Kodama, E., Uchiyama, S., Nakamura, S., Nakano, H., Tamamura, H., Kobayashi, Y., Matsuoka, M., Fujii, N., Remodeling of gp41-C34 peptide leads to highly effective inhibitors of the fusion of HIV-1 with target cell. *Angew. Chem. Int. Ed.*, 2002, **41**: 2937-2940
19. Root, M. J., Kay, M. S., Kim, P. S. Protein design of an HIV-1 entry inhibitor. *Science*, 2001, **291**: 884-888
20. Eckert, D. M., Kim, P. S., Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc. Natl. Acad. Sci. USA*, 2001, **98**: 11187-11192
21. Louis, J. M., Nesheiwatt, I., Chang, L. C., Clore, M. G., Bewley, C. A. Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelop-mediated cell fusion. *J. Biol. Chem.*, 2003, **278**: 20278-20285.
22. Louis, J. M., Bewley C.A., Clore G. M. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J Biol. Chem.*, 2001, **276** (31):29485-9
23. Gochin M., Kiplin Guy R., Case M. A. A Metallopeptide assembly of the HIV-1 gp41 coiled coil is an idea recptor in fluorescence detection of ligand binding. *Angew. Chem. Int. Ed.*; 2003, **42**: 5325-5328.
24. Bianchi E., Finotto M., Ingallinella P., Hrin R., Carella A.V., Hou X.S., Schleif W.A., Miller M., Geleziunas R., Pessi A. Covalent stabilization of coiled coil of the HIV gp41 N region yields extremely potent and broad inhibitors of viral infection. *Proc. Natl. Acad. Sci. USA*, 2005, **102**: 12903-12908.
25. Mutter M., Vuilleumier S. A Chemical approach to protein design - template-assembled synthetic proteins (TASP). *Angew. Chem. Int. Ed. Engl.*, 1989, **28**:535-554.
26. Sasaki T., Kaiser E.T. Synthesis and structural stability of helichrome as an artificial heme proteins. *Biopolymers*, 1990, **29**: 79-88.
27. Akerfeldt K.S., Kim R.M., Camac D., Groves J.T., Lear J.D., DeGrado W.F. Tetraphillin: a four-helix proton channel built on a tetraphenylporphyrin framework. *J. Am. Chem. Soc.*, 1992, **114**: 9656-9657.
28. Ghadiri, M. R., Soares, C., Choi, C. Design of an artificial four-helix bundle metalloprotein via a novel ruthenium (II)-assisted self-assembly process. *J. Am. Chem. Soc.*, 1992, **114**: 4000-4002.
29. Goodman, M., Feng, Y., Melacini, G., Taulane, J. A template-induced incipient collagen-like triple-helical structure. *J. Am. Chem. Soc.*, 1996, **118**: 5156-5157.

30. Kwak, J., De Capua, A., Locardi, E., Goodman, M. TREN (tris (2-aminoethyl) amine): an effective scaffold for the assembly of triple helical collagen mimetic structures. *J. Am. Chem. Soc.*, 2002, **124**: 14085-14091.
31. Brask, J., Jensen, K. J. Carboproteins: a 4- α -helix bundle protein model assembled on a D-galactopyranoside template. *Bioorg. Med. Chem. Lett.*, 2001, **11**: 697-700.
32. Kemp, D.S., Petrakis, K.S. Synthesis and conformational analysis of cis, cis-1, 3, 5-trimethylcyclohexane-1, 3, 5-tricarboxylic acid. *J. Org. Chem.*, 1981, **46**: 5140-5143.
33. Xu, W., Taylor, J.W. Template-assembled peptide models of the N-peptide Helix bundle from HIV-1 gp41. Proceedings of 19th American Peptide Society Symposium, San Diego, U.S.A., 2005: 563-564.

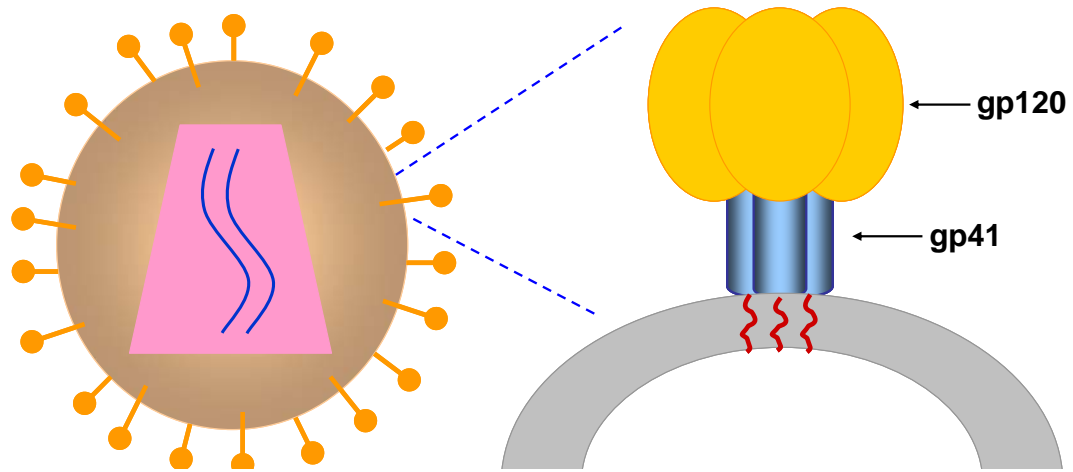


Figure 1.1.a Schematic graph of the glycoproteins (spikes) on the surface of HIV-1.

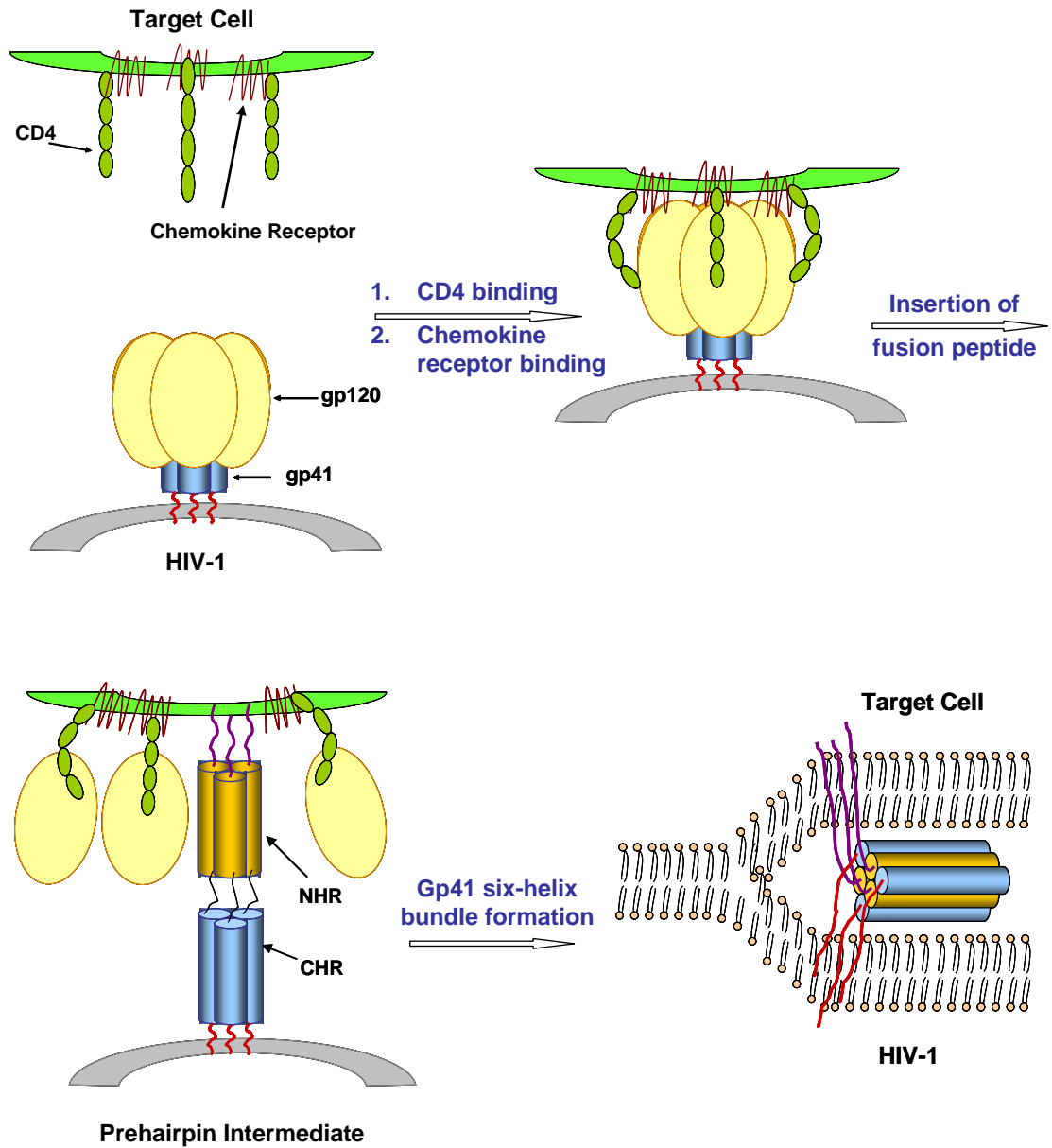
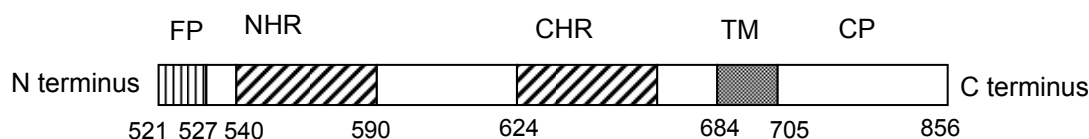


Figure 1.1.b Proposed mechanism of HIV entry. On CD4 binding, gp120 undergoes conformational changes. CD4-induced epitopes on gp120 can then bind to chemokine receptors. Thereafter, gp41 is released into a fusogenic conformation and its N-terminal (yellow) and C-terminal (blue) helices form a hairpin structure, leading to the approximation of viral and cellular membranes, which results in membrane fusion.

A.**B.**

Gp41 (553-581): NNLLRA IEAQQHL LQLTVWG IKQLQAR IL
 Cys-N29: ac-CNNLLRA IEAQQHL LQLTVWG IKQLQAR IL-nh₂
 Cys-N29b: ac-CNNLLRA IQAQQHL LQLTVWG IKQLQAR IL-nh₂

Gp41 (628-655): WMEWDRE INNYTSL IHSLEE SQNQKEK
 C28: ac-WXEWDR E INNYTSL IHSLEE SQNQKEK-nh₂ (**X**=Nle)

Figure 1.2 **A.** Structural regions of gp41. The residues are numbered according to the gp160 sequence of the HXB2 strain of HIV-1. gp41 consists of an ectodomain, a transmembrane domain (TM) and a cytoplasmic domain (CP). The ectodomain contains three major functional regions, i. e., fusion peptide (FP), N-terminal heptad repeat (NHR), and C-terminal heptad repeat (CHR). **B.** Peptide sequences used in the current study. Bold underlined letters highlight substitutions in the native sequence.

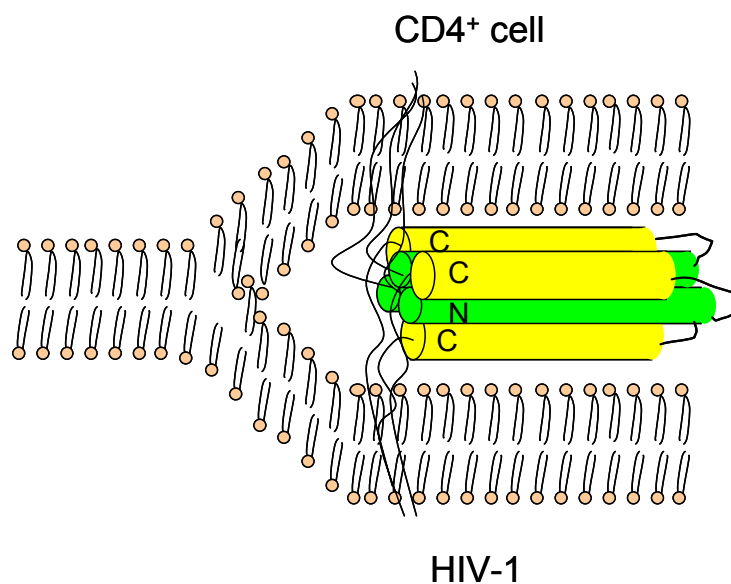


Figure 1.3.a Schematic presentation of core structure of gp41 in its fusion-active conformation, which is believed to be the end-point state of gp41 during the fusion process. Cylinders were used to represent helix conformation. The core structure of gp41 is a six-helix bundle. In the center, three NHR form a coiled-coil, three-helix bundle (green cylinders) and outside, three copies of CHR pack in the reverse direction into the three hydrophobic grooves on the surface of the NHR core (yellow cylinders) (10, 14, 15).

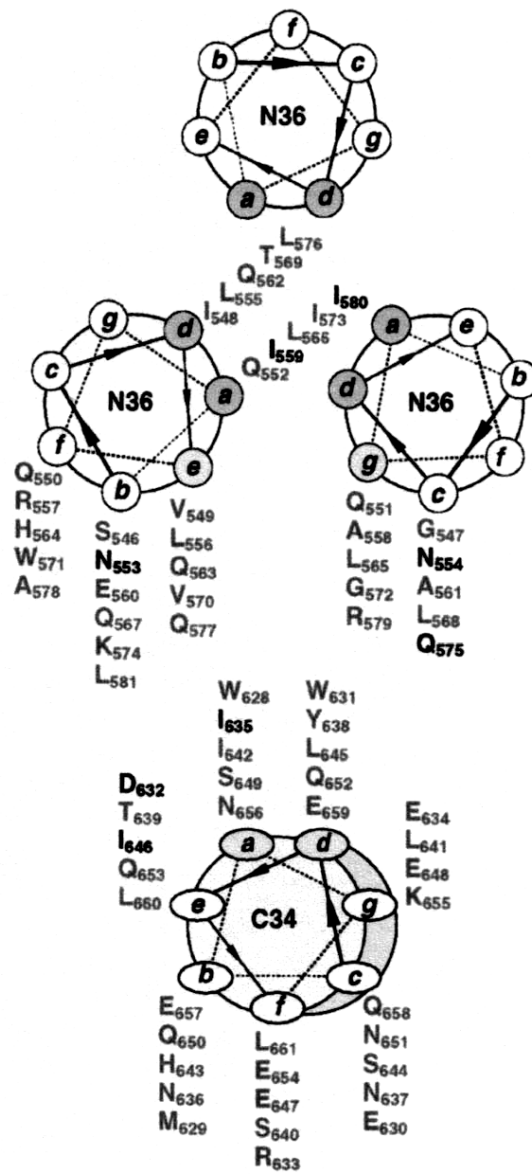


Figure 1.3.b Helical wheel presentation of N36 and C34 complex crystal (adapted from (10)). Three N36 helices and one C34 helix are presented as helical wheel projections. The view is from the top of the complex.

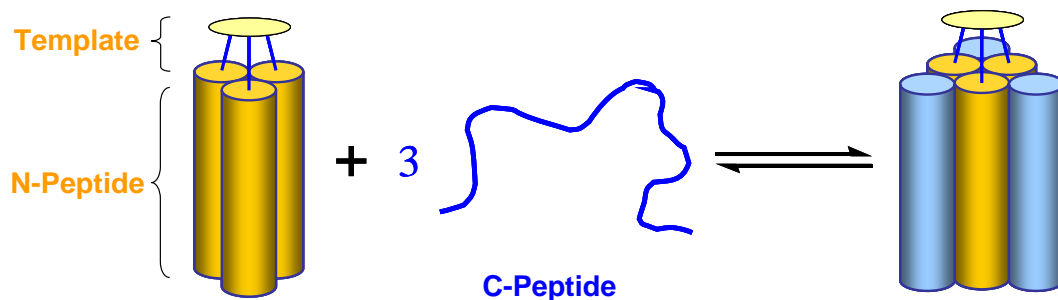


Figure 1.4 Overall design of this project: 1. develop models of the N-peptide triple-helical structure by assembling the N-peptide sequence onto a three-fold symmetric synthetic scaffold; 2. investigate the binding process of C-peptide to these model structures.

II. Literature Survey

In this chapter, literature in the related fields will be summarized with the intention of providing background information for this project and emphasizing its significance.

A. HIV-1 membrane fusion

In a brief overview, HIV entry into target cells may be summarized as shown in Figure 1.1.b. In this section, this membrane fusion procedure will be discussed in detail.

- *Synthesis of HIV-1 glycoproteins and their native pre-fusion state*

As for many other enveloped viruses, membrane fusion of HIV-1 with the target cellular membrane is mediated by the viral glycoproteins (1). In the infected cell, the major envelope glycoproteins are synthesized as an approximately 845- to 870-amino acid precursor in the rough endoplasmic reticulum. Asparagine-linked, high-mannose sugar chains are added to form the gp160 glycoprotein, which assembles into oligomers. Most evidence suggests that these oligomeric complexes are trimers (2)-(4). The gp160 trimers are transported to the Golgi apparatus, where they are cleaved by a cellular protease, generating the mature non-covalently associated envelope glycoproteins: gp120, the exterior envelope glycoprotein, and gp41, the transmembrane glycoprotein. Selected, presumably well-exposed, carbohydrates on the gp120 are modified in the Golgi apparatus by the addition of complex sugars (5). In fact, gp120 is one of the most heavily glycosylated proteins known, with about 50% of its molecular mass being carbohydrate. This sugar coating helps HIV to elude the immune system,

because the carbohydrate is derived from the host cell and is poorly recognized by antibodies (6). Gp120 and gp41 complexes are then transported to the infected cell surface, while maintaining their trimeric state. At the cell surface, a fraction of these complexes are incorporated into budding virus particles and exist as spikes on the surface of the new virions.

Crystal structures of gp41 in its post-fusion stage (3) (7) (8), and of the truncated gp120 core either in complex with fragments of CD4 and an antibody (9), or in its unliganded fully glycosylated state (10) have all been solved. However, an atomic-level structure of the intact gp120/gp41 complex has not yet been obtained. Therefore, information about the glycoprotein complexes in their native state is limited and still little is known about what gp41 looks like before the virus connects with the target cell.

Recently Zhu P. *et al*/ reported a low-resolution (at about 3 nm) structure of the spikes and their distribution on the surface of unfixed, unstained, frozen, hydrated HIV and SIV virus, using cryoelectron microscopy tomography (11). They observed that a mutant SIV displayed 73 ± 25 spikes covering the entire surface and normal HIV has only 14 ± 7 spikes, which form clusters on the viral surface. The average tomogram of a large number of spikes clearly showed images of gp120/gp41 complexes in their tripod shape where a head composed mostly of three gp120 molecules balances on three gp41 legs. By fitting the crystal structures of unliganded gp120 core to their surface density images, they deduced that the sugar faces of gp120 are at the top of the spike, and the CD4-binding site on gp120 is close to variable loop structures. This agrees with

current opinion that the spike displays on its surface poorly immunogenic sugars and immunogenic but variable loops, while restricting access to more conserved and vital parts of the complex, such as the CD4-binding site of gp120 and most of gp41 (5). It was also observed that, in contrast to most representations so far, the three gp41 legs are well separated and are well exposed for neutralizing antibodies, such as 2F5 and 4E10. This will encourage vaccine designers to target these regions.

- *Binding of glycoproteins to the CD4 receptor and the chemokine receptor*

When envelope glycoprotein complexes are incorporated into budding virions, many other cell surface proteins (12) (13), including adhesion molecules, are also incorporated. These host cell-derived molecules can assist the attachment of viruses to potential target cells. Virus attachment also involves the interaction of gp120 with specific receptors – the CD4 glycoprotein and members of the chemokine receptor family (CCR5 or CXCR4) (14-18).

Early investigations of gp120 revealed the existence of five variable (V1 to V5) regions interspersed with five conserved regions (19). The variable regions were proven to possess looplike structures and are well exposed on the surface of the gp120 glycoprotein (20) (21). The more conserved regions fold into the gp120 core, whose crystal structures, either in its unliganded form or CD4-bound form (9) (10), are available now. By comparing these two structures, molecular mechanism of the binding of gp120 to CD4 was deduced. The gp120 core is composed of two domains, a more conserved inner domain, facing the trimer axis, presumably gp41, and an outer domain that is mostly exposed on the

surface of the trimer, and a β sheet (the “bridging sheet”) that does not properly belong to either domain. Elements of both domains and the bridging sheet contribute to CD4 binding. Upon binding, the inner domain of gp120 undergoes extensive conformational rearrangement, resulting in a recessed pocket which contacts CD4 extensively. A major function of CD4 binding is to induce further conformational changes in gp120. The bridging sheet closes up to create the co-receptor binding surface, which is flanked by the V1-V2 and V3 loops, which contributes to the formation or exposure of the gp120 binding site for the chemokine receptors (22) (23).

For most HIV-1 isolates that are transmitted and that predominate during the early years of infection, CCR5 is the major co-receptor used for infection. HIV-1 isolates that arise later in the course of infection often use other chemokine receptors, frequently CXCR4, in addition to CCR5 (24). Studies have demonstrated that the V3 loop of gp120 is a major determinant of which chemokine receptor is used (25) (26), and conserved gp120 structures also appear to play an important role in chemokine receptor binding (9).

After the initial steps of fusion described above, namely binding of gp120 to cell surface CD4 molecules, followed by interaction with the co-receptors, additional reduction of gp120 disulfide bonds which is mediated by lymphocyte surface-associated protein-disulfide isomerase (27), and major conformational changes in gp120 are induced. This releases gp41 from its metastable constraints, resulting in its activation for mediating membrane fusion (28) (29).

- *Gp41 mediated membrane fusion*

Prevailing thoughts on the sequential events that take place after gp41 is activated may be summarized as follows (refer to Figure 1.1.b). Gp41 changes conformation to release its fusion peptide (FP) from a buried state and insert it into the target cell membrane, using a “spring-loaded” mechanism similar to that for the low pH-induced influenza hemagglutinin HA₂-mediated membrane fusion (30) (1). This results in a transient trimeric extended prehairpin intermediate (PHI) conformation that bridges both the viral and cellular membranes, with its FP in the target cell and its transmembrane region (TM) in the viral membrane. The PHI conformation, which is believed to exist for approximately 15 min (31), features exposed NHR and CHR regions. Collapse of the PHI into a six helix bundle (SHB), whereby three CHR helices pack in an antiparallel configuration into the highly conserved hydrophobic grooves on the surface of the central NHR coiled coil, as described by three different crystal structures representing the gp41 core (3) (7) (8). This conformational rearrangement is thought to be the driving force for drawing the viral and host membranes into close proximity. It is speculated (32) that, once the membranes are pulled together, the outer phospholipid monolayer (referred to as *cis*) leaflets that have been brought into close proximity then merge, but the distal (referred to as *trans*) leaflets remain intact, resulting in a hemifusion intermediate. The trans-monolayers making up the hemifusion diaphragm then rupture to form the complete fusion pore (33).

This predominant model was proposed based on the crystallographic studies on gp41 core and further findings in biophysical studies (3), including observations that the six-helix bundle structure folds in the absence of gp120;

that this isolated structure is exceedingly stable to thermal denaturation; and that mutations in gp41 that abolish infectivity and membrane fusion often map to residues that are expected to stabilize the gp41 core structure. In addition, the gp41 core structure is similar to the low-pH-induced conformation of influenza HA₂ and to the structure of Mo-MLV TM, each of which has been proposed to represent a fusion-active conformation. Therefore, it is believed that this six-helix bundle structure is the fusogenic (fusion-active) state of gp41 and above fusion process was proposed.

Numerous findings later also supported this sequential model described above in various aspects. For example, many peptides and small molecules designed to block bundle formation are proved to prevent fusion (34) (35) (as summarized in next part); mutations in gp41 that abolish infectivity and fusion often map to residues within the heptad repeats that are expected to stabilize the SHB (36)-(40).

Meanwhile, there are new findings that supplement this sequential model described above, by providing additional information in to the pre-fusion and post-SHB, pore formation stages, or modifying the actual role of PHI and/or SHB intermediates during the entire fusion process.

The first research of this kind was reported by Markosyan *et al*, who used a temperature jump technique to arrest pore formation at 4 °C and resume it at 37 °C (41). By adding peptides that inhibit SHB formation to cold-arrested fusion pore, they, for the 1st time, successfully characterized various transition states in the fusion process. Their results clearly demonstrated that fusion pores are

indeed formed before the folding of gp41 into bundles is complete, and therefore suggests that the actual function of the bundle is to stabilize the pore against collapse and to ensure that the pore may grow. Based on their observations, they also proposed a mechanism for pore formation and the fusion process (Figure 2.1), in which a late pre-bundle intermediate was formed after the extended PHI intermediate, but before the final SHB.

Later, Dimitrov et al. (42) investigated the fusion process by trapping conformational intermediates of gp41 with various monoclonal antibodies, including antibody NC-1, which was originally induced by a peptide analog of SHB by Jiang S. *et al.* (43), but was reevaluated in this report to target the N-peptide three-helix bundle region. Based on their observations, they concluded that two parallel pathways of gp41 conformational rearrangements co-exist in this process: one leading to SHB formation and the other leading to the generation of gp41 monomers (Figure 2.2). This proposal was reinforced by their previous finding (44) that polypeptide N36^{Mut(e,g)}, which was designed to completely abolish any N-helical binding to the C-helical region while maintaining the ability to self-associate into well-defined trimers, inhibits fusion about 50-fold more effectively than the parent N36 peptide ($IC_{50} \sim 308 \text{ nM}$ *versus* $\sim 16 \text{ }\mu\text{M}$). N36^{Mut(e,g)} peptide is proved to act by disrupting the homotrimeric coiled coil of N-terminal helices in the pre-hairpin intermediate to form heterotrimers.

Recently, Y. Shai et al. found that N36, a peptide derived from NHR, having strong affinity towards phospholipids, prominently shifts conformation from α -helix in an outer leaflet-like zwitterionic membrane to β -sheet in a

membrane mimicking the negatively charged inner leaflet environment, leading to pronounced fusion-activity, and that synergistically with the FP it induces formation of fusion pores in this environment (45). This activity can be inhibited by CHR-derived peptides that shield the membrane perturbing surfaces of N36. Taken together, their results suggest that, in addition to the main function of stabilizing the trimeric oligomeric state of the gp41 ectodomain, forming a central coiled coil at the PHI stage, which is crucial for increasing the FP's ability to fuse, the NHR also takes an active role in destabilizing membranes enriched in electronegative phospholipids, which may occur locally in the process of viral fusion due to the membrane perturbing effect of the FP. Based on their results, they also proposed a modified fusion mechanism, as described in Figure 2.3, in which a fusogenic pre-hairpin conformation of gp41 was believed to exist during the membrane fusion process.

To summarize briefly, fusion of the HIV-1 membrane with target cell membrane is initiated with binding of gp120 to the CD4 receptor on the target cell, and subsequently to a chemokine co-receptor. Gp41 then undergoes a series of conformational changes from a non-fusogenic state to a fusion-active state, leading to fusion of the viral envelope and the cellular membranes, and subsequent the delivery of the HIV-1 genome into the target cells. Even though the atomic structure of the Env proteins (gp120 and gp41) in their native state is still not solved and there are still controversies on the molecular mechanism of the post-SHB formation scenario, transient existence of PHI, that features exposed NHR and CHR regions, has been widely accepted. This proposed

intermediate provides scientists with an ideal target for the design of anti-HIV drugs. It is believed that any molecule that blocks the process of NHR homotrimerization and six-helix bundle formation by targeting NHR, NHR trimer and CHR may inhibit HIV-1 mediated membrane fusion.

B. Current developments of HIV-1 fusion inhibitors targeting gp41

Acquired immunodeficiency syndrome (AIDS), caused by the human immunodeficiency virus (HIV) infection, has rapidly become one of the most serious health problem facing the world since it was first discovered in the early 1980's (46). Though a vaccine is the last and best hope to solve the AIDS problem, an effective vaccine is still not available. Pioneering scientists in the field of gene therapy are also exploring an ambitious strategy which may provide people with immune cells that are more resistant to HIV or that can cripple the virus (47). However, all these efforts are still far away from clinical application. Therefore, developing effective, safe and affordable anti-HIV drugs is critical to save the lives of people who have been infected by HIV.

Since the first anti-HIV drug, zidovudine (AZT), was approved by FDA in 1987, twenty-one anti-HIV drugs and combinations of these drugs have been licensed. These drugs fall into three categories: 1) HIV reverse transcriptase inhibitors (RTIs), including nucleotide RTIs (NRTIs) and non-nucleotide RTIs (NNRTIs); 2) HIV protease inhibitors (PIs); and 3) HIV entry inhibitors (only T-20 has been approved) (48). Clinical applications of the anti-HIV drugs from the first two categories in various combinations, known as highly active antiretroviral therapy (HAART), have dramatically reduced the viral load to an undetectable

level and have significantly improved life expectancy for HIV-infected patients. However, substantial numbers of HIV/AIDS patients on HAART regimens have failed to respond to the current RTIs and PIs due to the emergence of drug-resistant HIV mutants (49) and serious adverse effects (50). Thus, the development of new classes of anti-HIV drugs targeting different stages of HIV replication, especially the HIV entry process, is still an urgent need.

HIV envelope glycoproteins play an important role in the entry of HIV into target cells, as discussed in the previous section. Its surface subunit, gp120, is responsible for virus binding to the cellular receptor (CD4) and coreceptors (CXCR4 or CCR5), and the transmembrane subunit gp41 mediates fusion of the virus with the target cell. In principle, all the viral and host cell components involved in the virus entry process may serve as targets for HIV entry inhibitors. Anti-HIV drugs targeting the cellular molecules, such as CD4, CCR5 and CXCR4, may have more chance to cause toxic or adverse effects, but less chance to induce drug-resistance than those targeting viral proteins. For the two Env subunits that are involved in HIV entry into target cells, gp41 is more conservative than gp120 in sequence. Therefore, gp41 is expected to be more suitable as a drug target for development of HIV entry inhibitors.

Gp41 is a glycoprotein with 345 amino acid residues (residues 512-856 of the HIV-1HXB2 gp160) and a molecular weight of 41 kD (Figure 1.2). It consists of an extracellular domain (residues 512-683), transmembrane domain (residues 684-704) and cytoplasm domain (residues 705-856). The extracellular domain contains four major functional regions, fusion peptide (FP, residues 512-

527), N-terminal heptad repeat (NHR or HR1, residues 542-592), C-terminal heptad repeat (CHR or HR2, residues 623-664) and a tryptophan-rich region (TR, residues 665-683). Both the NHR and CHR contain the 4-3 hydrophobic heptad repeat. During membrane fusion, gp41 fold its CHR back onto NHR in an anti-parallel orientation, resulting in the highly stable six-helix bundle (SHB) core, whose atomic structures have been solved by three groups in 1997 (3)(7)(8). A helical wheel representation of the interior side-chain packing of gp41 is shown in Figure 1.3.b. During the past decade, knowledge of the gp41 structure and function is rapidly growing, leading to a blooming discovery of new antiviral agents, from polypeptides to low molecular weight compounds, with potent HIV entry inhibitory activities. Some of these important discoveries will be discussed in the following section.

- *Polypeptide based inhibitors derived from gp41 NHR or CHR*

1. T-21, SJ-2176, T-20 and T-1249

In the early 1990s, before the crystal structure of the gp41 core was solved, several polypeptides with HIV-1 Env-mediated cell-cell fusion inhibitory activity had been discovered (Figure 2.4).

In 1992, Wild *et al.* reported the first synthetic peptide, derived from NHR region, designated as DP107 and later T-21 (residues 553-590), that blocked HIV-1 fusion at micromolar concentrations (51). In 1993, Jiang *et al.* found that a synthetic C-peptide corresponding to the gp41 sequence 630-659, designated as SJ-2176 (residues 638-673), inhibited HIV-1 replication at nanomolar concentrations (52) (53). Later, in 1994, Wild *et al.*, reported another C-peptide

fusion inhibitor, designated as DP178 and later T-20 (residues 638-673), that has inhibitory potency about 1000-fold more than DP107 (T-21) (54) (55).

Compared with the gp41 core six-helix bundle peptide complex formed by N36 (residues 546-581) and C34 (residues 628-661), whose crystal structure was later solved in 1997 (Figure 2.4), DP107 (T-21) has seven residues less towards the N-terminus and eight residues more towards the C-terminus of N36, while containing the hydrophobic pocket region of the internal NHR coiled-coils, which is believed to be critical for the formation of SHB and an ideal target for drug design; SJ-2176 is a part of the C34 sequence, with two residues less at the N- and C-termini, respectively; T-20 partially overlaps with the C34 sequence, lacking the hydrophobic pocket binding region (residues 628-635) of C34, which is important for the stability of the gp41 core conformation and gp41-mediated membrane fusion, but with an additional Trp rich sequence, WNWF, at its C-terminus, which is critical for the T-20-mediated anti-HIV-1 activity.

Though both C34 and T-20 peptide have potent anti-HIV activity, their applications are very distinct. T-20 was developed as a novel anti-HIV drug by Trimeris Inc. (Durham, NC) and Hoffmann-La Roche Ltd (Roche, Nutley, NJ), while C34 has been widely used as a laboratory tool to test the concept of peptide blocking of HIV entry and to study the mechanism of the HIV-1 Env-mediated membrane fusion.

Since, like C34, T-20 is also derived from the CHR region, it has been proposed that T-20 acts in the same way as C34 to inhibit viral fusion by interacting with the viral NHR region and blocking formation of the fusogenic core

of gp41 (56)-(59). However, recent investigations demonstrated that: (1) in contrast to C34, T-20 cannot form a stable SHB with N36, nor can it inhibit the SHB formation of the gp41 fusogenic core (60); (2) the tryptophan-rich domain within T-20 can bind to the membrane-spanning domain of gp41 (61)-(63); (3) T-20 is able to bind to the gp120 coreceptor-binding site in the conformation triggered by the gp120-CD4 interaction (64)-(66). These results suggest that T-20 probably inhibits HIV entry by interacting with multiple sites on gp120 and gp41 through different regions within T-20.

Though T-20 possesses potent anti-HIV activity, this peptide drug has at least three critical weaknesses, including lack of oral bioavailability, a high cost of production and a short half-life in vivo. Therefore, developing other HIV fusion inhibitors is still the major focus of scientists.

By combining the advantages of both C34 and T-20, Trimeris rationally designed a second-generation fusion inhibitor, T-1249, which is a 39-amino-acid peptide containing both the pocket-binding sequence and tryptophan-rich domain (Figure 2.4) (67). T-1249 was evaluated in Phase I/II clinical studies and showed more potent anti-HIV activity (2- to 100-fold) and longer half-life (about 2-fold) than T-20 and is active against most T-20 resistant viruses. However, its clinical development was suspended in 2004 because of formulation concerns (68).

2. Modified C-peptides with increased α -helicity

Compared with T-20, C34 is more potent in inhibiting HIV-1 infection and less susceptible to the evolution of resistant viruses. Therefore, developing potent

peptide fusion inhibitors by modifying C34 is a rational direction and has been the focus of many groups. These designs are based on the ideas that: since C34 is less soluble in aqueous solution than T-20, increasing its solubility will probably increase its effectiveness; since C34 changes from most disordered structure to a highly helical conformation upon the formation of SHB, increasing the inherent helicity of C-peptides will probably increase its NHR core binding affinity and potency, by lowering the energetic penalty arising from loss of entropy upon folding.

Otaka *et al.* replaced the non-conserved residues located at the solvent-accessible sites in the CHR (i.e., at the “*b*”, “*c*”, “*f*” and “*g*” positions in the α -helical wheel) with hydrophilic residues Glu or Lys (Figure 2.4) (69). These residues can form intrahelical salt bridges, which are expected to enhance the both solubility and α -helicity of the C34 peptide. One of their designed C-peptide analogs, SC34EK(Nle), has eight Glu/Lys ion pairs between *i* and *i* + 4 positions and shows an inhibitory activity about 3-fold higher than that of C34. As they expected, the solubility of SC34EK(Nle) in H₂O was increased more than 1000-fold over that of C34; SC34EK(Nle) showed a typical α -helical conformation in CD spectra, while C34 showed a disordered spectrum; the apparent melting temperature (T_m) of the N36/SC34EK(Nle) complex is about 20 °C higher than that of the N36/C34 complex, suggesting that the interaction between SC34EK(Nle) and N36 is much stronger than that between C34 and N36.

Judice *et al.* have designed and synthesized a 27-residue peptide analog, HIV35 (residue 638-664), corresponding to the sequence of T-20 with 9 residues

removed from its C-terminus (70). This shortened peptide has much lower anti-HIV-1 activity than T-20. Then they modified the peptide HIV35 by incorporating a covalent cross-linker between non-conserved residues at positions i and $i + 7$ to lock the intervening residues into an α -helical conformation, resulting in HIV24 and HIV31 analogs, which have one and two i and $i + 7$ crosslinked tethers, respectively (Figure 2.5). The peptide HIV31 with the highest α -helicity is much more potent in inhibiting HIV-1 infection than the peptides HIV24 and HIV35, almost comparable to T-20-mediated anti-HIV-1 activity.

Sia *et al.* have used a 14-residue C-peptide, C14 (residues 626-639), corresponding to the hydrophobic pocket binding region based on the crystal structure, as their original peptide for modification (71). The short peptide C14, itself, has a low helix content, low binding affinity for the gp41 pocket and very poor HIV inhibitory activity. They first substituted some residues of C14 with unnatural helix-favoring amino acids, and then crosslinked it chemically to stabilize the helix of C14. Among these C14 analogs, C14linkmid, was the most potent inhibitor. C14linkmid was synthesized by substituting residues at the “b” position (629 and 636) with Glu and crosslinking these residues *via* a (α , ω)-diaminoalkane group (Figure 2.5). C14linkmid could inhibit the HIV-1 induced cell-cell fusion with an IC_{50} of 35 μ M. Thermodynamic analysis of the binding affinities of the C14 analogs have shown that C14linkmid binds to the gp41 hydrophobic pocket with a dissociation constant (K_d) of 1.2 μ M. X-ray crystallography and NMR studies have revealed that C14linkmid binds to the

gp41 hydrophobic pocket in essentially the same conformation as the pocket-binding sequence of C34.

Recently, Deng *et al.*, designed a 52 residue recombinant peptide C52L (residue 624 to 675), that includes both C34 and the tryptophan-rich regions within T-20 (72). In addition, they replaced a buried polar residue (Gln 652) at a *d* position of CHR by a nonpolar Leu residue to present the C-peptide in a preformed α -helical conformation (proved by its CD spectrum) and fused a short coiled coil to the C-terminal of the gp41 C-peptide to increase its solubility. By calorimetry, C52L binds to a peptide mimic of the amino-terminal coiled coil with a K_d of 80 nM. The C52L peptide potently inhibits *in vitro* infection of human T cells by diverse primary HIV-1 isolates irrespective of coreceptor preference, with nanomolar IC₅₀ values. Significantly, C52L is fully active against T-20-resistant variants in a single-cycle HIV-1 infectivity assay.

3. Modified N-peptides

In principle, N-peptides derived from the HIV-1 gp41 NHR region should have potencies for inhibiting HIV-1 infection that are similar to C-peptides derived from the gp41 CHR region since either N- or C-peptide is expected to interact with the counterpart region of gp41 (CHR or NHR region, respectively) with a 1:1 ratio to form SHB. But in fact, C-peptides are much more potent in inhibiting HIV-1-mediated membrane fusion than N-peptides. The major reason is that N-peptides have a tendency to aggregate and cannot form a stable trimer coiled-coil conformation in the absence of C-peptides. Therefore, the current focus of N-

peptide amodification is to avoid the aggregation problem of N-peptide and increase water solubility.

Since N-peptides have been shown to aggregate in their isolated form (2) (4), one effective strategy to solve this problem is to build up N-peptide three-helix bundle structures, in which trimeric N-peptides are pre-organized into a soluble, nonaggregating coiled-coil conformation. To date, there have been several successful reports of this kind. Since these results are closely related to this project, they will be summarized in detail under a separated subtitle in the following section.

Another approach to modify N-peptide was reported by Berley *et al.* (44). In this study, they designed and synthesized two mutant N36 (residues 546-581) sequences, N36^{Mut(e,g)}, in which the 9 residues of N36 at positions “e” and “g” were replaced with residues at positions “e” and “g”, respectively, of C34 (Figure 2.4) and N36^{Mut(a,d)}, in which residues of N36 at positions “a” and “d” were replaced with residues at positions “f” and “c”, respectively, of C34. According to the crystal structure of gp41, the residues at the “a” and “d” positions of N36 are responsible for the formation of internal NHR coiled-coil, while the residues at the “e” and “g” positions of N36 are responsible for the interaction with those at the “a” and “d” positions of C34 in the SHB structure. Therefore, N36^{Mut(e,g)} is designed to completely lose its capability to bind to C34, but can still interact with NHR region, while N36^{Mut(a,d)} is designed to completely lose its capability to form homo-trimer of N-peptide, but can still bind to CHR region. As expected, N36^{Mut(a,d)} is monomeric, largely random coil, does not interact with the C34, and

does not inhibit fusion. This indicates that the trimeric coiled-coil structure is a prerequisite for interaction with the C-terminal region of gp41. However, N36^{Mut(e,g)} inhibited HIV-1 induced fusion about 50-fold more effectively than the parent N36 peptide (IC₅₀ ~308 nM *versus* ~16 μM). Since the N36^{Mut(e,g)} does not interact with C-peptides, it must act, in a way different from N36, by binding to viral gp41 NHR to form a hetero-trimer, thereby disrupting the formation of homotrimers. This design represents a new class of fusion inhibitor and also provided evidence for an alternative fusion mechanism (discussed in the previous section of this chapter).

4. Pocket-binding D-amino acid peptide

Peptide-based drugs usually have a problem of having a short half-life *in vivo*, mainly because they are readily proteolytically degraded. Using non-natural amino acids, such as D-amino acids and β-amino acids, to design a peptide drug is an effective way to avoid this problem, since their presence renders peptides resistant to proteolytic enzymes.

A very creative research program applying this idea to the design of anti-HIV drugs was reported by Eckert *et al.* in 1999 (73). In this study, they first designed a peptide chimera, IQN17, which is a stable and soluble trimeric coiled coil, composed of the N-terminal portion of GCN4-pIQI and the C-terminal fragment of N36 (N17). IQN17 successfully mimics the hydrophobic pocket of the gp41 core and therefore can be used as a model peptide to screen for the pocket-binding peptide with L-amino acids. Second, they synthesized a mirror image of IQN17 using D-amino acids (D-IQN17) and used it for screening a

phage display 12-L-aa peptide library. Last, D-peptides corresponding to the mirror images of the phage-displayed sequences that bound to D-IQN17 were synthesized and tested for anti-HIV-1 activity. Several cyclic D-peptides with anti-HIV activity were then identified. Among those, the most potent D-peptide, D10-p5-2K (Figure 2.5), inhibited HIV-1 infection with an IC_{50} in the low micromolar range. The binding of these peptides to the hydrophobic pocket in IQN17 was confirmed by X-ray crystallography and Nuclear Magnetic Resonance (NMR). These anti-HIV D-peptides are expected to be resistant to proteolytic degradation. Though they are less potent than T-20 and C34 in inhibiting HIV-1 infection, they can be used as leads to develop therapeutic anti-HIV agents with improved half-life and bioavailability.

- *Small compound inhibitors*

As mentioned before, although the first peptide anti-HIV drug T-20 has shown promise in treating HIV/AIDS patients, it has several limitations, including lack of oral availability, high cost of production and short half-life *in vivo*. The most appropriate approach to overcoming these problems is to identify nonpeptide small molecule HIV fusion inhibitors targeting gp41. However, the development of drugs of this kind has lagged behind that of HIV co-receptor antagonists targeting CXCR4 and CCR5. Many researchers believe that it is more difficult for a small molecule compound to block the formation of the gp41 fusion-active core, a process requiring protein-protein interactions involving multiple binding sites (3) (30) than it is to block the receptor-ligand interaction, because the size of the protein-protein interaction interface areas ($>1100 \text{ \AA}^2$) is generally greater than

those of receptor-ligand interactions (usually ranging from approximately 150 to 500 Å²) (74). Actually, some residues presented at the protein-protein interfaces may serve as key components providing the major free energy required for the protein-protein binding. These residues have a tendency to cluster at the center of the protein-protein interface to form “hot spots” (75). A compound binding to the “hot spot” may be able to block the interaction between two proteins.

According to the crystal structure of the gp41 core and computer modeling analysis, it has been suggested that the deep hydrophobic pocket in the groove on the surface of the gp41 internal trimer formed by the NHR domains may serve as a “hot pocket” in the gp41 coiled coil. Despite these difficulties, to date, several compounds have been discovered by applying this strategy that can disrupt gp41 SHB formation and inhibit membrane fusion (Figure 2.6).

Ferrer *et al.* developed a hybrid HIV fusion inhibitor that was composed of a 445-Dalton synthetic organic moiety attached to the N-terminus of a shorter C-peptide, P30 (residues 636-665), that lacked the pocket-binding domain (76) (Figure 2.6). The non-peptide moiety was designed to mimic the three highly conserved hydrophobic residues (Trp628, Trp631 and Ile635) of gp41 CHR that bind to the hydrophobic pocket on the surface of the internal N-trimer. They created a combinatorial library consisting of 61,275 compounds. Each of these compounds consisted of three building blocks (C-M1- M2) linked to P30, in which the building blocks at position C were selected to mimic the indole group of the Trp residue, whereas those at positions M1 and M2 were selected from non-natural amino acids that might bind to the hydrophobic residues in the deep

pocket. The most active hybrid molecule selected, C7-Mn34-Mn42-P30, inhibited HIV-1-mediated cell fusion with a potency about 20-fold higher than the P30 peptide alone, whereas the small non-peptide moiety itself had no anti-HIV activity. Further structural study of the hybrid molecule together with the coiled-coil core of gp41 (77) has demonstrated that the non-peptide component (C7-Mn34-Mn42) of the hybrid molecule bound to the hydrophobic pocket within the gp41 core; but unexpectedly, the moiety bound in two orientations, each with about 50% occupancy.

Jiang's group in the New York Blood Center has developed a series of high throughput screening assays, including the virtual screening method based on computer modeling, to screen HIV entry inhibitors targeting gp41 with low molecular weight (78)-(81). Using these assays, they have identified several small molecule organic compounds with potent inhibitory activity against HIV-1 Env-mediated cell-cell fusion and gp41 six-helix bundle formation. One of the compounds, ADS-J1 (Figure 2.6), interacts with the hydrophobic residues in the pocket, with one of its sulfonic acid groups in close proximity to K574, a basic residue in the NHR region located around the pocket. This compound inhibits HIV-1 mediated cell-cell fusion with an IC_{50} in the low micromolar range. Unexpectedly, they found that two dyes widely used in biological studies, XTT-F and Serva blue G, could inhibit HIV-1 replication and cell-cell fusion. XTT-F inhibited HIV-1-mediated cell-cell fusion with an IC_{50} of 7 μ M (82) and Serva blue G inhibited HIV-1-induced cell-cell fusion with an IC_{50} of 6 μ g/mL (83).

A major disadvantage of using ADS-J1, XTT formazan and Serva blue G as lead compounds for developing anti-HIV drugs is that they contain several sulfonic acid and/or azo groups that may potentially induce tumorigenesis. To avoid the selection of such groups, they then used the high throughput screening assays described above to screen candidates from a chemical library consisting of 33,040 “drug-like” compounds. They identified two N-substituted pyrroles, NB-2 and NB-64, that have potent anti-HIV activity (84) (85) (Figure 2.6). Both NB-2 and NB-64 effectively inhibited HIV-1 Env-mediated cell-cell fusion, and inhibited infection by both laboratory-adapted and primary HIV-1 strains with distinct genotypes and phenotypes at low micromolar levels. They also demonstrated that NB-2 and NB-64 may not significantly interact with gp120, CD4, CXCR4 or CCR5. However, they do interfere with the gp41 conformational changes by blocking the formation of the fusion-active gp41 SHB as demonstrated by several assay systems, such as ELISA, FN-PAGE and CD analysis. Optimization of these compounds to develop more potent small molecule fusion inhibitor is currently under investigation.

Recently, Jiang's group also discovered that theaflavin derivatives and catechin derivatives, the major polyphenols in black tea and green tea, respectively, inhibit HIV-1-induced membrane fusion by blocking the formation of the gp41 six-helix bundle (86) (Figure 2.6). They found that tea polyphenols at a concentration with significant fusion inhibitory activity on HIV-1 Env-mediated cell-cell fusion cannot block the interaction between gp120 and CD4 molecules, nor interact with CXCR4 or CCR5, but effectively inhibited the gp41 six-helix

bundle formation. Theaflavin-3,3'-digallate (TF3) inhibits HIV-1 replication, HIV-1 induced virus-cell fusion and cell-cell fusion, and the gp41 six-helix bundle formation with similar low concentrations (IC_{50} values range from 1 – 7 μ M). Computer-aided molecular docking analysis indicated that the TF3 molecule may fit in the hydrophobic pocket formed on the surface of the NHR coiled coil. Given this high anti-HIV potency and the fact that tea polyphenols have been consumed by human beings for a long history, these polyphenolic compounds, especially TF3, may be used as leads for designing more potent HIV entry inhibitors.

- *5-Helix and N36(L6)C34 Mutants*

One of the major limitations of using synthetic peptides, like T-20, as drugs for long-term treatment is the high cost of production. It is expected the cost of production can be significantly reduced if the peptides can be generated through recombinant techniques. 5-helix and mutants of N36(L6)C34 are recombinant proteins containing the sequences of the gp41 N- or C-peptides that may act in a similar way to the N- or C-peptides to block HIV-1 gp41-mediated membrane fusion and be developed as anti-HIV drugs (Figure 2.7).

5-helix is a novel recombinant protein designed and constructed by Kim's group in 2001 (87). It is composed of three N-peptides (N40, residues 543-582) and two C-peptides (C-38, residues 625-662), connected by -GGSGG- linkers (Figure 2.7). Under physiological conditions, 5-helix is well folded, soluble and extremely stable with a typical α -helical conformation. Since 5-Helix contains five of the six helices of the gp41 core and has one groove exposed on the surface, it can attract one of the gp41 CHR to fill in the gap and prevent the formation of

SHB, thus blocking HIV-mediated membrane fusion. It was observed that 5-Helix can inhibit HIV infection with a potency as high as T-20 (nanomolar IC₅₀).

Therefore, this polypeptide may be developed as a new anti-HIV drug. Later, they conjugated exotoxin protein, Pseudomonas, to 5-Helix (88). This chimeric protein specifically binds to the CHR region of the viral gp41 expressed on the HIV-1-infected cells. Therefore, the recombinant chimeric protein can specifically attack the HIV-1-infected cells by killing these cells with its toxin component, or block fusion of HIV-1 with uninfected cells with its 5-Helix.

N36(L6)C34 was originally designed and constructed by Min Lu's group, and its crystal structure was solved in 1997 (8). It is a single-chain recombinant protein composed of an N-peptide N36 and a C-peptide C34, connected by six hydrophilic residues (Figure 2.7). N36(L6)C34 can self-associate to form a stable SHB conformation (T_m value of 76 °C), that mimics the gp41 core. In 2002, Markosyan *et al.* proposed that introducing a single point mutation at one of the “g” positions of N36 in pocket forming region might reduce the stability of the SHB and expose some unbound fragments which might act as an isolated N- or C-peptides to inhibit the HIV-1 Env-mediated membrane fusion (89). They replaced the Ile of residue 571 with Val, Ala, Ser and Pro, respectively, and tested the α -helicity and anti-HIV-1 activity of these mutated N36(L6)C34 analogs. As they observed, the stability of the α -helical structure formed by these mutants significantly decreased, with T_m values as low as 40 °C, and the HIV-1 fusion inhibitory activities of these N36(L6)C34 variants increased as the α -helicity decreased. The most potent variant, with an I571P mutation, had an IC₅₀

value of 2.4 nM. This is comparable to that of T-20 and C34. These results indicate that N36(L6)C34 mutants may also be developed as HIV fusion inhibitors, and, furthermore provides strong evidence that the proteins inhibit fusion because they expose their C34 segments, rather than as fully folded six-helix bundles.

C. Current approaches to construct HIV-1 N-peptide three helix bundle analogs

As discussed above, characterization of the six-helix bundle structure formed by gp41, together with its proposed role in the viral fusion mechanism, has generated considerable interest in the design of analogs of the helical N-peptide trimer at the core of this structure. Appropriately designed analogs, that are nonaggregating, stable and soluble, are expected to mimic the native structure with proper conformation, as well as proper function, and therefore may have various applications. For example, a preoriented three-helix bundle formed by NHR may be used in HIV therapy, since it retains the ability of the isolated N-peptide to bind to the CHR region in gp41 and therefore interfere with the formation of the SHB and inhibit the HIV-1 Env-mediated cell-fusion process. These structural mimics of the internal NHR coiled coil, especially when it contains that hydrophobic pocket region, can also be used as a target molecule to screen small molecule fusion inhibitors that mimic the conformation of C-peptide. Furthermore, since several monoclonal antibodies have been found that specifically bind to the N-peptide trimers, such as NC-1 (42) (43) and D-5 (90), it is possible that these three-helix bundle mimics can be applied as immunogen

candidates in HIV vaccine development. For these various potential applications, several approaches have already been applied to the construction of N-helix-core trimeric models (Figure 2.8). These designs are summarized below.

Kim's group have constructed several chimeric peptides in which a partial N-pwptide sequence is fused with a peptide derived from either GCN4-pIQI' or IZ_m that can form a soluble trimeric coiled coil by itself (73) (91) (92). As we have mentioned above, IQN17 is such a chimeric peptide with its N-terminal 29 residues derived from a designed trimeric coiled coil, GCN4-pIQI', and the C-terminal 17 residues derived from the pocket forming sequence in the HIV-1 gp41 NHR region (N17: residues 565-581) (Figure 2.8 (a)). IQN17 adopts a folded conformation consisting of highly stable trimeric helices, and has good solubility. It inhibits HIV-1-mediated membrane fusion and is about two orders of magnitude more potent than the short N-peptide N17 (IC₅₀ 0.19 μ M *versus* 13 μ M). These data indicate that such approaches to increasing the solubility and decreasing aggregation of N-peptides may improve N-peptide-mediated inhibitory activities against HIV-1. Then, IQN17 was further modified by replacing N17 with longer N-peptides N23 (residues 559-581) and N36 (residues 546-581) to generate IQN23 and IQN36, respectively. As they observed, IQN23 and IQN36 have similar inhibitory activities with IC₅₀ of about 0.015 μ M and 0.088 μ M, respectively, and both of them are much more potent than IQN17. This suggests that the length of N-peptide in the chimeric peptide is not fully correlated with the inhibitory activity. Furthermore, by replacing the GCN4-pIQI' segment in the chimeric peptide with a 22-residue sequence derived from

another IZ_m, they constructed analogous chimeric N-peptides, IZN17, IZN23 and IZN36. All these structures showed anti-HIV-1 activities similar to that of IQN23 and IQN35. These data suggest that, after being fused with IZ_m, the antiviral activity of this class of chimeric N peptides could be recapitulated in the N17 portion of the N-helix, which is mostly the hydrophobic pocket forming region of the gp41 internal core.

Louis *et al.* designed and constructed a recombinant chimeric protein, N_{CCG}-gp41, by linking an N-peptide N35 (residues 546-580) to the N-terminus of N34(L6)C28, that represents the ectodomain of the gp41 core, and substituting Leu576, Gln577 and Ala578 of N35 with Cys, Cys and Gly, respectively (Figure 2.8 (b))(93). This trimeric coiled-coil is stabilized both by the intrinsic tendency of N34(L6)C28 to form a trimer and by the engineered intermolecular disulfide bonds. The trimeric nature of N_{CCG}-gp41 was confirmed by size-exclusion chromatography and sedimentation equilibrium studies. Circular dichroism (CD) spectra of N_{CCG}-gp41 displayed a typical α -helical conformation with an estimated α -helix content of 96% at pH 3. N_{CCG}-gp41 showed potent inhibitory activity against HIV-1 Env-mediated membrane fusion, with an IC₅₀ value of about 16 nM. The shortcoming of N_{CCG}-gp41 as an HIV fusion inhibitor is its low solubility at neutral pH, although it is fully soluble at low pH (< pH 4). Later on, the same group developed another two trimeric N-peptides, N35_{CCG}-N13 and N34_{CCG}, in which the N34(L6)C28 scaffold was simplified to a 13 residue N-peptide sequence (residue 546 to 558) or was deleted (Figure 2.8 (b)) (94). These two molecules also inhibited HIV-induced cell-cell fusion with IC₅₀ values

of 15 and 95 nM, respectively. Unfortunately, both molecules showed significantly less helix content (both around 70%) estimated by CD, and could not, by themselves, fold correctly into N-trimers.

Taking advantage of the above two ideas, Bianchi *et al.* recently combined the concept of fusion to a stable coiled-coil domain with the concept of covalent stabilization by using disulfide bonds, and designed a new covalent trimer (Figure 2.8 (c))(95). They chose the peptide sequence IZN17 originally designed by Kim's group (92), which has the minimum sequence that can recapitulate the activity of the N-helix trimer, and built a cysteine-knot on the N-terminus of this molecule. Formation of the covalent trimer, (CCIZN17)₃, was confirmed by HPLC-MS. Under physiological conditions (5 mM Hepes, pH 7.3 / 150 mM NaCl), the resultant (CCIZN17)₃ has an extraordinary thermodynamic stability with a T_m (measured in the presence of 2 M guanidinium hydrochloride) higher than 90 °C. It is also the most potent fusion inhibitor described to date, and inhibits fusion in a cell-cell fusion assay with an IC_{50} value of 260 pM, inhibits single-cycle infectivity against HIV_{HXB2}, HIV_{NL4-3}, and HIV_{MN-1} (IC_{50} = 40-380 pM), efficiently neutralizes acute viral infection in peripheral blood mononuclear cells, and displays a broad antiviral profile, being able to neutralize 100% of a large panel of HIV isolates, including R5, X4, and R5/X4 strains. Its potency is equal to or more than that of DP178 (T-20). Therefore, (CCIZN17)₃ is an attractive lead to develop as an antiviral drug, either alone or in combination with DP178.

A different approach, using a tris-bipyridyl metal complex to stabilize the N-peptide coiled-coil, was reported by Gochin *et al.* in 2003 (Figure 2.8 (d)) (96).

In this design, a metal-ion-binding bidentate ligand, 2,2'-bipyridyl (bpy) was attached to the N-terminus of a 31-residue peptide that consists of residues 565-584 of gp41 (the pocket forming region) at its C-terminal end and a sequence designed to stabilize three-helix bundle formation at its N-terminal end. Upon addition of Fe^{II} and Ni^{II} ions into the bpy-peptide solution, the CD spectrum showed that helix content increased from 76%, to 87% and 90%, respectively, and NMR showed that broadened lines generated by Trp and bpy resonances due to aggregation of the monomeric bpy-peptide were significantly sharpened. These results supported the formation of nonaggregating trimer in presence of metal ion. The Fe^{II} -bipyridyl complex is magenta, which corresponds to an absorbance maximum at 545 nm. This absorbance agrees well with the emission maximum of a dansylated fluorophore, which absorbs at 340 nm and emits at 542 nm. Therefore, binding of Fe^{II} -bound complex (env2.1) to the dansylated C-peptides can be directly followed by observing quenching of the dansyl fluorescence by fluorescence resonance energy transfer (FRET). Using the titration method, while monitoring the decrease of dansyl fluorescence intensity, they obtained binding isotherms of Fe-env 2.1 to dansylated C-peptide which gave a dissociation constant, K_d , of about 0.65 μM . They further demonstrated that compounds that are able to bind to the coiled coil and displace the C-peptide can also be measured in a competitive assay by following the increase in dansyl fluorescence intensity. This approach provides an efficient way to measure binding affinity of molecules to N-peptide model coiled coils and can be applied to screen small molecules as drugs. Later, the same group did report an

optimized assay based on this design (97), and then used it successfully screened several small compounds with IC_{50} values in low mM range (98).

D. Protein design of helix bundle structures

- *De novo protein design, general rules and examples*

De novo protein design, in its purest and most challenging form, is the approach in which proteins are designed and constructed to fold into a precisely defined 3-dimensional structure, with a sequence that is not directly related to that of any natural protein (99)(100). It has also been applied to a variety of strategies involving redesign of naturally occurring proteins. Since the first serious attempt was made to design a protein from scratch by DeGrado and coworkers (99), de novo protein design has proved to be a powerful method for addressing questions concerning protein stability and folding specificity. The advantage of this approach is that it critically tests our understanding of protein structures. Through an iterative process of design and characterization, the principles governing protein folding and function can be evaluated. Thus, “failures” highlight gaps in our understanding, whereas “successes” confirm the principles used in the design.

Before the idea of “de novo design” was applied to design proteins, it had been successfully applied to the study of peptides. Kaiser, Kezdy and coworkers had pioneered this field. Based on a hypothetical structure for the receptor-bound conformation of a peptide, analogs were designed that had minimal sequence homology to the parent, but retained the conformational and physical features thought to be responsible for activity. Their results demonstrated that amphiphilic

helices are important determinants of the properties of lipoproteins, peptide toxins and peptide hormones (101)-(103).

Peptides are different from proteins in their size and their conformational properties. Linear peptides, containing less than about 50 residues, generally exist, in water, as large ensembles of conformational states and adopt well-defined conformations only when bound to metal ions, surfaces or receptors (101)-(104). Thus a major strategy in peptide design is to decrease the flexibility of a peptide through the introduction of conformational constraints such as side-chain lactam bridges (105). Proteins can be considered as a number of covalently connected peptides that, although individually highly flexible and disordered, cooperatively assemble into a well-defined 3-dimensional structure. Thus, designing a specific molecular recognition surface from multiple, flexible segments that orient and assemble as desired is the challenging part of protein design.

As shown in (Figure 2.9), proteins with different free energies occupy different conformational states, with the native state as the lowest energy state, the unfolded states as a densely populated, nearly isoenergetic ensemble and the molten globule (non-native folded states) in between. To design a sequence that will fold into a unique, well-defined native structure, both conformational stability and specificity need to be taken into consideration. A successful design must contain (a) the elements of positive design (stability) to thermodynamically stabilize the desired fold; and (b) the elements of negative design (specificity) to create a large energy gap between the native fold and any other folded

conformation, otherwise the protein would adopt a molten globule-like ensemble of folds rather than a unique, native structure (106).

The folded conformations of natural proteins are generally stabilized by about 4 - 10 kcal/mol, relative to their unfolded states. A large number of interactions contribute to this stability of a protein, which, in general, include hydrophobic effects, hydrogen bonding, electrostatic interactions and van der Waals interactions. The overall stability of a protein is the sum of all types of favorable (stabilizing) interactions subtracting the sum of the unfavorable (destabilizing) interactions. Site-directed mutagenesis (107, 108) and studies with designed peptides (109) indicate that energies involved in burying residues in the interior of proteins are the major forces to determine proper folding. Relative to Ala, a buried Leu or Phe residue contributes $\sim 2 - 5$ kcal/mol per residue to stability, and buried H-bonds or salt bridges further increase stability by about 3 kcal/mol for each residue. Short-range (in sequence) H-bonds between small polar side chains and backbone amides can be stabilizing by 1–2 kcal/mol, as in helix N caps (110). Exposed salt bridges are less stabilizing, contributing ≤ 0.5 kcal/mol under physiological conditions (111)-(113). Finally, each residue in a protein contributes its intrinsic conformational preference for adopting a given set of ϕ , ψ torsional angles (114). With the exception of Pro, the energetic contribution of this effect spans a range of ~ 1 kcal/mol for the commonly occurring amino acids. The energetically unfavorable component during folding is that configuration entropy is lost, which accounts for about 1–2 kcal/mol per residue of destabilizing energy at room temperature (115). For folding to

proceed, the overall stability of a protein needs to be favorable and the information required for folding is distributed throughout the chain in a network of diverse interactions, each with different energetic contributions.

Whereas, as shown above, hydrophobic interactions provide a powerful driving force for formation of the desired structure, polar interactions are also critical (116), providing conformational specificity, often at the expense of thermodynamic stability (117). Examples of the importance of the use of polar residues in negative design include (a) the placement of hydrophilic groups that are solvent exposed in the desired fold but buried in alternatively folded structures; (b) inclusion of buried H-bonds and salt bridges that are less stabilizing than hydrophobic interactions, but can be formed only in the desired structure; (c) solvent-accessible electrostatic interactions that are very weakly stabilizing of the desired fold but strongly destabilize alternative folds. Other elements of negative design include conformational restraints from residues such as Gly and Pro that are often used to interrupt elements of secondary structure.

Since the concept of “de novo protein design” was first systematically introduced in late 1980s’ (99) (118), it has been successfully applied to design protein model structures that possess a variety of tertiary structures (such as coiled coil, helical hairpins, β -sandwich proteins), and desired functions (such as binding to metal-ions, coreceptors) (reviewed in (100) (119)-(124)). Recently, as understanding of the kinetics and thermodynamics of the folding of water-soluble proteins has been significantly advanced, scientists started to move further into the field of designing integral membrane proteins (reviewed in (133)). Due to the

limitations of space in this thesis, we will only briefly discuss examples of designed coiled coils that are closely related to the system we have worked on herein.

The α -helical coiled coil classically consists of a bundle of parallel or antiparallel α -helices with a left-handed superhelical twist. Coiled coils generally contain a 7-residue (heptad) repeat whose positions are designated *a–g*, with the *a* and *d* positions directed towards the interior of the structure (Figure 2.10). To be soluble in an aqueous solution, proteins intrinsically tend to bury the hydrophobic residues inside, occupying the *a* and *d* positions, while leaving major hydrophilic residues outside on the surface, interacting with water molecules. To date, a large number of coiled coils have been designed, including dimers (125), parallel trimers (126-129), antiparallel trimers (130), and parallel tetramers (131). And the atomic structures of some of them are now available.

An excellent example of coiled coil design that clearly demonstrated the necessity of combining positive and negative design (general rules we discussed in previous section) was reported by Alber, Kim and co-workers (127)(131), in which they examined the aggregation states and side-chain packing in a series of GCN4 analogs. The yeast transcriptional activator GCN4 is one of over 30 identified eukaryotic proteins containing the basic region leucine zipper (bZIP) DNA-binding motif. Crystallographic studies indicate that the basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices (132). In the studies of Kim and Alber, residues at ‘a’ and ‘d’ positions were systematically replaced with various hydrophobic residues while maintaining the identity of the

remaining residues. A peptide (p-IL) with isoleucine at each 'a' position and leucine at each 'd' position forms dimers in solution. Remarkably, reversal of this order of substitutions results in formation of tetramers – a peptide (p-LI) with isoleucine at each 'd' and leucine at each 'a' is tetrameric in solution and in the solid state (131). Comparison of the crystal structures of the dimer and tetramer revealed that in the dimer conformation (peptide p-IL), geometrical features of isoleucine and leucine residues determine their sequential preference in a two-stranded coiled coil; in the tetramer conformation, the local packing geometry of the two hydrophobic positions in the heptad repeat is reversed relative to that in the dimer, and as a result, p-LI peptide prefers the tetrameric structure as it preserves the same interaction patterns found in p-IL. These analyses predict that if isoleucine were placed at both 'a' and 'd' positions, both the dimeric and tetrameric states would be destabilized relative to other aggregation states. Indeed, the corresponding peptide (p-II) forms parallel trimers with 'acute packing' of the side chains at 'a' and 'd' positions (127). As shown in Figure 2.11, these hydrophobic side chains pack efficiently in the center of the structure and adopt well-defined, low-energy conformations. Thus, the steric interactions of apolar side chains dictate the fold of these peptides. It also appears that the trimer (isoleucine zipper) is the energetically most favorable state in the absence of more specific interactions. This trimeric motif was later applied by Kim in the construction of an N-peptide three-helix bundle chimeric model structure (73) (91), as we described in the previous section.

A buried, polar H-bonded interaction between two pseudosymmetrically disposed Asn side chains is important for maintaining the dimeric conformation of GCN4 (131). This is an excellent example of "negative design"; the H-bonded interaction is actually less stabilizing than a corresponding hydrophobic interaction. However, it more strongly destabilizes alternatively folded states, which would require full or partial burial of the polar Asn side chains in less favorable environments. Additionally, electrostatic interactions between partially exposed residues that interact across the helix-helix interface are important for conformational specificity. Under physiological conditions, these residues are more important for destabilizing alternatively folded structures than for stabilizing the desired fold.

- *Template assembled synthetic proteins (TASP)*

The concept of Template Assembled Synthetic Proteins (TASP) was introduced by Mutter and co-workers with the intention of circumventing the main hurdle in protein de novo design, i.e. the protein folding problem (134) (135). In this approach, linear peptide sequences were assembled onto a molecular template to achieve the desired folded structures (Figure 2.12). The multifunctional carrier molecule (template) plays a determining role by (a) favoring intramolecular interactions between the covalently fixed peptide blocks and (b) predetermining the packing topology of the TASP molecule by the number, type, and spatial arrangement of its functional groups (peptide blocks). In general, the covalent connection of multiple peptide chains to a directing template structure also has the advantage of reducing the unfavorable entropy loss associated with the

folding process, and this assembly of relatively short peptides onto a template also has the considerable advantage of bringing globular protein-like structures into the range of synthetic accessibility. Up to date, a wide variety of templating structures have been explored for the purposes of various investigations.

Linear and cyclic peptides with multifunctional groups have been used as templates to construct TASP by many groups (Figure 2.13). Mutter's group applied lysine-rich peptides (linear or cyclic) as templates in the preparation of various functional mimetics of proteins (Figure 2.13 (a)) (136)-(138). In these structures, peptide blocks were connected to the ϵ -amino groups of lysine side chains. By using different protecting strategies, they successfully constructed hetero-four-helix bundle structures. Mihara's group constructed a series of 3 α -helix bundle peptides to mimic receptor proteins by using linear cysteine-rich peptide as a template (Figure 2.13 (b)) (139). In Mihara's approach, the N-termini of component peptide segments were first bromoacetylated and then reacted with the thiol group on the side-chain of cysteine, resulting in a thiol-ether bond. By selectively removing the protecting groups on the cysteine side-chains, they also successfully prepared homo- and hetero- three-helix bundle model structures. Haehnel's group have designed and constructed hetero four-helix bundle model structures with bound cofactors to mimic hemeproteins (Figure 2.13(c))(140) (141). To connect a peptide helix component to the cyclic peptide template with four cysteine residues for the coupling reaction, they first added a 3-maleimidopropionyl (Mp) group to the N-terminus of each peptide segment, then they reacted this with the thiol group on the side-chain of cysteine. By

selectively removing the protecting groups on each thiol, they successfully prepared a four-helix bundle structure composed of three or four different peptide helices.

As an alternative strategy, small molecules with desired intrinsic symmetry were also developed as templates to build TASP (Figure 2.14). The first template of this kind, coproporphyrin, was applied by Sasaki and Kaiser (Figure 2.14 (a)) to construct helichrome, a four-helix bundle artificial hemeprotein (142).

Compared with peptide templates, this small molecule provides more rigidity while stabilizing the peptide blocks in the desired conformation. The unfolding transition of this structure induced by guanidine hydrochloride was then studied and the resulting parameters were comparable to that of native globular proteins.

DeGrado (143) then successfully constructed a four-helix bundle structure that mimicked ion channel proteins using tetraphenylporphyrin, another four-fold symmetric porphyrin derivative with even more rigidity than coproporphyrin, as the template (Figure 2.14 (b)). Other examples of organic small molecule templates include the “TRIPOD”, an organic trialdehyde, designed by Sasaki (144) to build a three-helix bundle model structure (Figure 2.14(c)), and the Kemp’s triacid (KTA) and Tris(2-aminoethyl)amine (TREN) derived templates designed by Goodman’s group for use in the formation of collagen model mimics (Figure 2.14 (d)) (145) (146). Templates used in our design to mimic the three-helix bundle formed by gp41 N-peptides are also derived from KTA and TREN, which represent rigid and flexible templates, respectively. Considering the structural differences between collagen, where three left-handed polyproline II-

like chains supercoil together to form a right-handed super helix with a one-residue shift from each strand to the next, and the classic 3-fold symmetric coiled coil formed by N-peptides, we speculate that our application of the rigid KTA-derived templates is more appropriate than that in Goodman's design. In contrast, for the flexible TREN-derived templates, since its intrinsic flexibility can accommodate some conformational variations by twisting or stretching its arms, we speculate that it should work effectively in both systems.

Metal ions are known to play important roles in defining the three-dimensional structure of native metalloproteins such as the DNA-binding zinc finger proteins. In 1991, Sasaki's group reported the synthesis of a three- α -helix-bundle protein using a tris-bipyridine metal complex as a template (Figure 2.15(a)) (147). In this approach, an amphiphilic α -helical peptide segment was covalently attached to the 4-position of bipyridine, and upon addition of a metal ion, such as Fe^{2+} , a three-helix bundle structure was then formed, assisted by formation of the Fe^{2+} complex. This approach was later applied by Gochin *et al.* in constructing a model of gp41 N-peptide three-helix bundle (Figure 2.8 (d)) (96). Ghadiri *et al.* reported a parallel four-helix bundle structure built via a similar approach (Figure 2.15 (b)) (148). They first connected a pyridyl moiety onto the N-terminus of each peptide segment, and then upon the addition of Ruthenium (II) ions, a four-helix bundle was formed with the assistance of the formation of the metal ion complex.

Jensen and co-workers have used carbohydrates as templates to construct protein mimics (reviewed in (149)). They believed that carbohydrates

were promising candidates as templates because monosaccharides are polyfunctional molecules and pyranose ring forms are relatively rigid, and epimers of sugars are often accessible. Furthermore, a vast literature has described the regiospecific manipulation of their functional groups. Examples of their constructs, chimeric polypeptide-carbohydrates, are shown in (Figure 2.16 (a)). Depending on the size and degree of folding, they proposed the terms “carbopeptides” and “carboproteins” for members of this class of chimeras. By this novel approach, they have successfully constructed a 4- α -helix bundle carboprotein.

Sherman's group have used cavitands as templates to construct helix-bundle structures (Figure 2.16 (b)) (150) (151). A cavitand is a rigid macrocycle with an enforced cavity. Cavitands can be synthesized with a variety of functional groups at the “rim” and “foot” positions. Specifically, in Sherman's design, the arylthiol-rimmed cavitand template with methyl “feet” was employed. The rim positions are spaced approximately 7 Å apart, and are within the range of natural interhelical distances, which are between 7 and 14 Å apart. Furthermore, a cavitand template can regulate the number of helices in a bundle as determined by the number of functional groups, and can direct the orientation of the helices to be in a parallel arrangement.

In a brief summary, the idea of connecting peptide sequences onto a multifunctional topological carrier molecule, which directs the spatial accommodation of the peptides to form a well-defined tertiary structure, has been proven to be a successful strategy to construct artificial proteins. To date, a

variety of template molecules have been developed, each of which has its own special features with advantages and disadvantages. Apparently, small molecule, metal complex, carbohydrate and cavitand templates are generally more synthetically accessible and are more efficient in directing the peptide blocks to adapt the desired conformation due to their relatively higher rigidity, whereas unconstrained peptide templates provide more efficient ways to build structures with hetero peptide segments due to the readily available methods to manipulate various protecting groups on the functional side-chains of the templates.

References

1. Weissenhorn W., Hinz A. Gaudin Y. Virus membrane fusion. *FEBS Letters* 2007, **581**: 2150-2155
2. Lu M, Kim PS. A trimeric structural subdomain of the HIV-1 transmembrane glycoprotein. *J Biomol Struct Dyn.* 1997, **15**:465-71.
3. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell.* 1997, **89**:263-73.
4. Lu M, Blacklow SC, Kim PS. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol.* 1995, **2**:1075-82.
5. Wyatt R., Sodroski J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science*, 1998, **280**:1884-1888.
6. Burton D. Images from the surface of HIV. *Nature*, 2006, **441**:817-818.
7. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., Wiley, D. C., Atomic structure of the ectodomain from HIV-1 gp41. *Nature*, 1997, **387**: 426-430
8. Tan, K., Liu, J., Wang, J., Shen, S., Lu, M., Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. USA*, 1997, **94**: 12303-12308
9. Kwong D, Wyatt R., Robinson J., Sweet R. W., Sodroski, J., Hendrickson W. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, 1998, **393**: 648-659.
10. Chen B., Vogan E. M., Gond H., Skehel J. J., Wiley D. C., Harrison S. C. Structure of an unliganded simian immunodeficiency virus gp120 core, *Nature*, 2005, **433**: 834-841.
11. Zhu P., Liu J., Bwess Jr J., Chertova E., Lifson J. D., Grise H., Ofek G. A., Taylor K. A., Roux K. H., Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature*, 2006, **441**: 847-852
12. G. Pantaleo, *et al.*, Human immunodeficiency virus (HIV) infection in CD4+ T lymphocytes genetically deficient in LFA1: LEA-1 is required for HIV mediate cell fusion but not for viral transmission. *J. Exp. Med.* 1991, **173**: 511-154.
13. M. Saifuddin, *et al.*, Role of virion-associated glycosylphosphatidylinositol-linked proteins CD55 and CD59 in complement resistance of cell line-derived and primary isolates of HIV-1. *J. Exp. Med.* 1995, **182**, 501- 509.
14. D. Klatzmann, *et al.*, T- Lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature*, 1984, **312**, 767-768
15. Y. Feng, C. C. Broder, P. E. Kennedy, E. A. Berger, HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane G protein-coupled receptor. *Science*, 1996, **272**, 872-877.

16. H. Choe, *et al.*, The β -Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell*, 1996, **85**, 1135-1148.
17. H. K. Deng, *et al.*, Identification of a major co-receptor for primary isolates of HIV-1, *Nature*, 1996, **381**, 661- 666.
18. Q. Sattentau and J. Moore, Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J. Exp. Med.* 1991, **174**, 407- 415.
19. B. R. Starcich, *et al.*, Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 1986, **45**, 637 – 648.
20. J. Moore, Q. Sattentau, R. Wyatt, J. Sodroski, Probing the structure of the human immunodeficiency virus surface glycoprotein gp120 with a panel of monoclonal antibodies. *J. Virol.* 1994, **68**, 469 – 484.
21. R. Wyatt, *et al.*, Functional and immunologic characterization of human immunodeficiency virus type 1 envelope glycoproteins containing deletions of the major variable regions. *J. Virol.* 1993, **67**, 4557- 4565.
22. A. Trkola, *et al.*, CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature*, 1996, **384**, 184-187.
23. J. C. Bandres, *et al.*, Human Immunodeficiency Virus (HIV) Envelope Binds to CXCR4 Independently of CD4, and Binding Can Be Enhanced by Interaction with Soluble CD4 or by HIV Envelope Deglycosylation. *J. Virol.* , 1998, **72**, 2500-2504.
24. R. I. Connor, K. Sheridan, D. Ceradini, S. Choe, N. Landau, Change in Coreceptor Use Correlates with Disease Progression in HIV-1-Infected Individuals. *J. Exp. Med.* 1997, **185**: 621 – 628
25. F. Cocchi, *et al.*, The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nature Med.* 1996: **2**: 1244 – 1247.
26. R. Speck, *et al.*, Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop. *J. Virol.* 1997, **71**: 7136 - 7139.
27. R. Barbouche, R. Miquelis, I.M. Jones and E. Fenouillet, Protein-disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion, *J. Biol. Chem.* 2003, **278**: 3131–3136.
28. F.D. Veronese, A.L. DeVico, T.D. Copeland, S. Oroszlan, R.C. Gallo and M.G. Sarngadharan, Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene, *Science* , 1985, **229**: 1402–1405.
29. M. Kowalski, J. Potz, L. Basiripour, T. Dorfman, W.C. Goh and E. Terwilliger *et al.*, Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1, *Science* , 1987, **237**: 1351–1355.
30. D. M. Eckert, P. S. Kim, Mechanism of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.* 2001, **70**: 777-810.

31. I. Munoz-Barroso, S. Durell, K. Sakaguchi, E. Appella and R. Blumenthal, Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41, *J. Cell Biol.* 1998, **140**: 315–323.
32. M.M. Kozlov, S.L. Leikin, L.V. Chernomordik, V.S. Markin and Y.A. Chizmadzhev, Stalk mechanism of vesicle fusion. Intermixing of aqueous contents, *Eur. Biophys. J.* 1989, **17**: 121–129.
33. L.V. Chernomordik, G.B. Melikyan and Y.A. Chizmadzhev, Biomembrane fusion: a new concept derived from model studies using two interacting planar lipid bilayers, *Biochim. Biophys. Acta* 1987, **906** : 309–352.
34. C.T. Wild, D.C. Shugars, T.K. Greenwell, C.B. McDanal and T.J. Matthews, Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection, *Proc. Natl Acad. Sci. USA* 1994, **91**: 9770–9774.
35. S. Jiang, K. Lin, N. Strick and A.R. Neurath, HIV-1 inhibition by a peptide, *Nature* 1993, **365**: 113.
36. J.W. Dubay, S.J. Roberts, B. Brody and E. Hunter, Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity, *J. Virol.* 1992, **66**: 4748–4756.
37. P. Pountourios, K.A. Wilson, R.J. Center, W. El Ahmar and B.E. Kemp, Human immunodeficiency virus type 1 envelope glycoprotein oligomerization requires the gp41 amphipathic alpha-helical/leucine zipper-like sequence, *J. Virol.* 1997, **71**: 2041–2049.
38. C. Wild, J.W. Dubay, T. Greenwell, T. Baird Jr, T.G. Oas and C. McDanal *et al.*, Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex, *Proc. Natl Acad. Sci. USA* 1994, **91**: 12676–12680.
39. S.S. Chen, C.N. Lee, W.R. Lee, K. McIntosh and T.H. Lee, Mutational analysis of the leucine zipper-like motif of the human immunodeficiency virus type 1 envelope transmembrane glycoprotein, *J. Virol.* 1993, **67**: 3615–3619.
40. S.S. Chen, Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane protein gp41, *J. Virol.* 1994, **68**: 2002–2010.
41. R.M. Markosyan, F.S. Cohen and G.B. Melikyan, HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation, *Mol. Biol. Cell*, 2003, **14**: 926–938.
42. A.S. Dimitrov, J.M. Louis, C.A. Bewley, G.M. Clore and R. Blumenthal, Conformational changes in HIV-1 gp41 in the course of HIV-1 envelope glycoprotein-mediated fusion and inactivation, *Biochemistry*, 2005, **44**: 12471–12479.
43. S. Jiang, K. Lin, and M. Lu, A conformation-specific monoclonal antibody reacting with fusion-active gp41 from the human immunodeficiency virus type 1 envelope glycoprotein, *J. Virol.* , 1998, **72**: 10213–10217.

44. C. A. Bewley, J. M. Louis, R. Ghirlando, and G. M. Clore Design of a novel peptide inhibitor of HIV fusion that disrupts the internal trimeric coiled-coil of gp41, *J. Biol. Chem.* , 2002, **277**: 14238-14245.
45. O. Korazim, K. Sackett, and Y. Shai Functional and structural characterization of HIV-1 gp41 ectodomain regions in phospholipids membranes suggest that the fusion-active conformation is extended. *J. Mol. Bio.* 2006, **364**: 1103-1117.
46. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, Palker TJ, Redfield R, Oleske J, Safai B, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*. 1984, **224**: 500-3.
47. J. Cohen Building an HIV-proof immune system. *Science*. 2007, **317**: 612-614.
48. E. De Clercq, Antiviral drugs in current clinical use. *J Clin Virol*, 2004, **30**: 115-33
49. V. A. Johnson, F. Brun-Vezinet, B. Clotet, B. Conway, R.T. D'Aquila, L. M. Demeter *et al.* Drug resistance mutations in HIV-1. *Top HIV Med* 2003, **11**: 215-21
50. A. Carr, D. A. Cooper. Adverse effects of antiretroviral therapy. *Lancet*, 2000, **356**: 1423-30.
51. Wild C, Oas T, McDanal C, Bolognesi D, Matthews T. A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. *Proc Natl Acad Sci USA* 1992, **89**: 10537-41.
52. Jiang S, Lin K, Strick N, Neurath AR. HIV-1 inhibition by a peptide. *Nature* 1993, **365**: 113.
53. Jiang S, Lin K, Strick N, Neurath AR. Inhibition of HIV-1 infection by a fusion domain binding peptide from HIV-1 envelope glycoprotein gp41. *Biochem Biophys Res Commun* 1993, **195**: 533-8
54. Wild C, Greenwell T, Matthews T. A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS Res Hum Retroviruses* 1993, **9**: 1051-3
55. Wild CT, Shugars DC, Greenwell TK, McDanal CB, Matthews TJ. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc Natl Acad Sci USA* 1994, **91**: 9770-4
56. Chan DC, Kim PS. HIV entry and its inhibition. *Cell* 1998, **93**: 681-4.
57. Kilby JM, Hopkins S, Venetta TM, DiMassimo B, Cloud GA, Lee JY, *et al.* Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat Med* 1998, **4**: 1302-7.
58. Kilby JM, Eron JJ. Novel therapies based on mechanisms of HIV-1 cell entry. *N Engl J Med* 2003, **348**: 2228-38.
59. Moore JP, Doms RW. The entry of entry inhibitors: a fusion of science and medicine. *Proc Natl Acad Sci USA* 2003, **100**: 10598-602.

60. Liu S, Lu H, Niu J, Xu Y, Wu S, Jiang S. Different from the HIV fusion inhibitor C34, the anti-HIV drug fuzeon (T-20) inhibits HIV-1 entry by targeting multiple sites in gp41 and gp120. *J Biol Chem* 2005, **280**: 11259-73.
61. Kliger Y, Gallo SA, Peisajovich SG, Munoz-Barroso I, Avkin S, Blumenthal R, *et al.* Mode of action of an antiviral peptide from HIV-1. Inhibition at a post-lipid mixing stage. *J Biol Chem* 2001, **276**: 1391-7.
62. Hildinger M, Dittmar MT, Schult-Dietrich P, Fehse B, Schnierle BS, Thaler S, *et al.* Membrane-anchored peptide inhibits human immunodeficiency virus entry. *J Virol* 2001, **75**: 3038-42.
63. Peisajovich SG, Gallo SA, Blumenthal R, Shai Y. C-terminal octylation rescues an inactive T20 mutant: implications for the mechanism of HIV/SIV-induced membrane fusion. *J Biol Chem* 2003, **278**: 21012-7.
64. Reeves JD, Miamidian JL, Biscone MJ, Lee FH, Ahmad N, Pierson TC, *et al.* Impact of mutations in the coreceptor binding site on human immunodeficiency virus type 1 fusion, infection, and entry inhibitor sensitivity. *J Virol* 2004, **78**: 5476-85.
65. Yuan W, Craig S, Si Z, Farzan M, Sodroski J. CD4-induced T-20 binding to human immunodeficiency virus type 1 gp120 blocks interaction with the CXCR4 coreceptor. *J Virol* 2004, **78**: 5448-57.
66. Alam SM, Paleos CA, Liao HX, Searce R, Robinson J, Haynes BF. An Inducible HIV Type 1 gp41 HR-2 Peptide-Binding Site on HIV Type 1 Envelope gp120. *AIDS Res Hum Retroviruses* 2004, **20**: 836-45.
67. Veiga AS, Santos NC, Loura LM, Fedorov A, Castanho MA. HIV fusion inhibitor peptide T-1249 is able to insert or adsorb to lipidic bilayers. Putative correlation with improved efficiency. *J Am Chem Soc* 2004, **126**: 14758-63.
68. Martin-Carbonero L. Discontinuation of the clinical development of fusion inhibitor T-1249. *AIDS Rev* 2004, **6**: 61.
69. Otaka A, Nakamura M, Nameki D, Kodama E, Uchiyama S, Nakamura S, *et al.* Remodeling of gp41-C34 peptide leads to highly effective inhibitors of the fusion of HIV-1 with target cells. *Angew Chem Int Ed* 2002, **41**: 2937-40.
70. Judice JK, Tom JY, Huang W, Wrin T, Vennari J, Petropoulos CJ, *et al.* Inhibition of HIV type 1 infectivity by constrained alpha-helical peptides: implications for the viral fusion mechanism. *Proc Natl Acad Sci USA* 1997, **94**: 13426-30.
71. Sia SK, Carr PA, Cochran AG, Malashkevich VN, Kim PS. Short constrained peptides that inhibit HIV-1 entry. *Proc Natl Acad Sci USA* 2002, **99**: 14664-9.
72. Deng Y., Zheng Q., Ketas T. J., Moore J. P., Lu M., Protein Design of a Bacterially Expressed HIV-1gp41 Fusion Inhibitor. *Biochemistry*, 2007, **46**: 4360-4369.
73. Eckert DM, Malashkevich VN, Hong LH, Carr PA, Kim PS. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. *Cell* 1999, **99**: 103-15.

74. Lo CL, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. *J Mol Biol* 1999, **285**: 2177-98.
75. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J Mol Biol* 1998, **280**: 1-9.
76. Ferrer M, Kapoor TM, Strassmaier T, Weissenhorn W, Skehel JJ, Oprian D, *et al.* Selection of gp41-mediated HIV-1 cell entry inhibitors from biased combinatorial libraries of non-natural binding elements. *Nat Struct Biol* 1999, **6**: 953-60.
77. Zhou G, Ferrer M, Chopra R, Kapoor TM, Strassmaier T, Weissenhorn W, *et al.* The structure of an HIV-1 specific cell entry inhibitor in complex with the HIV-1 gp41 trimeric core. *Bioorg Med Chem* 2000, **8**: 2219-27.
78. Jiang S, Lin K, Zhang L, Debnath AK. A screening assay for antiviral compounds targeted to the HIV-1 gp41 core structure using a conformation-specific monoclonal antibody. *J Virol Methods* 1999, **80**: 85-96.
79. Liu S, Jiang S. High throughput screening and characterization of HIV-1 entry inhibitors targeting gp41: theories and techniques. *Curr Pharm Des* 2004, **10**: 1827-43.
80. Liu S, Boyer-Chatenet L, Lu H, Jiang S. Rapid and automated fluorescence-linked immunosorbent assay for high throughput screening of HIV-1 fusion inhibitors targeting gp41. *J Biomol Screen* 2003, **8**: 685-93.
81. Liu S, Zhao Q, Jiang S. Determination of the HIV-1 gp41 postfusion conformation modeled by synthetic peptides: applicable for identification of the HIV-1 fusion inhibitors. *Peptides* 2003, **24**: 1303-13.
82. Zhao Q, Ernst JT, Hamilton AD, Debnath AK, Jiang S. XTT formazan widely used to detect cell viability inhibits HIV type 1 infection *in vitro* by targeting gp41. *AIDS Res Hum Retroviruses* 2002, **18**: 989-97.
83. Liu S, Lu H, Zhao Q, Jiang S. Serva blue derivatives inhibit HIV entry. XV International AIDS Conference 2004; E710F2527: 9-13.
84. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001, **46**: 3-26.
85. Jiang S, Lu H, Liu S, Zhao Q, He Y, Debnath AK. N-substituted pyrrole derivatives as novel human immunodeficiency virus type 1 entry inhibitors that interfere with the gp41 six-helix bundle formation and block virus fusion. *Antimicrob Agents Chemother* 2004, **48**: 4349-59.
86. Liu S, Lu H, Zhao Q, He Y, Niu J, Debnath AK, *et al.* Theaflavin derivatives in black tea and catechin derivatives in green tea inhibit HIV-1 entry by targeting gp41. *Biochim Biophys Acta* 2005, **1723**: 270-81.
87. Root MJ, Kay MS, Kim PS. Protein design of an HIV-1 entry inhibitor. *Science* 2001, **291**: 884-8.
88. Root MJ, Hamer DH. Targeting therapeutics to an exposed and conserved binding element of the HIV-1 fusion protein. *Proc Natl Acad Sci USA* 2003, **100**: 5016-21.

89. Markosyan RM, Ma X, Lu M, Cohen FS, Melikyan GB. The mechanism of inhibition of HIV-1 env-mediated cell-cell fusion by recombinant cores of gp41 ectodomain. *Virology* 2002, **302**: 174-84.
90. Miller, M.D., Geleziunas, R., Bianchi, E., Lennard, S., Hrin, R., Zhang, H., Lu, M., An, Z., Ingallinella, P., Finotto, M., Mattu, M., Finnefrock, A.C., Bramhill, D., Cook, J. Eckert, D.M., *et al.* A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope. *Proc. Natl. Acad. Sci. USA*, 2005, **102**: 14759-14764.
91. Eckert DM, Kim PS. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc Natl Acad Sci USA* 2001, **98**: 11187-92.
92. Suzuki K, Hiroaki H, Kohda D, Tanaka T. An isoleucine zipper peptide forms a native-like triple stranded coiled coil in solution. *Protein Eng* 1998, **11**: 1051-5.
93. Louis JM, Bewley CA, Clore GM. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J Biol Chem* 2001, **276**: 29485-9.
94. Louis JM, Nesheiwat I, Chang L, Clore GM, Bewley CA. Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope-mediated cell fusion. *J Biol Chem* 2003, **278**: 20278-85.
95. Bianchi E., Finotto M., Ingallinella P., Hrin R., Carella A.V., Hou X.S., Schleif W.A., Miller M., Geleziunas R., Pessi A. (2005) Covalent stabilization of coiled coil of the HIV gp41 N region yields extremely potent and broad inhibitors of viral infection. *Proc. Natl. Acad. Sci. USA* 102: 12903-12908.
96. Gochin M., Kiplin Guy R., Case M. A. A Metallopeptide assembly of the HIV-1 gp41 coiled coil is an idea receptor in fluorescence detection of ligand binding. *Angew. Chem. Int. Ed.*, 2003, **42**: 5325-5328.
97. Gochin, M., Savage, R., Hinckley, S. and Cai, L. (2006) A fluorescence assay for rapid detection of ligand binding affinity to HIV-1 gp41. *Biol. Chem.*, 2006, **387**: 477-483.
98. Cai L., Gochin M., A novel fluorescence intensity screening assay identifies new low-molecular-weight inhibitors of the gp41 coiled-coil domain of human immunodeficiency virus type 1. *Antimicrob. Agents Chemother.*, 2007, **51**: 2388-2395.
99. DeGrado W.F., Wasserman Z. R., Lear J. D., Protein design, a minimalist approach. *Science*, 1989, **243**: 622-628
100. DeGrado WF, Summa CM, Pavone V, Natri F, Lombardi A. De novo design and structural characterization of proteins and metalloproteins. *Annu Rev Biochem.* 1999, **68**:779-819.
101. Kaiser ET, Kezdy FJ. Secondary structures of proteins and peptides in amphiphilic environments. *Proc Natl Acad Sci U S A.* 1983, **80**(4):1137-43.
102. Kaiser ET, Kezdy FJ. Amphiphilic secondary structure: design of peptide hormones. *Science.* 1984, **223**(4633):249-55.

103. Taylor JW, Kaiser ET. The structural characterization of beta-endorphin and related peptide hormones and neurotransmitters. *Pharmacol Rev.* 1986, **38**(4):291-319.
104. Garipey J, Sykes BD, Hodges RS. Lanthanide-induced peptide folding: variations in lanthanide affinity and induced peptide conformation. *Biochemistry.* 1983, **22**(8):1765-72.
105. Taylor J.W., The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers (Peptide Science)*, 2002, **66**: 49-75.
106. Handel TM, Williams SA, DeGrado WF. Metal ion-dependent modulation of the dynamics of a designed protein. *Science.* 1993, **261**(5123):879-85.
107. Kellis JTJr , Nyberg K, Sali D, Fersht AR. Contribution of hydrophobic interactions to protein stability. 1988 *Nature* **333**: 784– 86
108. Matthews BW. Studies on protein stability with T4 lysozyme. 1995 .*Adv. Protein Chem.* **46**: 249– 78
109. Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HX, et al. Protein design: a hierarchic approach. *Science* 1995. **270**: 935– 41.
110. Aurora R, Rose GD. Helix capping. *Protein Sci.* 1998;**7**: 21– 38.
111. Lumb KJ, Kim PS. Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper. *Science* 1995.**268**: 436– 39.
112. Kenar KT, Garcia-Moreno B, Freire E. A calorimetric characterization of the salt dependence of the stability of the GCN4 leucine zipper. *Protein Sci.* 1995. **4**: 1934– 38
113. Yu Y, Monera OD, Hodges RS, Privalov PL. Ion Pairs Significantly Stabilize Coiled-coils in the Absence of Electrolyte. *J. Mol. Biol.* 1996. **255**: 367– 72
114. Dunbrack RLJr , Karplus M. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nat. Struct. Biol.* 1994.**1**: 334– 40.
115. D'Aquino JA, Gomez J, Hilser VJ, Lee KH, Amzel LM, Freire E. The magnitude of the backbone conformational entropy change in protein folding.. *Proteins* 1996, **25**: 143– 56.
116. Betz SF, Bryson JW, DeGrado WF. Native-like and structurally characterized designed alpha-helical bundles. *Curr Opin Struct Biol.* 1995, **5**(4):457-63.
117. Cordes MHJ, Davidson AR, Sauer RT. Sequence space, folding and protein design. *Curr. Opin. Struct. Biol.* 1996.**6**: 3– 10.
118. Regan L., DeGrado W.F., Characterization of a helical protein designed from first principles. *Science* 1988, **241**: 976-978.
119. Hill RB, Raleigh DP, Lombardi A, DeGrado WF. De novo design of helical bundles as models for understanding protein folding and function. *Acc Chem Res.* 2000, **33**(11):745-54.

120. Betz SF, Bryson JW, DeGrado WF. Native-like and structurally characterized designed alpha-helical bundles. *Curr Opin Struct Biol.* 1995, **5**(4):457-63.
121. Bryson JW, Betz SF, Lu HS, Suich DJ, Zhou HX, et al. Protein design: a hierarchic approach. *Science* 1995. **270**: 935– 41.
122. Beasley JR, Hecht MH. Protein design: the choice of de novo sequences. *J. Biol. Chem.* 1997. **272**: 2031– 34.
123. Richardson J, Richardson DC, Tweedy NB, Gernert KM, Quinn TP, et al. Looking at proteins: representations, folding, packing, and design. Biophysical Society National Lecture, 1992. *Biophys. J.* 1992. **63**: 1186– 209.
124. Baltzer L. Functionalization of designed folded polypeptides. *Curr. Opin. Struct. Biol.* 1998. **8**: 466– 70.
125. Zhu BY, Zhou NE, Kay CM, Hodges RS. Packing and hydrophobicity effects on protein folding and stability: effects of beta-branched amino acids, valine and isoleucine, on the formation and stability of two-stranded alpha-helical coiled coils/leucine zippers. *Protein Sci.* 1993, **2**(3):383-94.
126. Gonzalez LJr , Brown RA, Richardson D, Alber T. Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. *Nat. Struct. Biol.* 1996. **3**: 1002– 9.
127. Harbury PB, Kim PS, Alber T. Crystal structure of an isoleucine-zipper trimer. *Nature* 1994. **371**: 80– 83.
128. Gonzalez LJr , Plecs JJ, Alber T. An engineered allosteric switch in leucine-zipper oligomerization. *Nat. Struct. Biol.* 1996. **3**: 510– 15.
129. Gonzalez LJr , Woolfson DN, Alber T. Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* 1996. **3**: 1011– 18.
130. Lovejoy B, Choe S, Cascio D, McRorie DK, DeGrado WF, Eisenberg D. Crystal structure of a synthetic triple-stranded alpha-helical bundle. *Science* 1993. **259**: 1288– 93.
131. Harbury PB, Zhang T, Kim PS, Alber T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 1993. **262**: 1401– 7.
132. Ellenberger, T., Brandl, C., Struhl, K., and Harrison, S. The GCN4 Basic Region Leucine Zipper Binds DNA as a Dimer of Uninterrupted Alpha Helices: Crystal Structure of the Protein-DNA Complex. *Cell.* 1992. **71**: 1223-1237.
133. DeGrado WF, Gratkowski H, Lear JD. How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Sci.* 2003, **12**(4):647-65.
134. Mutter M., Vuilleumier S. A Chemical approach to protein design - template-assembled synthetic proteins (TASP). *Angew. Chem. Int. Ed. Engl.* 1989, **28**:535-554.

135. Tuchscherer G, Scheibler L, Dumy P, Mutter M. Protein design: on the threshold of functional properties. *Biopolymers*. 1998, **47**(1):63-73.
136. Pawlak M, Meseth U, Dhanapal B, Mutter M, Vogel H. Template-assembled melittin: structural and functional characterization of a designed, synthetic channel-forming protein. *Protein Sci*. 1994, **3**(10):1788-805.
137. Mutter M., Tuchscherer G.G., Miller C., Altmann K., Carey R. I., Wyss D. F., Labhardt A. M., Rivier J. E., Template- Assembled Synthetic Proteins with Four-Helix-Bundle Topology. Total Chemical Synthesis and Conformational Studies. *J. Am. Chem. SOC*. 1992, **114**: 1463-1470.
138. Mutter M., Dumy P., Garrouste P., Lehmann C., Mathieu M., Peggion C., Peluso S., Razaname A., Tuchscherer G. Template Assembled Synthetic Proteins (TASP) as Functional Mimetics of Proteins. *Angew. Chem. Int. Ed. Engl*. 1996, **35**: 1482-1485.
139. Obataya I., Sakamoto S., Ueno A., Mihara H. Design and synthesis of 3 α -helix peptides forming a cavity for a fluorescent ligand. *Biopolymers*. 2001, **59**: 65-71.
140. Rau HK, DeJonge N, Haehnel W. Modular synthesis of de novo-designed metalloproteins for light-induced electron transfer. *Proc Natl Acad Sci U S A*. 1998, **95**(20):11526-31.
141. Rau HK, DeJonge N, Haehnel W. Combinatorial Synthesis of Four-Helix Bundle Hemoproteins for Tuning of Cofactor Properties. *Angew Chem Int Ed Engl*. 2000, **39**(1):250-253.
142. Sasaki T., Kaiser E. T. Synthesis and structural stability of helichrome as an artificial hemeoproteins. *Biopolymers*. 1990, **29**(1):79-88.
143. Akerfeldt K.S., Kim R.M., Camac D., Groves J.T., Lear J.D., DeGrado W.F. Tetraphillin: a four-helix proton channel built on a tetraphenylporphyrin framework. *J. Am. Chem. Soc*. 1992, **114**: 9656-9657.
144. Tahmassebi DC, Sasaki T. Synthesis of a Three-Helix Bundle Protein by Reductive Amination. *J Org Chem*. 1998, **63**(3):728-731.
145. Goodman, M., Feng, Y., Melacini, G., Taulane, J. A template-induced incipient collagen-like triple-helical structure. *J. Am. Chem. Soc*. 1996, **118**: 5156-5157.
146. Kwak, J., De Capua, A., Locardi, E., Goodman, M. (2002) TREN (tris (2-aminoethyl) amine): an effective scaffold for the assembly of triple helical collagen mimetic structures. *J. Am. Chem. Soc*. **124**: 14085-14091.
147. Lieberman M., Sasaki T., Iron(II) organizes a synthetic peptide into three-helix bundles. *J. Am. Chem. Soc*. 1991, **113**: 1470-1471.
148. Ghadiri, M. R., Soares, C., Choi, C. Design of an artificial four-helix bundle metalloprotein via a novel ruthenium (II)-assisted self-assembly process. *J. Am. Chem. Soc*. 1992, **114**: 4000-4002.
149. Jensen KJ, Brask J. Carbohydrates as templates for control of distance-geometry in de novo-designed proteins. *Cell Mol Life Sci*. 2002, **59**(5):859-69.

150. Seo ES, Sherman JC. Analysis of peptide design in four-, five- and six-helix bundle TASP molecules. *Biopolymers*. 2007 Jun 6; [Epub ahead of print]
151. Seo ES, Scott WR, Straus SK, Sherman JC. Optimal attachment position and linker length promote native-like character of cavitand-based template-assembled synthetic proteins (TASPs). *Chemistry*. 2007, **13**(13):3596-605.

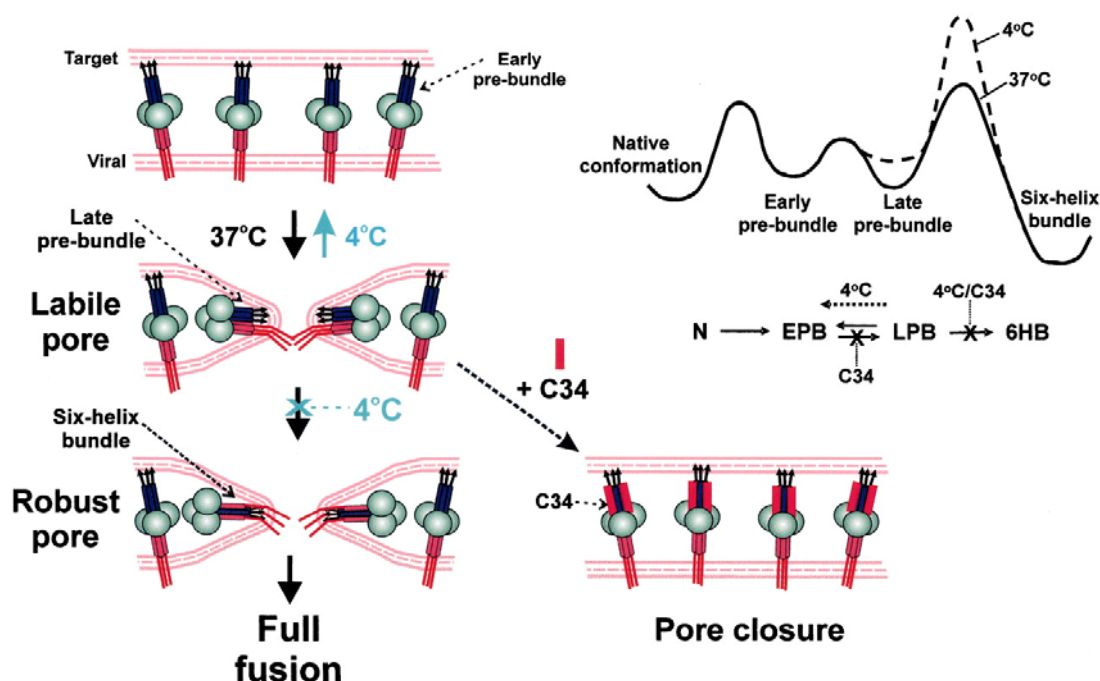


Figure 2.1 A proposed mechanism for the conformational changes of gp41 in pore formation and stabilization (reprint from (41)). Left panels: Conversion of Env in early prebundles (top) at a fusion site into late prebundles (middle) creates a labile pore, and subsequent folding into the final six-helix bundle configuration establishes the robust pore (bottom). Early prebundles are shown as an extended conformation with N- and C-terminal helices well separated from each other; late prebundles are depicted with one C-terminal helix bound into a groove of the triple-stranded coiled-coil. Arresting a labile pore by 4°C leads to reduction in pore size and eventual closure (transition from middle to top). The addition of 6HB blockers (C34, red rectangle) to the labile pore results in rapid pore closure (dashed arrow). Late prebundles stabilize the pore, but only when the six-helix bundle is formed is the pore permanently stabilized against closure. Top right: The proposed energy profile of gp41 during the fusion reaction at 37°C (solid line) and at 4°C (dashed line).

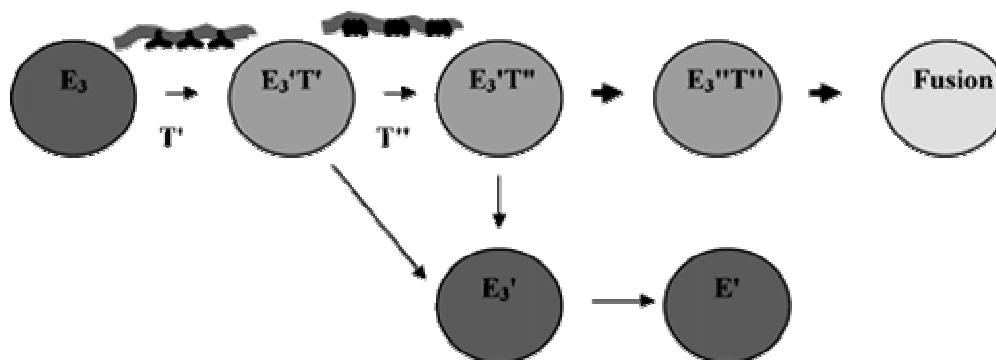


Figure 2.2 Model of intermediates during HIV-1 Env-mediated fusion (reprint from (42)).

The interactions of HIV-1 Env trimers (E_3) with CD4 clusters (T') induce conformational changes that result in the exposure of the binding site on gp120 that can engage CXCR4 and the exposure of the gp41 prehairpin trimer (E_3') in an Env-target membrane complex ($E_3'T'$). This is followed by engagement of CXCR4 clusters (T'') to form an $E_3'T''$ complex that results in 6-helix bundle formation ($E_3''T''$) and fusion. In a parallel reaction, Env-target complexes dissociate to expose triggered Env (E_3') on the surface, which is susceptible to NC-1 MAb binding. The triggered Env trimer (E_3') can then dissociate to gp41 monomers, a process that is enhanced by entry inhibitors such as N36^{Mut(e,g)}.

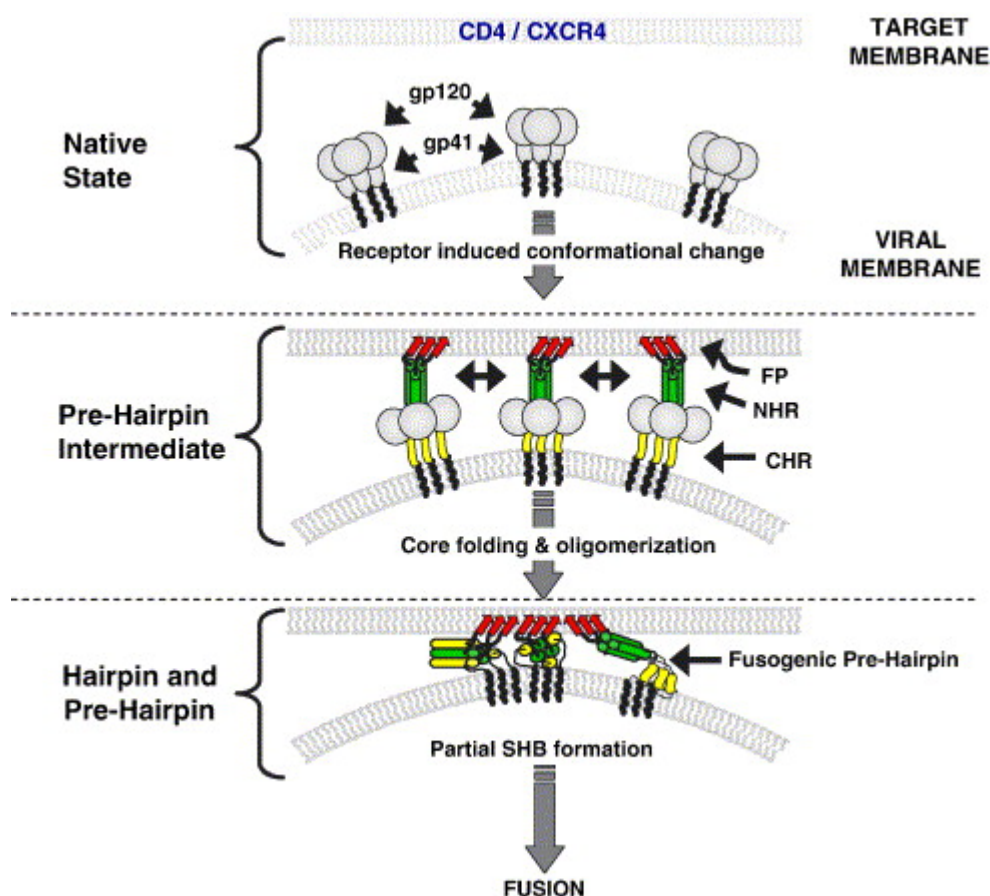


Figure 2.3 The HIV fusion model proposed by Y. Shai (reprint from (45)). Gp41 is held in a metastable conformation by gp120 in the native state. The binding of HIV-1 Env with CD4 induces conformational changes that result in the exposure of the binding site of CXCR4 on gp120 and the exposure of the gp41 prehairpin intermediate trimer. After binding to CXCR4 clusters complex, six-helix bundle was formed to a fraction of the complex. The remaining pre-hairpin conformation exposed on the surface of the target membrane and susceptible to NC-1 MAb binding, can also initiate pore formation and fusion. Inhibition of fusion by synthetic CHR can be achieved by inhibiting the formation of the SHB, which prevents membrane apposition, and/or alternatively by directly binding the NHR and inhibiting its fusogenic activity.



Figure 2.4 Original (in boxes) and modified peptides derived from the gp41 N-helix and C-helix. In the modified N- and C-peptides, the residues that are identical to those in the HIV-1HXB2 gp41 sequence were marked in black, while the residues that are different from those of the gp41 sequence were marked in red. Characters “a” to “g” in lower case corresponds to positions in helical wheel representation.

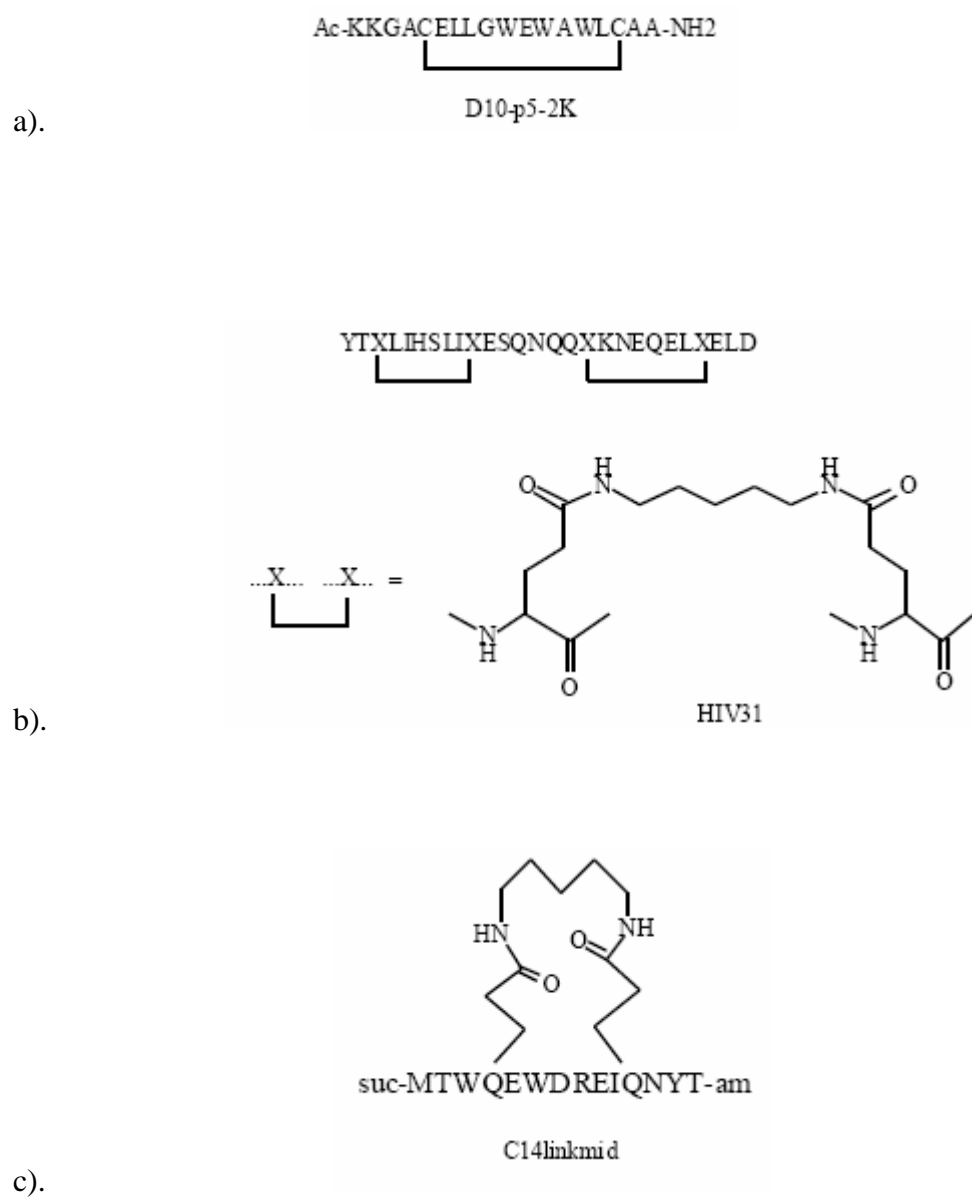


Figure 2.5 Modified short peptides with anti-HIV activities. a) pocket-binding D-amino acid peptide (D-peptide) D10-p5-2K, b) HIV31, and c) C14linkmid.

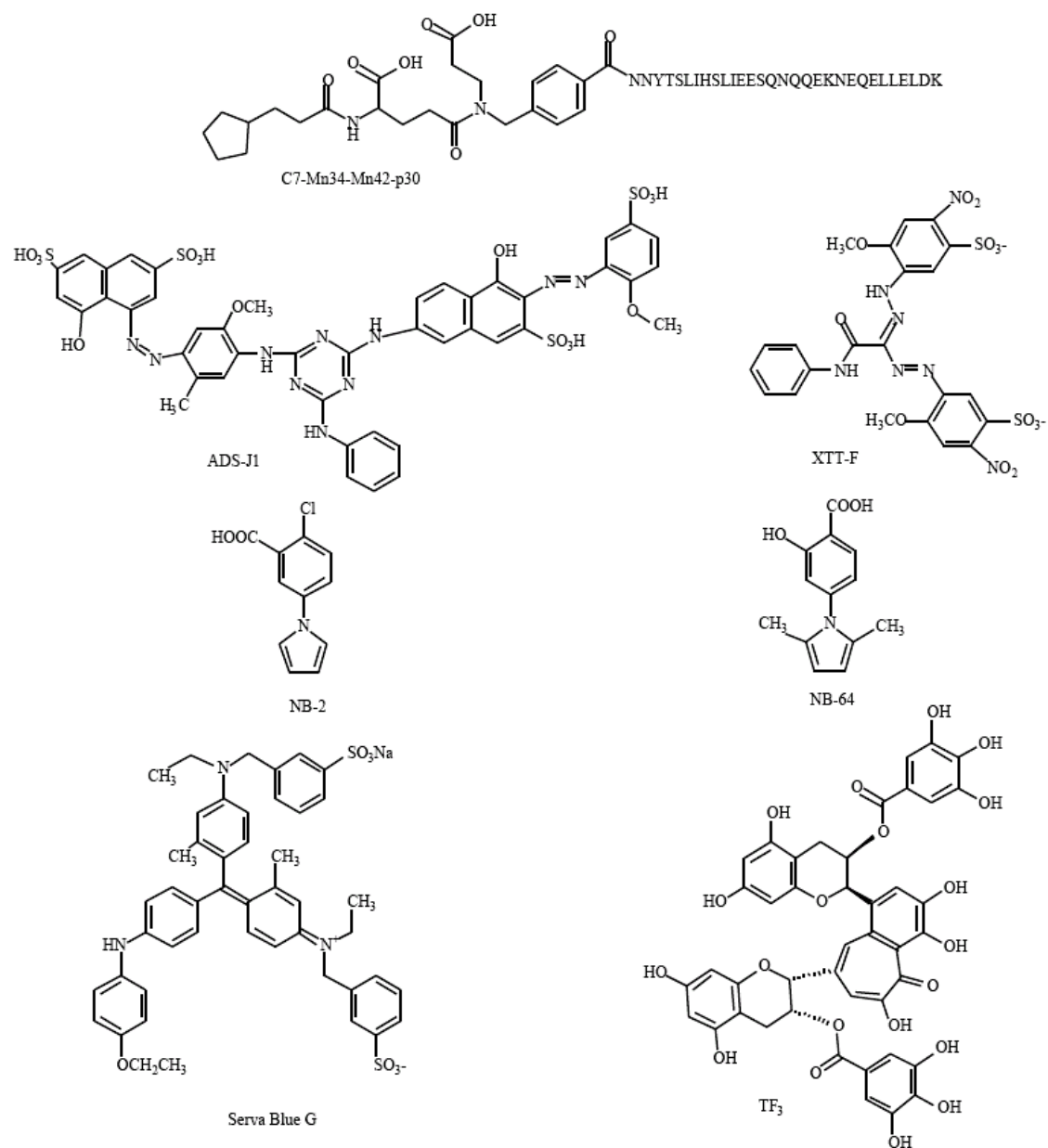


Figure 2.6 Several examples of anti-HIV compounds targeting gp41.

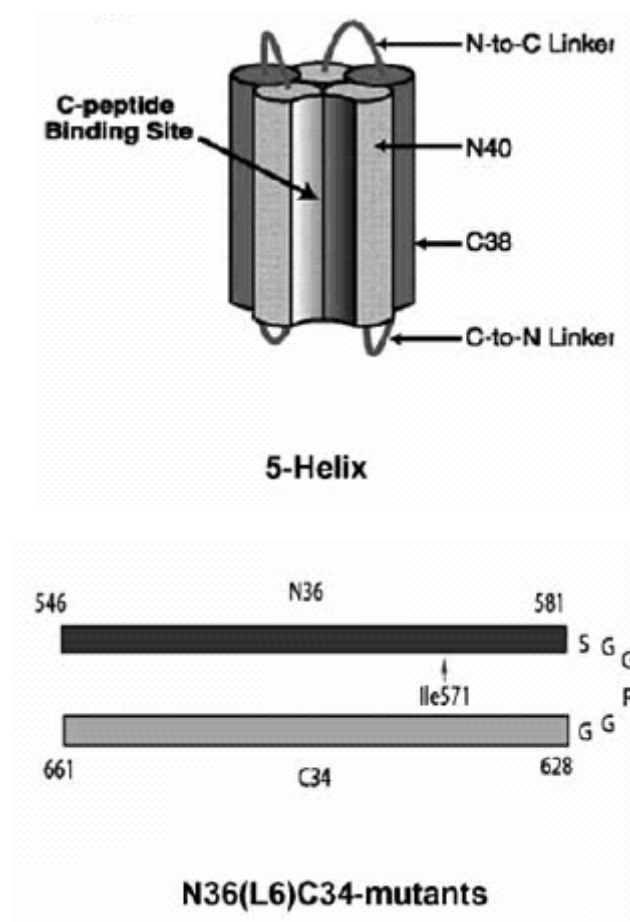


Figure 2.7 Schematic representation of 5-helix and N36(L6)C34 mutants.

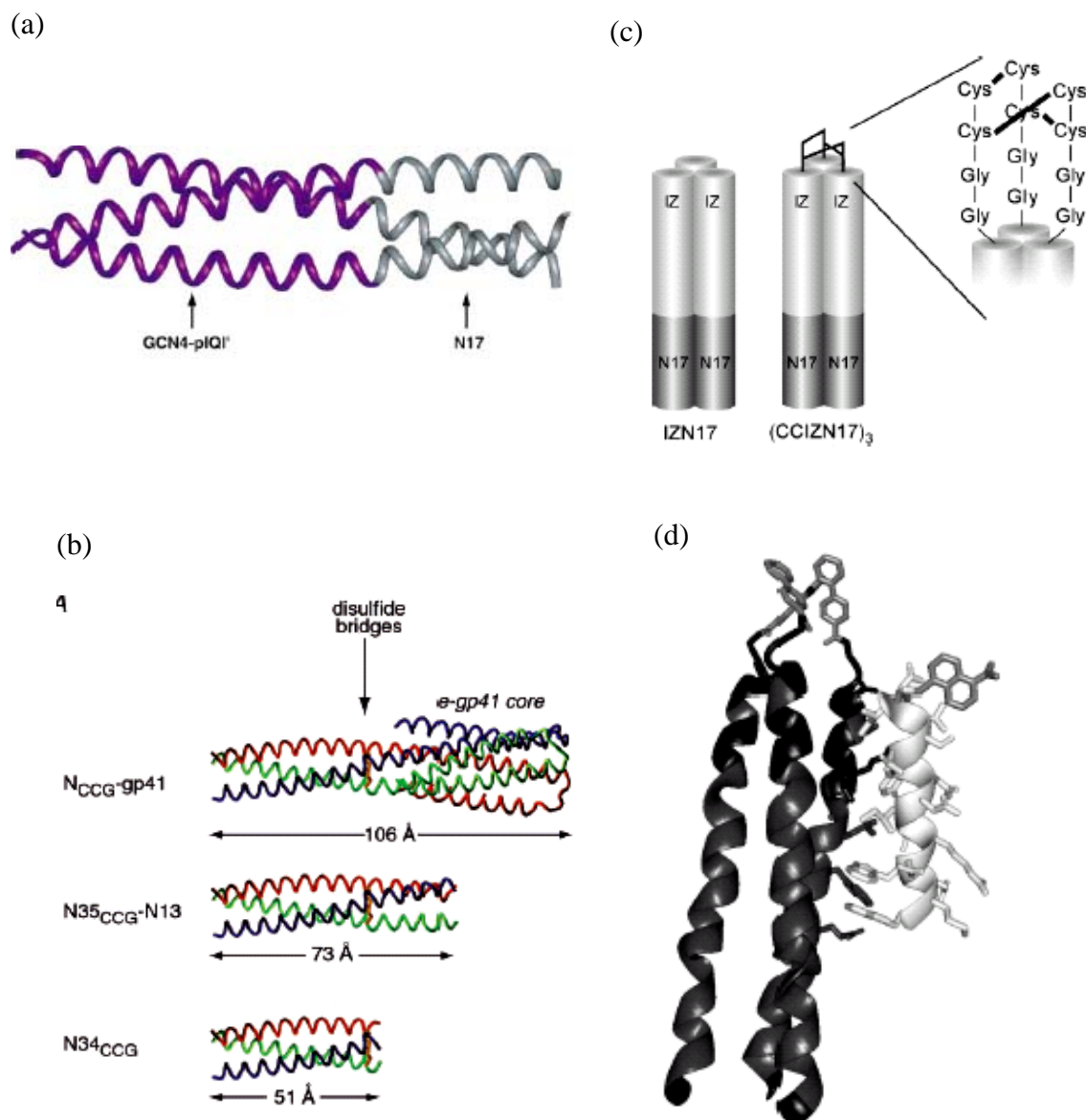


Figure 2.8 Schematic representations of current approaches to construct N-peptide three helix bundle mimics. (a) IQN17; (b) Nccg-gp41, N35ccg-N13 and N34ccg; (c) (CCIZN17)₃; and (d) Fe-env2.1 – dansylated C-peptide interaction.

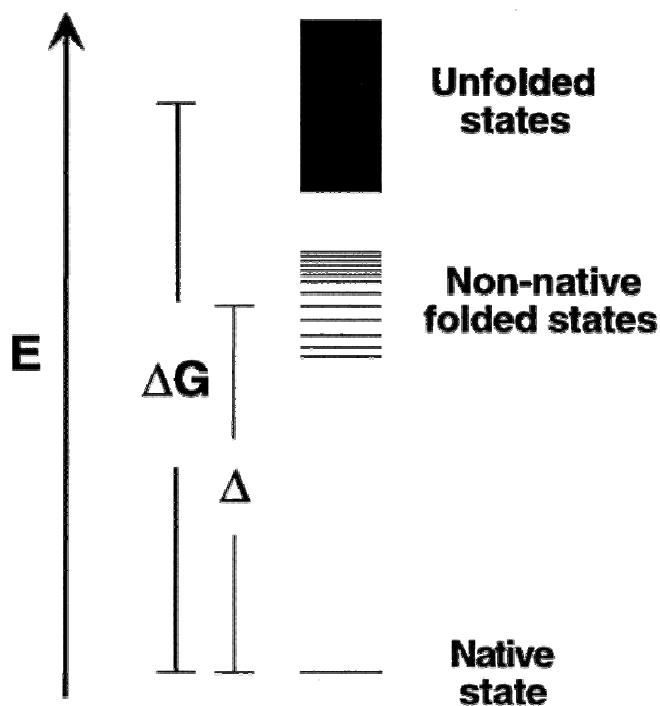


Figure 2.9 Hypothetical free energy diagram for a protein. The native state is the lowest energy state. The unfolded states are a densely populated, nearly isoenergetic ensemble. Between these extremes are non-native folded states, often referred to as the molten globule ensemble. The population of each state is dictated by the Boltzmann distribution. For native protein structure the free energy gap, Δ , must be large enough to significantly populate a distinct native state.

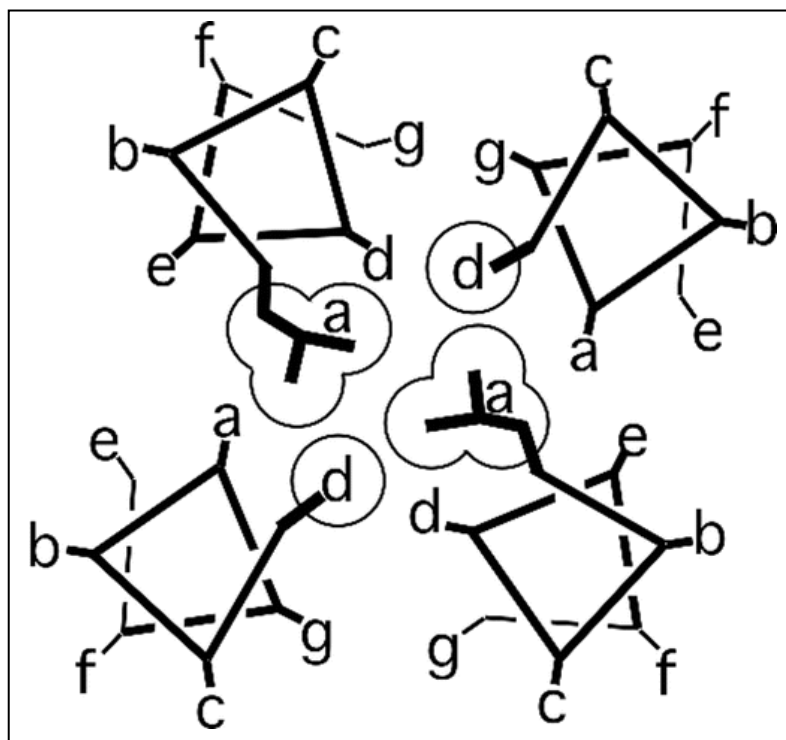


Figure 2.10 Schematic of the heptad repeat of an antiparallel coiled coil (reprinted from (120)). The letters *a* through *g* corresponds to a seven-residue repeat (heptad). The *a* and *d* residues are hydrophobic residues leucine and alanine, respectively, that are well packed in the interior of the coiled coil.

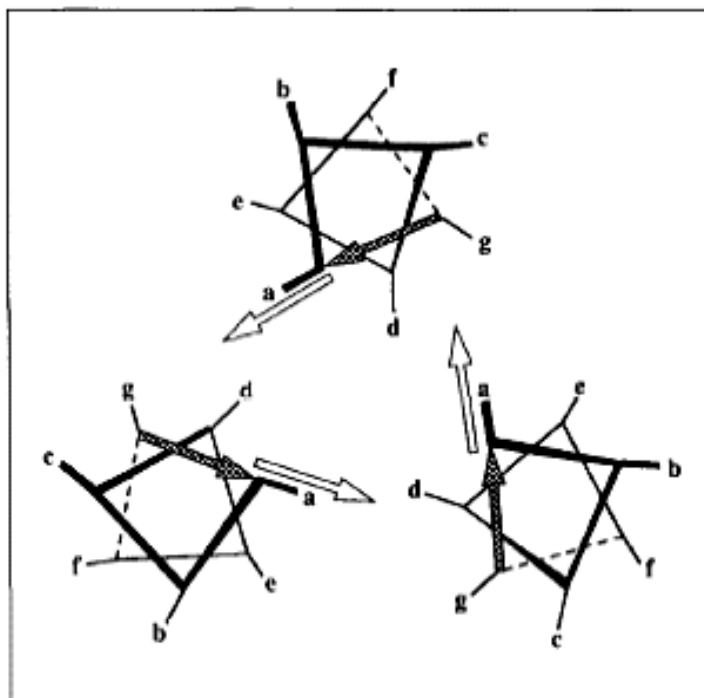


Figure 2.11 Acute packing in the three-stranded coiled coil GCN4-pII (reprinted from (120)). Open arrows describe the C_{α} - C_{β} bond vector of the 'a' residues; cross-hatched arrows describe the C_{α} - C_{α} vectors of the residues against which they pack. The packing at the 'd' positions is similar.

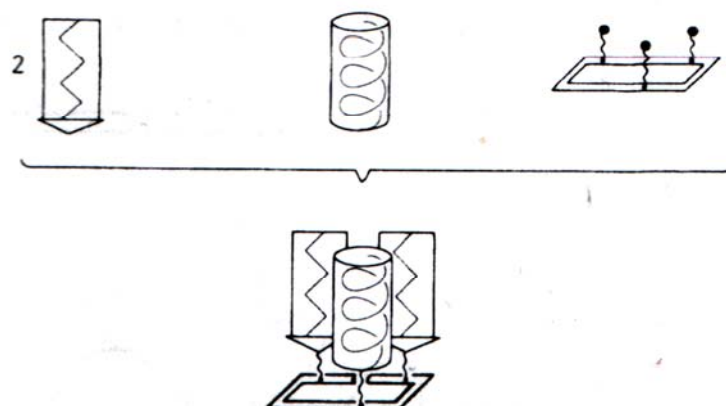


Figure 2.12 Schematic representation of template-assembled synthetic proteins (TASP). In this approach, amphiphilic secondary-structure-forming peptide blocks (α -helices, β -sheets) are grafted onto an adequate carrier molecule (template). In above example, two β -sheet blocks and an α -helix block are assembled on a trifunctional template molecule to a $\beta\alpha\beta$ -like folding unit.

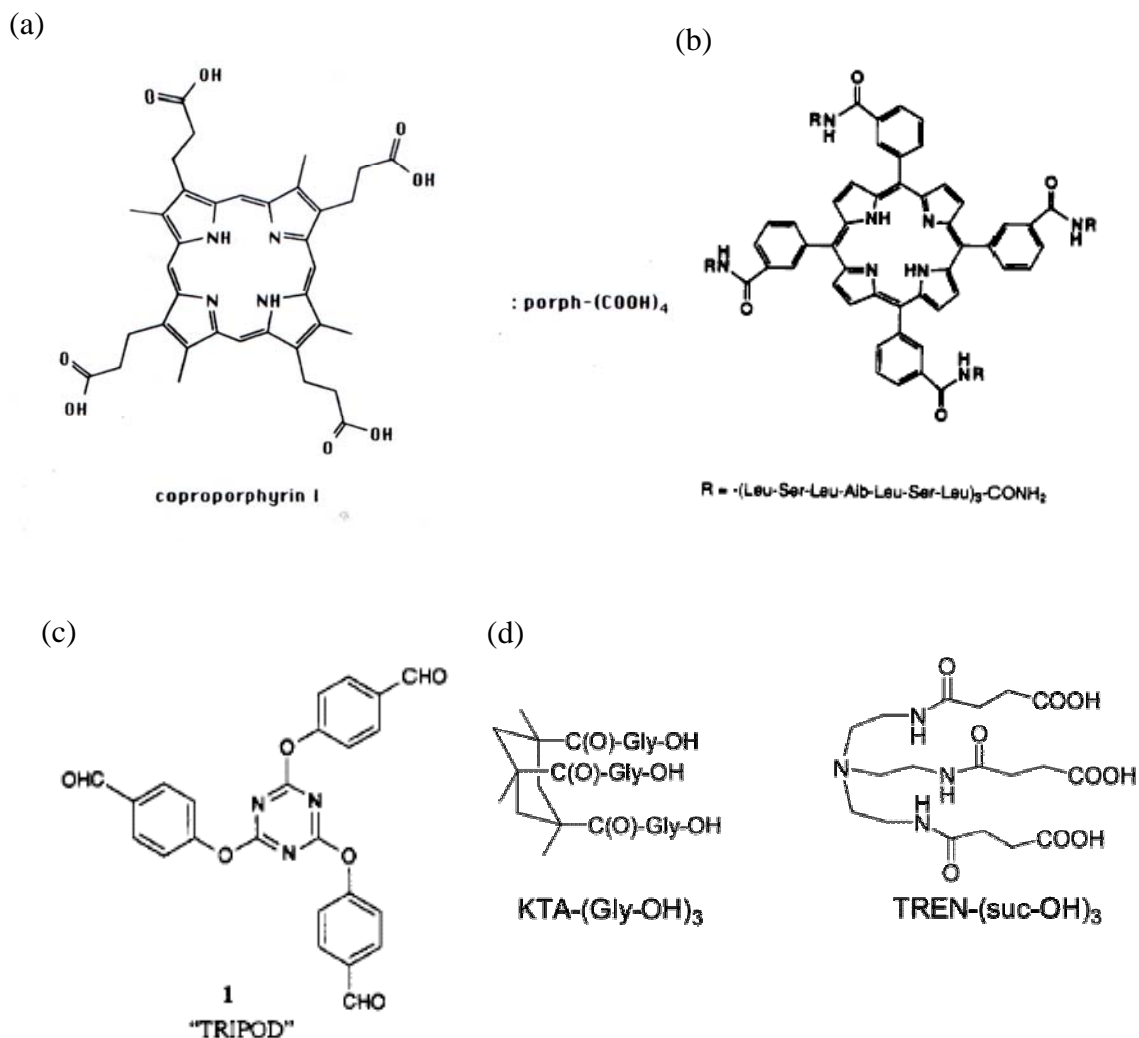


Figure 2.14 Examples of branched organic molecules served as templates in TASP. (a) Coproporphyrin molecule used to construct four-helix bundle hemeprotein model structure designed by Sasaki and Kaiser's group (142); (b) Tetraphenylporphyrin template used to construct four-helix bundle mimic of ion channel protein designed by DeGrado's group (143); (c) "TRIPOD" template used to prepare a three-helix bundle model protein designed by Sasaki's group (144); (d) Small molecule templates used to construct collagen mimetic structures designed by Goodman's group (146).

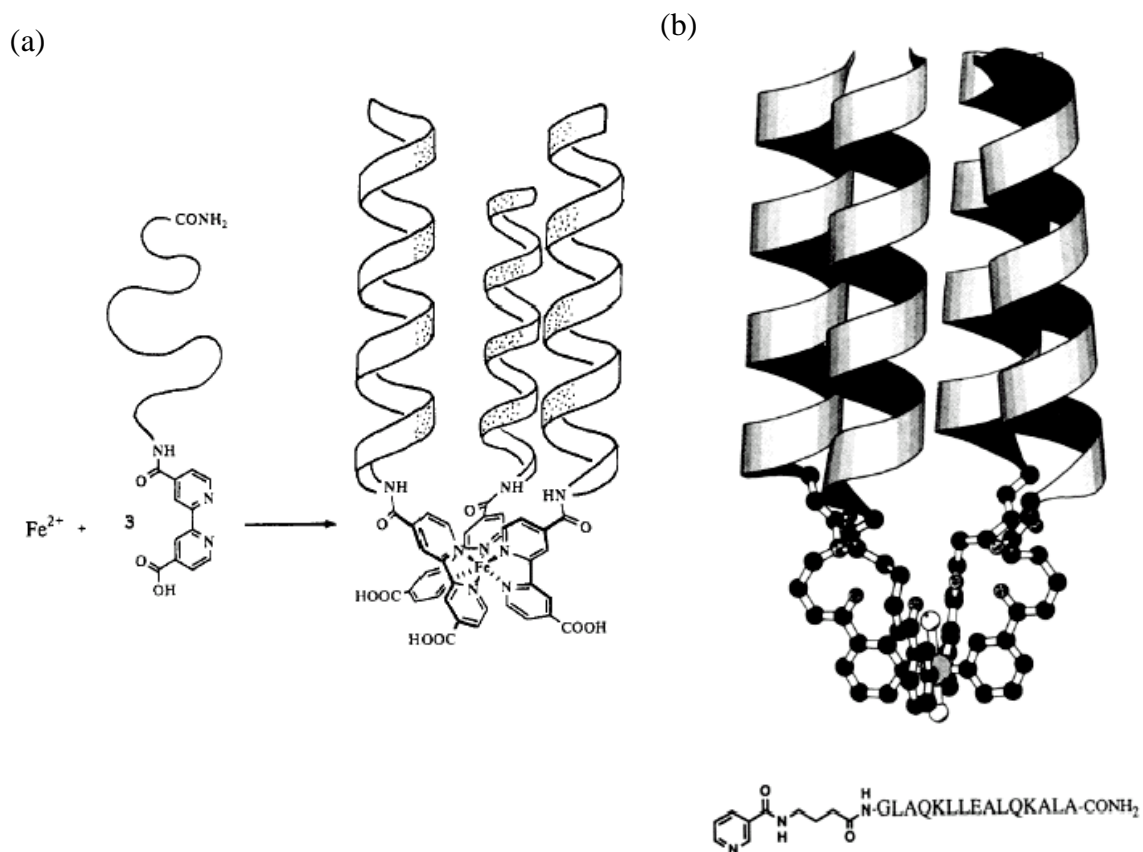
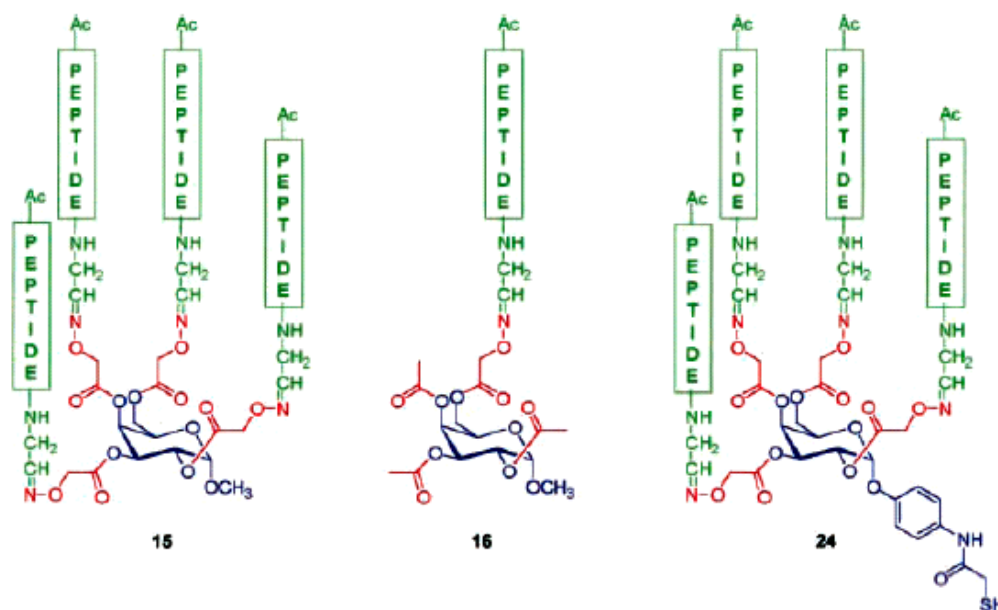


Figure 2.15 Examples of TASP formed via metal-assisted self-assembly process. (a) a three-helix bundle structure using Iron (II) induced complex as directing template (147). (b) model of an artificial four-helix bundle metalloprotein using Ruthenium(II) complex as template (148).

(a)



(b)

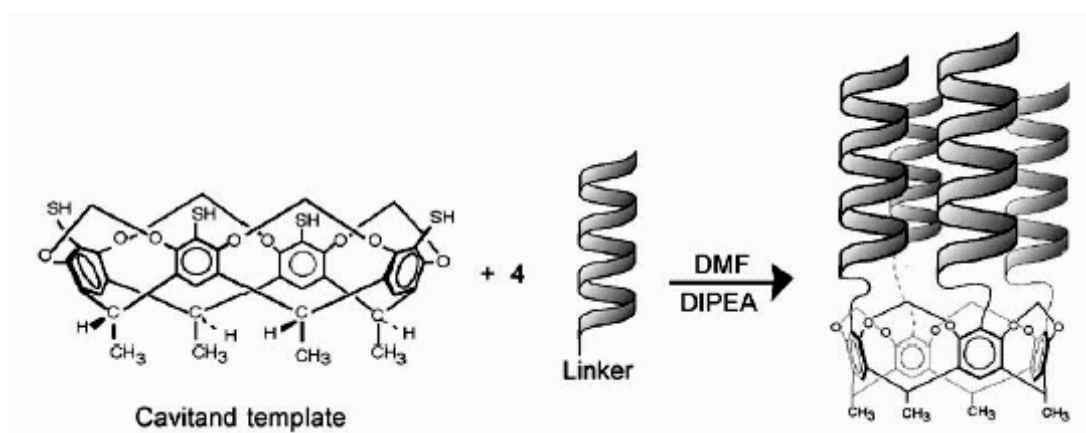


Figure 2.16 Other templates used in construction of TASP. (a) Jensen's group used carbohydrates as templates in designed proteins (149). As shown, are the structures of 4-helix bundle, single helix, and 4-helix thiol-functionalized carboproteins. (b) Sherman's group use cavitand as templates to construct helix bundle structures (150-151).

III. Experimental

In this chapter, all experiments are described in detail, including the optimized synthesis procedure, characterization methods and other instrumentation conditions applied during measurements.

A. Materials and Instruments

Starting materials for synthesis of templates: tris-(2-aminoethyl)-amine (TREN), with a purity of 96%; *cis,cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid (Kemp's Triacid, KTA), with a purity of 99%; *tert*-butyl N-(2-aminoethyl) carbamate, with a purity of 98%; and bromoacetic acid, with a purity of 99%, were all purchased from Aldrich.

Reagents for synthesis and purification of templates and model structures: tris-[hydroxymethyl]-aminomethane (TRIZMA BASE), reagent grade, with a purity of 99.9% purity, was purchased from Sigma. Silical gel for flash chromatography was purchased from J.T. Baker Chemical Co. 1-(3-Dimethylaminopropyl)-3-ethyl carbodiimide hydrochloride (EDC) for coupling reactions, with a purity of 98+%, was purchased from Aldrich. Citric acid and NaCl for work-up of reaction mixture were from Aldrich, and NaHCO₃ was from Fisher.

Reagents for sample preparation for CD analysis: Phosphoric acid (85 wt.%), with a purity of 99.999% and anhydrous KF were purchased from Aldrich. NaOH and HCl were purchased from Fisher. Guanidine hydrochloride, either as white solid or 8M solution, was purchased from Sigma. Urea was purchased from Bio-Rad.

Deuterated solvents for NMR were purchased from Acros and Aldrich.

Resin: 4-(2', 4'-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetamido-norleucyl-MBHA resin (RINK amide MBHA resin) for standard solid-phase Fmoc peptide synthesis was purchased from Novabiochem.

Protected L-amino acids: Boc-Lys(Fmoc)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Glu(O^tBu)-OH, Fmoc-Asn(Trt)-OH, and Fmoc-Thr(^tBu)-OH were purchased from Novabiochem; Fmoc-Trp(Boc)-OH, Fmoc-Ala-OH, Fmoc-Leu-OH, and Fmoc-Ser(^tBu)-OH were purchased from AnaSpec. All other L-protected amino acids were purchased from Advanced ChemTech.

Coupling reagents: PyBOP was purchased from Novabiochem; HOBt and DCC were purchased from Advanced ChemTech. DIEA was purchased from Aldrich and was redistilled from ninhydrin before use (bp:126-127 °C).

Reagents for peptide deprotection and cleavage: Piperidine was purchased from Aldrich; TFA was purchased from Halocarbon and was redistilled before use (bp: 71-72 °C); TIS was purchased from Aldrich. Diethyl ether, acetic acid (AcOH) and HCl were from Fisher.

Reagents for peptide modification: DTT and iodoacetamide were from Sigma.

Reagents for the Kaiser test: Ninhydrin, KCN and phenol were from Aldrich.

Solvents: DMF, DCM, iPrOH, MeOH, EtOH, acetonitrile (HPLC grade) and hexane were purchased from Aldrich. EtOAc, Et₂O and Ac₂O were purchased from Fisher. Deionized ultra filtered (DIUF) water was purchased from Fisher.

Instruments: Reverse-phase HPLC for analytical and semi-preparative purposes was carried out on a RAININ HPXL SOLVENT DELIVERY SYSTEM equipped

with a DYNAMAX UV-D detector. LC/ESI-MS was carried out on a ThermoQuest, FINNIGAN LCQ Duo ESI-MS equipped with ThermSpectraSYSTEM (SP) P4000 pumps, an SP AS3000 autosampler and an SP UV2000 dual wavelength UV detector. Circular Dichroism studies were performed on AVIV model 215 spectropolarimeter. ^1H NMR spectra were collected on a Varian Unity spectrometer at 400MHz. UV-Vis spectra or data were collected either on Beckman DU-65, or Beckman DU-7400, or Varian Cary Model 50 spectrophotometers, with tested and proved measurement consistency between each instrument.

General methods of LC/MS analysis and HPLC separation: For the purpose of analysis, a Waters XTerra MSC₁₈ 5 μM , 4.6 x 50 mm column was used. For the purpose of separation, Dynamax C18-300A, 10 x 250 mm (for template molecules) or Waters XTerra Prep MSC₁₈ 5 μM , 10 x 100 mm column were used. Unless otherwise mentioned or specially developed, the buffer system used for LC/MS and HPLC (reverse phase) was as follows: Buffer A, 0.05%TFA in H₂O, Buffer B, 0.05% TFA in CH₃CN. Eluting absorbance was monitored at 220nm and 280nm.

Exact molecular masses of peptides and theoretical ESI-MS monoisotopic mass series were calculated with "Molecular mass calculator for peptides v1.2" software at the website http://thor.prohosting.com/~jef_r/cgi-bin/molpep.cgi

Exact molecular masses and average masses of template molecules and their theoretical ESI-MS mass series were calculated either with "Mass Spec Software

Tools” provided at website <http://www.sisweb.com/mstools.htm> or “Molecular Mass Calculator v2.0” at website <http://library.med.utah.edu/masspec/mole.htm>.

CD data analysis software used: Data conversion from AVIV 215 file format to AVIV 62 file format was performed on CDSO (operation control software of AVIV 215); spectrum smoothing, subtracting, and data transformations were performed on AVIV PLOT; spectrum overlays were performed on ORIGIN 6.0 PROFESSIONAL; linear and non-linear regression of melting-curves and binding curves were performed on SIGMAPLOT 2001.

B. Synthesis Methods

B.1. Synthesis of template molecules

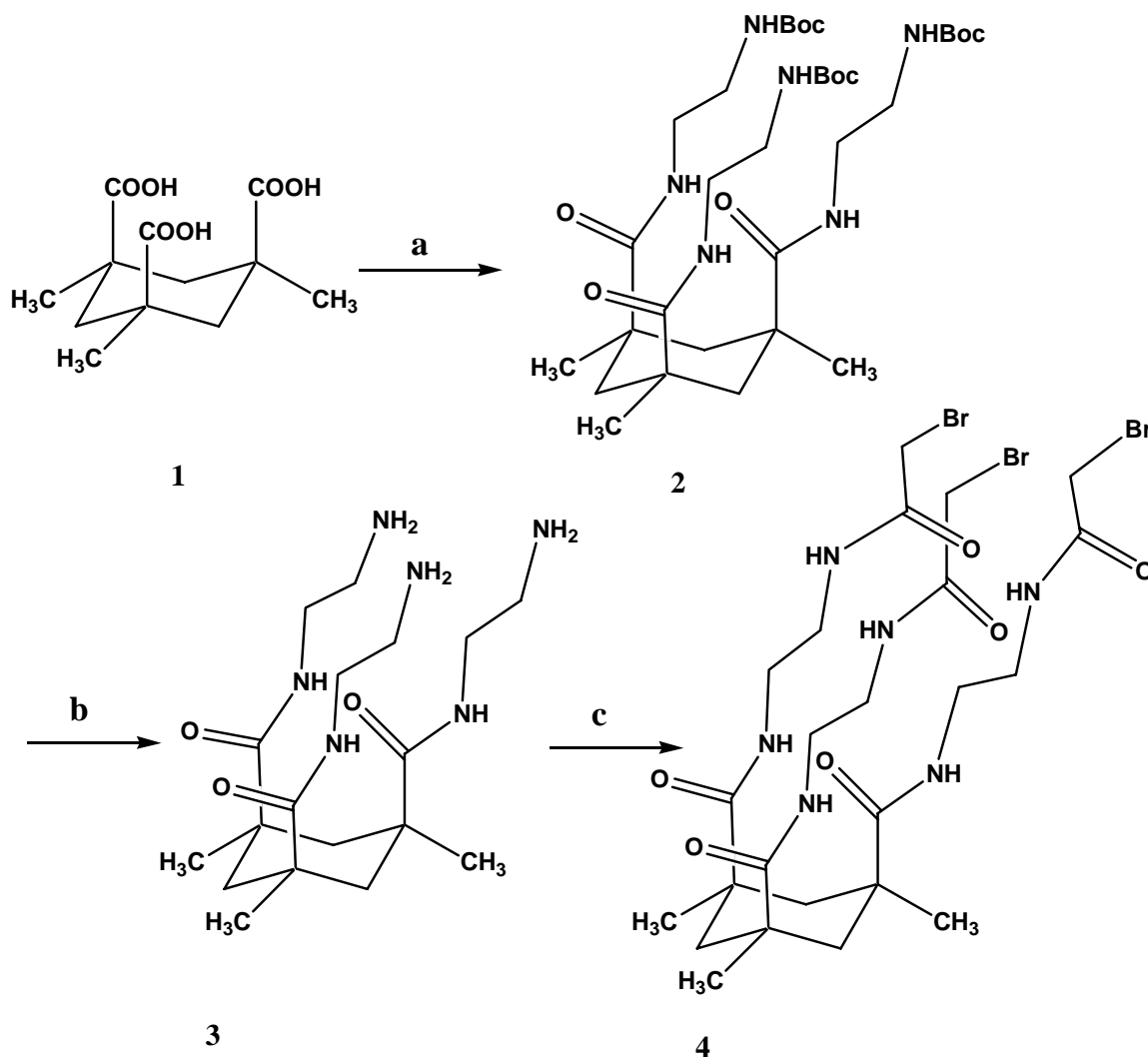
a. KTA-Br

Starting from KTA (**1**), a three-step synthesis was carried out to obtain the KTA-Br (**4**) template molecule (Scheme 3.1.). Each reaction has been optimized in terms of reaction time, reaction solvent and coupling reagent. The detailed procedure is as follows.

Kemp's triacid (KTA, **1**, 103 mg, 0.4 mmol) and HOBt (245 mg, 1.6 mmol) were dissolved in a minimum volume of DMF (4 mL). After stirring at 0 °C for 10 min., EDC (307 mg, 1.6 mmol) was added to the solution. A small amount of DMF (about 1.5 mL) was added as needed to achieve a clear solution, and then the mixture was stirred at 0 °C for a further 1 h. After this pre-activation step, N-Boc-1,2-diaminoethane (500 mg, 3.12 mmol), DIEA (280 µL, 1.6 mmol) and DMF (1 mL) were added to the mixture. The reaction was stirred for another 40 h and was monitored by LC/MS. The solvent was then evaporated under vacuum. The

oily residue was dissolved in 15 mL ethyl acetate, and washed with citric acid (3 x 15 mL), brine (1 x 15 mL), saturated NaHCO₃ solution (3 x 15 mL) and then brine (1 x 15 mL). After drying the organic layer over anhydrous Na₂SO₄, the ethyl acetate was evaporated under vacuum. The resultant crude product was characterized by LC/MS and indicated about 90% purity (based on UV absorbance at 220 nm), for the major component, **2**, [M + H]⁺ observed = 685.3 (calc. 685.5, Δ= 0.2). From the weight of the crude product, **2**, (0.28 g) and the purity measured by UV, the yield of this reaction was 97%, based on KTA.

Deprotection of the above crude product of, **2**, was achieved using 50% (v/v) TFA/ DCM (8 mL) added to solid **2**. After stirring at RT for 30 min., the reaction mixture was evaporated to dryness under a stream of N₂ gas. The solid residue was washed with diethyl ether (3 x 15 mL). The resulting dry residue was dissolved in 1% AcOH in water (12.5 mL), and then lyophilized. LC/MS analysis indicated about 90% purity (based on UV absorbance at 220nm). The major UV absorbance peak gave the molecular ion expected for the product, **3**, observed at [M + H]⁺ = 385.4 (calc. 385.3, Δ= 0.1). The yield for this reaction was quantitative.



Scheme 3.1 Synthesis of the KTA-Br template. **a.** HOBt, DMF, 0°C 10mins; EDC, DMF, 0°C 1hr; $\text{H}_2\text{N}(\text{CH}_2)_2\text{NHBoc}$, DIEA, 40 h **b.** 50% TFA/DCM, 30 min. **c.** BrCH_2COOH , DCC, DMF, 3 h.

Bromoacetic acid (300 mg, 2.16 mmol, 18 equivalents) was reacted with DCC (213 mg, 1.08 mmol) in DCM (3 mL) to form the symmetric anhydride. After 20 min. stirring at 0°C , the DCU precipitate was removed by gravity filtration and then the solvent was evaporated under vacuum. The solid anhydride residue was dissolved in DMF, and then a solution of **3** (60 mg, 0.12 mmol) in DMF (0.5

mL) was slowly added. The reaction was stirred at RT for another 3 h, and monitored by LC/MS. After evaporating the DMF under vacuum, the residue was precipitated and washed with diethyl ether (3 x 10 mL), giving a white solid (52 mg). LC/MS analysis showed that the major component in the crude product was about 71% pure (based on the UV absorbance at 220 nm) and corresponded to compound **4**, $[M + H]^+$ observed = 745.7 (calc. = 745.1, Δ = 0.7). From the weight of the crude product and the purity measured by UV, the yield of this reaction was about 41%, based on **3**. The final product, KTA-Br, was then purified by reverse phase HPLC in micromolar aliquots, as needed for the final coupling reactions (Figure 3.1).

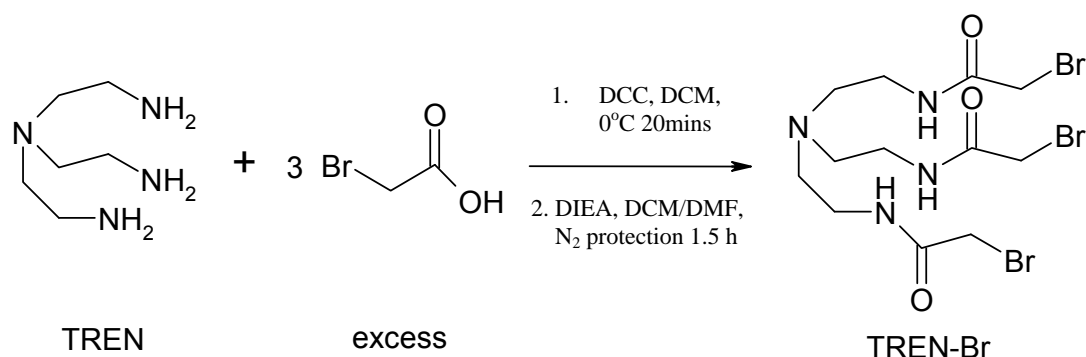
The ^1H NMR spectrum of the purified KTA-Br (9mg in 1mL 1:1 $\text{CD}_3\text{OD}/\text{CD}_3\text{CN}$) was then measured. Details of the resonance peaks were as follows: δ (ppm) 1.16-1.20 (d, 1H, J 15.2Hz), 1.23 (s, 3H), 2.71-2.75 (d, 1H, J 15.2Hz), 3.23-3.26 (m, 2H), 3.32-3.35 (m, 2H), 3.90 (s, 2H). Assignment of each peak was indicated in (Figure 3.1.b).

b. TREN-Br

The following reaction was carried out in solution phase to synthesize the TREN-Br template molecule (Scheme 3.2). The procedure has been optimized in terms of reaction time, and was improved by the use of an inert atmosphere (N_2).

Bromoacetic acid (9 mmol, 1.25 g) was pre-activated by DCC (4.5 mmol, 0.928 g) in DCM (20 mL) for 20 min. at 0 °C to form the symmetric anhydride. The DCU byproduct was filtered out as a white precipitate, and then the solvent

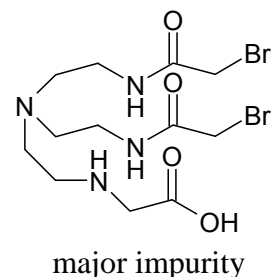
DCM was removed by rotary evaporation. The resultant solid residue was dissolved in 14 mL DMF with addition of 5 mL DCM to achieve a clear solution. The flask was capped with a rubber septum and purged with N₂ from a connected balloon for about 5 min., then TREN (0.5mmol, 75μL) was added to start the reaction. This reaction was stirred under N₂, and during the first 15 minutes, DIEA (2.3 mmol, 400 μL) was slowly added to the reaction mixture. After stirring for 1.5 h at RT, some solid residue was filtered out. The solvent was then evaporated under vacuum.



Scheme 3.2 Synthesis of TREN-Br template.

The orange yellow oily residue resultant from the above reaction was precipitated and washed by anhydrous diethyl ether (2 x 10 mL), followed by ethyl acetate (2 x 10 mL), yielding a light orange solid.

About 1.0 mL of buffer B (0.05% TFA in CH₃CN) and 2.0 mL buffer A (0.05% TFA in water) was used to dissolve the solid. After centrifugation, the clear supernatant was applied to semi-preparative reverse-



phase HPLC for purification (the solid residue left was later analyzed and proved to contain no product). Product fractions from the HPLC were collected and

lyophilized, yielding 160 mg final purified material. LC/MS showed that this product was about 80% pure (based on the UV absorbance at 220 nm) with the expected peak of "TREN-Br" at $[M + H]^+ = 507.0$ (cal. = 507, $\Delta = 0$). One major impurity, showing in above picture, was found by MS analysis, having a value for $[M + H]^+ = 445.2$ (cal. = 446, $\Delta = 0.8$). The yield of this reaction was about 63% based on TREN (Figure 3.2).

The ^1H NMR spectrum of the purified TREN-Br in D_2O was obtained on a Varian Unity at 400MHz. Details of the resonance peaks were as follows: (ppm) 3.51-3.54 (t, 2H, J 5.6Hz), 3.67-3.70 (t, 2H, J 5.6Hz), 4.00 (s, 2H). Assignment of each peak was indicated in (Figure 3.2.c).

c. TREN-3[K(Br)-NH₃⁺]

Starting from TREN, a four-step synthesis was carried out in solution phase to obtain the TREN-3[K(NH₃⁺)-Br] (**8**) template (Scheme 3.3) (K is the single letter code for amino acid Lys).

Fmoc-Lys(Boc)-OH (0.90 mmol, 0.422 g) and HOBt (1.20 mmol, 0.189 g) were first dissolved in a minimum volume of DMF (about 4.5 mL). After stirring at 0 °C for 10 min., EDC (1.20 mmol, 0.231 g) solid were added to the solution. A small amount of DMF was needed to achieve a clear solution (about 0.5 mL), and then the mixture was stirred at 0 °C for another 40 min. After this pre-activation step, the reaction flask was purged with N₂ from a connected balloon for about 5 min., and under this N₂ atmosphere, 0.10 mmol of TREN (15.0 μL) was added to the mixture, using a syringe, to start the reaction. (No DIEA was used to avoid loss of Fmoc protecting groups.) This reaction was

stirred for another 48 h under N_2 and was monitored by LC/MS. The solvent was then evaporated under vacuum. The oily residue was dissolved in 9.0 mL EtOAc followed by washing with H_2O (3 x 3 mL) and brine (1 x 3 mL), respectively. After EtOAc was evaporated under vacuum, the crude product was left as a white sticky solid residue, and was characterized by LC/MS (about 1.2 mg, dissolved in 50 μ L buffer A + 50 μ L buffer B and 50 μ L MeOH). This analysis showed major UV peaks of unreacted (excessive of) protected amino acid ($[M + H]^+ = 468.7$) and product (**5**) ($[M + H]^+ = 1497.5$, cal. = 1497.8, $\Delta = 0.3$).

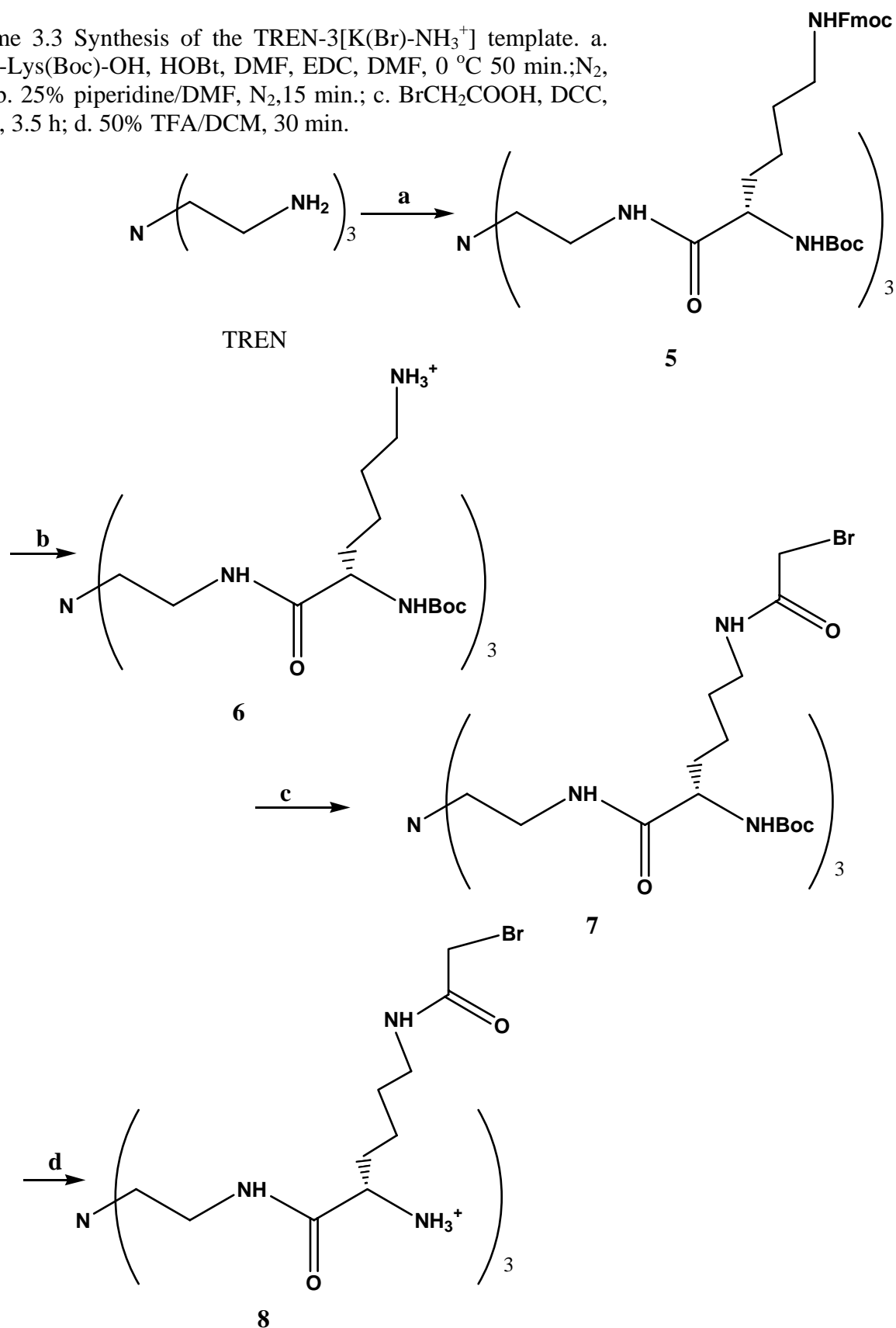
Crude product, **5**, was then dissolved in 1.0 mL DCM and 0.5 mL of EtOAc to achieve a clear solution, which was then applied to a silica-gel column (1 cm x 23 cm) packed with about 18 mL silica-gel slurry (in DCM). The following eluents were used to separate product, which came out last, from excessive starting material, which come out earlier: 50 mL DCM + 25 mL EtOAc + 3 mL MeOH; 100 mL DCM + 100 mL EtOAc + 8 mL MeOH. After evaporating solvent from the combined product fractions under vacuum, the resultant solid was analyzed by LC/MS, and showed a single pure product peak ($[M + H]^+ = 1497.5$, cal. = 1497.8, $\Delta = 0.3$). From the weight of purified product, **5**, the yield of this step (**a** in scheme 3.3), was about 77% based on TREN.

Removal of the Fmoc protecting groups from **5** was carried out by adding a small amount of 25%(v/v) piperidine/DMF solution to the above purified solid residue. After stirring under N_2 for about 20 min., the reaction mixture was evaporated to dryness under vacuum. Cold ether was then used to precipitate the product and wash away by-products, resulting in a sticky solid residue after

decanting. About 7.0 mL 0.1% TFA/H₂O was then used to dissolve this residue. Lyophilization of the resulting clear solution gave a white fluffy solid (**6**). LC/MS analysis of this product showed a single product peak (very pure) with $[M + H]^+ = 831.4$ (cal. = 831.6, $\Delta = 0.2$). From the weight of product, **6**, the yield of this step (**b** in scheme 3.3) was 80% based on **5**.

Bromoacetic acid (1.07 mmol, 0.148 g) was first converted to its corresponding symmetric anhydride by stirring in DCM (1.8 mL) for 20 min. at 0 °C in the presence of DCC (0.53 mmol, 0.11 g). The DCU byproduct was removed as a white precipitate by vacuum filtration and then the DCM was evaporated down to about 4.0 mL under vacuum for later use. Solid **6** (0.0426 mmol, 0.05 g) was put into a separate round bottom flask. Above solution containing pre-activated symmetric anhydride was mixed with this solid, and then the flask was immediately capped with a septum. After purging this reaction flask for about 5 min. by connecting to a balloon filled with N₂, DIEA (0.256 mmol, 44.5 μ L) was slowly injected into the reaction mixture with stirring. The Reaction was stirred for another 3 - 4 h under N₂. The DCM was then evaporated under vacuum. Buffer A (3.0mL) and buffer B (1.5mL) were added to the resultant solid residue. Some white solid was filtered out, which was proved to containing no product and the left solution was lyophilized, yielding a sticky yellow residue. LC/MS analysis of this residue showed a single strong UV absorbance peak corresponding to product, **7**, with $[M + H]^+ = 1191.5$ (calc. = 1191.4, $\Delta = 0.1$).

Scheme 3.3 Synthesis of the TREN-3[K(Br)-NH₃⁺] template. a. Fmoc-Lys(Boc)-OH, HOBt, DMF, EDC, DMF, 0 °C 50 min.; N₂, 48 h; b. 25% piperidine/DMF, N₂, 15 min.; c. BrCH₂COOH, DCC, DCM, 3.5 h; d. 50% TFA/DCM, 30 min.



Removal of the Boc protecting group from **7** was achieved by adding about 5.0 mL 50% TFA/ DCM to above resultant solid **7**. After 30 mins' stirring, the reaction mixture was evaporated to dryness under a stream of N₂ gas and then triturated by diethyl ether (3 x 20 mL). The resulting white powder was analyzed by LC/MS, using a very small amount of the solid dissolved in 0.1% TFA/H₂O. This analysis showed a very pure single peak of product, **8**, with $[M + H]^+ = 891.2$ (cal. = 891.2, $\Delta = 0$). From the weight of the final product **8**, the overall yield of steps **c** through **d** (scheme 3.3) was about 80%, based on **6**. Due to its high purity, no further separation was needed and this material was then used directly in the coupling reaction with N-peptide sequences.

The ¹H NMR spectrum of TREN-3K(Br)-NH₃⁺ (2.4 mg in 0.8 mL D₂O) was obtained on a Varian Unity at 400 MHz. Resonance peaks were observed as follows: δ (ppm) 1.40-1.52 (m, 2H), 1.68-1.90 (m, 4H), 2.69-2.74 (t, 2H), 2.99-3.03 (t, 2H, J 8.0Hz), 3.29-3.36 (m, 2H), 3.96 (s, 2H), 4.28-4.32 (t, 1H, J 8.8Hz). Assignment of each peak was indicated in (Figure 3.3).

d. TREN-3[K(NH₃⁺)-Br]

Starting from TREN, a four-step synthesis was carried out in solution phase to obtain the TREN-3[K(NH₃⁺)-Br] (**12**) template molecule (Scheme 3.4) (K is the single letter code for amino acid Lys).

Fmoc-Lys(Boc)-OH (0.45 mmol, 0.211 g) and HOBt (0.45 mmol, 0.069 g) were first dissolved in a minimum volume of DMF (about 1.5 mL). After stirring at 0 °C for 10mins, EDC (0.45 mmol, 0.086 g) solid was added to the solution. A

small amount of DMF (about 0.5 mL) was needed to achieve a clear solution, and then the mixture was stirred at 0 °C for another 40 min. After this pre-activation step, the reaction flask was purged with N₂ from a connected balloon for about 5 min., and under this N₂ atmosphere, 0.05 mmol of TREN (7.5 µL) was injected to the mixture to start the reaction. (No DIEA was used to avoid loss of the Fmoc protecting groups.) This reaction was stirred under N₂ for 24 h and was monitored by LC/MS. The solvent was then evaporated under vacuum. The oily residue was dissolved in 4.5 mL EtOAc, followed by washing with H₂O (2 x 3.0 mL) and brine (1 x 3.0 mL). After the EtOAc was evaporated under vacuum, the crude product was obtained as a sticky white solid. This product was characterized by LC/MS, using a small amount of the dried residue, about 1/100 of total dissolved in 80 µL buffer A + 80 µL buffer B and 40 µL MeOH. LC/MS showed major peaks in the m/z and UV chromatography corresponding to excess unreacted protected amino acid ([M + H]⁺=468.9) and product (**9**) with [M + H]⁺=1497.4 (cal. = 1497.8, Δ = 0.4).

Above crude product, **9**, was then dissolved in DCM (1.0 mL) and EtOAc (0.5 mL) to give a clear solution. This solution was applied onto a silica-gel column (1 cm x 23 cm) packed with about 18 mL of silicagel slurry (in DCM). The following eluents were used to separate the product, which came out last, from excess starting material, which come out earlier: 50 mL DCM + 25 mL EtOAc + 3 mL MeOH; 75 mL DCM + 75 mL EtOAc + 6 mL MeOH. After evaporating the solvent from the combined product fractions under vacuum, the resultant solid was analyzed by LC/MS, and gave a single peak corresponding to the product ,

9, ($[M + H]^+ = 1497.4$, cal. = 1497.8, $\Delta = 0.4$). From the weight of purified product, **9**, the yield of this step (**a** in scheme 3.4), was 80% based on TREN.

Removal of the Fmoc protecting groups was carried out by adding a small amount of 25% piperidine/DMF solution to the above purified residue. After stirring under N_2 for about 20 min., the reaction mixture was evaporated to dryness under vacuum. Cold ether was then used to precipitate product and wash away by-products, resulting in a sticky solid residue after decanting. About 6.0 mL 0.1% TFA/ H_2O was then used to dissolve this residue. Lyophilization of the resultant clear solution gave a fluffy white solid (**10**). LC/MS analysis of this product showed a single product peak (very pure) with $[M + H]^+ = 831.4$ (cal. = 831.7, $\Delta = 0.3$). From the weight of the product, **10**, (0.04 mmol, 47 mg), conversion from **9** to **10** (**b** in scheme 3.4) was quantitative.

About 1.0 mmol of bromoacetic acid (0.139 g, about 25-fold excess over to **10**) was first converted to its corresponding symmetric anhydride by stirring in DCM (4.0 mL) for 20 min. at 0 °C in the presence of 0.5 mmol of DCC (0.103 g). The DCU byproduct was filtered out as a white precipitate and then the solvent was evaporated down to about 2.0 mL under vacuum. A solution of **10** (about 0.04 mmol, 0.047 g) in 0.8 mL DCM was then added dropwise into above symmetric anhydride/DCM solution under N_2 with stirring. Another small amount of DCM was used to wash and transfer all of the solution of **10** into the reaction flask. During the first 15 min. of reaction, of DIEA (0.24 mmol, 41.8 μ L) solution in about 0.5 mL DCM was also slowly added into the reaction flask under N_2 with stirring. The reaction was then allowed to proceed for another 3 – 4 h. The

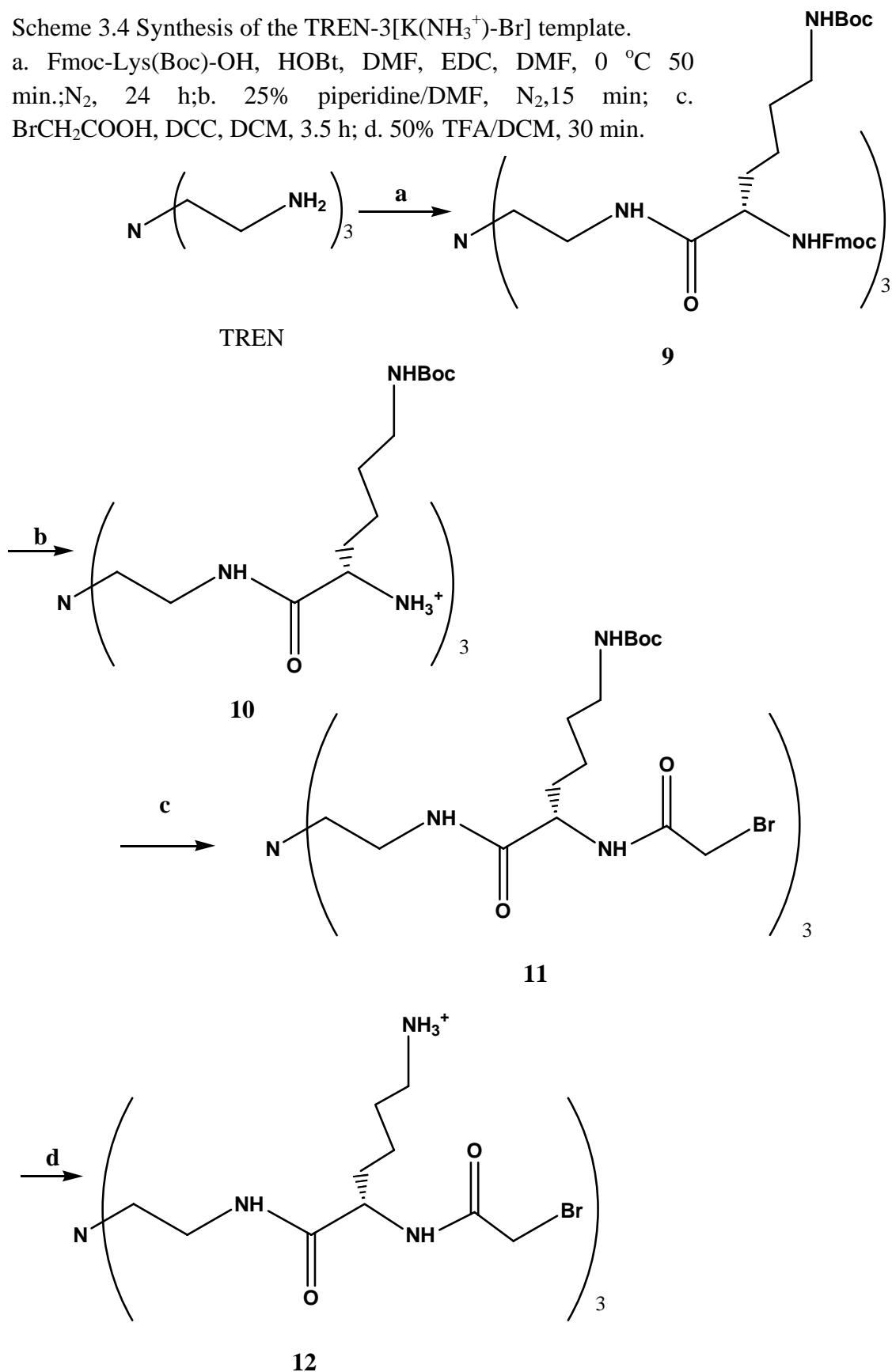
solvent DCM was then evaporated under vacuum. Buffer A (3.0 mL) and buffer B (1.5 mL) were added to the resultant solid residue. Some white solid was filtered out, which was proved to containing no product and the left solution was lyophilized, yielding a sticky yellow residue. LC/MS analysis of the product showed a single strong UV peak corresponding to the major product, **11**, with $[M + H]^+ = 1191.0$ (cal. = 1191.0, $\Delta = 0$).

Removal of the Boc protecting group from **11** was achieved by adding about 4.0 mL 50% TFA/ DCM to solid **11**. After stirring for 30 min., the reaction mixture was evaporated to dryness under a stream of N₂ gas and then washed and triturated with diethyl ether (3 x 8 mL). The resulting white solid (powder) was analyzed by LC/MS using a very small amount of solid dissolved in 0.1% TFA/H₂O. This analysis showed a very pure single peak of product, **12**, with $[M + H]^+ = 891.1$ (cal. = 891.3, $\Delta = 0.2$). From the weight of the final product **12**, the yield of step **c** through **d** (scheme 3.4) was about 90% based on **10**. Due to its high purity, no further separation was considered necessary, and this material was then used directly in the coupling reaction with N-peptide sequences.

The ¹H NMR spectrum of TREN-3K(NH₃⁺)-Br (2.7mg in 1.0mL D₂O) was obtained on a Varian Unity at 400 MHz. Details of the resonance peaks were as follows: δ (ppm) 1.35-1.40 (m, 2H), 1.56-1.62 (m, 2H), 1.85-1.92 (m, 2H), 2.74-2.77(t, 2H, *J* 7.2Hz), 3.23-3.27 (t, 2H, *J* 6.8Hz), 3.32-3.45 (m, 2H), 3.90(s, 2H), 3.93-3.97 (t, 1H, *J* 6.4Hz). Assignment of each peak was indicated in (Figure 3.4).

Scheme 3.4 Synthesis of the TREN-3[K(NH₃⁺)-Br] template.

a. Fmoc-Lys(Boc)-OH, HOBT, DMF, EDC, DMF, 0 °C 50 min.; N₂, 24 h; b. 25% piperidine/DMF, N₂, 15 min; c. BrCH₂COOH, DCC, DCM, 3.5 h; d. 50% TFA/DCM, 30 min.



B.2. Peptide synthesis

a. Standard procedure for Fmoc solid-phase peptide synthesis

All peptide synthesis was carried out in a glass reaction vessel with a built-in frit, and started with 1 - 2 g of Rink amide MBHA resin (0.64mmol/g). Before starting, the resin was first washed and swollen with DMF and DCM (~10 mL/g resin each). Fmoc protecting groups on the resin were then removed with 25% piperidine/DMF (~10 mL/g resin) to produce free amino groups, which was confirmed by the Kaiser test for amines. After the resin was washed with DMF (~10 mL/g resin) for four to five times to neutral pH (checked with wet pH paper), the first Fmoc protected amino acid (0.40 mmol/g resin), PyBOP (0.40 mmol/g resin), HOBt (0.40 mmol/g resin) and DIEA (0.60 mmol/g resin) were added. The first coupling reaction usually was left overnight. Then the substitution level was determined for triplicate samples of 3 – 5 mg vacuum-dried resin by measuring UV absorbance at 300 nm of the released 9-methylene fluorene using a Beckman DU-65 UV spectrometer. The extinction coefficient of the deprotection product, fluorene, in solution was previously determined to be $7300\text{M}^{-1}\text{cm}^{-1}$ by cleaving Fmoc from a known amount of Fmoc-Ser(tBu)-OH, which has no other UV absorbent group, by a known volume of 25% piperidine/DMF and then measuring its UV absorbance at 300nm in EtOH after 20-fold dilution. Substitution levels for peptide synthesis were controlled to be around 0.3 - 0.4 mmol/g resin. The unreacted free amine groups were then capped by adding a 20-fold excess of Ac_2O in DCM and 1-fold of DIEA and reacting for 40 min. The completion of the reaction was confirmed by the Kaiser test.

For coupling of the subsequent amino acids, standard reaction cycle was repeated for each residue, which includes deprotection, washing, coupling, and washing. Re-coupling (with another washing) was performed only if necessary. Deprotection was performed with 25% piperidine/DMF (~10 mL/g resin) followed by washing with DMF (~10 mL/g resin each time) until neutral pH was attained. Coupling reactions were carried out for 3 - 4 h after adding the Fmoc-protected amino acid, PyBOP, HOBT and DIEA in the amount of 3, 3, 3, and 4.5 equivalents to the free amine groups on the peptidyl-resin, and using a minimal amount of DMF solvent (~5 mL/g resin). Completion of each coupling reaction was confirmed by the Kaiser test, showing a clear light yellow solution and transparent beads indicating the absence of free amines. If coupling was not completed after 4 h, the old coupling solution was drained out and fresh reagents with half amount of what had been used in the previous coupling were added into the reaction vessel. The recoupling reactions then proceeded for another 2-3 hours.

To confirm that solid resin beads was not lost during the repetitive washing and reaction steps, after each day's work (usually coupling of two or three residues), the peptidyl-resin was dried in a vacuum desiccator overnight. The following morning, the first deprotection solution with released Fmoc groups in it was checked by UV at 300nm for total absorbance. For the synthesis performed in glass reaction vessel with built-in fine frit, no obvious decrease of total absorbance was observed through the whole synthesis. While for that using reaction vessel with course frit, total absorbance decreased gradually to about

20% of the starting point after 22 residues were coupled, indicating a constant loss of available amine for coupling, which is probably due to the loss of solid resin.

All amino acids except for cysteine were coupled by the PyBOP/HOBt method described above. Since cysteine racemizes easily in the presence of an organic base like DIEA, the symmetrical anhydride method at neutral pH was used instead: 4 eq. of Fmoc protected cysteine was dissolved in DCM with addition of a minimum amount of DMF, to which 2 eq. of DCC were then added. The mixture was stirred at 0 °C for 15 min. After the white precipitate of DCU was filtered off, the cysteinyl symmetrical anhydride was transferred to the reaction vessel and allowed to react with the peptidyl-resin for another 2 h.

For an end-capped (Ac-) sequence, after deprotection of the last amino acid residue at the N-terminus, acetylation was performed by adding 20 eq. of Ac₂O and 1 eq. of DIEA (relative to expected total amount of amine at the end) in DCM (~5 mL/g resin) and shaking for 40 min. followed by washing and the Kaiser test.

The final sequence was cleaved from the resin and the side-chains were deprotected simultaneously by treatment with a mixture of 95% (v/v) TFA, 2.5% (v/v) TIS and 2.5% (v/v) H₂O for about 2 – 4 h. TIS was added to scavenge carbocations generated from the deprotection of Trt, tBu and Boc groups. After the resin was filtered off, TFA was removed using a gentle stream of N₂ flow. Ice-cold diethyl ether (15 – 20 mL) was added to precipitate the peptide. After mixing and centrifugation, diethyl ether was decanted. Residual cleavage reagents were

washed away with two more portions of diethyl ether (15 – 20 mL each time). Crude peptides were then dissolved in 5% AcOH followed by lyophilization, usually resulting in white fluffy powders.

The crude lyophilized product was usually dissolved in a combination of HPLC buffers A and B, and then purified by reverse-phase HPLC after centrifugation to yield a clear solution.

b. N-peptide

In our initial design, N-peptide analog N^α-acetylated-peptidyl-amide, Cys-N29, corresponding to native sequence of gp41 (553-581) (shown below) was investigated. A Cys was added to the N-terminal to allow final coupling to our designed template molecule via the thiol side-chain.

Cys-N29: ac-**C**NNLLRA IEAQQHL LQLTVWG IKQLQAR IL-nh₂

Cys-N29b: ac-**C**NNLLRA **I**QAQQHL LQLTVWG IKQLQAR IL-nh₂

Given the long synthesis time and the low yields anticipated due to serious aggregation tendencies of the sequence itself, this peptide was purchased from ANASPEC. LC/MS analysis of the purchased product indicated that it is 80% pure based on UV absorbance chromatography at 222 nm. The product showed a molecular ion peaks at $[M + 3H]^{3+} = 1176.3$ (cal. = 1175.7, $\Delta = 0.6$). Without further purification, this purchased peptide product was used in the coupling reactions described below.

In our modified design, N-peptide analog N^α-acetylated-peptidyl-amide, Cys-N29b, corresponding to [Gln⁵⁶⁰]-gp41(553-581) (shown above), was investigated. Compared with Cys-N29, in Cys-N29b, the only negative charged

residue within this sequence, Glu⁵⁶⁰ was replaced by a Gln in an attempt to increase the solubility of the final product at neutral pH by increasing the net charge (positive) of the resultant molecule. This peptide was also purchased from ANASPEC. LC/MS analysis of the purchased product indicated that it is 80% pure based on UV absorbance chromatography at 222 nm. The product showed a molecular ion peak at $[M + 3H]^{3+} = 1175.9$ (cal. = 1175.7, $\Delta = 0.2$). Without further purification, this purchased peptide product was used in the coupling reactions described below.

c. C-peptide

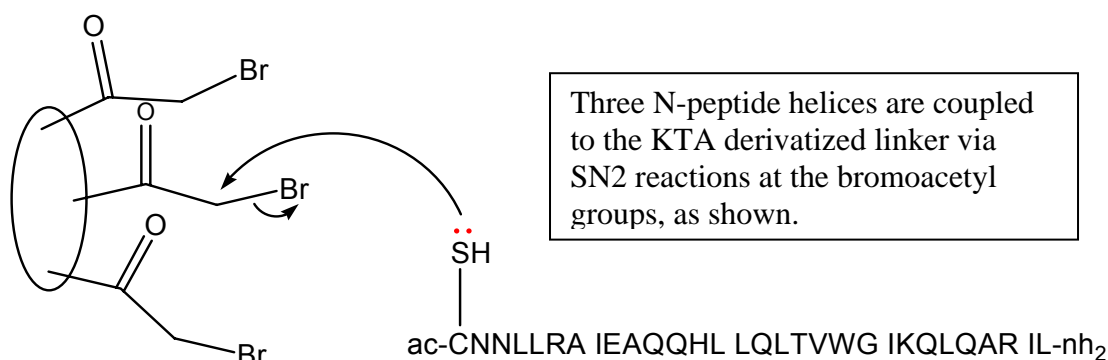
Standard Fmoc chemistry (as described above) was applied to synthesize the end-capped C-peptide analog corresponding to N^α-acetyl-[Nle⁶²⁹]-gp41(628-655)-amide (C28). M⁶²⁹ was replaced by Nle to avoid oxidation problems during handling and characterization.

2.0 g of Rink amide MBHA resin (0.64mmol/g) was used. A substitution level of 0.41 ± 0.2 mmol/g was then determined by Fmoc deprotection and UV analysis. Reactive side chains of amino acids were protected as follows: Lys(Boc), Glu(OBu^t), Gln(Trt), Asn(Trt), Ser(Bu^t), His(Trt), Thr(Bu^t), Tyr(Bu^t), Arg(Pbf), Asp(OBu^t), and Trp(Boc). Peptide assembly followed the general methods described above. After cleavage and deprotection, ether precipitation and washing, dissolution of crude peptide, HPLC purification and lyophilization, a white solid was obtained, which showed only product peaks on ESI-MS as $[M + 2H]^{2+} = 1773.0$ (cal. = 1172.6, $\Delta = 0.4$) , and $[M + 3H]^{3+} = 1182.6$ (cal. = 1181.9, Δ

= 0.7). Stock solutions in 20 mM Tris.HCl, 100 mM KF, pH = 7.40 were then prepared for further analysis and study.

B.3. Synthesis of N-peptide three-helix bundle model structures

Coupling of three copies of thiol containing peptide chains onto the three bromide-substituted arms of each template molecule via a simple SN2 mechanism results in the desired three-helix bundle model structures, with a thiol-ether peptide link in each arm (Scheme 3.5).



Scheme 3.5 Coupling of the template molecule to three copies of N-peptide.

All coupling reactions to construct the final three-helix bundle model structures were carried out in screw-cap vials (1.5 mL) connected to a balloon filled with N₂. Small scale synthesis, for the purpose of testing reaction conditions, was performed using 0.5 – 1.0 mg N-peptide analog. Large scale synthesis, for the purpose of preparation, was performing using 4.0 mg N-peptide analog for each bunch. The following procedure was optimized in terms of the pH value, the reaction time and additional organic solvent needed to solublize the mixture.

Solution of the tribrominated template in DMF (~3 – 5 mg/mL) was prepared first. A small amount of this solution containing 1 equivalent of the template molecule was then added to 4 equivalents of N-peptide solid (either Cys-N29 or Cys-N29b) followed by immediate addition of degassed 20% (v/v) DMF/100mM Tris.HCl buffer (pH= 8.00) (about 100-200 μ L/mg peptide) to the mixture. After purging the reaction vial with N₂ (from connected balloon) for about 5 mins., the mixture was stirred under N₂ for 3 - 4h and then quenched by adding 1N HCl to pH 2 or below. This reaction mixture was then transferred to a 1.5 mL centrifuge tube. After centrifugation, the clear solution of reaction mixture was applied onto reverse phase HPLC Waters XTerra column (10x100mm) for purification followed by lyophilization of the product fractions. Any solid left in the centrifuge tube was later identified to be the unreacted N-peptide analog.

Figure 3.5 is a typical HPLC/MS characterization of the final coupling reaction mixture. On ESI-MS, chromatography peaks corresponding to the product showed clear serial molecular ion mass peaks with $[M + 6H]^{6+}$, $[M + 7H]^{7+}$, $[M + 8H]^{8+}$ and $[M + 9H]^{9+}$. Table 3.1 summarizes the accurate MS identification of the correct products.

Three-helix bundle model structures constructed by assembling N-peptide analog Cys-N29 and Cys-N29b onto KTA-Br template are denoted as KTA-3N29 and KTA-3N29b, respectively. Three-helix bundle model structures constructed by assembling N-peptide analog Cys-N29 onto TREN-Br template is denoted as TREN-3N29. Three-helix bundle model structures constructed by assembling N-peptide analog Cys-N29b onto TREN-3[K(NH³⁺)-Br] template is denoted as

TREN-3K-N29b. These denotations are used in the rest part of this thesis for the purpose of simplicity.

For HPLC analysis and purification, a combination of buffer A, 0.05% TFA in H₂O and buffer B, 0.05% TFA in CH₃CN, was used for all model structures, except for TREN-3K-N29b, in which the buffer combination had to be changed to buffer A, 0.1% TFA in H₂O and buffer B, 0.1% TFA in 1:1 i-PrOH:CH₃CN, to get adequate separation.

Table 3.1 Comparison of Experimental and Calculated Mass of Model Structures

Model Structures	M		[M+6H] ⁶⁺			
	Mono. (cal.)	Ave. (cal.)	Mono. (cal.)	Ave. (cal.)	Ave. (exp.)	ΔAve.
KTA-3N29	11076	11083.2	1847.0	1848.2	1847.8	0.4
KTA-3N29b	11073	11080.2	1846.5	1847.7	1847.3	0.4
TREN-3N29	10838	10844.9	1807.3	1808.5	1808.1	0.4
TREN-3K-N29b	11219	11226.5	1870.8	1872.1	1871.7	0.4

B.4. Synthesis of thiol-blocked N-peptide sequences

The free thiols of the cysteine residues on N-peptides were blocked for CD studies of these peptides alone, unconnected to template. Degassed 20%DMF/100mM Tris.HCl (pH 8.0) was used as the reaction buffer. About 2.0 mg peptide was dissolved in about 300μL buffer. A 10-fold molar excess of DTT relative to the amount of thiol group on peptide was dissolved in about 80 μL buffer and then added to the peptide solution. The resulting solution was stirred under N₂ for about 3 h. A 25-fold molar excess of iodoacetamide relative to the amount of thiol group on peptide was dissolved in 80 μL buffer and then was

added into the above reaction mixture. The reaction was allowed to proceed for another 5 h before it was quenched by adding 1 M HCl to pH 2.0 or below.

HPLC buffer A (200 μ L, 0.05%TFA in H₂O, v/v) and buffer B (150 μ L, 0.05%TFA in CH₃CN, v/v) were added into the reaction mixture to dissolve most of the solid formed during the reaction. After centrifugation, this solution was applied onto a semi-preparative HPLC Waters XTerra column (10x100mm) for separation. Product fractions were confirmed by ESI-MS, showing $[M + 2H]^{2+} = 1791.9$ (cal. = 1791, $\Delta = 0.9$), $[M + 3H]^{3+} = 1194.9$ (cal. = 1194.3, $\Delta = 0.6$) and $[M + 4H]^{4+} = 896.5$ (cal. = 896, $\Delta = 0.5$) serial peaks for thiol blocked Cys-N29b. After lyophilization, a fluffy white power was obtained.

C. Methods for CD Characterization

C.1. Sample preparation

a. Buffer used for CD studies

CD characterizations were carried out under two sets of conditions. Acidic conditions were achieved by titrating 10 mM H₃PO₄ solution to pH 2.50 with NaOH solution at 25 °C. Neutral conditions were achieved by titrating either 20 mM TRIZMA BASE solution alone, or 20 mM TRIZMA BASE /100 mM KF solution to pH = 7.40 with HCl at 25 °C.

b. Preparation of stock solutions

Stock solutions of thiol-blocked N-peptides (Cys-N29 and Cys-N29b) and the synthetic, templated three-helix model structures were all prepared in 1 mM HCl.

Due to the low solubility of C28 under acidic conditions, stock solutions of C28 were prepared in 20 mM Tris.HCl/100 mM KF, pH 7.40.

Concentrations of stock solutions were determined based on the UV absorbance of Trp at 280nm, using an extinction coefficient of $5500 \text{ M}^{-1}\text{cm}^{-1}$ per Trp.

C.2. Collecting CD spectra

All CD studies were performed on an AVIV model 215 or 400 spectropolarimeter at 25 °C, unless otherwise noted. Cells having 1.0 mm or 10.0 mm pathlengths were used, as appropriate for the concentration under study. Data collected at dynode voltages greater than 480 V were discarded. Spectra were recorded every 0.5 nm in the range 190-260 nm, using a 1.0 nm bandwidth and a signal averaging time of 2.0 to 4.0 s, depending on the signal-noise ratio.

C.3. Melting curves

CD melting curves were all collected in acidic buffer (10 mM phosphate pH 2.50). CD absorbance at 222nm, which corresponds to the minimum signal characteristic of helix structure, was monitored over the temperature range of 0-85 °C with 0.20 °C temperature increment steps and a 0.15 °C temperature dead band. For each point, the signal averaging time was set to be 5.0 s and a 1.0 nm bandwidth was used.

C.4. Urea denaturation experiments

Urea denaturation curves were determined using a series of solutions prepared individually with constant concentration of three-helix bundle model structure (6.6 μM) and increasing concentrations of urea from 0 M to about 8 M in the neutral

buffer (20 mM Tris.HCl or 20 mM Tris.HCl/100 mM KF, pH 7.40). Partial spectra in the range 210-260 nm were recorded for each sample every 0.5 nm using a 1.0 nm bandwidth and a signal averaging time of 2.0 to 4.0 s, depending on the signal-noise ratio. After converting each spectrum to mean residue ellipticity, $[\theta]$, the values at 222 nm ($[\theta]_{222}$) and an average of the values between 250 nm and 260 nm ($[\theta]_{260}$) were obtained. The denaturation curve was drawn by plotting the difference between $[\theta]_{222}$ and $[\theta]_{260}$, against urea concentration.

The reversibility of urea denaturation was tested by diluting a 6.6 μ M solution of the model structure in 4.8 M urea to a fixed concentration of 1.7 μ M peptide in 1.2 M urea, by adding a three-fold larger volume of the buffer. The partial spectrum of this diluted solution was then analyzed by the same method described above and then compared with that of a separate solution that had been directly prepared to have the same composition. Identical or approximately identical spectra of the diluted and the directly prepared solutions would confirm the reversibility of the urea denaturation process.

C.5. C-peptide binding studies

Preliminary CD studies of the binding of C-peptide (C28) to the N-peptide, three-helix bundle model structures under neutral conditions were carried out by collecting the CD spectra of 10 μ M C28 solution alone, and 3.3 μ M model structure solution alone, and CD spectra of solutions containing both 10 μ M C28 solution and 3.3 μ M model structure mixed together.

CD titrations in 20 mM Tris.HCl, 100 mM KF, pH = 7.40 buffer were also applied to study the binding properties of C-peptide (C28) to our model structures. During these experiments, C28 solutions at low concentration (4.00 μ M binding assays with KTA-3N29b and 3.00 μ M for binding assays with TREN-3K-N29b) were titrated with increasing amounts of a stock solution containing 20 μ M of the templated three-helix bundle structures premixed with the same concentration of C28 (4.00 μ M for the case of KTA-3N29b and 3.00 μ M for the case of TREN-3K-N29b), to maintain a constant C28 concentration during the titration. Blank titrations were performed by titrating a blank buffer solution with about 20 μ M templated three-helix bundle structures solution only. After each addition of titrant, 10 min. was allowed for the solution to reach equilibrium. For each titration point, either the mixture titration or blank titration, kinetic traces at 222 nm and 250 nm were recorded over 5-10 min. Readings at 222 nm and 250 nm was obtained by taking the average of each trace, respectively and the difference between these two readings was recorded as the ellipticity (or CD absorbance) of this titration point at 222 nm.

After normalization of 222nm-ellipticity of each titration point to the 1mm pathlength cell (at low concentrations, 1 cm cell was used for best results), experimental binding isotherms were drawn by plotting the 222nm CD absorbance difference ($\Delta\theta_i = \theta_{i, \text{model structure}} + \theta_{i, \text{C28}} - \theta_{i, \text{mixture}}$) versus the increasing concentration of templated three-helix bundle structure. (For detailed calculations, please refer to Chapter IV.C.2)

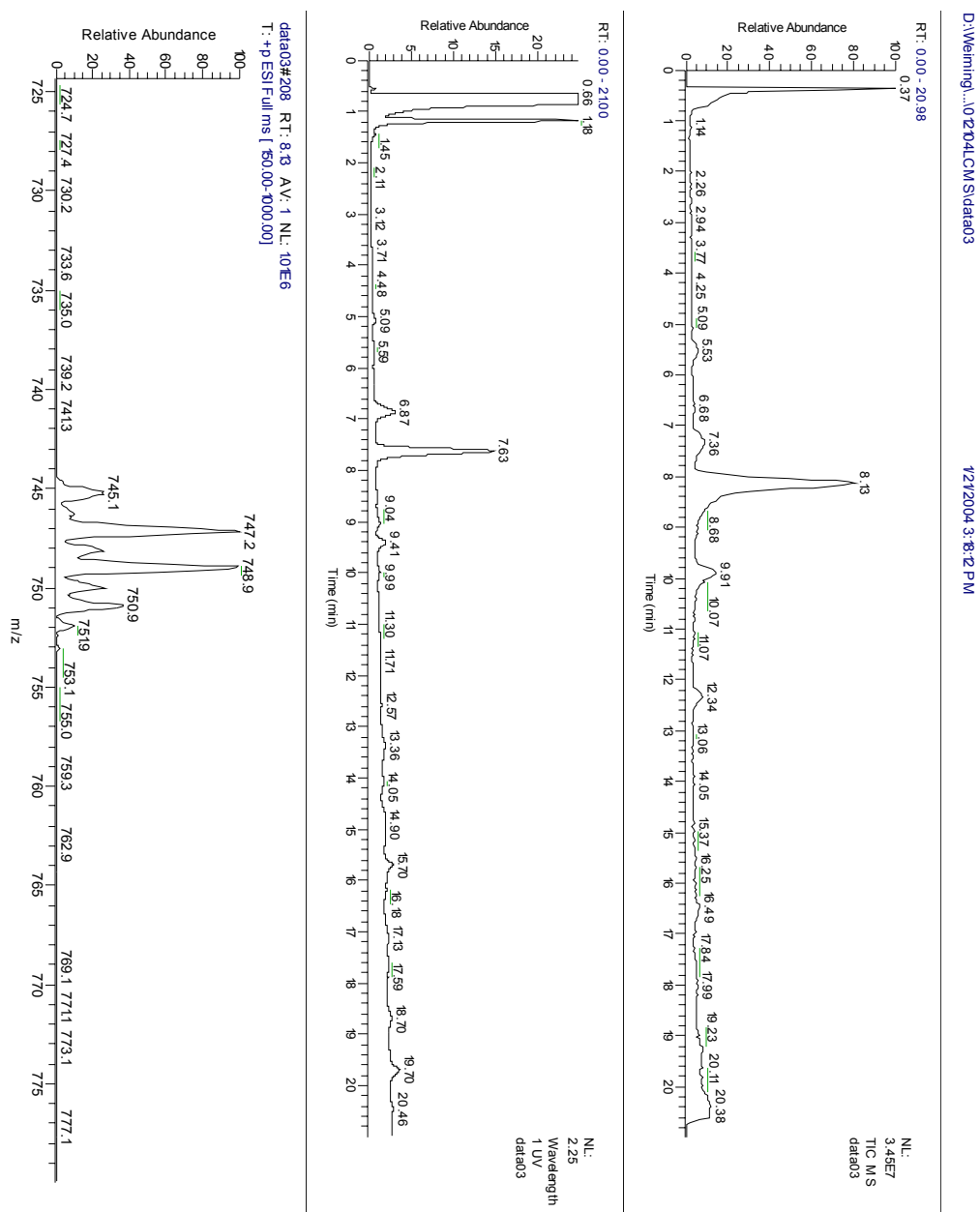


Figure 3.1.a LC/MS characterization of KTA-Br template (reaction mixture of the step c in Scheme 3.1). LC conditions: Buffer A: 0.05% TFA in H₂O, buffer B: 0.05% TFA in CH₃CN, 10% - 60% buffer B over 20 min., 1.0 mL/min, using Waters XTerra MSC₁₈ 5 μ M 4.6 x 50 mm column. From the top panel, shown are the TIC, UV (220 nm) and the mass spectrum corresponding to the product peak (retention time = 7.63 min.), respectively.

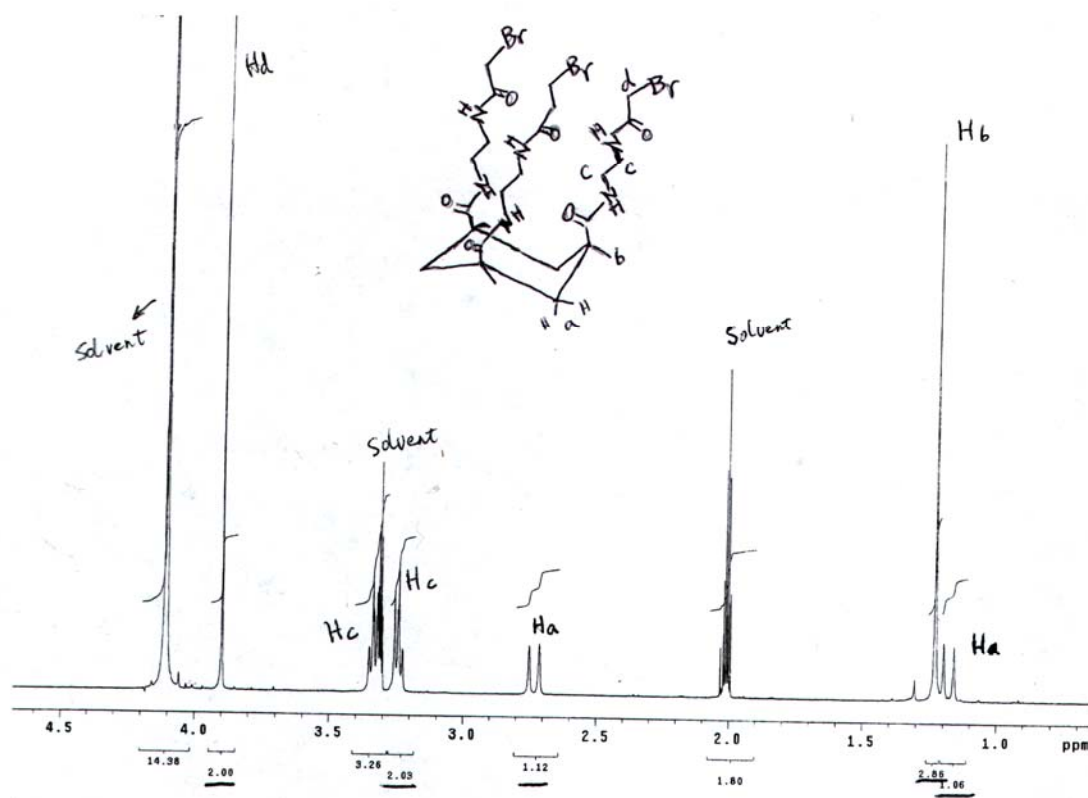


Figure 3.1.b The ^1H NMR spectrum of the purified KTA-Br recorded at RT on a Varian Unity 400 MHz (9mg in 1mL 1:1 $\text{CD}_3\text{OD}/\text{CD}_3\text{CN}$).

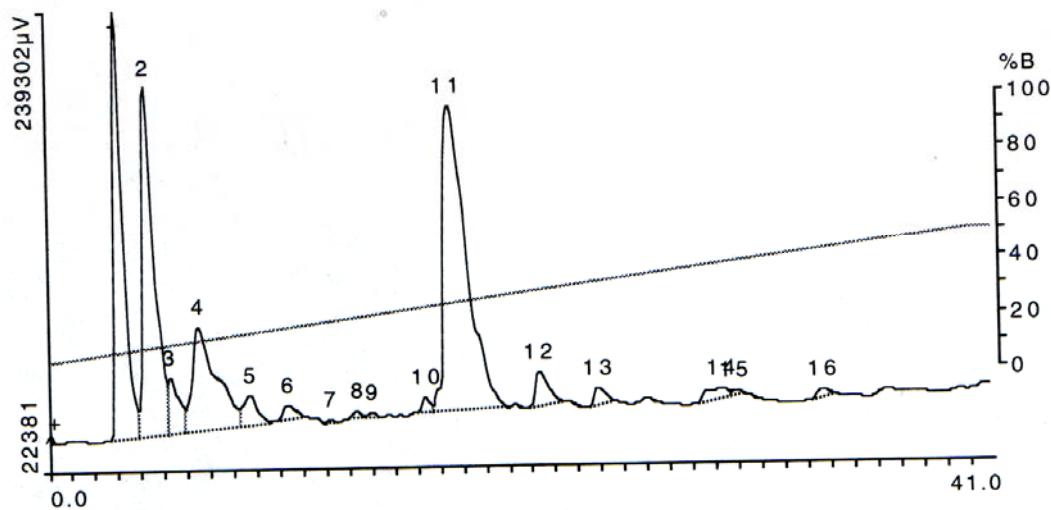


Figure 3.2.a Reverse-phase HPLC of the crude reaction mixture used to prepare TREN-Br template (Scheme 3.2). HPLC conditions: Buffer A: 0.05% TFA in H_2O , buffer B: 0.05% TFA in CH_3CN . 5% - 50% buffer B over 40 min., 3.5 mL/min, using Dynamax C18-300A, 10 x 250 mm column. Solid line (left axis) is the UV absorbance trace at 220 nm. Dotted line (right axis) indicates the eluting gradient used. Peak 11 corresponds to product TREN-Br.

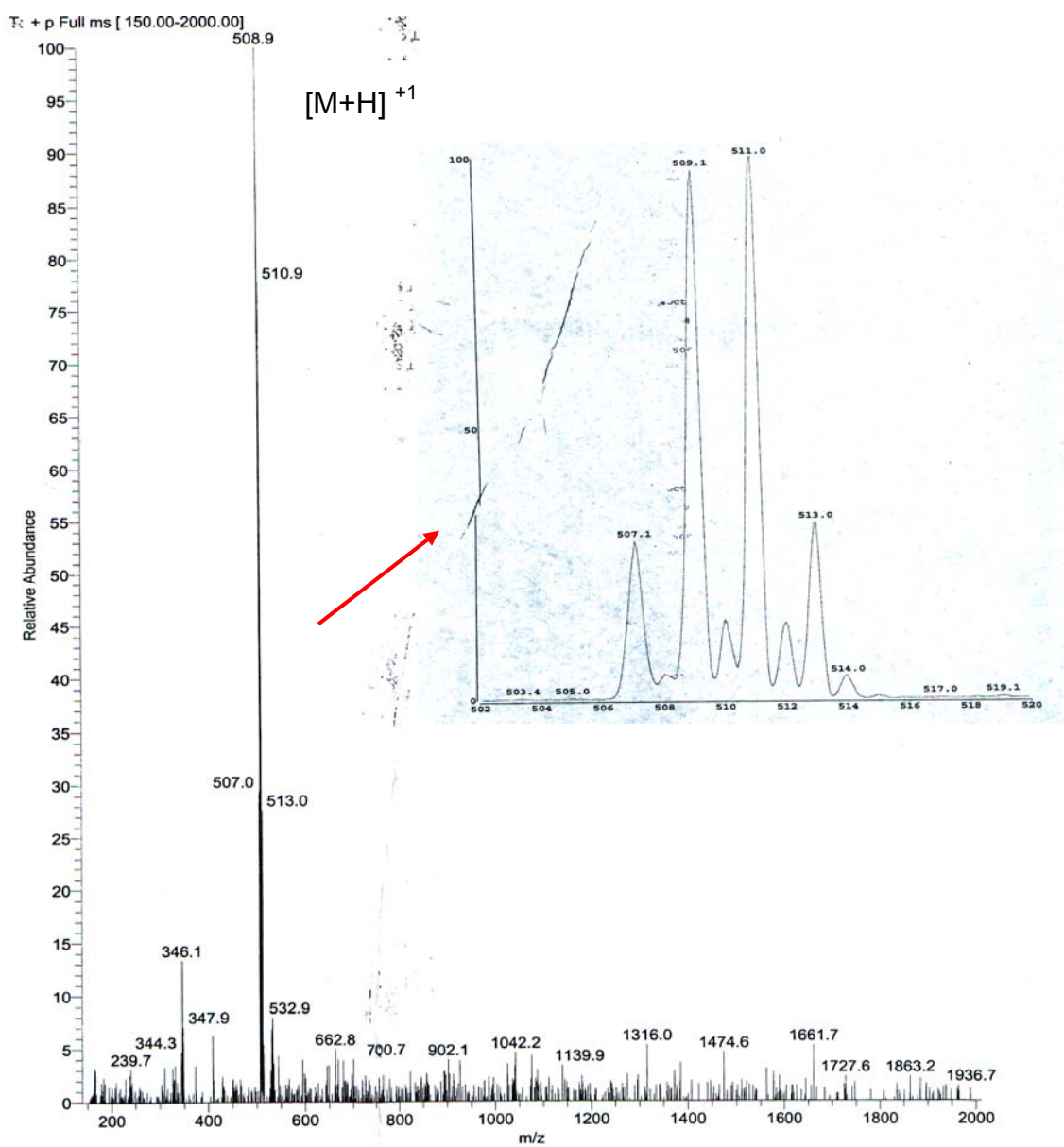


Figure 3.2.b ESI-MS of HPLC-purified TREN-Br template (peak 11 from HPLC shown in Figure 3.2.a). Inserted panel is the expended view of the molecular ion peak.

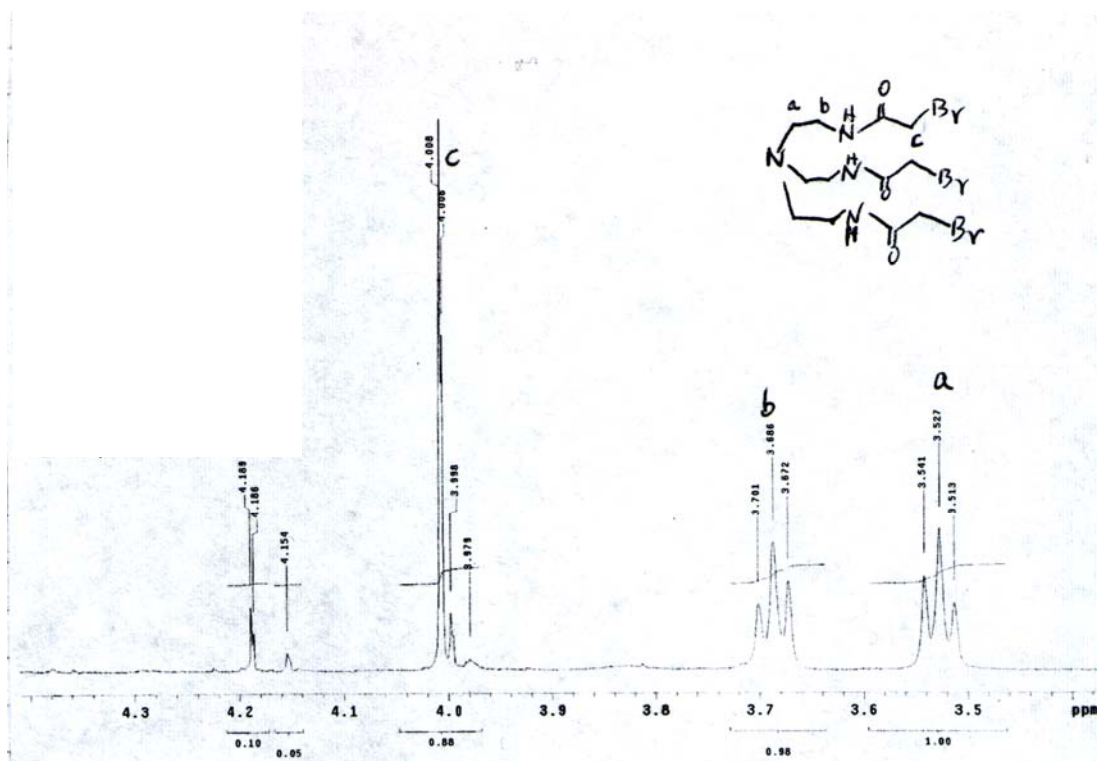


Figure 3.2.c The ^1H NMR spectrum of the purified TREN-Br template in D_2O recorded at RT on a Varian Unity 400 MHz.

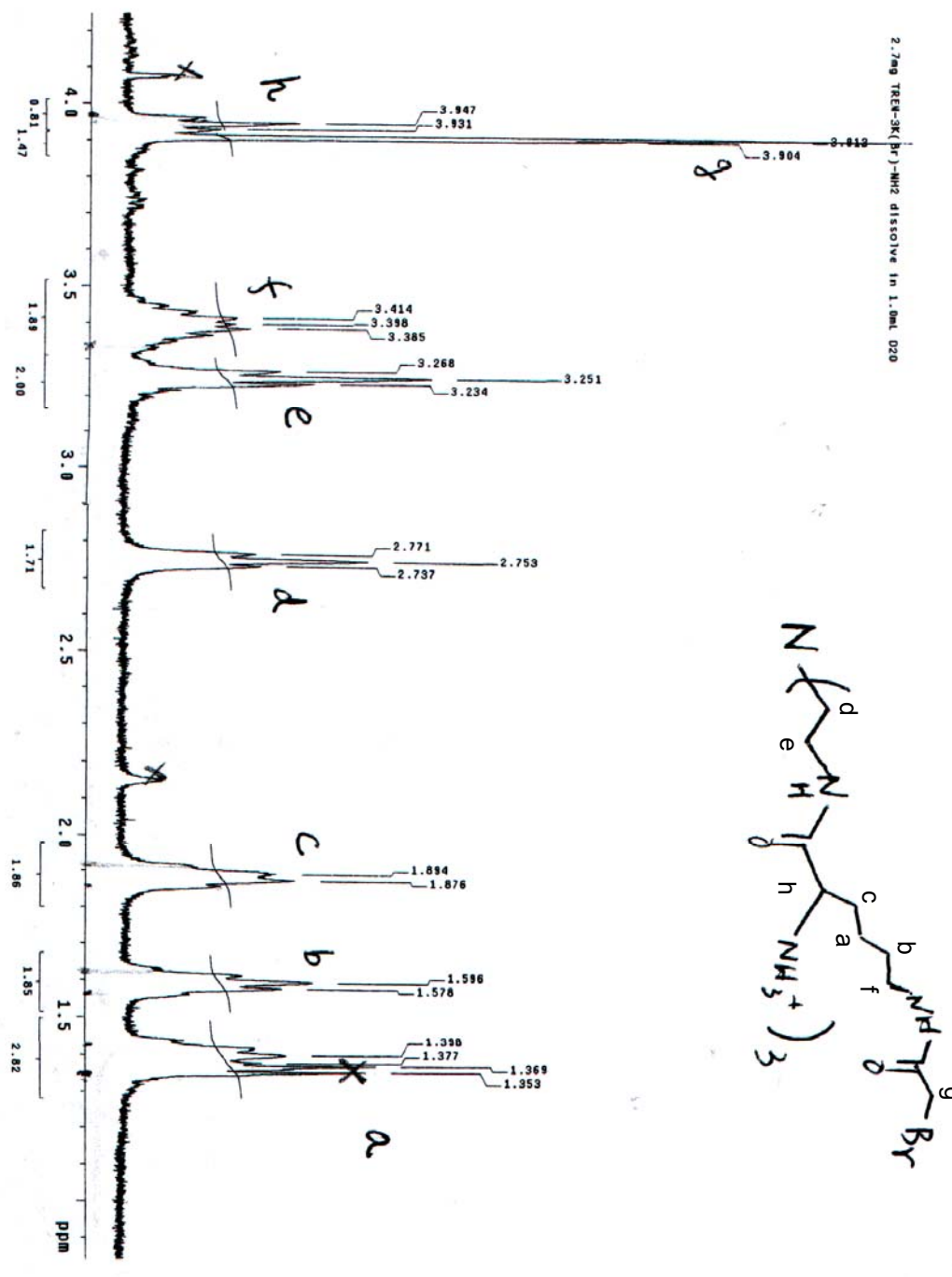


Figure 3.3 The ^1H NMR spectrum of purified template TREN-3 $[\text{K}(\text{Br})\text{-NH}_3^+]$ recorded at RT on a Varian Unity 400 MHz. (2.4 mg in 1.0mL D_2O)

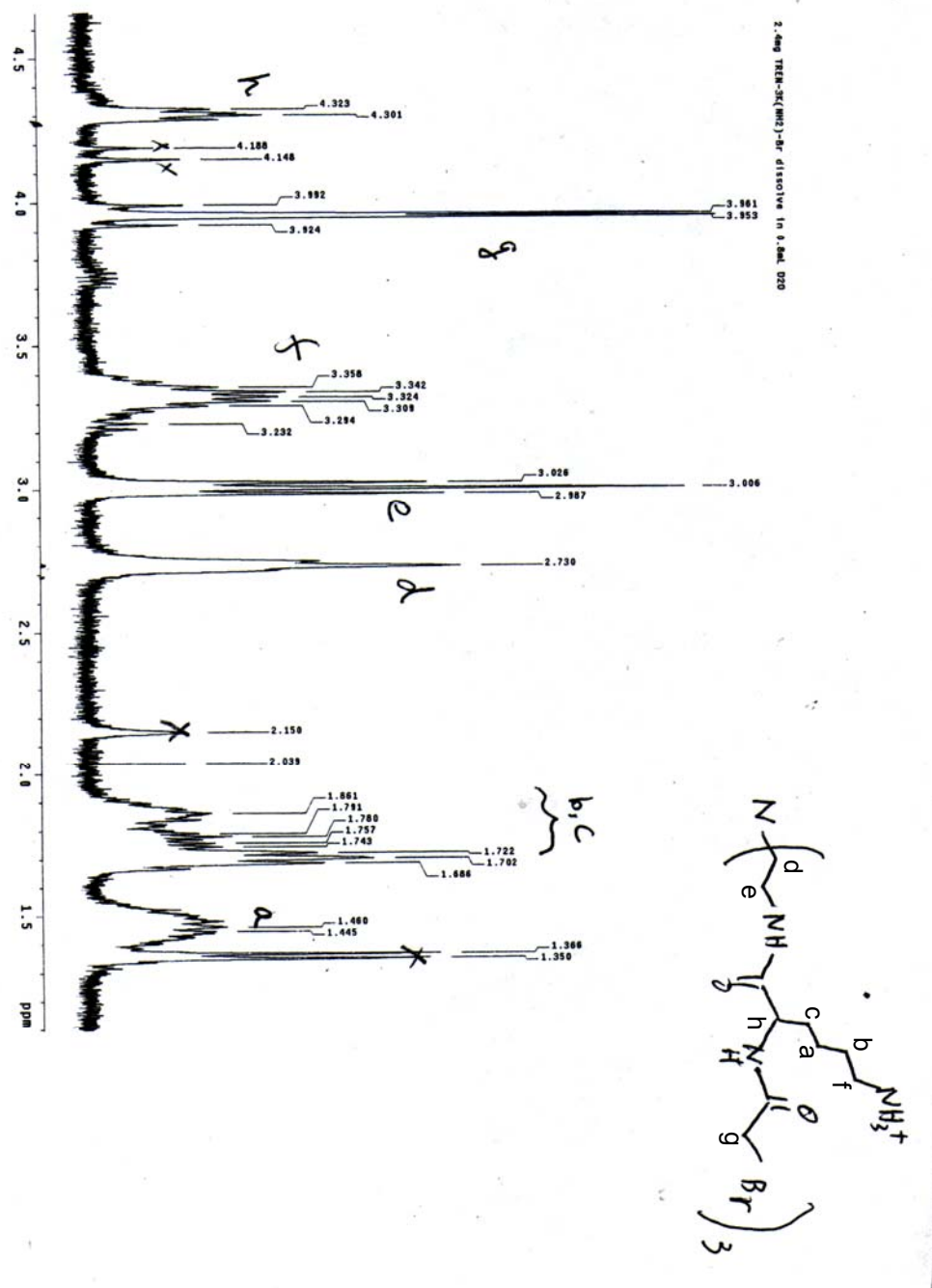


Figure 3.4 The ^1H NMR spectrum of purified template TREN-3[K(NH₃⁺)-Br] recorded at RT on a Varian Unity 400 MHz. (2.7 mg in 1.0mL D₂O)

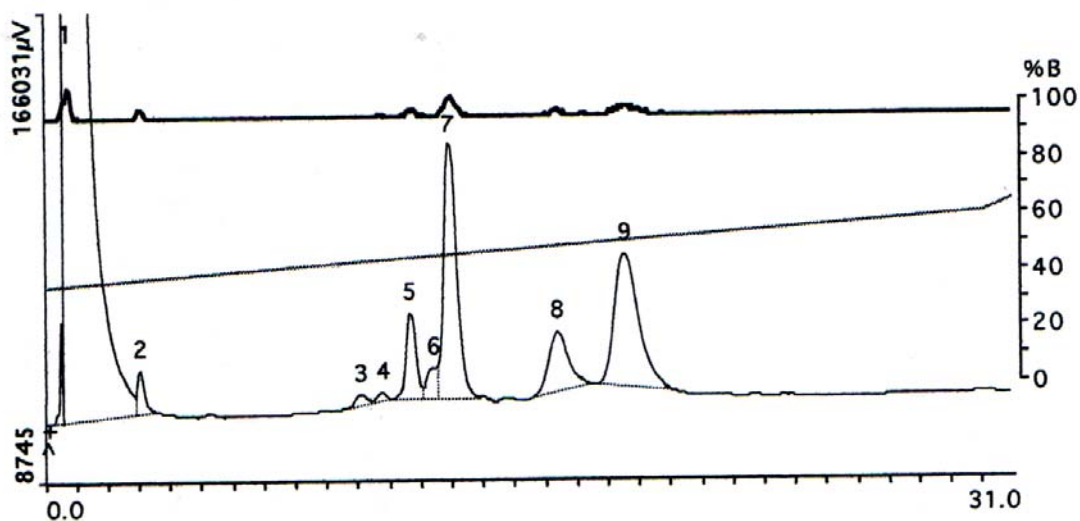


Figure 3.5.a. HPLC (reverse phase) UV trace of the reaction mixture used to prepare KTA-3N29. Column used is Waters, XTerra, MSC₁₈, 5 μM, 4.6 x 50mm. Buffer A: 0.05% TFA in H₂O, buffer B: 0.05% TFA in CH₃CN. 35% to 60% buffer B over 30 min. 1 mL/min. Solid line is the UV absorbance at 220 nm. Bold solid line is the UV absorbance at 280 nm. Dotted line indicates the eluting gradient used. Peak 9 is identified to be product KTA-3N29.

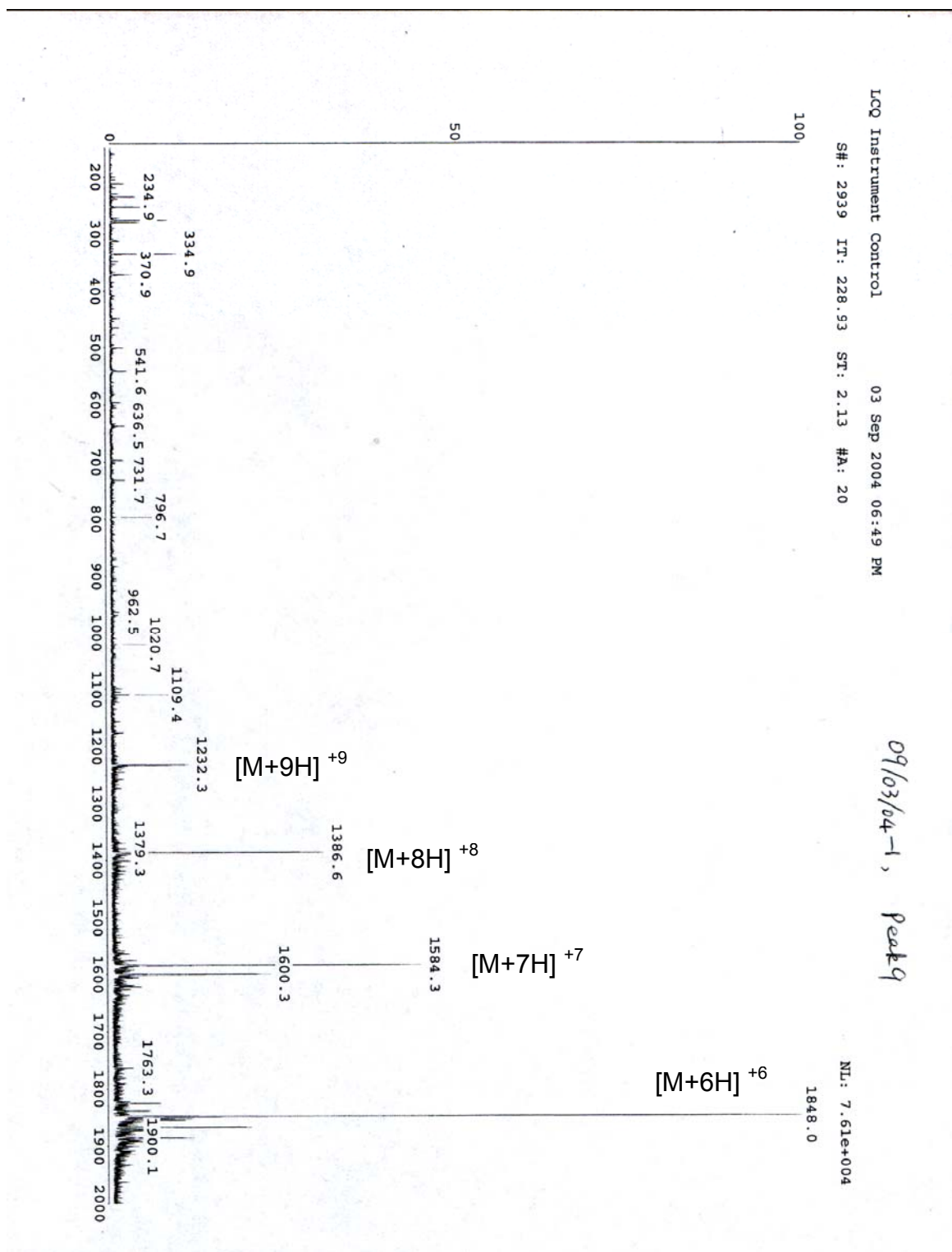


Figure 3.5.b ESI-MS of purified KTA-3N29 (peak 9 of the HPLC shown in Figure 3.5.a).

IV. Characterization of Three Helix-Bundle Model Structures

This chapter described how three-helix bundle model structures built using designed template molecules linked N-peptides, were further characterized. First, their properties, including secondary structure and conformational stability, were compared with the corresponding N-peptide sequence alone, to demonstrate the helix-stabilizing and organizational folding effects of the template molecules. Second, the monomeric state of the model structures was confirmed over a wide concentration range. Third, binding of C-peptide (C28) to selected model structures was investigated in detail, to demonstrate that our approach provides a straightforward, synthetically accessible way to assemble three-helix bundle model structures which can mimic not only the correct conformation of native N-peptide three-helix bundle, but also their proper function *in vitro*, formation of a six-helix bundle of N- and C-peptides as the end product of HIV-1 gp41 in the cell fusion process.

A. Initial design: helix-bundle model structures built using Cys-N29 (KTA-3N29 and TREN-3N29)

In this part, our initial designs, TREN-3N29 and KTA-3N29, composed of either KTA-Br (compound **4** in Scheme 3.1) or TREN-Br (Scheme 3.2), representing rigid and flexible templates, respectively, connected to peptide Cys-N29, which corresponds to the native gp41 553-581 N-peptide region, were characterized.

Untemplated Cys-N29 with a blocked thiol group on the Cys side chain, was used as a control for all comparisons.

A.1. CD spectrum and curve fitting results under neutral conditions

CD spectra of each three-helix bundle model structure and the N-peptide alone were collected in 20 mM Tris.HCl (pH = 7.40) buffer at 25 °C, using instrument conditions described in section III.C.2. Samples with 10-fold different concentrations were studied. The raw data, θ (experimental CD signal in millidegree), were then converted to mean residue ellipticity, $[\theta]$ (deg.cm².dmol⁻¹) (Figure 4.1.a-c), based on equation (4.1), where C is the sample concentration in μ M, l is the pathlength within cell in mm, and n is the number of residues in each molecule.

$$[\theta] = \frac{\theta}{C \cdot l \cdot n} \quad (4.1)$$

Comparing spectra in each data set, it can be seen that samples with about 10-fold concentration difference showed almost identical spectra, indicating that no obvious change in aggregation state occurred within the examined concentration range for KTA-3N29, TREN-3N29 and Cys-N29 alone.

Figure 4.2 shows the overlaid CD spectra of these three compounds at similar peptide concentration (the concentration of Cys-N29 was about three times that of the three-helix bundles). It can be clearly seen that, compared with the N-peptide alone, template-assembled three helix bundle model structures are much more helical. This is indicated by the greater magnitude of the maximum at

195 nm, the minimums at 208 nm and 222 nm for these templated-structures than Cys-N29 alone.

To further evaluate the secondary structures of each compound, the mean residue ellipticity (normalized) spectra were fitted by least-squares methods to linear combinations of the standard spectra of Brahms and Brahms by using LINCOMB software (1). In Table 4.1, the curve fitting results are listed for comparison. It can be seen that the helix content of the N-peptide sequence has been great increased from about 50% to over 95% when KTA-Br was used as the template molecule to create a three-helix bundle, and to about 85% when TREN-Br template was used. This obvious increase clearly demonstrates the helix stabilizing effect of our designed templates.

From the curve fitting data, it is also clear that the more rigid template, KTA-Br, is slightly better at directing helical conformation under neutral conditions than the flexible one, TREN-Br, since the KTA-Br template resulted in a model structure with more helical content and less disordered structure.

Table 4.1 Curve fitting results for CD spectra in Figure 4.2. ^{a,b}

	<i>Concentration</i>	<i>% α-helix</i>	<i>% β-beta</i>	<i>% random</i>	<i>% β-turn</i>
Cys-N29	56.9 μ M	56%	/	44%	/
		43%	16%	41%	/
TREN-3N29	16.5 μ M	83%	/	17%	/
KTA-3N29	18.5 μ M	97%	/	3%	/

^a. CD spectra were fitted to the standard spectra for individual secondary structures of Brahms and Brahms by using LINCOMB software.

^b. Samples were dissolved in 20 mM Tris.HCl buffer, pH = 7.40.

A.2. Urea denaturation and curve-fitting under neutral conditions

To compare the conformational stability of the template-assembled model structures with the untemplated N-peptide under neutral conditions (20 mM Tris.HCl, pH=7.40), their reversible folding-unfolding (native-denatured) process could be examined using standard denaturation methods, including thermal denaturation and chemical denaturation. During denaturation, a molecule with higher conformational stability will require more stringent conditions to become unfolded (denatured).

As described in literature (2), reversibility of the folding-unfolding process needs to be confirmed first, to allow an investigation of the thermodynamic parameters associated with this equilibrium. Unfortunately, thermal denaturation of all three of these templated three-helix bundles was irreversible. Figure 4.3 compares the CD spectra of each compound at 0 °C with those of the same compound after heating up to 80 °C and cooling back to 0 °C. It is clear that in each pair of spectra compared, the spectra are different, indicating that the peptide structures had irreversibly changed from a mainly helical conformation in each case to other nonhelical conformations and that therefore, thermal denaturation under neutral conditions could not be used to evaluate thermodynamics parameters associated with the folding-unfolding process.

Chemical denaturation methods were then tried. Guanidine hydrochloride (Gdn.HCl) denaturation was proved to be irreversible (not shown). However, urea was proved to reversibly denature the three-helix bundle model structures. Figure 4.4 compares the CD spectra of samples of KTA-3N29 to test whether dilution of

a solution with high urea concentration and mainly denatured (unfolded) model helix bundle would recover the native state (folded) of the molecule. From the figure, it can be seen that the diluted sample and the directly prepared sample gave approximately identical CD spectra within experimental errors, indicating that diluting 4.8 M urea, which is higher than the $[\text{Urea}]_{1/2}$ mentioned later in this section, to 1.2 M urea clearly recovered the folded conformation of the peptide to a similar extent as the directly prepared sample, and that therefore, urea denaturation of the three helix bundle model structure, KTA-3N29, is reversible.

A detailed stepwise urea denaturation experiment was then carried out as described in section III.C.4. Denaturation curves were obtained by plotting the mean residue ellipticity difference between 222 nm and 260 nm versus increasing concentration of urea, as shown in Figure 4.5, for Cys-N29, TREN-3N29 and KTA-3N29. For Cys-N29, the mean residue ellipticity at 222 nm increased gradually from about $-10,000 \text{ deg.cm}^2.\text{dmol}^{-1}$ at low urea concentrations to about $-5,000 \text{ deg.cm}^2.\text{dmol}^{-1}$ with 8 M urea, but no indication of a cooperative transition was observed. In contrast, for the three-helix bundle model structures, TREN-3N29 and KTA-3N29, a cooperative transition from a folded to an unfolded state was clearly observed with increasing urea concentration. This distinct difference between the urea denaturation of N-peptide alone and those of the template-assembled model structures demonstrates the utility of both templates, TREN-Br and KTA-Br, for organization of the three-helix bundle in TREN-3N29 and KTA-3N29, respectively, into a stable, folded peptide assembly.

The experimental data were then fitted to a model describing a two-state cooperative unfolding process, applying the linear extrapolation method (LEM) (3), and a value of the free energy of denaturation of each helix bundle structure was obtained. As proposed by Green and Pace (2, 3) (equation 4.2), the standard free energy of denaturation, ΔG_D ($\Delta G_D = -RT \ln K$), is linearly dependent on the urea concentration:

$$\Delta G_D = \Delta G_D^{H_2O} - m \bullet [\text{denaturant}] \quad (4.2)$$

where $\Delta G_D^{H_2O}$ is the extrapolated conformational stability of a protein in the absence of denaturant, and m is a measure of the dependence of ΔG_D on urea concentration. The value if m has been shown to have a good correlation with the change in accessible surface area upon unfolding (4). Curve-fitting of both three-helix bundle model structures' unfolding curves based on the above model, assuming that there is linear dependence of the CD signal on the urea concentration at both the folded state and the unfolded state, gave the results of $\Delta G_D^{H_2O}$ and m that are listed in Table 4.2, and the theoretical curves shown in Figure 4.5 (solid lines).

Table 4.2 Curve fitting results of urea denaturation studies in 20mM Tris.HCl, pH=7.40 and 25°C (unfolding process)

RESULTS	TREN-3N29	KTA-3N29
ΔG_{H_2O} (kcal.mol ⁻¹)	4.48	5.78
m (kcal.mol ⁻¹ .M ⁻¹)	0.894	1.15
[Urea] _{1/2} (M)	5.01	5.00

From the curve-fitting results, it can be seen that KTA-3N29 and TREN-3N29 have a similar tolerance to the increasing concentration of urea denaturant with both $[Urea]_{1/2}$ at 5.0 M, which is defined as the urea concentration when 50% of the molecules are unfolded and therefore is the point where the equilibrium constant of the folding-unfolding reaction, K , is 1. KTA-3N29 has a slightly bigger m value than TREN-3N29, indicating that more surface area is exposed to water during its unfolding process. Since the unfolded states of each helix bundle are likely to have the same overall peptide conformations, this indicates that KTA-3N29 has more buried hydrophobic surface in its folded state. Therefore, rigid KTA-Br template is better at inducing a more conformationally stable model structure.

A.3. Thermal denaturation studies and curve-fitting under acidic conditions

Even though thermal denaturation of Cys-N29, TREN-3N29 and KTA-3N29 under neutral conditions was proved to be irreversible (Figure 4.3), the same tests performed under acidic conditions (10 mM H_3PO_4 .NaOH buffer, pH 2.50), (Figure 4.6), proved that thermal denaturation at pH 2.50 is reversible and therefore it could be used to evaluate thermodynamics parameters associated with the folding-unfolding process.

Melting curves for the three-helix bundle model structures, KTA-3N29 and TREN-3N29, and N-peptide alone, Cys-N29, were recorded over the temperature range 0-85 °C as shown in Figure 4.7 (experiment details described in section III.C.3).

From the experiment data, presented as solid dots in Figure 4.7, the CD signals at 222 nm of all three compounds, which are the characteristic minimum indicating helix conformation and therefore is usually used as a measurement of helix, increased to about $-5,000 \text{ deg.cm}^2.\text{dmol}^{-1}$, indicating that these three molecules, composed of mostly the same peptide sequence, all unfolded to adopt similar disordered, denatured states at high temperature.

Cys-N29 gradually changed to its unfolded state with no cooperative transition as temperature increased. However, the template-assembled structures, KTA-3N29 and TREN-3N29 both changed from their fully folded states, as shown with 222 nm signal of about $-28,000 \text{ deg.cm}^2.\text{dmol}^{-1}$, to their fully unfolded states with clear cooperative transition. For both of these model structures, the melting temperature of this process (t_m), defined as the temperature at which 50% of the molecules are denatured, is as high as about 70 - 80 °C. This distinct difference between untemplated peptide and the template-assembled structures also strongly demonstrates the organizing and helix stabilizing effects of our designed template molecules.

The reversible melting curves were then fitted to the van't Hoff equation (4.3), which describes the thermodynamic equilibrium constant, K , in relation to absolute temperature, T , for a two-state transition.

$$\frac{d \ln K}{dT} = \frac{\Delta H^0}{RT^2} \quad (4.3)$$

Assuming that the enthalpy change (ΔH^0) is independent of temperature, equation (4.3) can be converted to equation (4.4), where K_1 is the equilibrium constant at T_1 , and K_2 is the equilibrium constant at T_2 :

$$\ln \frac{K_2}{K_1} = \frac{-\Delta H^0}{R} \left[\frac{1}{T_2} - \frac{1}{T_1} \right] \quad (4.4)$$

If T_1 is T_m , the melting temperature (absolute temperature), at which $K_1 = 1$ (50% denatured), equation (4.4) could be further simplified as equation (4.5) and that is the one used in all the following curve-fitting of our reversible thermal denaturatuion curves to get thermodynamic properties of the folding-unfolding process.

$$K = \exp \left[\frac{\Delta H^0}{1.987 \times T} \cdot \left(\frac{T}{T_m} - 1 \right) \right] \quad (4.5)$$

Specifically, curve fitting was carried out by assuming: 1) the signal of the fully folded state is linearly dependent on temperature; 2) the signal of the fully unfolded state shows a fixed mean residue ellipticity signal of -4680 deg.cm².dmol⁻¹. This was the maximum value obtained by fitting the shape of the melting curve of Cys-N29 to a standard function describing exponential increase to a maximum ($f=y_0+a*(1-\exp(-b*x))$), as indicated by the dotted line in Figure 4.7. At each temperature, the equilibrium constant K could therefore be obtained from the theoretical best fit curve by taking the ratio of signals resulting from these two states. The best fit of theoretical curve to experimental data by numerical analysis, shown in Figure 4.7 as solid and dash lines, was used to obtain values of the enthalpy change (ΔH^0) and the absolute melting temperature (T_m).

Table 4.3 and 4.4 list the curve fitting results for the folding process of the three-helix bundle structures, KTA-3N29 and TREN-3N29, respectively, obtained by using equation (4.5) under the above assumptions. Different temperature ranges of the data were tested in the curve fitting to examine possible minor

conformational change at low temperature. Those analyses that are based on part of the experimental data are shown in Figure 4.8, and the results are also listed in the corresponding table for comparison.

Table 4.3 Curve-fitting results for thermal denaturation of 16.4 μM KTA-3N29 at pH 2.50 analyzed using different data ranges.

	<i>0-85°C</i>	<i>20-85°C</i>	<i>35-85°C</i>
ΔH (kcal.mol ⁻¹)	-27.3	-29.3	-29.8
t_m (°C)	78.4	79.2	79.5
$[\theta]_{222}(0^\circ\text{C})$ (deg.cm ² .dmol ⁻¹)	-29206.7	-29822.4	-30120.6
$\Delta[\theta]/\Delta t$ (0°C) (deg.cm ² .dmol ⁻¹ .K ⁻¹)	95.3	113.5	120.7
ΔS (cal.mol ⁻¹ . K ⁻¹)	-77.7	-83.1	-84.6

Table 4.4 Curve-fitting results for thermal denaturation of 16.3 μM TREN-3N29 at pH 2.50 analyzed using different data range.

	<i>0-85°C</i>	<i>20-85°C</i>	<i>35-85°C</i>
ΔH (kcal.mol ⁻¹)	-31.3	-35.2	-38.6
t_m (°C)	82.6	83.2	83.7
$[\theta]_{222}(0^\circ\text{C})$ (deg.cm ² .dmol ⁻¹)	-27636.4	-28315.9	-29247.2
$\Delta[\theta]/\Delta t$ (0°C) (deg.cm ² .dmol ⁻¹ .K ⁻¹)	87.4	105.5	125.6
ΔS (cal.mol ⁻¹ . K ⁻¹)	-87.9	-98.7	-108.3

From these results (Table 4.3 and 4.4), it can be seen that:

- At pH 2.50, both model structures undergo reversible thermal denaturation with very high melting temperatures (around 70 - 80 °C) and their folding-unfolding processes fit a two-state mechanism at high temperatures. When the experiment data over different temperature ranges of 0 – 85 °C, 20 – 85 °C and 35 – 85 °C were analyzed, different results of ΔH , t_m , ΔS were obtained, suggesting that additional minor conformational changes may also occur at low temperature.
- At 0 °C, KTA-3N29 showed a greater mean residue ellipticity at 222 nm than TREN-3N29 (-29,207 versus -27,636 deg.cm².dmol⁻¹), indicating that the more rigid KTA-Br template is more efficient than TREN-Br template in generating a more helical three-helix bundle structure, KTA-3N29.
- At pH 2.50, folding of both structures is strongly enthalpy-driven with the enthalpy changes, ΔH , of about -30 kcal.mol⁻¹ or greater. As shown, during the folding process, the favorable enthalpy change of KTA-3N29 (-27.3 kcal.mol⁻¹) is smaller than that of TREN-3N29 (-31.3 kcal.mol⁻¹), and the unfavorable entropy loss of KTA-3N29 (-77.7 cal.mol⁻¹.K⁻¹) is also smaller than that of TREN-3N29 (-87.9 cal.mol⁻¹.K⁻¹). These differences indicate that KTA-Br template is better at minimizing entropy losses upon folding since it is preorganized, while the TREN-Br template is better at maximizing the enthalpy changes upon folding because of its flexibility.

A.4. Concentration dependence of the melting temperature under acidic conditions

The native N-peptide three-helix bundle is the internal core of the final, folded six-helix bundle gp41. Inside this N-peptide three-helix bundle structure, there is a hydrophobic domain to form coiled-coil. On the surface of this three-helix bundle, in contrast to the typical amphiphilic helical structures that form bundles, where hydrophilic residues help to stabilize the conformation by contacting the polar solvent, mainly H₂O, this N-peptide bundle possesses many hydrophobic residues instead, including Leu⁵⁵⁶, Ala⁵⁵⁸, Ala⁵⁶¹, Leu⁵⁶⁵, Leu⁵⁶⁸, Val⁵⁷⁰, Trp⁵⁷¹, Ala⁵⁷⁸, and Leu⁵⁸¹ on its surface to provide three hydrophobic grooves for the anti-parallel binding of three copies of C-peptides (as indicated in the crystal structure, Figure 1.3). With this many hydrophobic residues within the sequence and naturally a low solvent-exposed surface area, N-peptide by itself is expected to have only marginal amphiphilic character and intrinsically possess a high tendency to aggregate (5).

To answer the question of whether our designed template molecules indeed enhance the intramolecular interaction and generate monomeric helix bundles in solution, as intended, instead of alternatively promoting intermolecular interaction resulting in higher ordered multimers, the concentration dependence of the melting temperatures (obtained from thermal denaturation) was examined over a 6-fold range of concentrations. Experimental data were then overlaid to compare the shapes of the curves, as shown in Figure 4.9. Melting temperatures

for each compound obtained by curve fitting, over the temperature range of 0 – 85 °C, are listed in Table 4.5.

From each comparison in Figure 4.9, it can be seen that the melting curves of samples with different concentrations are superimposable and, as listed in Table 4.5, the melting temperatures for all curves are approximately the same. All these results suggest that, under these experiment conditions, the aggregation state of each three-helix bundle model structure does not change over this concentration range and KTA-3N29 and TREN-3N29 exist mostly or entirely in their monomeric states in solution.

Table 4.5 Melting temperatures (t_m) calculated from curve-fitting (0-85 °C) of melting curves obtained for samples with a wide range of concentrations

	KTA-3N29			TREN-3N29		
	40.4μM	16.5μM	6.1μM	40.8μM	16.3μM	7.0μM
t_m (°C)	78.0	77.0	76.0	82.4	82.3	81.4

B. Modified design: three-helix bundle model structures built using Cys-N29b (KTA-3N29b, and TREN-3K-N29b)

Problems in our initial designs observed during the investigation include that first, they did not allow us to investigate the properties of the three-helix bundle model structures under physiological conditions (neutral pH in the presence of salt) due to their low solubility under these conditions; second, the yield of the final coupling reaction was very low (about 18%) due to precipitation of the unreacted

peptides. In an attempt to increase the solubility at neutral buffer pH 7.40 for characterization and in basic buffer (pH 8.0) for the coupling reaction, our designs were further modified by changing the template molecules and altering the N-peptide sequence.

In our initial designs, KTA-3N29 and TREN-3N29 are more soluble in acidic buffer than in neutral buffer, indicating that changes of the net charge of the compounds may play a role in solubilizing the three-helix bundle structures. Analysis of the native N-peptide sequence (Cys-N29) shows that, the net charge of this peptide is +4.0 at pH 2.50 and +2.11 at pH 7.40. Among all the residues with charged side chains (Arg⁵⁵⁷, Glu⁵⁶⁰, His⁵⁶⁴, Lys⁵⁷⁴, and Arg⁵⁷⁹), Glu⁵⁶⁰ is the only one bearing a negative charge at neutral pH, which lowers the net charge of the entire molecule when the solution pH changes from acidic to neutral or basic. Therefore, in the modified sequence under investigations, the nonconserved residue Glu⁵⁶⁰ (at **b** position in the coiled-coil structure, refer to Figure 1.3) was replaced by Gln, in an attempt to increasing the net charge of resultant structures at neutral pH, and all the other residues were retained. This modified peptide sequence was named as Cys-N29b.

In terms of the template molecule, the template derived from TREN was further modified by inserting a Lys residue onto each arm. A free amine on either the backbone, or the side chain gave the new template TREN-3[K(Br)-NH₃⁺] or TREN-3[K(NH₃⁺)-Br], respectively, both of which would bring three extra positive charges into the final model structure and the bromide groups on either the side

chain or the backbone provide functional sites necessary for the final coupling to the thiol groups on the N-peptide analogs (Cys-N29b).

The resultant three-helix bundle structures by linking the above modified templates to the modified N-peptide sequence included KTA-3N29b, TREN-3[K(N29b)-NH₃⁺] and TREN-3[K(NH₃⁺)-N29b]. Compared to the initial designs, both coupling reactions to prepare KTA-3N29b and TREN-3[K(NH₃⁺)-N29b] resulted in clear liquid and showed increased yield (about 25%-30% after HPLC purification). TREN-3[K(NH₃⁺)-N29b] was denoted as TREN-3K-N29b later in this thesis for the purpose of simplicity. However, coupling reaction to prepare TREN-3[K(N29b)-NH₃⁺] still generated gel-like solution with low yield. Therefore this molecule was not used for further analysis.

Solubility tests confirmed that, for Cys-N29b, KTA-3N29b and TREN-3K-N29b, 100 mM KF could be introduced into the sample solutions in 20 mM Tris.HCl at pH 7.40 to mimic physiological ionic strength. KF was chosen, because it interferes less with the CD signal, and also because the use of NaCl resulted in the precipitation of three-helix bundle model structures.

In this section, the characterization of KTA-3N29b and TREN-3K-N29b was presented. N-peptide alone, Cys-N29b with a blocked thiol on the Cys side chain, was used as a control for all comparisons.

B.1. CD spectra and curve fitting under neutral conditions

CD spectra of the second –generation model structures, KTA-3N29b and TREN-3K-N29b, and the N-peptide alone, Cys-N29b, were collected in 20 mM Tris.HCl,

100 mM KF (pH = 7.40) buffer at 25 °C, using instrument conditions described in section III.C.2.

Figure 4.10 a-c show the CD spectra of each compound with the concentrations indicated in the graph. Comparing the spectra in each data set, it can be seen that, over a 10-fold concentration difference, these templated peptide model structures showed almost identical mean residue ellipticity spectra, indicating that no obvious changes in the aggregation states occurred within examined range for KTA-3N29b and TREN-3K-N29b. Slight changes of spectral shape were observed for Cys-N29b control indicating probable changes in the aggregation state of the untemplated N-peptide alone.

Figure 4.11 shows the overlaid CD spectra of the three samples with similar peptide concentrations (the concentration of Cys-N29b is about three times that of three-helix bundle concentrations). Compared with N-peptide alone, the template-assembled three helix bundle model structures are clearly much more helical.

Table 4.6 lists conformational composition results obtained by fitting the model peptide spectra to the standard data set of Brahms and Brahms using LINCONB software. For each compound, curve-fitting results with the least square root errors and positive coefficients for all components are listed. Despite slight differences in the composition between each result, overall it can be seen that the helix content has been greatly increased from about 40% for the N-peptide alone, Cys-N29b, to over 85% in both template organized structures.

Similar to our initial design, this obvious increase clearly demonstrates the folding and helix stabilizing effects of our designed templates.

Table 4.6 Curve fitting results for CD spectra in Figure 4.10 using LINCOMB software and the standard data set of Brahms and Brahms.^a

	% α -helix	% β -sheet	% random	% β -turn
Cys-N29b	27%	21%	52%	/
	40%	/	47%	13%
	43%	/	57%	/
TREN-3K-N29b	82%	/	12%	6%
	84%	/	16%	/
KTA-3N29b	87%	/	12%	1%
	87%	1%	12%	/
	87%	/	13%	/

^a. Samples were dissolved in 20 mM Tris.HCl, 100 mM KF buffer, pH = 7.40.

Compared with data listed in Table 4.1 where curve-fitting results indicated about 50% helix content for Cys-N29 and a better stabilizing effect of the KTA-Br template than the TREN-Br template, it could be seen that:

- The Gln-substituted N-peptide Cys-N29b, showed less helix content than Cys-N29 (about 40% versus about 50%). Considering the difference between these two sequences (only one substitution Glu to Gln) and buffer conditions (with or without KF), this may possibly due to the interaction between the peptide and solvent and/or a salt/ionic effect in solution, which will be discussed later. This difference in the conformation of the untemplated N-

peptide is consistent with that observed between the three-helix bundle structures built using the same template, KTA-Br, but different N-peptide sequences, KTA-3N29b and KTA-3N29 (87% versus 97%).

- By extending the arms of TREN-Br with one Lys residue, the effect of this flexible template to induce helix conformation was increased to the level observed for KTA-Br. This shows that the modified template, TREB-3K-N29b is a better template than TREN-Br, probably due to the dimension or the length of each arm of the template fitting the three-helix bundle structure better.

B.2. Urea denaturation and curve-fitting under neutral conditions

Denaturation experiments were then carried out to compare the conformational stabilities of the template-assembled model structure with that of the N-peptide alone under neutral conditions with physiological ionic strength (20 mM Tris.HCl, 100 mM KF, pH=7.40).

As found for the peptides structures having our initial design, thermal denaturation of all the second generation compounds, KTA-3N29b, TREN-3K-N29b and Cys-N29b, was demonstrated to be irreversible. However, urea was shown to reversibly denature each of the three-helix bundle model structures. Figure 4.12 presents the results of the reversibility test of KTA-3N29b. From this figure, it can be seen that spectrum of a solution diluted from a sample containing 6.76 μ M KTA-3N29b and 4.8 M urea by adding 3 fold volume of buffer and that of a solution directly prepared containing similar compositions showed approximately identical CD spectra, indicating that diluting 4.8 M urea, which is

higher than the $[\text{Urea}]_{1/2}$ mentioned later in this section, to 1.2 M urea clearly recovered the conformation of molecule to a similar extent as the directly prepared sample, and therefore the urea denaturation of the three helix bundle model structure, KTA-3N29b, is reversible.

Then, a complete urea denaturation curve was determined as described in section III.C.4. Denaturation curves were obtained by plotting the mean residue ellipticity difference between 222 nm and 260 nm versus increasing concentration of urea, as shown in Figure 4.13, for Cys-N29b, TREN-3K-N29b and KTA-3N29b. Similar results to those for our initial design were observed. For the N-peptide alone, Cys-N29b, the mean residue ellipticity at 222 nm increased gradually from about $-10,000 \text{ deg.cm}^2.\text{dmol}^{-1}$ at low urea concentrations to about $-5,000 \text{ deg.cm}^2.\text{dmol}^{-1}$ in 8 M urea, but no indication of a cooperative transition was observed. In contrast, for the three-helix bundle model structures, TREN-3K-N29b and KTA-3N29b, a two-state transition from a folded to an unfolded state was clearly observed with increasing urea concentrations. This distinct difference between the urea denaturation of N-peptide alone and that of the template-assembled model structures demonstrates the utility of both templates for organization of the three-helix bundle in TREN-3K-N29b and KTA-3N29b.

To allow direct comparisons with the structures of our initial design under the same buffer conditions and also to examine possible salt effects, urea denaturation of KTA-3N29b was also carried out in 20 mM Tris.HCl, pH 7.40 alone (without added KF). The urea denaturation curves are compared in Figure 4.14.

Best-fits for the experimental data were obtained by applying the linear extrapolation method (LEM) (3) as described in the previous section (IV.A.2). The solid lines in Figure 4.13, the solid lines and the dash lines in Figure 4.14 a-b represent the best-fit curves obtained by fitting the experimental data of the model structures (solid points) to this two-state unfolding model, assuming that there is a linear dependence of the CD signal on urea concentration for both the folded and unfolded states.

For KTA-3N29b, some conformational stabilization was observed at low urea concentrations (< 2 M), which introduced a significant uncertainty into the curve-fitting due to deviation from the assumption of a two-state transition model. Therefore, an alternative curve-fitting method was tried, where KTA-3N29b data was fitted to the two-state model assuming fixed signals for the folded and unfolded states of $-27,600 \text{ deg.cm}^2.\text{dmol}^{-1}$ (the average of the first three data points) and $-2,800 \text{ deg.cm}^2.\text{dmol}^{-1}$ (the average of the last two data points), respectively (dash-dot-dot lines in Figure 4.13 and Figure 4.14 (a)). In this way, minor conformational changes in the low urea concentration range were averaged out, so that the curve-fitting represented the properties of the major transition and the results could be compared more directly to the other systems.

The $\Delta G_D^{H_2O}$ and m values obtained from these curve-fitting analysis of all denaturation curves under investigation, including those of our initial design, are all listed in Table 4.7.

Table 4.7 Curve-fitting results for urea denaturation of 6.7 μM three-helix bundle model structure at 25 $^{\circ}\text{C}$ in 20 mM Tris.HCl., pH 7.40

	TREN-3K-N29b	KTA-3N29b		TREN-3N29	KTA-3N29
Salt [KF] (mM)	100	100	0	0	0
ΔG_{H2O} (kcal.mol $^{-1}$)	2.37	3.50 [*] (1.50)	5.07	4.48	5.78
m (kcal.mol $^{-1}$.M $^{-1}$)	0.740	0.806 [*] (0.376)	1.15	0.894	1.15
[Urea] $_{1/2}$ (M)	3.20	4.38 [*] (4.00)	4.41	5.01	5.00

* Data were obtained by fitting experimental data to the two-state model assuming fixed signals for the folded and unfolded states of -27,600 deg.cm 2 .dmol $^{-1}$ and -2,800 deg.cm 2 .dmol $^{-1}$ for KTA-3N29b, respectively. Data of KTA-3N29b in parenthesis were obtained assuming linear dependence of the CD signal on urea concentration for both the folded and the unfolded states.

From the urea denaturation curves and the curve-fitting results showing in Figure 4.13 and 4.14 and Table 4.7, it can be seen that:

- Under neutral conditions (20 mM Tris.HCl, pH 7.40) in the presence of 100 mM of KF (Figure 4.13), KTA-3N29b is much more stable toward increasing concentrations of urea denaturant than TREN-3K-N29b, with [Urea] $_{1/2}$ of 4.38 M compared with 3.20 M. The m values of KTA-3N29b and TREN-3K-N29b are essentially the same (0.803 and 0.740 kcal.mol $^{-1}$.M $^{-1}$), indicating similar buried hydrophobic surface areas in the folded structures of these two molecules.

- KTA-3N29b showed a very different denaturation curve under neutral conditions (20 mM Tris.HCl buffer pH 7.40), in the presence of KF (solid triangles), from that obtained in the absence of KF (solid circles), as in Figure 4.14 panel (a). In the low urea concentration range, a significantly bigger signal at 222 nm was observed when KF was added (about -27,500 versus about -21,000 deg.cm².dmol⁻¹), suggesting that some helix conformations of partially denatured KTA-3N29b was stabilized. In the high urea concentration range, these two curves reach the same plateau, suggesting that KF salt does not affect the fully denatured (unfolded) state of KTA-3N29b.
- Under identical neutral conditions (20 mM Tris.HCl, pH 7.40, without added KF), KTA-3N29 and KTA-3N29b, which were built by linking the same KTA-Br template to the native gp41 sequence Cys-N29, and the modified sequence Cys-N29b, respectively, denatured in different ways as shown in Figure 4.14 panel (b). The *m* values obtained by curve-fitting analysis of these two experiments are identical to each other, demonstrating that the hydrophobic cores of these two structures are essentially the same, and that the substitution of Glu by Gln does not affect the hydrophobic core of the N-peptide coiled-coil. However, the $\Delta G_D^{H_2O}$ and [Urea]_{1/2} values of KTA-3N29 are higher than those of KTA-3N29b (Table 4.7), and in the absence of urea denaturant, KTA-3N29 is observed to have greater absorbance at 222 nm than KTA-3N29b (about -25,000 versus about -22,000 deg.cm².dmol⁻¹). These results indicate that KTA-3N29b is less helical and less stable toward urea denaturation than KTA-3N29.

B.3. Thermal denaturation studies and curve-fitting under acidic conditions

For our modified designs, KTA-3N29b and TREN-3K-N29b, unfortunately, thermal denaturation under neutral conditions still proved to be irreversible. However, as for our initial designs based in the native N-peptide sequence, thermal denaturation of KTA-3N29b and TREN-3K-N29b performed under acidic conditions at pH 2.50 (10 mM $\text{H}_3\text{PO}_4\text{NaOH}$ buffer), as shown in Figure 4.15, proved to be reversible. Therefore the folding-unfolding process for these new structures could be examined at pH 2.50.

Melting curves for the three-helix bundle model structures, KTA-3N29b and TREN-3K-N29b, and for N-peptide alone, Cys-N29b, were then recorded over the temperature range 0-85 °C as shown in Figure 4.16 (experiment details described in section III.C.3).

From the experiment data, presented as overlapping solid dots in Figure 4.16, Cys-N29b gradually changed to its unfolded state with no apparent cooperative transition. However, template-assembled structures, KTA-3N29b and TREN-3K-N29b, both changed from their fully folded states to their fully unfolded states with clear cooperative transition. For both of the model structures, melting temperature of this process (t_m), is as high as about (65-75 °C). This distinct difference between untemplated peptide alone and the template-assembled model structures again strongly demonstrates the organizing and helix stabilizing effects of our designed template molecules.

Experimental data was then fitted to the van't Hoff equation, as described in equations 4.3-4.5, under the assumptions that: 1) the signal of the fully folded

state is linearly dependent on temperature; 2) the signal of the fully unfolded state shows a fixed signal of $-4650 \text{ deg.cm}^2.\text{dmol}^{-1}$, which was the maximum obtained by fitting the shape of the melting curve obtained for Cys-N29b to a standard function describing exponential increase to a maximum ($f=y_0+a*(1-\exp(-b*x))$), as indicated in Figure 4.16 by the solid line. Best fits of the theoretical curve to the experimental data by numerical analysis, shown in Figure 4.16 as dashed lines, gave values of the standard enthalpy change (ΔH^0) and the absolute melting temperature (T_m).

Table 4.8 and 4.9 list the curve fitting results for the folding process of the three helix bundle structures KTA-3N29b and TREN-3K-N29b, respectively, obtained by using the van't Hoff equation (4.5) under above assumptions. Different ranges of data were subjected to analysis to examine possible additional conformational changes at low temperature, and all results are listed in the same table for comparison.

Table 4.8 Curve-fitting results for thermal denaturation of $16.4 \mu\text{M}$ KTA-3N29b at pH 2.50 using different temperature ranges.

	<i>0-85°C</i>	<i>20-85°C</i>	<i>35-85°C</i>
<i>ΔH (kcal.mol⁻¹)</i>	-32.9	-35.3	-36.6
<i>t_m (°C)</i>	74.7	75.5	76.0
<i>$[\theta]_{222}(0^\circ\text{C})$ (deg.cm².dmol⁻¹)</i>	-28232.9	-29058.8	-29846.9
<i>$\Delta[\theta]/\Delta t$ (0°C) (deg.cm².dmol⁻¹.K⁻¹)</i>	86.0	109.0	126.9
<i>ΔS (cal.mol⁻¹. K⁻¹)</i>	-94.5	-101.3	-104.8

Table 4.9 Curve-fitting results for thermal denaturation of 16.9 μM TREN-3K-N29b at pH 2.50 using different temperature ranges.

	<i>0-85°C</i>	<i>20-85°C</i>	<i>35-85°C</i>
ΔH (kcal.mol ⁻¹)	-37.6	-40.1	-40.9
t_m (°C)	66.0	66.9	67.3
$[\theta]_{222}(0^\circ\text{C})$ (deg.cm ² .dmol ⁻¹)	-25102.6	-26107.8	-26706.5
$\Delta[\theta]/\Delta t$ (0°C) (deg.cm ² .dmol ⁻¹ .K ⁻¹)	64.8	94.6	109.4
ΔS (cal.mol ⁻¹ . K ⁻¹)	-110.9	-118.1	-120.3

From Table 4.8 and 4.9, results similar to those found for peptides forming our initial design could clearly be seen for the newer modified structures based on Cys-N29b:

- At pH 2.50, both model structures undergo reversible thermal denaturation with very high melting temperatures (65 – 70 °C) and their folding-unfolding processes fit a two-state mechanism at high temperatures. When the experiment data over different temperature ranges of 0 – 85 °C, 20 – 85 °C and 35 – 85 °C were analyzed, different results of ΔH , t_m , ΔS were obtained, suggesting that additional minor conformational changes may also occur at low temperature.
- At 0 °C, KTA-3N29b showed a greater mean residue ellipticity at 222 nm than TREN-3K-N29b (-28,232 versus -25,102 deg.cm².dmol⁻¹), indicating that the more rigid KTA-Br template is more efficient in generating a more helical three-helix bundle structure, KTA-3N29..

- At pH 2.50, folding of both structures is strongly enthalpy-driven with the enthalpy changes, ΔH , of about $-33 \text{ kcal.mol}^{-1}$ or greater. As shown, during the folding process, the favorable enthalpy change of KTA-3N29b ($-32.9 \text{ kcal.mol}^{-1}$) is smaller than that of TREN-3N29 ($-37.6 \text{ kcal.mol}^{-1}$), and the unfavorable entropy loss of KTA-3N29 ($-94.5 \text{ cal.mol}^{-1}.\text{K}^{-1}$) is also smaller than that of TREN-3N29 ($-110.9 \text{ cal.mol}^{-1}.\text{K}^{-1}$). These differences indicate that the more rigid KTA-Br template is better at minimizing entropy losses upon folding since it is preorganized, while the TREN-3K-Br template is better at maximizing the enthalpy changes upon folding because of its flexibility.

Table 4.10 Complete comparison of thermodynamic parameters of folding process of three-helix bundle model structures at pH 2.50 (using 0-85°C data)

	KTA-3N29	TREN-3N29	KTA-3N29b	TREN-3K-N29b
$\Delta H (\text{kcal.mol}^{-1})$	-27.3	-31.3	-32.9	-37.6
$t_m (^{\circ}\text{C})$	78.4	82.6	74.7	66.0
$\Delta S (\text{cal.mol}^{-1}.\text{K}^{-1})$	-77.7	-87.9	-94.5	-110.9
$\Delta G (25^{\circ}\text{C}) (\text{kcal.mol}^{-1})$	-4.2	-5.1	-4.7	-4.6

Table 4.10 lists the thermodynamic properties calculated for all of the three-helix bundle model structures by fitting their thermal denaturation curves to the van't Hoff equation. This table clearly demonstrates the difference between the more rigid template, KTA-Br, and the more flexible templates, TREN-Br and TREN-3K-Br, in terms of their stabilizing effects on the three-helix bundle model structures, as discussed above. These observation are also consistent with the

results reported by M. Goodman, in which similar strategy was used to form collagen model structures (6).

B.4. Concentration dependence of the melting temperature under acidic conditions

Once again, to demonstrate that our designed template molecules indeed enhance the intramolecular interaction and generate monomers in solution as intended, instead of alternatively promoting intermolecular interaction resulting higher ordered multimer, the concentration dependence of their melting temperatures (obtained by thermal denaturation) were examined over a 6-fold range of concentrations. The experimental data were then overlaid to compare the shapes of the curves, as shown in Figure 4.17. The melting temperatures for each compound obtained by curve fitting over the temperature range of 0 – 85 °C are listed in Table 4.11.

From each comparison of Figure 4.17, it can be seen that the melting curves of samples of KTA-3N29b and TREN-3K-N29b with different concentrations are almost superimposable, and as listed in Table 4.11, the melting temperatures of the separate samples of each structure are approximately the same (71.2 – 74.7 °C for KTA-3N29b and 63.3 – 66 °C for TREN-3K-N29b). These results suggest that, under these acidic conditions (pH 2.50), the aggregation states of each three-helix bundle model structure do not change over this concentration range and that KTA-3N29b and TREN-3K-N29b probably exist mostly or entirely in their monomeric states in solution. Furthermore, considering that there are no titratable side chains within these

modified N-peptide bundle structures (from pH 2.50 to pH 7.40), the low pH experiments support our conclusion that these three-helix bundle both exist predominantly in their monomeric forms at neutral pH also.

Table 4.11 List of t_m 's calculated from curve-fitting (0-85°C) of melting curves obtained from samples with a wide range of concentration

	KTA-3N29b			TREN-3K-N29b	
	39.8μM	16.4μM	6.5μM	16.9μM	3.8μM
t_m (°C)	74.7	74.7	71.2	66.0	63.3

C. Studies on binding of C-peptide to KTA-3N29b and TREN-3K-N29b under neutral pH conditions

After demonstrating that our designs successfully mimic the folded conformation of the gp41 core N-peptide three-helix bundle structure, we wanted to further examine if they are also able to mimic the function of the N-peptide in vitro. In this section, assays of the binding of the gp41 C-peptide analog, C28, to three-helix bundle model structures, KTA-3N29b and TREN-3K-N29b, were investigated in detail, in order to determine if these constructed model structures are able to recognize and bind to the C-peptide and form the six-helix bundles mimicking the end product of gp41 in the membrane fusion process. All experiments were performed in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer at 25 °C, using CD spectropolarimetry.

C.1. Preliminary experiments

In a preliminary experiment shown in Figure 4.18 and 4.19, the CD spectra of C28 (10 μ M) and the three-helix bundle model structure, KTA-3N29b (3.3 μ M) and TREN-3K-N29b (3.3 μ M), were collected separately, and also as a mixture of C28 and each of the templated structures at the same concentrations. The difference CD spectra (panel b of Figure 4.18 and 4.19), that were calculated by subtracting the sum of the CD spectra of C28 and each model structure from the CD spectrum of the mixture, showed substantial signal increases between 210 and 215 nm, indicating significant increases in helical conformation in each case. Given that the C28 alone (Figure 4.20) and the three-helix bundle model structure alone (Figure 4.13 and 4.16) have proved to be stable under experimental conditions, this increase in signal is probably arising from the binding interaction between these components. Therefore, the designed three-helix bundle model structure, KTA-3N29b and TREN-3K-N29b, bind C28 in aqueous buffer at physiological pH and ionic strength. Furthermore, since C28 alone gives a weak CD spectrum in aqueous buffer that is indicative of little or no helical structure (solid squares in both figures), and the templated model structures alone are nearly fully folded in an α -helical conformation (Table 4.6), the increases in helical structure observed for the mixtures provide strong evidence that this binding interaction is accompanied by induction of substantial helical conformation in C28.

Moreover, comparison of Figure 4.18 with Figure 4.19 demonstrates that a more significant helix-increasing effect occurs upon binding of C28 to KTA-

3N29b than to TREN-3K-N29b under similar conditions. This suggests that KTA-3N29b binding by C28 occurs with higher affinity and induction of a greater amount of helical structure.

C.2. Derivation of a binding model for analysis of binding isotherms

A quantitative analysis of the binding of C-peptide to our template-assembled N-peptide three-helix bundle structures, KTA-3N29b and TREN-3K-N29b, was then performed in the same physiological buffer at 25 °C.

In this assay, the CD signals that are indicative of helix content ($\theta_{222} - \theta_{260}$) were measured for mixtures of a fixed concentration of C28 (4.00 μM for KTA-3N29b and 3.00 μM for TREN-3K-N29b) and increasing concentrations of each three-helix bundle model structure. Then the corresponding CD signals for C28 alone and the model structures alone were subtracted from the CD signal of the mixture to give difference CD signals, $\Delta\theta$, as a function of the concentration of the N-peptide three-helix bundle model structure (refer to III.C.5).

The CD signal for C28 was determined to be small and constant at low concentrations, within experimental error (Figure 4.20). Therefore, we assume that unbound C28 is predominantly monomeric at all concentrations present in our binding assays (0 - 4 μM). We have also demonstrated that the three helix bundle model structures, KTA-3N29b and TREN-3K-N29b, are most likely to be monomeric in the concentration ranges used in these assays (see earlier discussion, IV.B.4). Given the exclusive presence of the monomeric forms of unbound C28 and each of the three-helix bundle structures, KTA-3N29b and TREN-3K-N29b, it is appropriate to assume that the binding of C28 to our

designed model structures follows a simple model binding isotherm in which each three-helix bundle model molecule provides n independent binding sites for C28, and that dissociation of C28 from these sites was governed by a single equilibrium (dissociation constant, K_D .) between unbound C28 and its model-structure-bound state. Therefore a theoretical equation describing this process was derived as follows.

Binding of C28 (**L**) to a three-helix bundle model structure (**P**) was studied using a constant total concentration of C28 $[\mathbf{L}]_0$, and measuring a difference CD signal at 222 nm, $\Delta\theta_i$, as a function of the total concentration of model structure $[\mathbf{P}]_i$, in each mixture. The difference CD signal at each $[\mathbf{P}]_i$, $\Delta\theta_i$, was calculated as the positive value defined by equation (4.6), where $\theta(\mathbf{mix}_i)$ is the ellipticity signal of the mixture of C28 and the model structure, $\theta(\mathbf{P}_i)$ is the ellipticity signal due to corresponding model structure alone, and $\theta(\mathbf{L}_0)$ is the ellipticity signal due to C28 alone. All ellipticity measurements were measured in mdeg.mm⁻¹.

$$\Delta\theta_i = \theta(\mathbf{L}_0) + \theta(\mathbf{P}_i) - \theta(\mathbf{mix}_i) \quad (4.6)$$

Each three-helix bundle model structure, either KTA3N29b or TREN-3K-N29b, was assumed to have an unknown number, n , of independent binding sites for C28 per molecule. With this assumption, the overall binding equilibrium would be governed by a single equilibrium dissociation constant, K_D , as described in equation 4.7, where $[\mathbf{L}]_b$ is the concentration of bound C28, and the term $n[\mathbf{P}]_i - [\mathbf{L}]_b$ is the concentration of unoccupied binding sites on the model structure.

$$K_D = \frac{(n[\mathbf{P}]_i - [\mathbf{L}]_b) \cdot ([\mathbf{L}]_0 - [\mathbf{L}]_b)}{[\mathbf{L}]_b} \quad (4.7)$$

Also, the difference CD signal, $\Delta\theta_i$, will be proportional to the fraction of total C28 that is bound to KTA3N29b, and will rise to a maximum value, $\Delta\theta_{\max}$, when all of the C28 is bound to KTA3N29b. This is described by equation 4.8.

$$\frac{\Delta\theta_i}{\Delta\theta_{\max}} = \frac{[\mathbf{L}]_b}{[\mathbf{L}]_0} \quad (4.8)$$

Using equation 4.8 to substitute for $[\mathbf{L}]_b$ in equation 4.7, gives

$$K_D = \frac{\left(n[\mathbf{P}]_i - \frac{\Delta\theta_i}{\Delta\theta_{\max}} \cdot [\mathbf{L}]_0 \right) \cdot (\Delta\theta_{\max} - \Delta\theta_i)}{\Delta\theta_i} \quad (4.9)$$

Solving equation 4.9 for $\Delta\theta_i$, the difference CD signal measured as a function of the concentration of the corresponding model structure, gives equation 4.10. Values for K_D and $\Delta\theta_{\max}$ were then determined by nonlinear least-squares fit of equation 4.10 to the experimental data, using fixed integer values of 2, 3 or 4 for n .

$$\Delta\theta_i = \Delta\theta_{\max} \cdot \left\{ \frac{[\mathbf{L}]_0 + n[\mathbf{P}]_i + K_D}{2[\mathbf{L}]_0} - \sqrt{\frac{([\mathbf{L}]_0 + n[\mathbf{P}]_i + K_D)^2 - 4 \cdot n \cdot [\mathbf{L}]_0 \cdot [\mathbf{P}]_i}{4 \cdot [\mathbf{L}]_0^2}} \right\} \quad (4.10)$$

C.3. Results of binding isotherms

After obtaining the difference-CD signals as described in III.C.5, these data were plotted in Figures 4.21 and 4.22 (solid circles), and indicate helix induction as a function of C28 binding to the corresponding model structure, KTA-3N29b (Figure 4.21) or TREN-3K-N29b (Figure 4.22).

Both plots show that C28 binding increases with the three-helix bundle model structure concentration, until all of the C28 is in the bound form, and a maximum difference-CD signal is reached. In the case of KTA-3N29b, this maximum is reached at a concentration of about 2.5 μM or higher (Figure 4.21), and in the case of TREN-3K-N29b, it reached at a concentration higher than 8.0 μM (Figure 4.22).

The binding data in both figures were then fit to the simple binding model described above (Equation 4.10), using integer values for n of 2, 3 or 4 in each case. The resulting parameters are listed in Table 4.12 and 4.13.

Table 4.12 Curve-fitting results for binding isotherm of KTA-3N29b

	$n = 3$	$n = 2$	$n = 4$
R	0.9914	0.9919	0.9840
R²	0.9828	0.9838	0.9683
K_d (μM)	0.26	0.0006	1.10
$\Delta\theta_{\text{max}}$ (mdeg.mm⁻¹)	2.84	2.83	3.04

Table 4.13 Curve-fitting results for binding isotherm of TREN-3K-N29b

	$n = 3$	$n = 2$	$n = 4$
R	0.9829	0.9858	0.9814
R²	0.9661	0.9718	0.9631
K_d (μM)	2.56	1.00	4.21
$\Delta\theta_{\text{max}}$ (mdeg.mm⁻¹)	1.79	1.68	1.84

For KTA-3N29b (Figure 4.21 and Table 4.12), the most reasonable non-linear least-squares fit was found for $n = 3$ independent binding sites for C28 per KTA-3N29b molecule, indicating that the complex formed at saturation is, at least predominantly, a six-helix bundle comprising the templated three-helix bundle core, KTA-3N29b and three copies of C28. According to this model, the dissociation constant, K_D , for each C28 binding equilibrium was 260 nM, and the maximum difference-CD signal was 2.84 mdeg.mm⁻¹. This result is represented by the solid curve shown in Figure 4.21. Small deviations of this fitted curve from the experimental data points suggest that C28 binding to KTA-3N29b may involve some positive cooperativity between the three binding sites on each N-peptide helix bundle.

For TREN-3K-N29b, the best fit was found for $n = 2$ independent binding sites on each model molecule, with very close and reasonable fits for $n = 3$, as well as $n = 4$ (Figure 4.22 and Table 4.13). In all of these three fittings, the calculated dissociation constant, K_D , for each C28 binding equilibrium was in the low-micromolar range, and the maximum difference-CD signal was around 1.8 mdeg.mm⁻¹. The difficulty in getting accurate estimate of the parameters from the curve fitting was largely because that the titration experiment had to be carried out using a concentration similar to the dissociation constant, instead of about 10-20 times higher concentrations, when more accurate information of K_D would be obtained from a binding isotherm. However, due to the poor solubility of TREN-3K-N29b and the concern that C28 would become aggregate at higher

concentration, titration experiments could not be performed with higher concentrations of C28.

According to our assumed model, given that the core three-helix bundle model structures are very stable under experimental conditions, the CD signal change observed upon C28 binding might be attributed to the conformational changes of the C-peptide exclusively. Assuming each three-helix bundle model structure possesses three independent binding sites for C28, then, the maximum difference-CD signal observed corresponds to a change in $[\theta]_{222}$ for C28 from about $-4,000 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its unbound state to $-28,800 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its KTA-3N29b-bound state and $-25,600 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its TREN-3K-N29b-bound state (Table 4.14). This change of the average $[\theta_{\text{C28}}]$ in solution during the titration experiment has also been plotted as an added curve using the right y-axis in Figures 4.21 and 4.22.

Table 4.14 Comparison of the conformation change of peptide C28 during binding to KTA-3N29b or TREN-3K-N29b (assume $n = 3$)

	<i>Bind to KTA-3N29b</i>	<i>Bind to TREN-3K-N29b</i>
$[\theta_{\text{C28}}]$ before binding ($\text{deg.cm}^2.\text{dmol}^{-1}$)	-3,400	-4,300
$[\theta_{\text{C28}}]$ after binding ($\text{deg.cm}^2.\text{dmol}^{-1}$)	-28,800	-25.600

In a brief summary, the results of binding studies of C28 to our designed three-helix bundle model structures, KTA-3N29b and TREN-3K-N29b,

representing structures linked by rigid and flexible templates, respectively, clearly demonstrated that:

- Our designed model structures have high binding affinities to the C-peptide analog, C28, with the dissociation constants (K_D) in the low-micromolar range for TREN-3K-N29b and the sub-micromolar range for KTA-3N29b.
- During the binding process, C28 changed from its mostly disordered conformation to a mostly helical conformation, contributing to the significant increase in helix content in the titration mixture.
- The more rigid template KTA-Br is a better template than the flexible template TREN-3K-Br as judged by the observations that it generates a better structured and more stable (see previous results) three-helix bundle model structure, which, binds with higher affinity to three copies of C-peptide, and generates predominantly six-helix bundle structures mimicking the native gp41 core structure.

References

1. Greenfield, N.J. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry*. 1996, **235**: 1-10.
2. Pace, C.N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* 1986, **131**: 266-280.
3. Pace, C.N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131: 266-280.
4. Pace, C.N., Shaw, K. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins: Structure, Function, and Genetics Suppl.* 2000, **4**: 1-7.
5. Lu, M., Blacklow, S.C., Kim, P.S. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature Structural Bio.* 1995, **2**: 1075-1082.
6. Kwak, J., De Capua, A., Locardi, E., Goodman, M. TREN (tris (2-aminoethyl) amine): an effective scaffold for the assembly of triple helical collagen mimetic structures. *J. Am. Chem. Soc.* 2002, **124**: 14085-14091.
7. Gochin M., Kiplin Guy R., Case M. A. A Metallopeptide assembly of the HIV-1 gp41 coiled coil is an idea recptor in fluorescence detection of ligand binding. *Angew. Chem. Int. Ed.* 2003, **42**: 5325-5328.
8. Gochin, M., Savage, R., Hinckley, S. and Cai, L. A fluorescence assay for rapid detection of ligand binding affinity to HIV-1 gp41. *Biol. Chem.* 2006, **387**: 477-483.

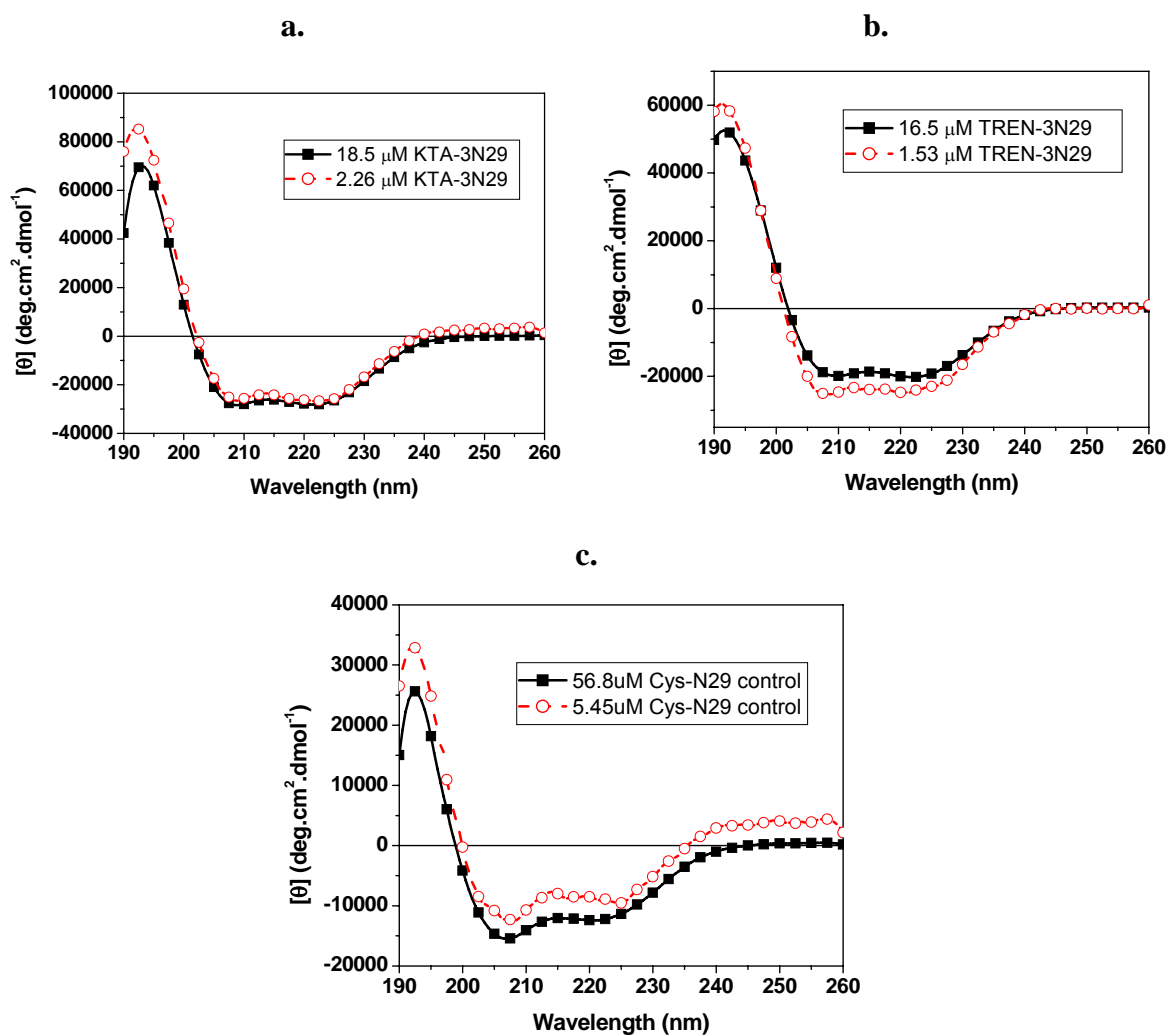


Figure 4.1 CD spectra collected at 25 °C of a) KTA-3N29, b) TREN-3N29 and c) Cys-N29 alone in 20 mM Tris.HCl, pH 7.40 with the sample concentrations indicated in each graph.

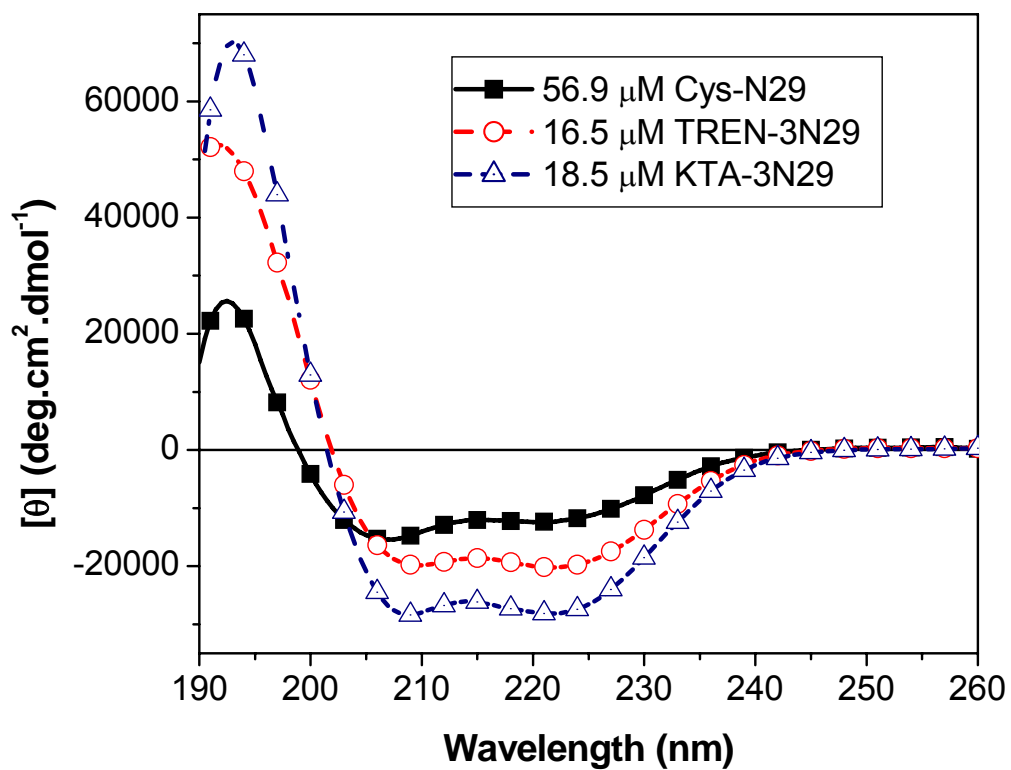


Figure 4.2 Comparison of CD spectra of model structures, KTA-3N29, TREN-3N29 and Cys-N29.

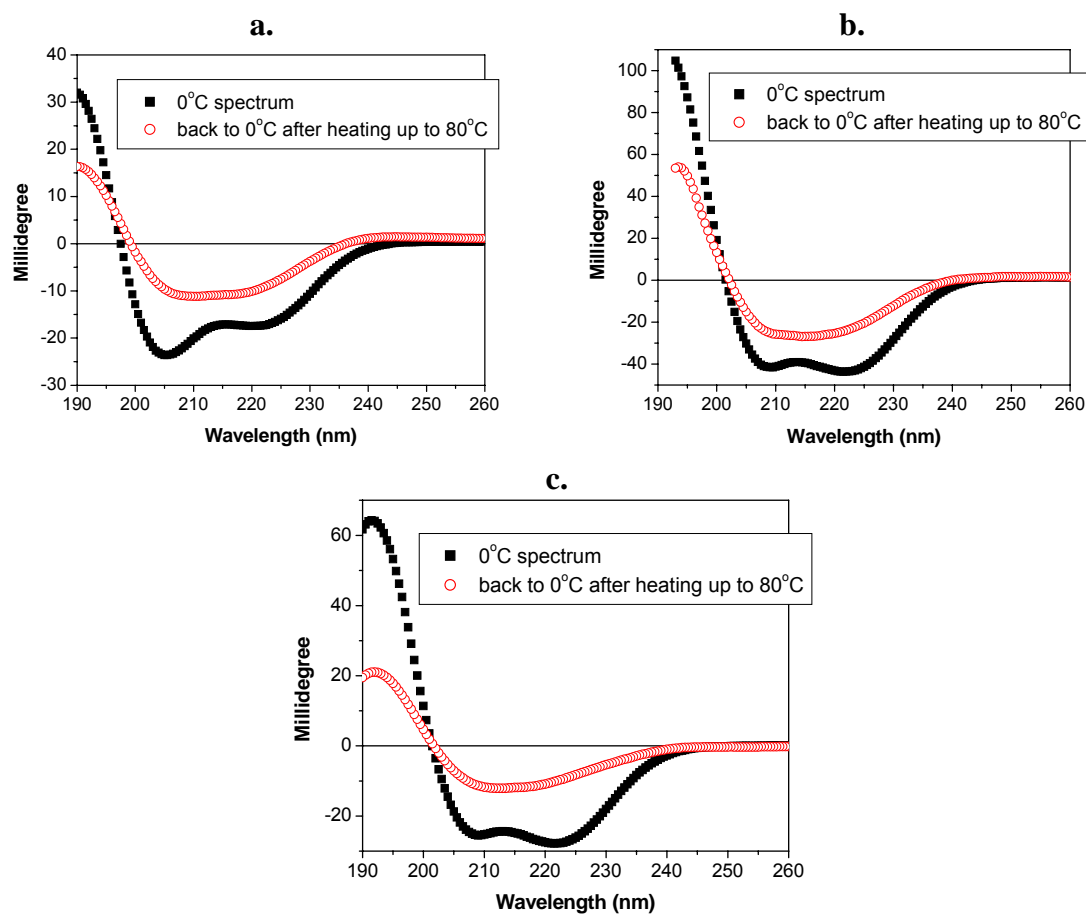


Figure 4.3 Comparison of the CD spectra collected at 0 °C before and after heating up to 80 °C, to test the reversibility of thermal denaturation under neutral conditions. a) 49.2 μM Cys-N29, b) 18.5 μM KTA-3N29b, and c) 16.5 μM TREN-3N29.

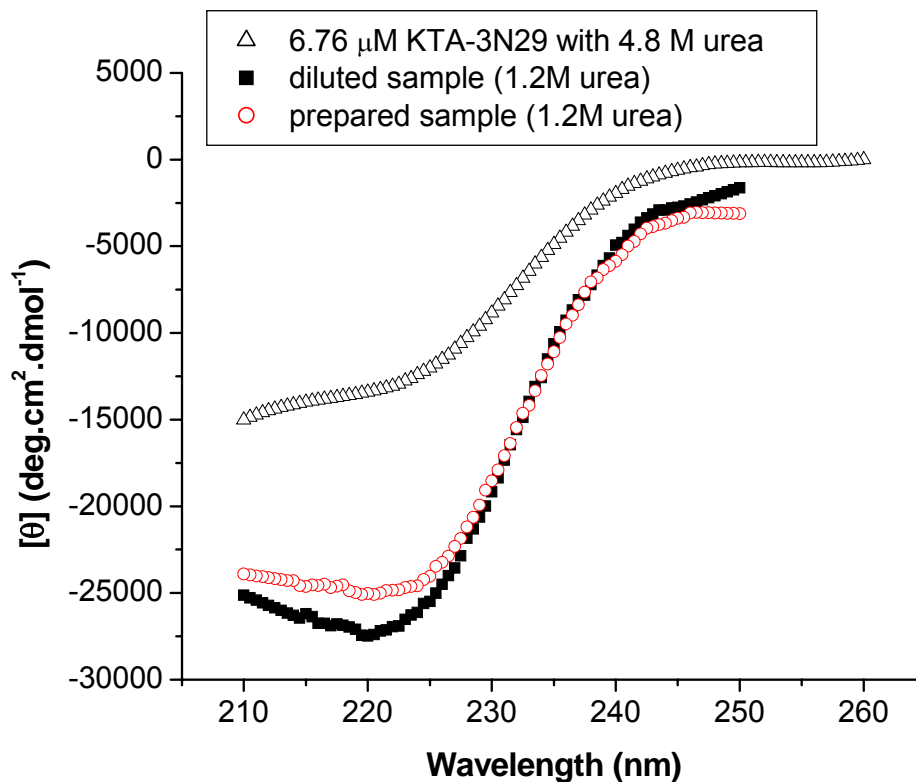


Figure 4.4 Reversibility of urea denaturation of KTA-3N29 in 20 mM Tris.HCl, pH 7.40.

Open triangles, spectrum of a sample prepared with 6.76 μM KTA-3N29 and 4.8 M.

Solid squares, spectrum of a sample diluted from a solution containing 6.76 μM KTA-3N29 and 4.8 M to about 1.69 μM KTA-3N29 with 1.2 M urea by adding three volumes of buffer. Open circles, spectrum of a sample directly prepared with 1.69 μM KTA-3N29 and 1.2 M urea.

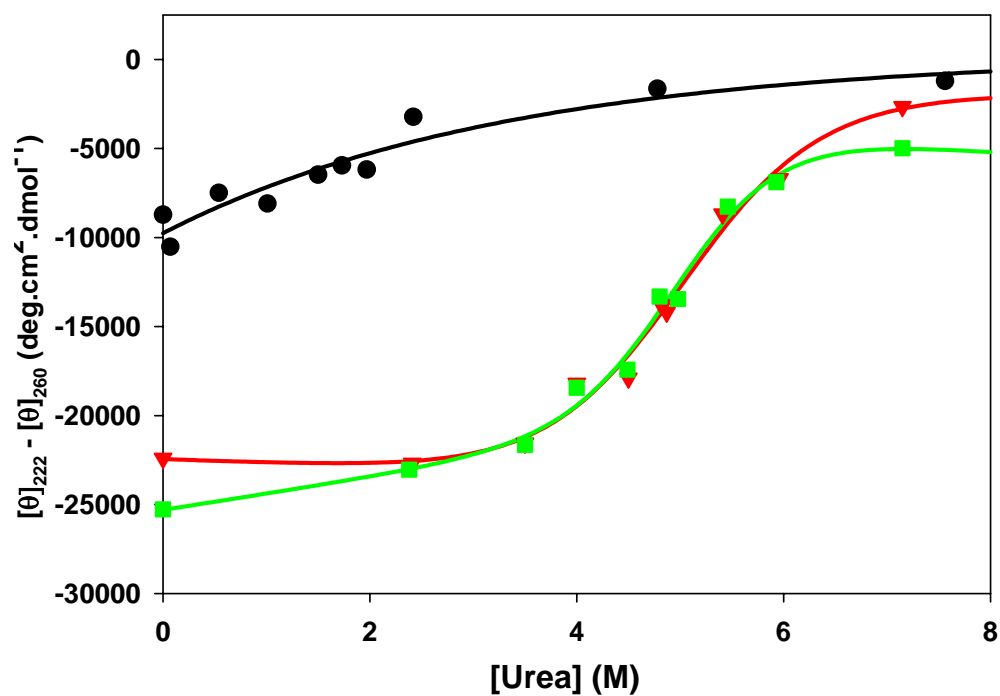


Figure 4.5 Urea denaturation studies carried out as described in III.C.4. Black circles, 20 μM Cys-N29; Red triangles, 6.7 μM TREN-3N29 and green squares, 6.7 μM KTA-3N29. The lines represent the theoretical curves of best fit (refer to text IV.A.2).

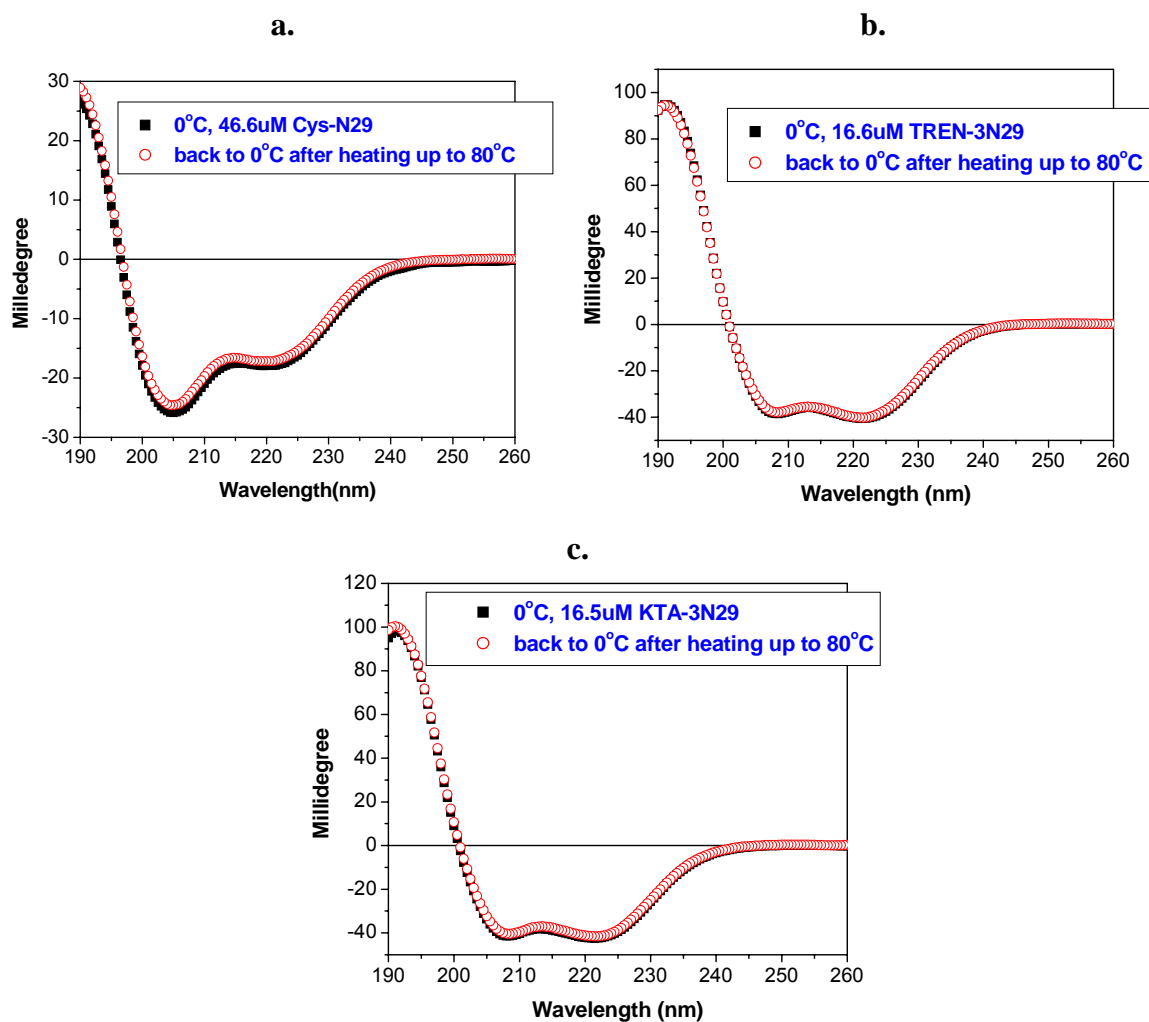


Figure 4.6 Comparison of CD spectra collected at 0°C before and after heating up to 80°C to test the reversibility of thermal denaturation under acidic conditions. a) 46.6 μ M Cys-N29, b) 16.6 μ M TREN-3N29, and c) 16.5 μ M KTA-3N29.

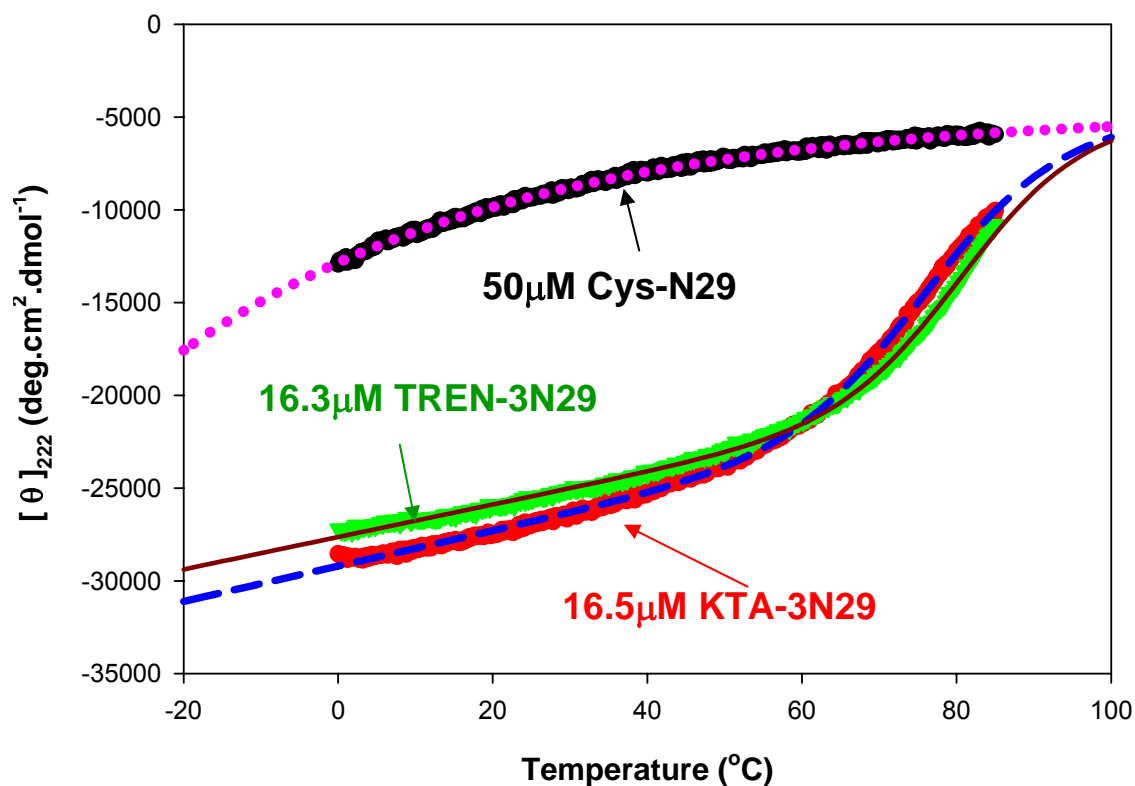
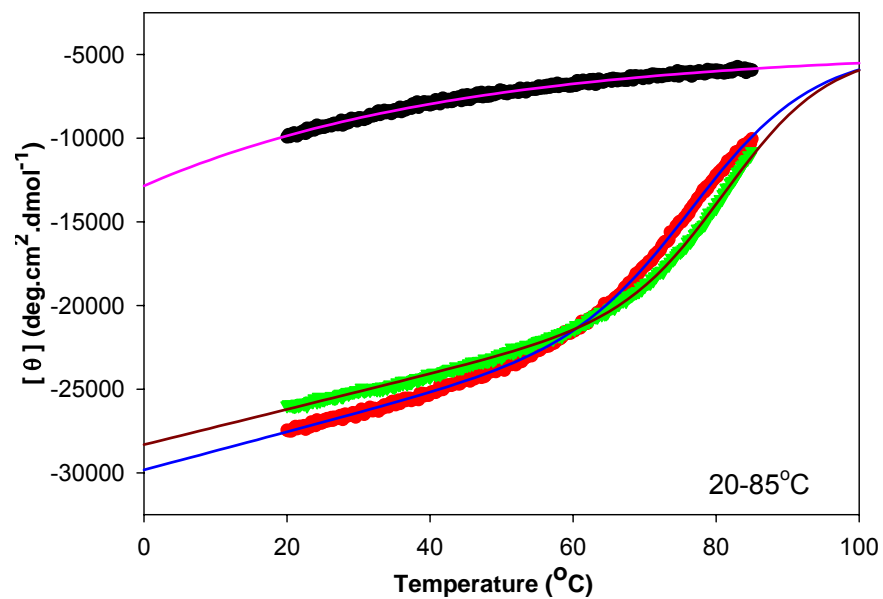


Figure 4.7 Thermal denaturation of KTA-3N29, TREN-3N29 and Cys-N29 under acidic conditions (phosphate buffer, pH = 2.50). Sample concentrations are as indicated in the graph. Solid dots are the experimental data, and lines are the theoretical curves obtained from the best fit over the temperature range of 0 - 85 $^{\circ}\text{C}$.

a.



b.

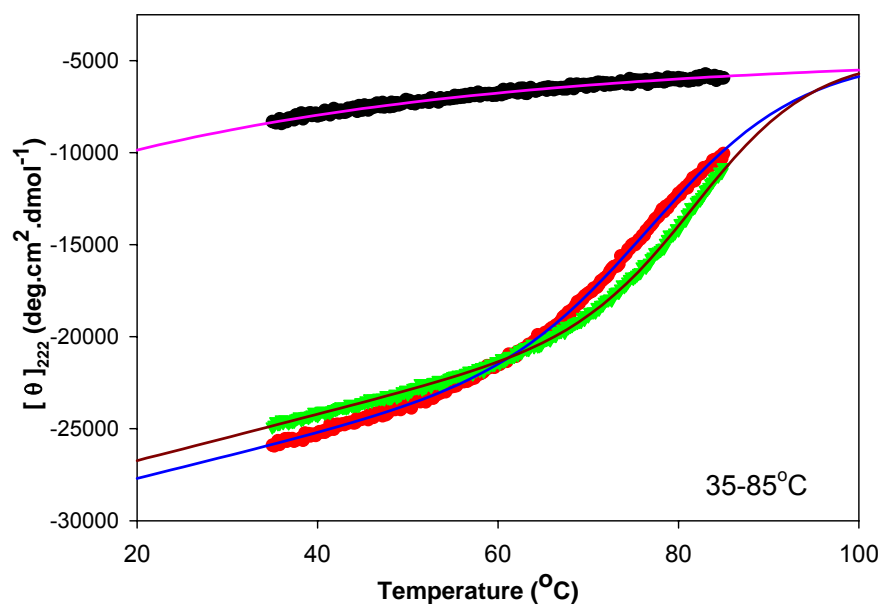


Figure 4.8 Curve fitting results in narrow temperature ranges. a) 20-85°C, b) 35-85°C

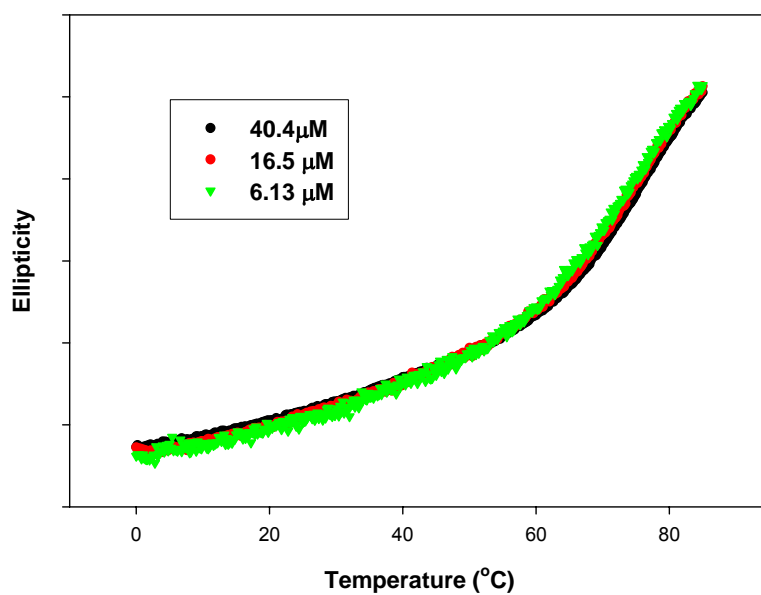
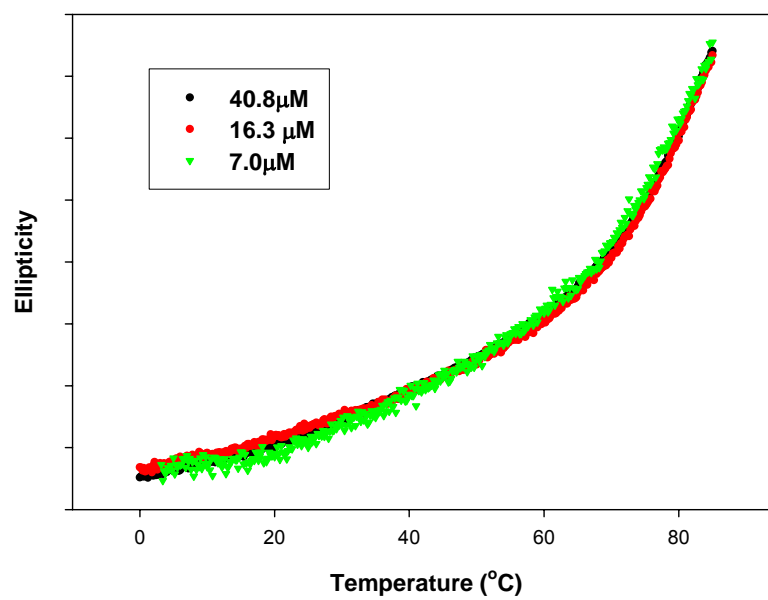
a.**b.**

Figure 4.9 Melting curves of compounds with the concentrations indicated in the graph.

For each compound, data of solutions with different concentrations were overlaid to

compare the shape of the curves. a) KTA-3N29, b) TREN-3N29

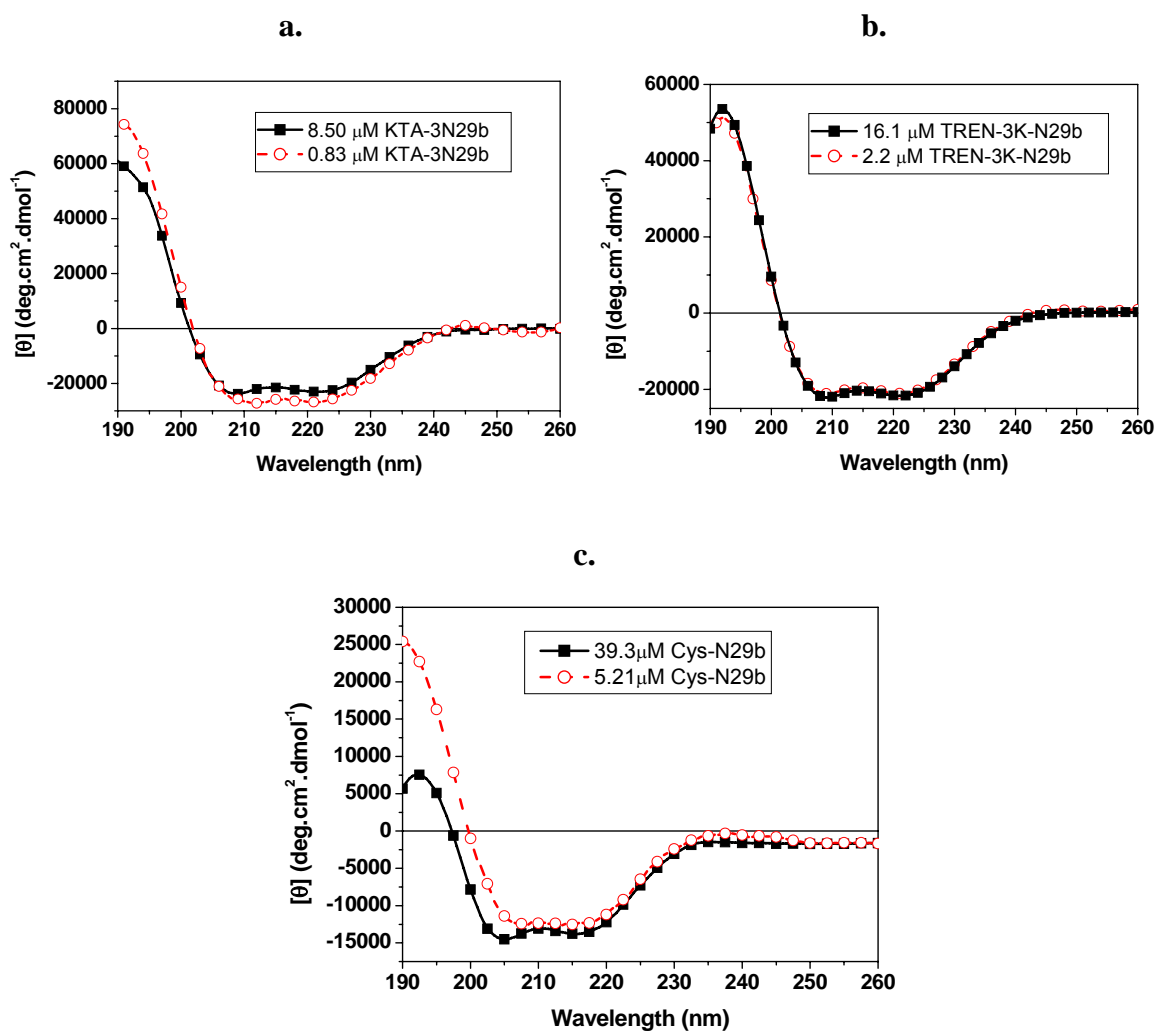


Figure 4.10 CD spectra collected at 25 °C of a) KTA-3N29b, b) TREN-3K-N29b and c) Cys-N29b alone in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer, with the sample concentrations indicated in each graph.

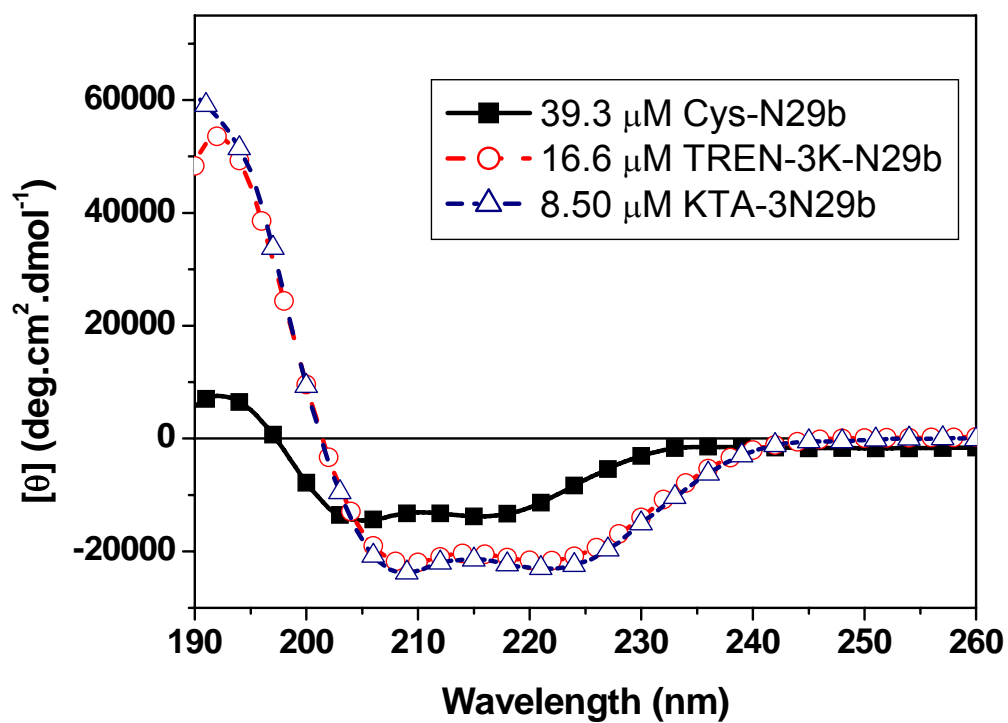


Figure 4.11 Comparison of CD spectra collected at 25 °C, in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer, of KTA-3N29b, TREN-3K-N29b and Cys-N29b.

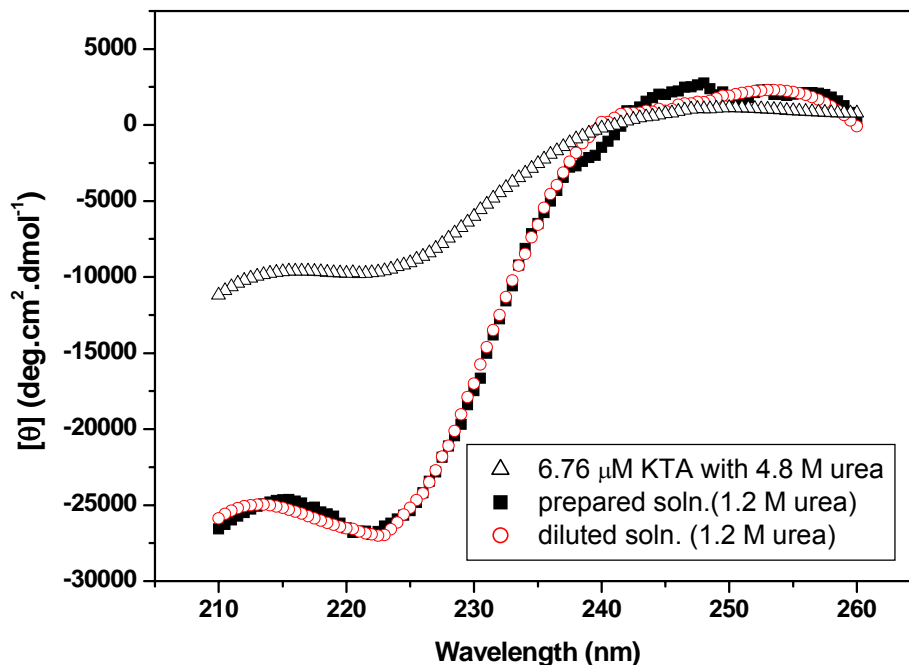


Figure 4.12 Reversibility of urea denaturation of KTA-3N29b in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer. Open triangles, spectrum of a sample containing 6.76 μ M KTA-3N29b and 4.8 M urea. Solid squares, spectrum of a sample directly prepared containing 1.69 μ M KTA-3N29b and 1.2 M urea. Open circles, spectrum of a sample diluted from a solution containing 6.76 μ M KTA-3N29b and 4.8 M to about 1.69 μ M KTA-3N29b and 1.2 M urea by adding three volumes of buffer. The spectra of the diluted and directly prepared samples are overlaid to compare the shape of the curves.

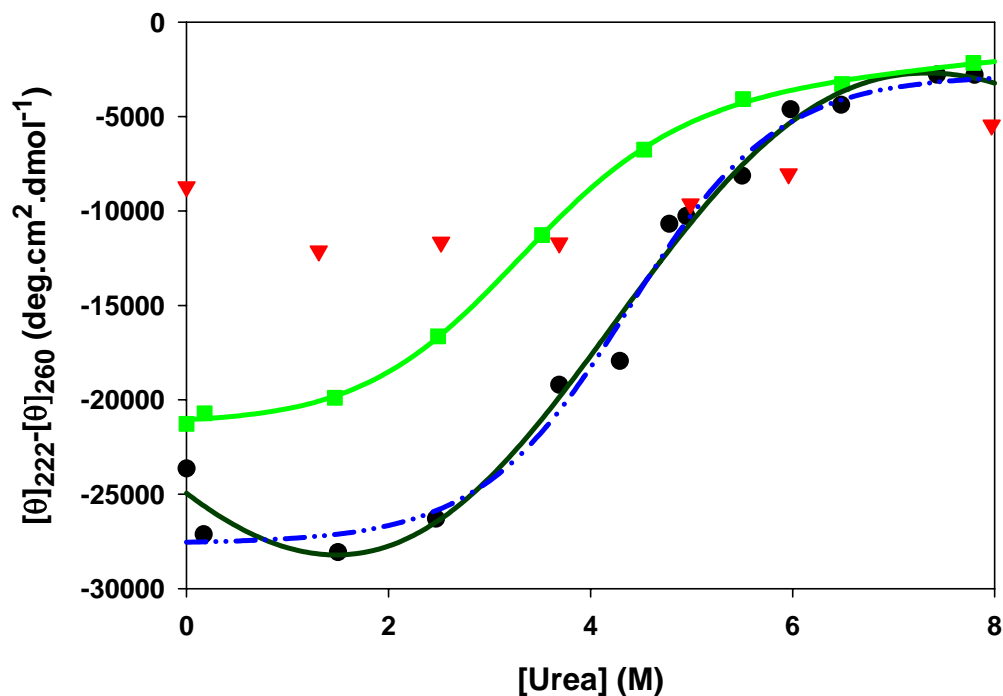
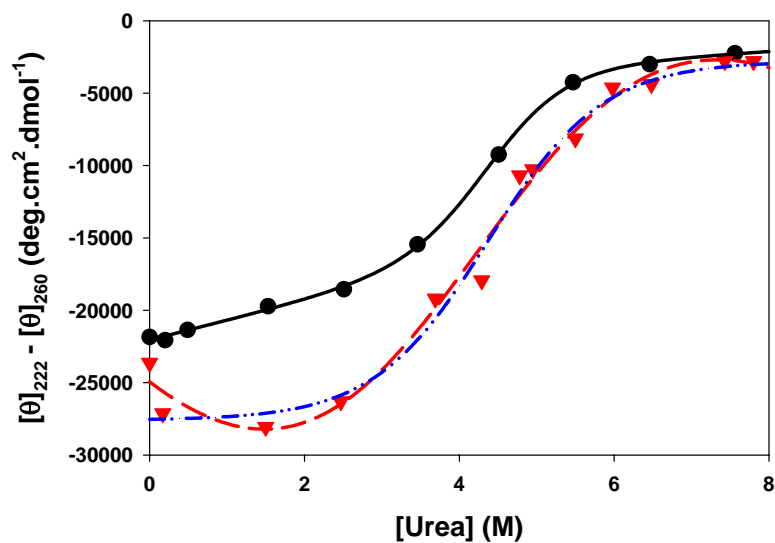


Figure 4.13 Urea denaturation studies carried out as described above in section III.C.4.

Red triangles, 20 μM Cys-N29b; green squares, 6.7 μM TREN-3K-N29b and black circles, 6.7 μM KTA-3N29b. The solid lines represent the data fit to a two-state unfolding model that assumes a linear dependence of the CD signal on urea concentration for both the folded and unfolded states (see text IV.B.2). The dash-dot-dot line represents the KTA-3N29b data fit to the two-state model assuming fixed signals for the folded and unfolded states of $-27,600 \text{ deg.cm}^2.\text{dmol}^{-1}$ and $-2,800 \text{ deg.cm}^2.\text{dmol}^{-1}$, respectively.

Resultant ΔG^{H_2O} , m, and $[\text{Urea}]_{1/2}$ values for each fit are listed in Table 4.7.

a.



b.

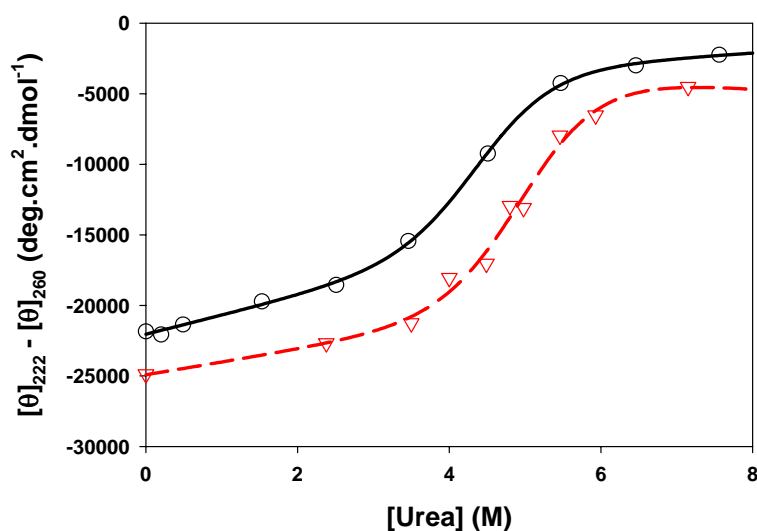


Figure 4.14 Comparison of urea denaturation curves. a) 6.6 μ M of KTA-3N29b in 20 mM Tris.HCl, pH 7.40 buffer, with 100 mM KF (solid triangles) and without added salt (solid circles). Solid line and dash line represent the theoretical curves obtained from the best-fit. Dash-dot-dot line represents the best-fit curve obtained using the alternative curve-fitting method as described in the text (IV.B.2). b) 6.6 μ M of the modified structure, KTA-3N29b (open circles), and the initial design, KTA-3N29 (open triangles), in 20 mM Tris.HCl buffer, pH 7.40.

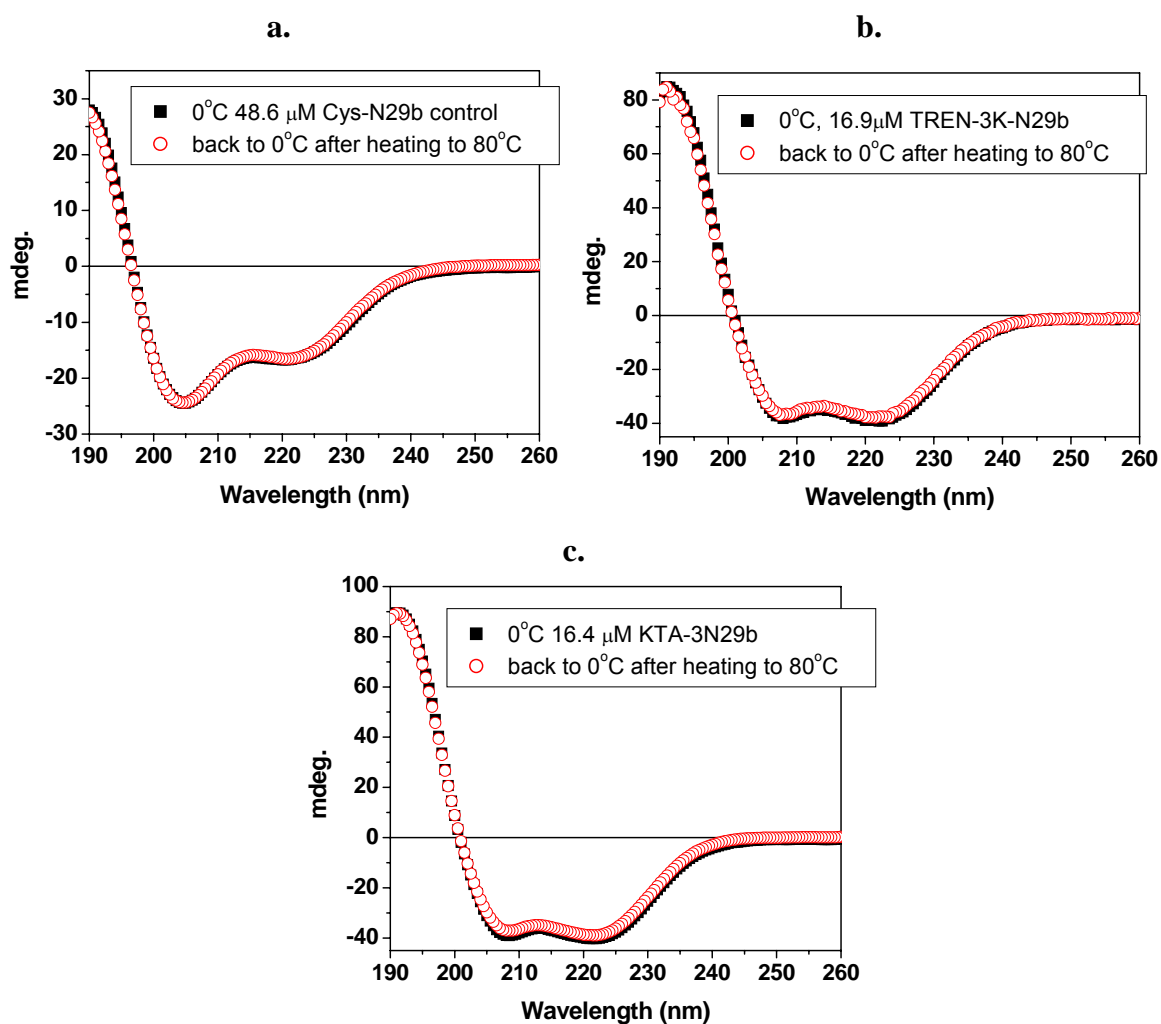


Figure 4.15 Comparison of CD spectra collected at 0 °C before and after heating up to 80 °C to test the reversibility of thermal denaturation under acidic conditions. a) 48.6 μ M Cys-N29b control, b) 16.9 μ M TREN-3K-N29b, and c) 16.4 μ M KTA-3N29b.

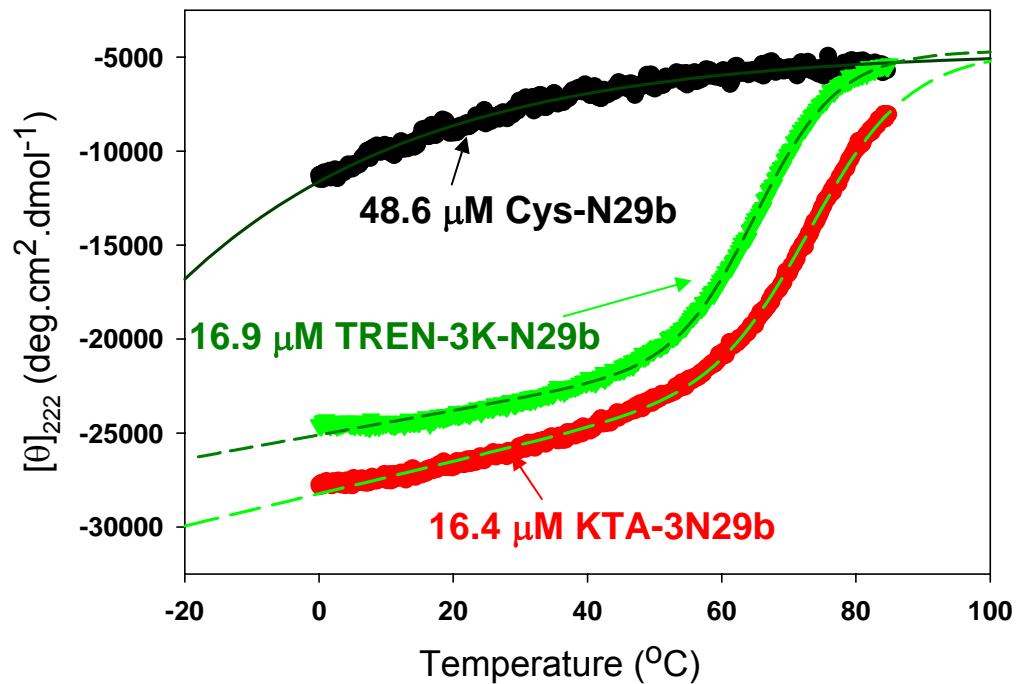


Figure 4.16 Thermal denaturation of KTA-3N29b, TREN-3K-N29b and Cys-N29b under acidic conditions (phosphate buffer, pH = 2.50). Sample concentrations are as indicated in the graph. Solid dots are the experimental data, and lines are the theoretical curves obtained from the best fit over the temperature range of 0-85 $^{\circ}\text{C}$.

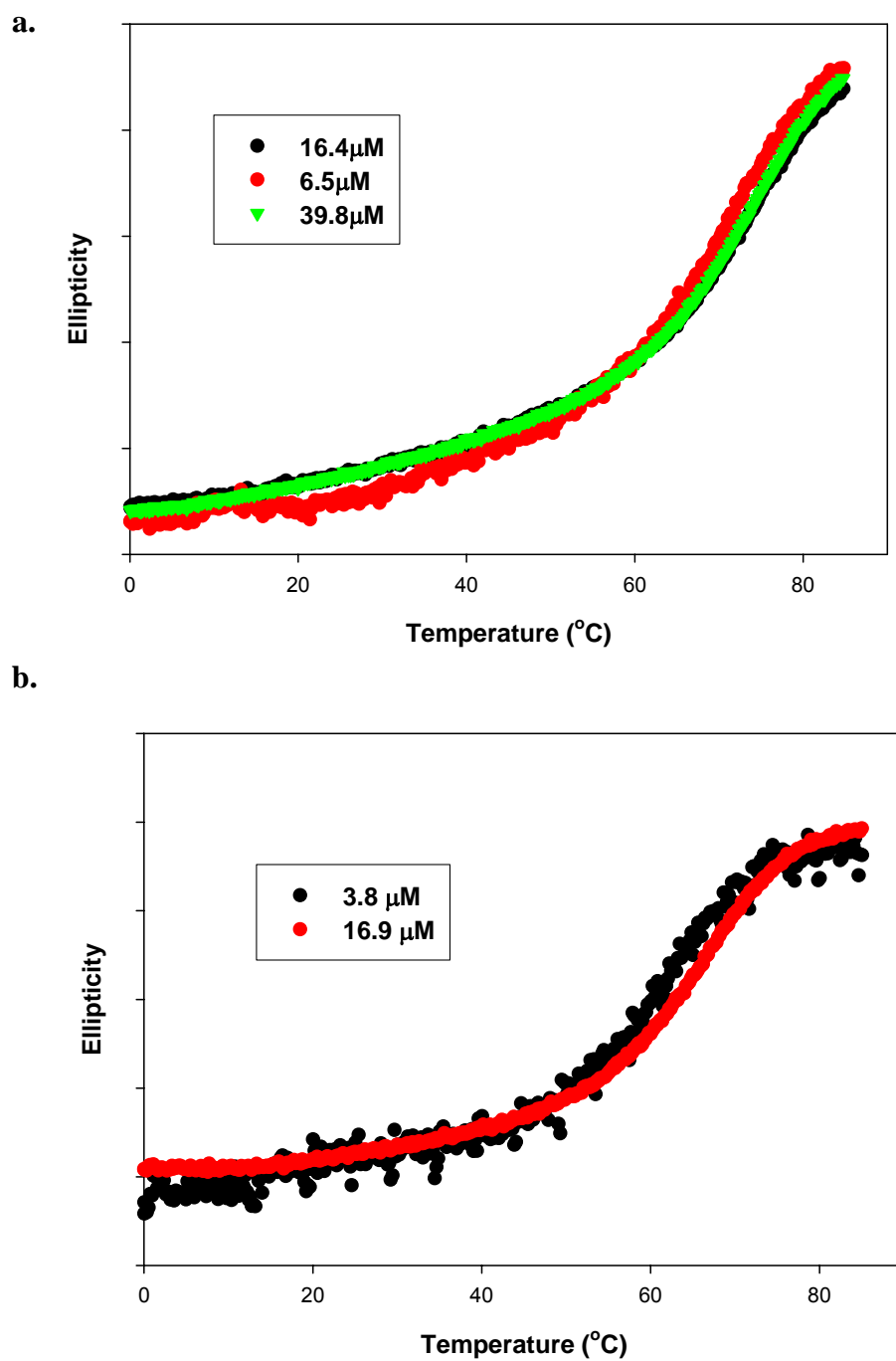


Figure 4.17 Melting curves of compounds with the concentrations indicated in the graph.

For each structure, data of samples with different concentrations were overlaid to

compare the shapes of the curves. a) KTA-3N29b, b) TREN-3K-N29b.

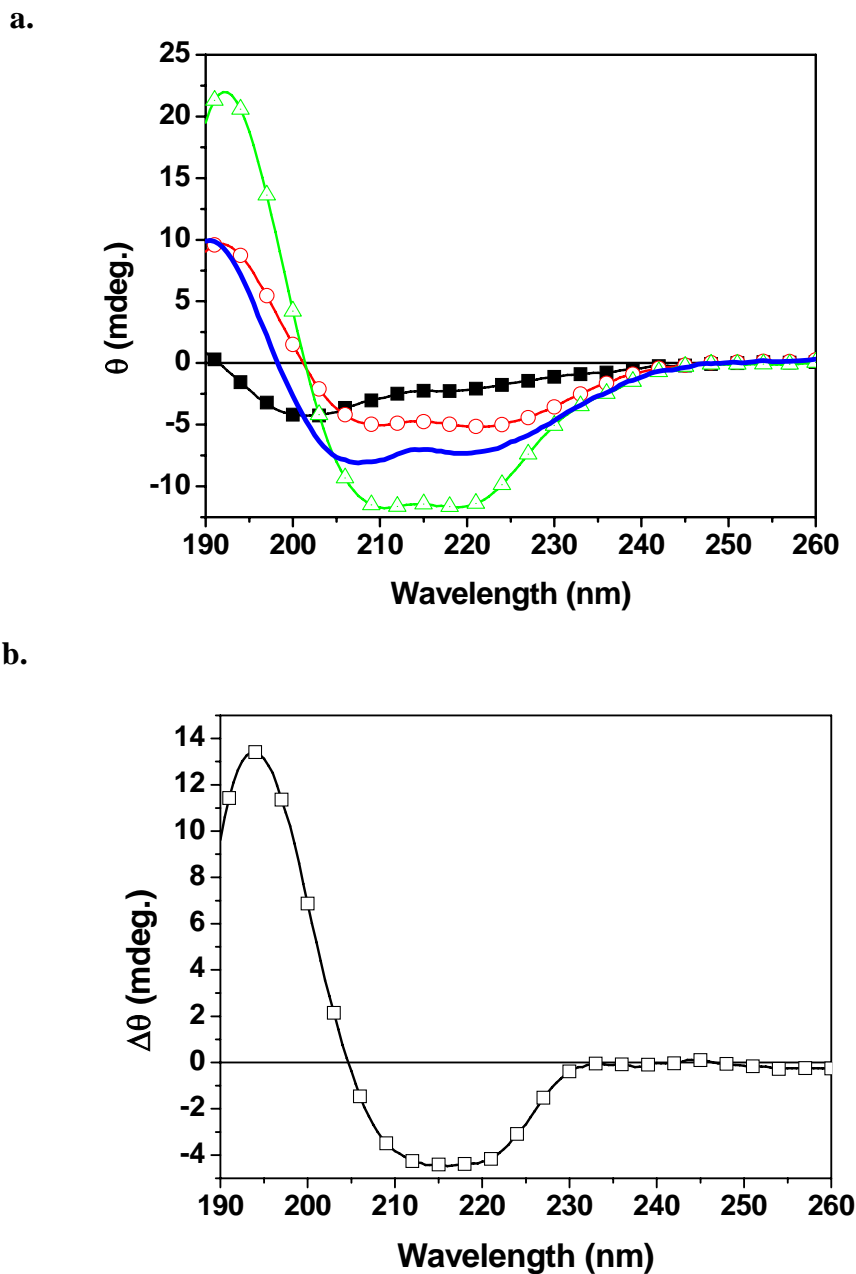
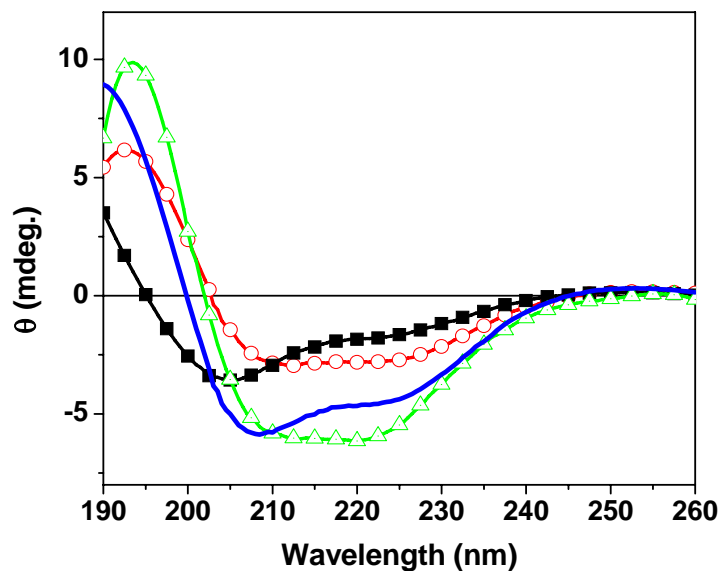


Figure 4.18 CD spectra showing C28 binding to KTA-3N29b. CD spectra were measured at 25 °C in 20 mM Tris.HCl, 100 mM KF, pH 7.40. Panel A shows (a) C28 (9.94 μ M; solid squares); (b) KTA-3N29b (3.19 μ M; open circles); (c) a mixture of 10.2 μ M C28 and 3.16 μ M KTA-3N29b (open triangles); and (d) the sum of the spectra for C28 (a) and KTA-3N29b (b) measured separately (solid line). Panel B shows the difference spectrum ($[\Delta\theta]$) calculated by subtracting spectrum (d) from spectrum (c) in panel A.

a.



b.

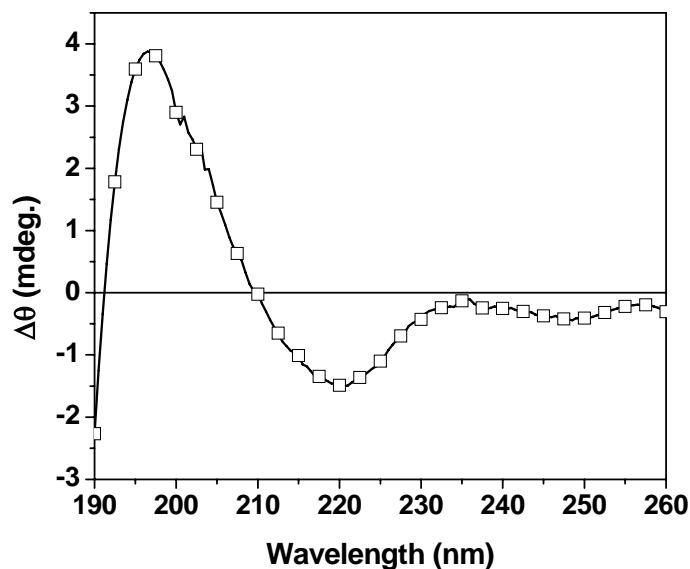


Figure 4.19 CD spectra showing C28 binding to TREN-3K-N29b. CD spectra were measured at 25 °C in 20 mM Tris.HCl, 100 mM KF, pH 7.40. Panel A shows (a) C28 (9.85 μM ; solid squares); (b) TREN-3K-N29b (3.24 μM ; open circles); (c) a mixture of 9.85 μM C28 and 3.25 μM KTA-3N29b (open triangles); and (d) the sum of the spectra for C28 (a) and KTA-3N29b (b) measured separately (solid line). Panel B shows the difference spectrum ($[\Delta\theta]$) calculated by subtracting spectrum (d) from spectrum (c) in panel A.

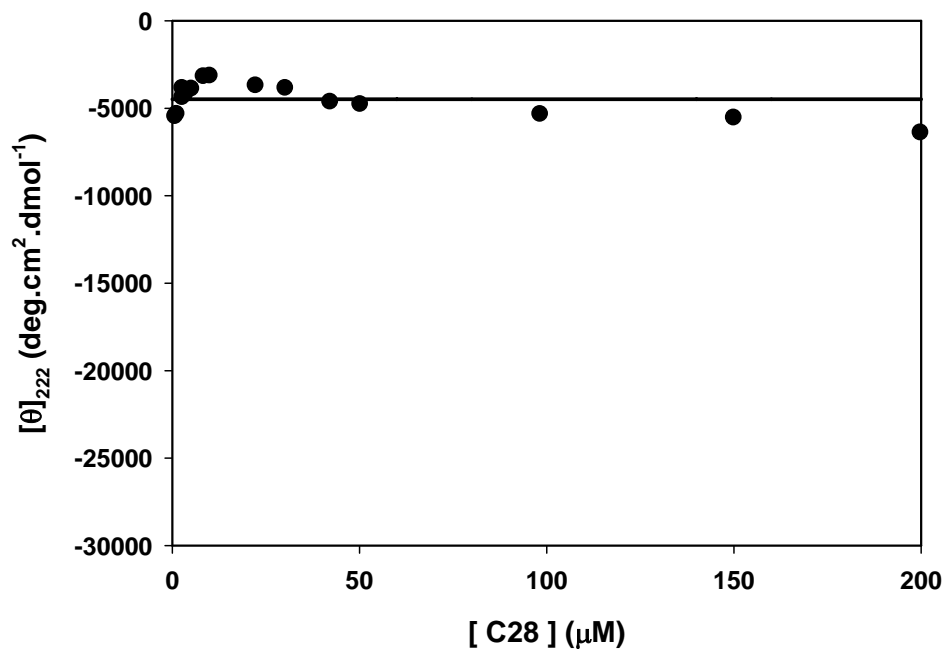


Figure 4.20 Concentration dependence of the CD signal at 222 nm of C-peptide analog C28. CD spectra were measured at 25 °C in 20 mM Tris.HCl, 100 mM KF, pH 7.40. The experimental values of $[\theta]_{222}$ are plotted as a function of C28 concentration (solid circles), together with the average value for all experiments (solid line; $[\theta]_{222} = -4,480$ deg.cm².dmol⁻¹).

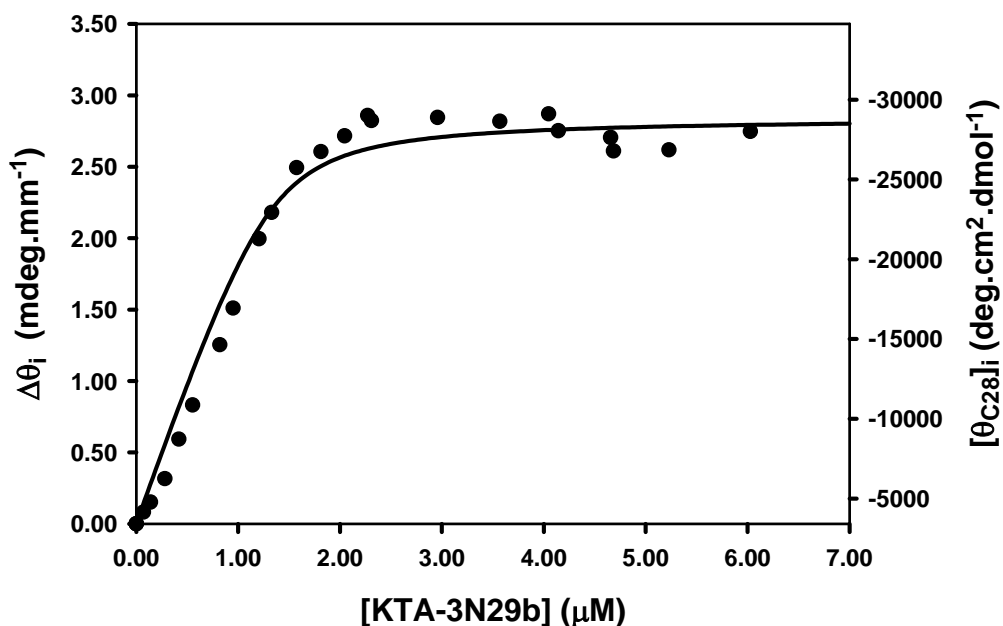


Figure 4.21 Assay of C28 binding to KTA-3N29b. Difference CD signals, $\Delta\theta_i$ (mdeg.mm⁻¹) were measured at 25 °C in 20 mM Tris-HCl, 100 mM KF, pH 7.40, for equilibrated solutions containing mixtures of a constant concentration of C28 (4.00 μM) and increasing concentrations of KTA-3N29b (solid circles). The solid line shows the calculated best fit of these data to a binding model that assumes three independent binding sites for C28 per KTA-3N29b molecule (see text).

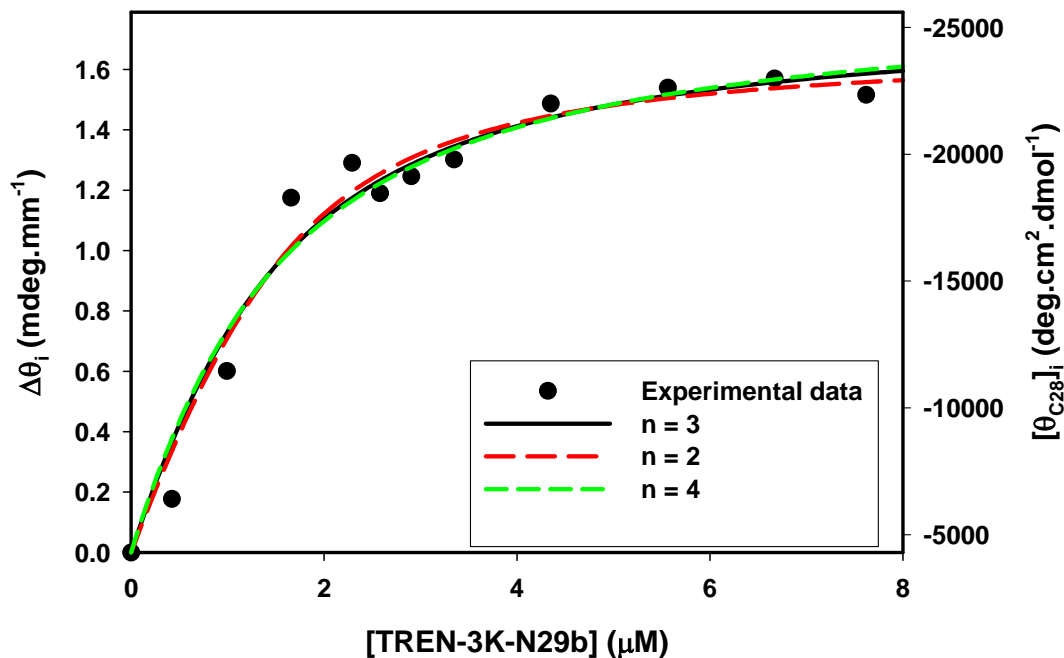


Figure 4.22 Assay of C28 binding to TREN-3K-N29b. Difference CD signals, $\Delta\theta_i$ (mdeg.mm⁻¹) were measured at 25 °C in 20 mM Tris-HCl, 100 mM KF, pH 7.40, for equilibrated solutions containing mixtures of a constant concentration of C28 (3.00 μM) and increasing concentrations of TREN-3K-N29b (solid circles). The solid line shows the calculated fit of these data to a binding model that assumes three independent binding sites ($n = 3$) for C28 per TREN-3K-N29b molecule (see text). Short dash shows the calculated curve for $n = 2$, and the long dash line shows the calculated curve for $n = 4$.

V. Discussion

In this chapter, our experimental results summarized in Chapter IV are further discussed, with respect to the conformational studies, stability studies and binding assays, with emphasis on the overall comparison between all constructed three-helix bundle model structures, KTA-3N29, TREN-3N29, KTA-3N29b, and TREN-3K-N29b . Effects of designed template molecules on directing the formation of an N-peptide three-helix bundle will also be related and compared with other groups' results.

A. Conformational studies

In the current investigation, with the intention being to build a system as small as possible for possible application of these model structures as fusion inhibitors, the peptide sequence N29 (gp41 residues 553-581) was chosen. This sequence corresponds to the minimum region of the N-peptide helix bundle that overlaps with C-peptide in the crystal structures of the six-helix bundle formed by gp41 (1-3). This peptide is shorter than the proteolysis resistant N-peptide sequences, N51 (540-590) and N36 (546-581) previously reported by Peter Kim's group, which have been proven to be highly helical, according to the X-ray crystal structures of the gp41 six-helix bundle (1-3) and biophysical analyses (4-6). Less extensive internal hydrophobic interactions are possible within the shorter N-peptides, Cys-N29 and Cys-N29b, studied here, and these peptides were only about 40% to 50% helical, as analyzed by curve-fitting, in Table 4.1 and 4.6.

In our approach, analogs of the trimeric N-peptide α -helical core of the six-helix bundle were synthesized by assembling the short N-peptide sequences, Cys-N29 or Cys-N29b onto designed three-fold symmetric synthetic scaffolds, including KTA-Br, TREN-Br and TREN-3K-Br. The rigid template, “KTA-Br” (compound **4** in Scheme 3.1) was derived from *cis,cis*-1,3,5-trimethylcyclohexane-1,3,5-tricarboxylic acid (Kemp’s triacid, KTA) (7). KTA itself was shown to favor the chair conformation in which the three carboxylic acid groups in the cyclohexane ring occupy the axial positions, while the three methyl groups are equatorial. Therefore, derivatization of each of these axial carboxylic acids by connection to the HIV-1 N-peptide was expected to provide an ideal orienting effect favoring the formation of a three-fold symmetric three-helix bundle structure that would mimic that formed by the N-peptide in the gp41 six-helix bundle (8). The flexible templates, “TREN-Br” (Scheme 3.2) and “TREN-3K-Br” (compound **12** in Scheme 3.4) were derived from *tris* (2-aminoethyl) amine (TREN), which is a three-fold symmetric molecule with a flexible straight chain of atoms on each arm. Derivatization of each of the flexible arms of TREN was expected to provide an alternative way to facilitate the formation of the three helix bundle structure.

As shown in Figure 4.2 and Figure 4.11, the resultant template-assembled three helix-bundle model structures KTA-2N29, TREN-3N29, KTA-3N29 and TREN-3k-N29b were all much more helical than the corresponding untemplated N-peptide. As compared in Table 4.1 and Table 4.6 (obtained from curve-fitting of experimental data to polypeptide standards of Brahms and Brahms), the peptide

helix content has been increased to 97% for KTA-3N29, and 83% for TREN-3N29 in 20 mM Tris buffer, pH 7.40 and about 87% for KTA-3N29b, and 84% for TREN-3K-N29b in 20 mM Tris/ 100 mM KF buffer, pH 7.40. This significant increase clearly demonstrates the utility of our designed template structures for linking and thereby facilitating the formation of three-helix bundle structures from the N-terminal of the N-peptide sequences.

As summarized in II.C and in Figure 2.8, previous approaches to constructing three-helix bundle model structures of gp41 N-peptides include (a) chimeric variants, IQN17 and IZN17 *et al.*, prepared by Kim and coworkers, in which soluble trimeric coiled coils are fused to portions of the gp41 N peptide (9, 10); (b) Nccg-gp41, N35ccg-N13 and N34ccg, prepared by Clore and coworkers, in which three intermolecular disulfide bridges were formed by cysteines inserted into the native sequence of gp41 (11, 12); (c) (CCIZN17)₃ prepared by Bianchi and coworkers, in which both a fused coiled coil and intermolecular disulfide bonds were introduced in order to direct the assembly of the N-peptide three-helix bundle (13); and (d) Metal-env 2.1 prepared by Gochin *et al.*, in which a metal-ion induced complex was formed by tris-bipyridyl groups added onto the N-terminal end of the N-peptide sequence (14, 15).

The helix content of our designed model structures is higher than that of N35ccg-N13 and N34ccg (only about 70% at pH 3), is comparable to that of metal-env 2.1 (about 90% helical) and Nccg-gp41 (about 90% helical), and is slightly less than that of IZN17 and (CCIZN17)₃ (about 99% helical). Even though estimations of helix content may not be done by the same analysis method,

these differences, to some extent, suggest that our designed template is better at stabilizing helix conformation than systems where N-peptide are covalently linked using a cysteine knot alone, and comparable to the metal coordination complex linker, which may also be considered as a templated system.

Then, which type of template works better, the rigid one or the flexible ones? In general, rigidity and flexibility contribute to different aspects of the function of the templated structures and they both have advantages and disadvantages. Rigidity provides better orienting effects on directing the formation of the desired symmetry, and yet more constraints when the ideal dimension is not achieved. Flexibility allows more freedom for the molecules to adjust their relative positions, and yet deviation from the desired orientation is also concomitant with it. Therefore, to design a good template molecule, a balance between rigidity and flexibility needs to be achieved.

From the results in Table 4.1 and Table 4.6, it could be seen that, in our case, the template KTA-Br derived from the rigid KTA molecule is better at inducing helical structure than the flexible templates under neutral conditions (20 mM Tris.HCl, pH 7.40) with or without added KF salt. In our design, the rigid cyclohexane ring of Kemp's triacid (KTA) provides the ideal three-fold symmetric orientation for the model structure. Meanwhile, extension of the three axial carboxylic acid groups out from the cyclohexane ring by insertion of 1,2-diaminoethane as a spacing group allows some flexibility for proper folding of the N-peptide sequences connected later.

It was also noticed that TREN-3K-N29b has similar helix content to KTA-3N29b (both are about 85% in Table 4.6), while TREN-3N29 has less helix content than KTA-3N29 (83% versus 97% in Table 4.1). These results suggest that template TREN-3K-Br (similar to KTA-Br) is better at generating a helical structure than TREN-Br (worse than KTA-Br). This is probably because the longer arms of TREN-3K-Br match the dimensions of the helical conformation better and therefore facilitate the formation of a more helical three-helix bundle.

B. Stability studies

In this study, the stabilities of our designed three-helix bundle model structures were evaluated by standard denaturation methods, including thermal denaturation (in 10 mM $\text{H}_3\text{PO}_4\cdot\text{NaOH}$ buffer, pH 2.50) and urea denaturation (in 20 mM Tris.HCl buffer, pH 7.40 with or without 100 mM KF). The reversible folding-unfolding equilibrium of each model structure and untemplated N-peptide control were followed by measuring their CD signals at 222 nm (Figures 4.5, 4.7, 4.13, 4.14 and 4.16). The experimental data were then fitted to the van't Hoff equation and, from the best fit, thermodynamic properties, including ΔH^0 , ΔS^0 and ΔG^0 , were obtained (summarized in Tables 4.7 and 4.10).

Due to the irreversible conformational changes in our model structures, which formed β -sheet structures under neutral conditions (20 mM Tris.HCl, pH 7.40) upon heating to over 70°C, thermal denaturation studies could only be carried out under acidic conditions (10 mM $\text{H}_3\text{PO}_4\cdot\text{NaOH}$, pH 2.50). It has previously been shown that (a) the structure and oligomerization state of gp41 are not affected by low pH (16); (b) N36 and C34 form a similar gp41 six-helix

bundle core structure at both low pH (17) and at physiologic pH (18) (by Sia *et al.*); and (c) the analogous gp41 of simian immunodeficiency virus (SIV) folds into a six-helix bundle core in either pH 7.5 (19) or 3.0 (20). Therefore, it is very likely that our data obtained under acidic conditions also represents similar properties of the same molecule under neutral conditions. In addition, our modified sequence Cys-N29b, in which the only nonconserved residue bearing negative charge, Glu⁵⁶⁰, was replaced by Gln, has no titratable side-chain in the pH range 2.5 – 7.4 and, therefore, differences in buffer pH are not expected to affect the properties of templated molecules incorporating this peptide during their folding-unfolding reaction.

According to the melting curves obtained between 0 °C and 85°C , presented as solid dots in Figures 4.7 and 4.16, both of the N-peptides (Cys-N29 and Cys-N29b) gradually changed to their unfolded states with no cooperative transition. However, all template-assembled structures changed from their fully folded states to their fully unfolded states with a clear cooperative transition and melting temperatures (t_m) as high as about 70 – 80 °C. This distinct difference between untemplated N-peptide alone and the template-assembled model structures strongly demonstrates the helix-bundle organizing and stabilizing effects of our designed template molecules.

Fitting these melting curves to the van't Hoff equation, as summarized in Table 4.10, indicated that folding of all the templated model structures is strongly enthalpy-driven. The rigid KTA-Br template was better at minimizing entropy loss upon folding since it is capable of preorganizing the three N-peptide chains due

to its peptide orienting effects, while the flexible templates derived from TREN were better at maximizing the enthalpy changes upon folding because their flexibility allows more relaxed accommodation of the peptide chains into their lower-energy conformation. This observation is consistent with Goodman's discussion on the difference between the rigid (KTA-derived) versus the flexible (TREN-derived) linkers that they used to construct triple helical collagen mimics (21). In their study, they demonstrated that TREN-derived flexible templates were more energetically favored templates than KTA-derived rigid templates at stabilizing a 30-residue collagen-like triple helix. Rigid templates did not work well in their case, probably because that collagen possess a triple helical conformation having a one-residue shift between each chain, and such that a flexible linker is more accommodating of the asymmetric factor arising from this shift.

From thermal denaturation studies (pH 2.50), the melting temperature, t_m , defined as the temperature when 50% of the molecules are denatured, is very high for all of the model structures. These were over 75 °C for KTA-3N29, TREN-3N29 and KTA-3N29b, and 66.0 °C for TREN-3K-N29b (Table 4.10). Among other constructions of N-peptide models reported previously in the literature (9 – 15) (also refer to II.C and Figure 2.8), only the thermal denaturation of the IQN/IZN17 series and (CCIZN17)₃ series have been characterized showing t_m 's over 90 °C. Based on these results, the authors have claimed that their structures have extraordinary stability (13). The slightly lower t_m values measured for our structures indicates that our model structures have a thermal stability

similar to those reported for the IQN/IZN17 series and (CCIZN17)₃ series. It has also been reported that a N51/C43 of gp41 complex has an apparent melting temperature of approximately 90 °C (4). The close thermal stability of our structure to this gp41 six-helix bundle is also an evidence that our structures are likely folded into a conformation mimicking that of the native gp41 core.

According to the urea denaturation curves obtained between 0 and 8 M urea, presented in Figures 4.5 and 4.13, the mean residue ellipticity at 222 nm of the N- peptides alone (Cys-N29 and Cys-N29b) increased gradually from about -10,000 deg.cm².dmol⁻¹ at low urea concentration to about -5,000 deg.cm².dmol⁻¹ in 8 M urea, but no indication of a cooperative transition was observed. In contrast, for the three-helix bundle model structures, a two-state transition from the folded to the unfolded state was clearly observed with increasing urea concentration. This distinct difference between the urea denaturation of N-peptide alone and that of the template-assembled model structures, KTA-3N29, TREN-3N29, KTA-3N29b and TREN-3K-N29b, demonstrates the helix-bundle organizing and stabilizing effects of our designed templates, KTA-Br, TREN-Br and TREN-3K-Br.

According to the urea denaturation experiments analyzed by linear extrapolation to zero urea concentration (Table 4.7), the rigid template KTA-Br is better at generating a stable three-helix bundle structure than are the flexible templates. Under the same buffer conditions (20 mM Tris.HCl, with or without added KF salt, pH 7.40), the free energy changes (ΔG_{H_2O}) for the unfolding equilibrium reaction of KTA-3N29 and KTA-3N29b are greater than those of

TREN-3N29 and TREN-3K-N29b, respectively, indicating that KTA-3N29 and KTA-3N29b have more stable structures toward urea denaturation than their TREN-based counterparts.

In Figure 4.14 panel (b), urea denaturation of KTA-3N29 and KTA-3N29b under the same conditions, 20 mM Tris.HCl pH 7.40, was compared. Since the KTA-Br template was used in both of these templated structures, this comparison directly reflects the influence of using different N-peptide analogs, Cys-N29 or Cys-N29b, on the stabilities of the resultant three-helix bundle structures. In this study (see results in section IV.B and IV.C), the gp41 N-peptide native sequence Cys-N29 was modified to the sequence Cys-N29b by replacing the nonconserved (located at the **b** position of coiled-coil, refer to Figure 1.3 and (1)), and negatively charged Glu⁵⁶⁰ to Gln residue. The intention of this modification is to increase the solubility of the three-helix bundle structures, at neutral pH in the presence of physiological salt, by increasing their net charges under these conditions. The second generation of the three-helix bundle structures, incorporating the sequence Cys-N29b, KTA-3N29b and TREN-3K-N29b, indeed, showed improved properties as we designed to do. For example, the yield of the coupling reactions (in 100 mM Tris.HCl pH 8.0) used to prepare the three-helix bundles (Scheme 3.5) was increase from about 20% for KTA-3N29 to about 30% for KTA-3N29b (both are calculated after HPLC purification), probably by decreasing the aggregation and precipitation of the unreacted N-peptide analogs under the reaction conditions (pH 8.0). However, Figure 4.14.b shows that KTA-3N29b is less helical and less stable toward urea denaturation than KTA-3N29

(refer to section IV.B.2). Since these two compounds were constructed using the same template KTA-Br, these differences in the shape of urea denaturation curve (Figure 4.14.b), as well as the results from curve-fitting (Table 4.7), suggest that Cys-N29b is not as good as the native sequence Cys-N29 at forming a helical and stable three-helix bundle structure. Therefore, our modification of replacing Glu⁵⁶⁰ with Gln apparently increases the solubility of the peptide models at the expenses of decreasing their helicities (at zero urea concentration, mean residue ellipticity at 222 nm, about -25,000 for KTA-3N29 versus about -22,000 deg.cm².dmol⁻¹ for KTA-3N29b) and stabilities slightly (Table 4.7). According to the crystal structure of gp41 (Figure 1.3.b), under neutral pH, positively charged residues Arg⁵⁵⁷ and His⁵⁶⁴ residues are at the positions adjacent to negatively charged Glu⁵⁶⁰. It is very likely that the ionic interactions between these residues also contribute to the stability of the gp41 core. Therefore, when we replace the Glu⁵⁶⁰ residue, under neutral conditions, although the solubilities of the resultant peptides were improved due to the increase of the net charge of these compounds (refer to the design section in IV.B), helicity and stability of these peptides were decreased due to disappearing of the ionic interactions between Arg⁵⁵⁷, His⁵⁶⁴ and Glu⁵⁶⁰.

As listed in (Table 4.10), the over free energy changes (ΔG^0) of the thermal unfolding of all the model structures (thermal denaturation, pH 2.50) were between -4.2 to -5.1 kcal/mol at 25 °C. Quite consistent with this range are those of the unfolding process, induced by urea (pH 7.40), which were between -2.4 to -5.8 kcal/mol for all the structures (Table 4.7). In particular, when no KF is

used in the buffer, and the ionic strength of the buffer is low, the ΔG values obtained by these two independent types of denaturation study of the same molecule are very close to each other (for KTA-3N29b at 25 °C, ΔG values obtained from the thermal and urea denaturation were 4.7 and 5.07 kcal/mol, respectively). Similar agreement between $\Delta G^0(T)$ values from thermal denaturation studies and $\Delta G^0_{H_2O}$ values from urea denaturation (not GdnHCl denaturation) studies that were analyzed by the linear extrapolation method (refer to section IV.A.2) has been reported previously (22,23) and it has been concluded that these two methods can be used to reliably measure the conformational stability of a protein.

As has been reviewed elsewhere (24), that the folded conformations of natural proteins are generally stabilized by about 4-10 kcal/mol, relative to their unfolded states. In addition, $\Delta G_D^{H_2O}$ and m values of urea denaturation of a selection of small proteins containing 62 to 131 residues have been measured in various buffer conditions (most in the absence of added salt in buffer). The Gibbs free-energy values for such proteins typically fall in the range of 3-10 kcal.mol⁻¹, while m values in the range of 770-1770 cal.mol⁻¹.M⁻¹ have been reported (reviewed in 23, 25). Compared with these results for natural proteins, the corresponding values of these parameters obtained for our designed structures as estimated by either method are within the range of those formed for natural proteins, but at the relatively low end of the range. This indicates that our model structures form a folded structure that has a discrete, but limited stability relative to the unfolded state, and a limited hydrophobic core. This is consistent with the

folding of these model structures to form a three-helix bundle that is structurally similar to the N-peptide core observed in the crystal structures of the six-helix bundles formed by gp41 N- and C-peptide constructs. According to the crystal structure of gp41 (1-3), within the N-peptide coiled coil, the buried core comprises only a few hydrophobic side chains, including those of Leu⁵⁵⁵, Ile⁵⁵⁹, Leu⁵⁶⁶, Ile⁵⁷³, Leu⁵⁷⁶, and Ile⁵⁸⁰, as well as the relatively hydrophilic side chains of Gln⁵⁶² and Thr⁵⁶⁹. This limited hydrophobic core is consistent with a low *m* value being obtained from our urea-denaturation study. In addition, significant hydrophobic surface is organized by the formation of the N-helix bundle in the gp41 crystal structures so that it faces outwards from the core, to provide hydrophobic binding grooves for the three C-peptide helices that pack in antiparallel fashion onto the surface. This includes the side-chains of Leu⁵⁵⁶, Ala⁵⁵⁸, Ala⁵⁶¹, Leu⁵⁶⁵, Leu⁵⁶⁸, Val⁵⁷⁰, Trp⁵⁷¹, Ala⁵⁷⁸, and Leu⁵⁸¹. The presence of so many hydrophobic elements on the solvent-exposed surface of this structure results in helices that have a marginal amphiphilic character driving the organization of the core N-peptide helix bundle. This explains both the relatively low thermodynamic stability we observe for these three-helix bundle model structures relative to their denatured state, and also their tendency to denature in irreversible fashion to form insoluble aggregates and β -sheet structures (see section IV.A.2 and Figure 4.3).

The urea denaturation curve for KTA-3N29b in the presence of 100 mM KF (Figure 4.13) shows that, in the low concentration range (< 2 M) of urea, the helical CD signal for both Cys-N29b and KTA-3N29b appeared to be enhanced.

Similar effects have been observed previously for highly α -helical peptides (26) and templated peptide bundles (27). However, as shown in Figure 14.a, a similar signal enhancement was not observed during the denaturation of KTA-3N29b in the absence of KF. Based on these observations, we speculate that this signal enhancement might be attributed to the effect of the KF salt stabilizing partially denaturated structures that possess a more-helical conformation, possibly by forming soluble aggregates in solution. The fact that this signal change was not observed at the low urea concentration range for the denaturation of TREN-3K-N29b could be because that positively charged template (with three free Lys side-chains) used in this structure does not favor the formation of soluble aggregates. The presence of inorganic salts in solution has effects on the solubility and stability of proteins. Such effects exhibit a reoccurring trend described by the Hofmeister series (28), which is generally more pronounced for anions than for cations. The typical ordering of the anion series and some of its related properties are shown in Figure 5.1. The species to the left of Cl^- are referred to as kosmotropes, which have stabilizing and salting-out effects on proteins, and those to the right of Cl^- are called chaotropes, which are known to destabilize folded proteins and give rise to salting-in behavior. Important theoretical and experimental contributions have identified some key factors relating to the influence of salt on protein stability, such as protein-cosolute interaction and the excluded volume effect (29, 30). However, the mechanisms by which ions stabilize proteins are not yet totally clear. Fluoride (F^-) is one of the kosmotropes, and its stabilizing effects, at concentration of about 100 mM, on the

leucine zipper region of GCN4 (31) and on the B1 domain of protein L (ProL) (32) have been reported. Even though the exact interaction between KF and our KTA-3N29b molecule during the urea denaturation study needs to be further clarified, these previous reports support our speculation that added 100 mM KF may play a role in stabilizing a partially denatured structure in this process.

C. Aggregation state studies

The N-peptide sequence by itself is known for its tendency to aggregate under neutral conditions (4) due to the amphiphilic and hydrophobic nature of this sequence (mentioned in the above section). Therefore, it was critical to determine if the directing effects of our designed templates were promoting helical structure in the covalently linked N- peptides through formation of discrete, monomeric, three-helix bundle structures, as our design intended. Alternatively, attaching N-peptide to a template may have promoted its aggregation into multiple higher-order forms, possibly including parallel and anti-parallel helical alignments.

In an initial attempt to distinguish between these possibilities, we examined the concentration dependence of the CD spectra of the template-assembled model structures over a 10-fold concentration range (from about 2.0 μM to 18.0 μM) for each constructed model structure. Figure 4.1 and 4.10 show that, over the concentration range examined, the molar spectra of each molecule are approximately identical within experimental error. This suggests that, under these conditions, the templated helix-bundle structure does not undergo a change in its aggregation state. This result is most simply explained if these

model structures, KTA-3N29, TREN-3N29, KTA-3N29b and TREN-3K-N29b, exist in their monomeric, 3-helix bundle form in these solutions, as they were designed to do.

Consistent with the conclusion stated above, thermal denaturation curves determined for each structure by CD have also been measured at pH 2.50, and show no significant changes in shape (Figure 4.9 and 4.17) or T_m over a 6-fold concentration range (Table 4.5 and 4.11). This provides a more definitive demonstration that these molecules exist predominantly in their monomeric form, at least in acidic solutions. Similar experiments, performed at pH 7.40, showed that the templated-assembled model structures underwent an irreversible change in structure upon heating to temperatures above 70 °C, adopting a β -sheet conformation (Figure 4.3) and precipitating out of solution. Nevertheless, as mentioned in the above section, the low pH experiment supports our conclusion that the templated helix bundles that incorporate the Cys-N29b sequence into their structures exist predominantly in their monomeric form at neutral pH also, because there are no titratable side-chain groups in these structures in the pH range 2.5 – 7.4.

D. Binding studies

To examine if our designed structures can mimic the function of the native gp41 NHR core, assays of the binding of the gp41 C-peptide analog, C28, to our modified model structures, KTA-3N29b and TREN-3K-N29b, were performed. These binding assays were carried out in 20 mM Tris.HCl, 100 mM KF, pH 7.40 buffer at 25 °C, and monitored by CD difference spectropolarimetry.

In preliminary experiments, shown in Figures 4.18 and 4.19, the CD spectra of C28 alone (10 μ M) and the template-assembled model structure alone (3 μ M KTA-3N29b in Figure 4.18 and 3 μ M TREN-3K-N29b in Figure 4.19) were collected separately, and also as mixtures at these same concentrations. When the CD spectrum of these mixture are compared to the sum of the CD spectra of each component collected separately, significant increases in helical structure are indicated for the mixtures. These differences in spectra (in panel b of Figure 4.18 and 4.19) clearly demonstrate that KTA-3N29b and TREN-3K-N29b bind C28 in aqueous buffer at physiological pH and ionic strength. Furthermore, since C28 alone gives a weak CD spectrum in aqueous buffer that is indicative of little or no helical structure (in panel a of Figure 4.18 and 4.19), and KTA-3N29b alone or TREN-3K-N29b alone is nearly fully folded in an α -helical conformation (Table 4.6), the increases in helical structure observed for the mixture provides strong evidence that these binding interactions are accompanied by formation of a substantial helical conformation in C28 in each case. These observations are all consistent with the formation of a three-helix bundle structure by each of our model structures that has a parallel coiled coil of N-peptide helices with functional binding sites on its surface that recognize C-peptide helices. This result is similar to the reports by Gochin *et al.* on the formation of six-helix bundle structures resulting from the binding of C-peptide segments onto a model constructed to mimic the folded core structure of gp41 N-peptide (14, 15).

Quantitative analyses of the binding of C-peptide to our template-assembled N-peptide model structures, KTA-3N29b and TREN-3K-N29b, were

then performed in the same physiological buffer at 25 °C. In these assays, the CD signals that are indicative of helix content ($\theta_{222} - \theta_{260}$) were measured for mixtures of a fixed concentration of C28 (4.00 μM for the test with KTA-3N29b and 3.00 μM for the test with TREN-3K-N29b) and increasing concentrations of each model structure respectively. Then the corresponding CD signals for C28 alone and the appropriate N-peptide bundle model structure alone were subtracted to give difference CD signals as a function of the concentration of the N-peptide bundle structure. These data are plotted in Figures 4.21 and 4.22 as solid circles. These plots show that C28 binding increases with increasing concentration of each model structure, until all of the C28 is in the bound form, and a maximum difference in CD signal is reached. This maximum is reached at about 2.5 μM or higher KTA-3N29b concentrations (Figure 4.21), and at higher than 8.0 μM TREN-3K-N29b concentrations (Figure 4.22), respectively.

The binding data in Figure 4.21 and 4.22 (solid dots) were then fitted to the binding isotherm derived under the assumptions that each templated N-peptide helix bundle has n independent binding sites on its surface for C28, and that the binding equilibrium of C28 with each of these sites was governed by a single equilibrium dissociation constant, K_D (refer to section IV.C.2, Equation 4.10). Curve fitting was then performed using integer values for n of 2, 3 or 4. The curve-fitting results for KTA-3N29b and TREN-3K-N29b, including the binding stoichiometry, n , calculated dissociation constant K_D , and the maximum CD signal change, $\Delta\theta_{\text{max}}$, are listed in Tables 4.12 and 4.13, respectively.

Both of the constructed model structures, KTA-3N29b and TREN-3K-N29b, were found to bind to C28, the gp41 C-peptide analog, tightly, with a dissociation constant (for each C28 binding), K_D , in sub-micromolar range for KTA-3N29b and in the low-micromolar range for TREN-3K-N29b. This high binding affinity is similar to that reported by Gochin *et al.* (650 nM) for their N-peptide three-helix core structure which was assembled by ligand coordination to nickel ions (14).

We have assumed an integer value for the fitted parameter n , which represents the number of independent C28 binding sites on the surface of each constructed N-peptide helix bundle. For KTA-3N29b, which is assembled on a rigid template, the best non-linear least-squares fit was found for $n = 3$, indicating that the complex formed at saturation is, at least predominantly, a six-helix bundle comprising the templated three-helix bundle core, KTA-3N29b and three copies of C28. However, for TREN-3K-N29b, which is assembled on a flexible template, similar reasonable fits were obtained when n is either 2, 3 or 4, indicating that an accurate estimation of the C28 binding stoichiometry could not be achieved in this case, by the current experimental method. This result is consistent with the low-micromolar K_D of TREN-3K-N29b: due to the poor solubility problem of TREN-3K-N29b, the binding titration experiment could not be extended to higher substrate concentrations in the range 10- to 20-fold greater than K_D , where n might be evaluated more accurately.

The maximum CD signal change ($\Delta\theta_{\max}$) corresponds to the signal change when 100% of the C28 is in its bound state (equation 4.8). Since our templated

helix bundle is already about 85% helical in the absence of C-peptide (Table 4.6), we can assume that the difference-CD signal observed during C28 binding is predominantly due to changes in the C-peptide conformation. Therefore, if we assume that both KTA-2N29b and TREN-K-3N29b possess 3 binding sites for C28, it is calculated that $[\theta]_{222}$ for C28 changes from about $-4,000 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its unbound state to $-28,800 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its KTA-3N29b-bound state, and to $-25,600 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ in its TREN-3K-3N29b-bound state (Table 4.14). These ellipticity values for bound C28 may be compared to values for $[\theta]_{222}$ of $1,000 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ and $-34,000 \text{ deg.cm}^{-2}.\text{dmol}^{-1}$ that are calculated for 0% α -helix and 100% α -helix, respectively, for a peptide having the same length (28 residues) at 25 °C, by the method of Baldwin and coworkers (33). Based on these values, the bound C-peptide is estimated to be 85% and 76% α -helical in the complex it forms with KTA-3N29b and TREN-3K-N29b, respectively. Therefore KTA-3N29b, which is constructed using a rigid template, KTA-Br, results in a more helical complex with C28, that is likely to be substantially identical to the six-helix bundle observed in gp41 crystal structures (1-3), and which is assumed to be the final folded product of HIV-1 fusion with CD4⁺ cells.

To summarize briefly, these binding studies using CD as a detector to monitor the conformational changes during C28-binding provide strong evidences that we successfully constructed gp41 N-peptide three-helix bundle model structures by using a template-assembly approach. The resultant model structures, KTA-3N29b and TREN-3K-N29b, possess functional binding sites on their surfaces for gp41 C-peptides, with binding affinities similar to previously

reported structures that are using a different templating assembly approach (14). Our designed template, KTA-Br, successfully incorporates desired rigidity and orienting effects into its cyclohexane ring, and suitable flexibility in its extended arms and, therefore, generates a well folded and stable mimic of the native gp41 N-peptide three-helix bundle.

References:

9. Chan, D. C., Fass, D., Berger, J. M., Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 1997, **89**: 263-273.
10. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature*, 1997, **387**: 426-430.
11. Tan, K., Liu, J., Wang, J., Shen, S., Lu, M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. USA*, 1997, **94**: 12303-12308.
12. Lu M, Blacklow SC, Kim PS. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol.* 1995, **2**(12):1075-82.
13. Blacklow SC, Lu M, Kim PS. A trimeric subdomain of the simian immunodeficiency virus envelope glycoprotein. *Biochemistry.* 1995, **34**(46):14955-62.
14. Lu M, Kim PS. A trimeric structural subdomain of the HIV-1 transmembrane glycoprotein. *J Biomol Struct Dyn.* 1997, **15**(3):465-71.
15. Kemp, D.S., Petrakis, K.S. Synthesis and conformational analysis of cis, cis-1, 3, 5-trimethylcyclohexane-1, 3, 5-tricarboxylic acid. *J. Org. Chem.* 1981, **46**: 5140-5143.
16. Xu, W., Taylor, J.W. Template-assembled peptide models of the N-peptide Helix bundle from HIV-1 gp41. Proceedings of 19th American Peptide Society, San Diego, U.S.A., 2005: 563-564.
17. Eckert DM, Kim PS. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc Natl Acad Sci USA* 2001, **98**: 11187-92.
18. Eckert DM, Malashkevich VN, Hong LH, Carr PA, Kim PS. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. *Cell* 1999, **99**: 103-15.
19. Louis JM, Bewley CA, Clore GM. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J Biol Chem* 2001, **276**: 29485-9.
20. Louis JM, Nesheiwat I, Chang L, Clore GM, Bewley CA. Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope-mediated cell fusion. *J Biol Chem* 2003, **278**: 20278-85.
21. Bianchi E., Finotto M., Ingallinella P., Hrin R., Carella A.V., Hou X.S., Schleif W.A., Miller M., Geleziunas R., Pessi A. (2005) Covalent stabilization of coiled coil of the HIV gp41 N region yields extremely potent and broad inhibitors of viral infection. *Proc. Natl. Acad. Sci. USA* 102: 12903-12908.
22. Gochin M., Kiplin Guy R., Case M. A. A Metallopeptide assembly of the HIV-1 gp41 coiled coil is an idea receptor in fluorescence detection of ligand binding. *Angew. Chem. Int. Ed.*, 2003, **42**: 5325-5328.

23. Gochin, M., Savage, R., Hinckley, S. and Cai, L. A fluorescence assay for rapid detection of ligand binding affinity to HIV-1 gp41. *Biol. Chem.* 2006, **387**: 477-483.
24. Caffrey, M., Kaufman, J., Stahl, S., Wingfield, P., Gronenborn, A. M., and Clore, G. M. Monomer-trimer equilibrium of the ectodomain of SIV gp41: insight into the mechanism of peptide inhibition of HIV infection. *Protein Sci.* 1999, **8**:1904–1907.
25. K. Sackett, Y. Wexler-Cohen, Y. Shai. Characterization of the HIV N-terminal Fusion Peptide-containing Region in Context of Key gp41 Fusion Conformations. *J. Biol. Chem.*, 2006, **281**: 21755-21762.
26. Gallo, S. A., Sackett, K., Rawat, S. S., Shai, Y., and Blumenthal, R. The Stability of the Intact Envelope Glycoproteins is a Major Determinant of Sensitivity of HIV/SIV to Peptidic Fusion Inhibitors *J. Mol. Biol.* 2004, **340**: 9–14.
27. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: Conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc. Natl. Acad. Sci. U. S. A.* 1998, **95**: 9134–9139.
28. Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41 *EMBO J.* 1998, **17**: 4572–4584.
29. Kwak, J., De Capua, A., Locardi, E., Goodman, M. TREN (tris (2-aminoethyl) amine): an effective scaffold for the assembly of triple helical collagen mimetic structures. *J. Am. Chem. Soc.* 2002, **124**: 14085-14091.
30. Pace, C.N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* 1986, **131**: 266-280.
31. Pace, C.N., Shaw, K. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins: Structure, Function, and Genetics Suppl.* 2000, **4**: 1-7.
32. DeGrado WF, Summa CM, Pavone V, Nastri F, Lombardi A. De novo design and structural characterization of proteins and metalloproteins. *Annu Rev Biochem.* 1999, **68**:779-819.
33. Myers, J.K., Pace, C.N., Scholtz, J.M. Denaturant m values and heat capacity changes: relation the changes in accessible surface areas of proteins unfolding. *Protein Science*, 1995, **4**: 2138-2148.
34. Osapay, G., Taylor, J.W. Multicyclic polypeptide model compounds. 2. Synthesis and conformational properties of a highly α -helical uncosapeptide constrained by three side-chain to side-chain lactam bridges. *J. Amer. Chem. Soc.* 1992, **114**: 6966-6973.
35. Mutter, M., Tuchscherer, G.G., Miller, C., Altmann, K.-H., Carey, R.I., Wyss, D.F., Labhardt, A.M. and Rivier, J.E. Template-assembled synthetic peptides with four-helix-bundle topology. Total chemical synthesis and conformational studies. *J. Amer. Chem. Soc.* 1992, **114**: 1463-1470.
36. Hofmeister, F. Zur lehre der Wirkung der Salze. Zweite Mittheilung, *Arch. Exp. Pathol. Pharmacol.* 1888, **24**: 247-260.

37. Timasheff, S. N. Protein solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components, *Proc. Natl. Acad. Sci. U.S.A.* 2002, **99**: 9721-9726.
38. Record, M. T., Zhang, W., and Anderson, C. F. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects and osmotic effects of salts, *Adv. Protein Chem.* 1998, **51**: 281-353.
39. Kenar K. T., Garcia-Moreno B., Freire E., A calorimetric characterization of the salt dependence of the stability of the GCN4 leucine zipper. *Protein Sci.* 1995, **4**: 1934-1938.
40. Tadeo X., Pons M., Millet O., Influence of the Hofmeister Anions on Protein Stability As Studied by Thermal Denaturation and Chemical Shift Perturbation. *Biochemistry*, 2007, **46**: 917 -923.
41. Lou, P., Baldwin, L. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 1997, **36**: 8413- 8421.

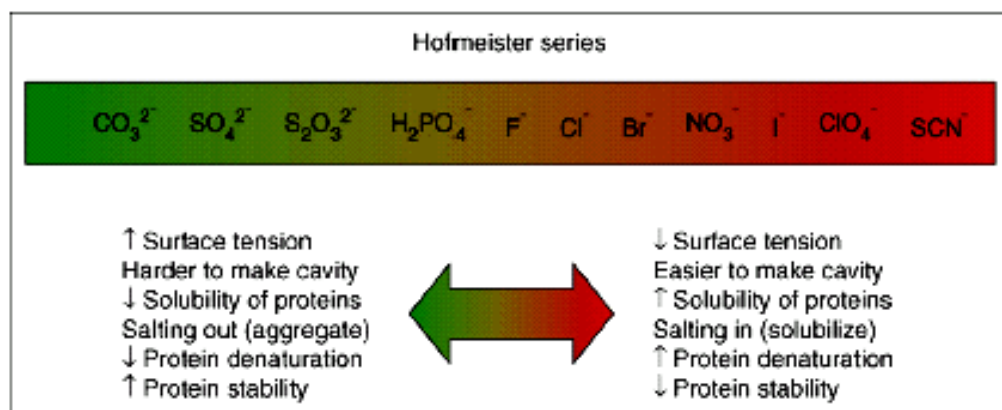


Figure 5.1 The Hofmeister series.

VI. Conclusion

In this project, small-molecule templates, KTA-Br, derived from the rigid Kemp's triacid structure, and TREN-Br and TREN-3K(NH³⁺)-Br, derived from the flexible TREN structure, have been designed and synthesized. By assembling these templates to either the native gp41 sequence, Cys-N29, or the modified sequence, Cys-N29b, template-assembled N-peptide three-helix bundles, KTA-3N29, TREN-3N29, KTA-3N29b, and TREN-3K-N29b, were then synthesized that mimic the corresponding N-peptide core structure of the six-helix bundles formed by HIV-1 gp41 (1-3). In neutral buffer, the compounds all exist as stable and folded monomers with high α -helix contents. In addition, the N-peptide sequence-modified structures KTA-3N29b and TREN-3K-N29b, which have increased solubility under neutral conditions in the presence of physiological salt, can serve as functional mimics for the N-peptide helices of gp41, as demonstrated by their ability to bind the native C-peptide. KTA-Br is the best of these designed templates as judged by the observation that it generates a better folded monomeric structure that was proved to bind three copies of C-peptide per N-peptide bundle with a dissociation constant of 260 nM, resulting a six-helix bundle that is estimated to be about 85% α -helical overall.

The use of a template, such as KTA-Br, which has three-fold symmetry and is readily synthesized from Kemp's triacid, provides a straightforward way to direct the parallel folding of N-peptide analogs so that they form monomeric three-helix bundles with the same three-fold symmetry. In these aspects, our template design and results are similar to those reported by Gochin *et al.* (4-5),

who have reported an N-peptide helix bundle that is organized for trimerization by metal-ion chelation. However, the KTA-Br template imposes a stronger directional orientation on the N-peptides than does the metal-ion chelation design. The trimerizing template of Gochin *et al.* includes a fluorescence energy transfer reporter group for binding assays that our KTA template currently lacks, but does not direct a parallel arrangement of the N-peptide helices.

Organized N-peptide helix bundles have potential utility in a variety of studies. For example, both N-peptide three-helix bundles and C-peptide analogs have the potential to neutralize HIV-1 by interfering with the gp41 folding process that is thought to drive cell fusion and results in formation of the six-helix bundle (1). Indeed, the all-protein constructs of Louis *et al.* (6-7) and Bianchi *et al.* (8), in which folding of the N-peptide trimers is organized by leucine zipper motifs and/or disulfide knots rather than an organizing small-molecule template, have been shown to inhibit HIV-1 envelope-mediated cell fusion.

C-peptide binding by N-peptide helix bundle model structures has been directly demonstrated only for our approach (Figure 4.21 and 4.22) and for that of Gochin *et al.* (4). Such assays may be useful to further optimize either the template or the N-peptide structure in our approach, or the structures of C-peptide analogs, such that higher affinity binding is attained. Alternatively, these binding assays may be used to screen small-molecules as inhibitors of N-peptide and C-peptide binding interactions. Each of these approaches has the potential for development of new HIV-1 fusion inhibitors.

Finally, it has been reported that a human monoclonal antibody can neutralize diverse HIV-1 isolates by binding to the N-peptide sequence of gp41 (9). This suggests that template-assembled N-peptide models, such as KTA-3N29b, could be investigated as primary or boosting immunogens that might raise a neutralizing immune response to HIV-1 (7).

References:

42. Chan, D. C., Fass, D., Berger, J. M., Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 1997, **89**: 263-273.
43. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature*, 1997, **387**: 426-430.
44. Tan, K., Liu, J., Wang, J., Shen, S., Lu, M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. USA* , 1997, **94**: 12303-12308.
45. Gochin M., Kiplin Guy R., Case M. A. A Metallopeptide assembly of the HIV-1 gp41 coiled coil is an idea rector in fluorescence detection of ligand binding. *Angew. Chem. Int. Ed.*, 2003, **42**: 5325-5328.
46. Gochin, M., Savage, R., Hinckley, S. and Cai, L. A fluorescence assay for rapid detection of ligand binding affinity to HIV-1 gp41. *Biol. Chem.* 2006, **387**: 477-483.
47. Louis JM, Bewley CA, Clore GM. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J Biol Chem* 2001, **276**: 29485-9.
48. Louis JM, Nesheiwat I, Chang L, Clore GM, Bewley CA. Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope- mediated cell fusion. *J Biol Chem* 2003, **278**: 20278-85.
49. Bianchi E., Finotto M., Ingallinella P., Hrin R., Carella A.V., Hou X.S., Schleif W.A., Miller M., Geleziunas R., Pessi A. (2005) Covalent stabilization of coiled coil of the HIV gp41 N region yields extremely potent and broad inhibitors of viral infection. *Proc. Natl. Acad. Sci. USA* 102: 12903-12908.
50. Miller, M.D., Geleziunas, R., Bianchi, E., Lennard, S., Hrin, R., Zhang, H., Lu, M., An, Z., Ingallinella, P., Finotto, M., Mattu, M., Finnefrock, A.C., Bramhill, D., Cook, J. Eckert, D.M., *et al.* A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope. *Proc. Natl. Acad. Sci. USA* 2005, **102**: 14759-14764.

Curriculum Vitae

Weiming Xu

EDUCATION

- Oct., 2007 PhD in Chemistry
Department of Chemistry and Chemical Biology, Rutgers University,
Piscataway, New Jersey, U.S.A.
- Jun., 2001 MS in Chemistry
College of Chemistry and Molecular Engineering, Peking University,
Beijing, P.R. China
- Jul., 1998 BS in Chemistry
College of Chemistry and Molecular Engineering, Peking University,
Beijing, P.R. China

PUBLICATIONS

1. **Xu, W.**, Taylor, J., W., "Template-assembled peptide model of the N-peptide helix bundle from HIV-1 Gp41 with high affinity to C-peptide", *Chemical Biology and Drug Design*, 2007, accepted
2. **Xu, W.**, Taylor, J., W., "Template-assembled peptide models of the N-peptide helix bundle from HIV-1 Gp41", *Proceedings of 19th American Peptide Society Symposium, San Diego, U.S.A., 2005*, 563-564
3. **Xu, W.**, Xu, J., Wu, N., Yan, J., Zhu, Y., Huang, Y., He, W., Xie, Y., "Study of the diffusing behavior of MoO₃ and ZnO on oxide thin films by SR-TXRF", *Surface and Interface Analysis*, **32** (2001), 301-305
4. **Xu, W.**, Yan, J., Wu, N., Zhang, H., Xie, Y., Tang, Y., Zhu, Y., Yao, W., "Diffusing behavior of MoO₃ on Al₂O₃ and SiO₂ thin films", *Surface Science*, **470** (2000), 121-130