DISSERTATION

PREPARATION AND INVESTIGATION OF UNNATURAL POLAR SIDECHAINS IN DESIGNED COILED-COILS

Submitted by

Maria L. Diss

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2008

COLORADO STATE UNIVERSITY

May 30, 2008

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY MARIA LUCIA DISS ENTITLED PREPARATION AND INVESTIGATION OF UNNATURAL POLAR SIDECHAINS IN DESIGNED COILED-COILS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Prof. Debbie Crans

Prof. Olve Peersen

Prof. Dawn Rickey

Prof. Yian Shi

Adviser: Prof. Alan Kennan

Department Head/Director: Prof. Anthony Rappe

ABSTRACT OF DISSERTATION

PREPARATION AND INVESTIGATION OF UNNATURAL POLAR SIDECHAINS IN DESIGNED COILED-COILS

Programmed self-assembly of discrete molecular species to form complex aggregates provides the opportunity to both refine and exploit current knowledge of molecular recognition patterns. Self-assembly events dominate the precise control and understanding of macromolecular structure, placing a premium on development and discovery of molecular recognition motifs. One such motif is the alpha helical coiled coil. A key natural strategy for controlling specific assembly is the burial of polar side chains (particularly asparagines) at central hydrophobic core positions. Alignment of the polar side chains opposite each other, rather than opposite hydrophobic alternatives, drives formation of the intended complex. Utilizing this polar contact, we sought to control coiled-coil formation through the use of unnatural polar side chains. These unnatural side chains included various chain length arginine derivatives and the corresponding urea analog, citrulline, along with its chain length variants.

Synthetic methodology compatible with solid phase peptide synthesis was developed to form the desired functionalized primary amine side chain. Guanidinylation occurred in one step through the use of a di-2-Cl-Z protected pyrazole derivative. Urea formation also proceeded in one step through the use of a preformed *p*-methoxybenzyl-*p*-nitrophenyl carbamate. Deprotection to give the desired functional group occurred through standard cleavage conditions.

With the establishment of the synthetic methodology, numerous peptides were synthesized incorporating these new polar groups into the hydrophobic core.

Additionally, asparagine, aspartic acid, and glutamic acid were used as core residues. Heterodimeric mixtures of these sequences with guanidine, urea, amide and carboxylic acid binding partners form a large number of reasonably stable coiled coils ($T_m \geq 60$ °C), allowing for application-specific tuning. A number of four-component selective recognition systems are also presented, in which two distinct heterodimers form from an input of four different peptides. More impressively, examples of six-component systems are demonstrated. Control mixtures establish subtle structural requirements for successful recognition.

Having demonstrated successful recognition motifs within a dimeric system, a trimeric system has been investigated. Once again, these polar contacts allow for heterotrimerization. However, the relatively narrow stability range of these trimers does not allow for successful exchange experiments (T_m = 43 to 59 °C).

Maria Lucia Diss
Department of Chemistry
Colorado State University
Fort Collins, CO 80523
Summer 2008

Acknowledgments

I would like to take this opportunity to first of all thank my graduate advisor, Alan Kennan, a gracious mentor, whose constant encouragement has allowed me to accomplish more than I ever thought possible. Additionally, I would like to thank my co-workers both past and present for their helpful discussions and support. In particularly, I would like to thank Brent for his *Pon da Replay* dance which always seemed to brighten my dreariest of grad school days, Ashley for helping me accomplish my running goals and Ross for his sweet mixed CDs that helped pump me up for the long days of research.

I would also like to thank my girls who stuck it out with me in the good and bad times. Andrea, I really do appreciate all of those "stress relieving lunches" with our green tea. Cherisse and Mary, for all of those late night phone calls and random visits, you both helped out more than you can ever imagine.

To my family in Colorado, thank you for opening your homes to me over the past five years. It has been wonderful to have some many homes away from home. Grandpa, thanks for always being around to talk science and encourage me with my research. T, although you are no longer in Colorado, thank you for all of your encouragement and support.

I would like to thank my parents, Bill and Cathy, and my brother, Billy, for their love, support and understanding during the long years of my education. Rents, thanks for teaching me the value of education. For all of the sacrifices you both have made to get me to this point I am eternally grateful. Billy, thank you for always finding the humor in things especially our family.

Lastly, I want like to thank my sister, Christina, my best friend who has been the greatest cheerleader, drill sergeant, voice of reason, counselor and life raft. Thank you for always being there for me.

TABLE OF CONTENTS

Chapter 1

Introduction

1-1: Molecular Recognition and Biological Assemblies

Precise control and understanding of biologically inspired macromolecules is an elusive and ever-present target in research. Self-assembly events dominate such investigations, inspiring the discovery and development of molecular recognition motifs. Uncovering these new routes to specific recognition increases design potential, thus improving the odds of success in an array of applications. Recently, protein recognition elements, such as β-sheets and α-helical coiled coils, have shown considerable promise for employment of such designed motifs.[1]

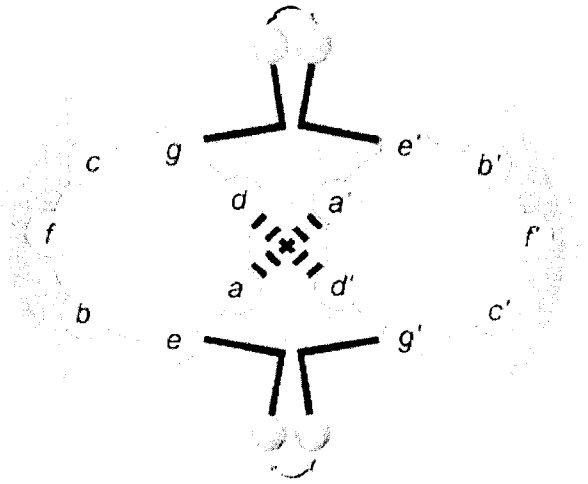


Figure 1.1. Graphic representation of a parallel, dimeric coiled coil. Positions *a* and *d* joined by dashes illustrate the hydrophobic core, while arrows connecting positions *e* and *g* illustrate hydrophilic interactions. Residues *b*, *c* and *f* are exposed to solvent in lower order oligomers.

1-2: Alpha Helical Coiled Coils

The α-helical coiled coil is a prominent structural motif, adopted by 3-5% of all amino acids in proteins.[2] These complexes are formed by the super-coiling of two or more helical strands and exhibit a heptad sequence repeat $(abcdefg)_n$ that generally features non-polar *a* and *d* position residues, and polar residues at the *e* and *g* positions (Figure 1.1). The remaining residues have varying degrees of solvent exposure based on

the oligomerization state of the coiled coil. These residues are generally helix-promoting (such as alanine and lysine), and in lower order oligomers, do not play a specific role in helix recognition.[3] Assembly of coiled coils is driven by packing of the *a* and *d* residues to form a hydrophobic interface. Additionally, the assembly of coiled coils is augmented by electrostatic contacts between the *e* and *g* residues.

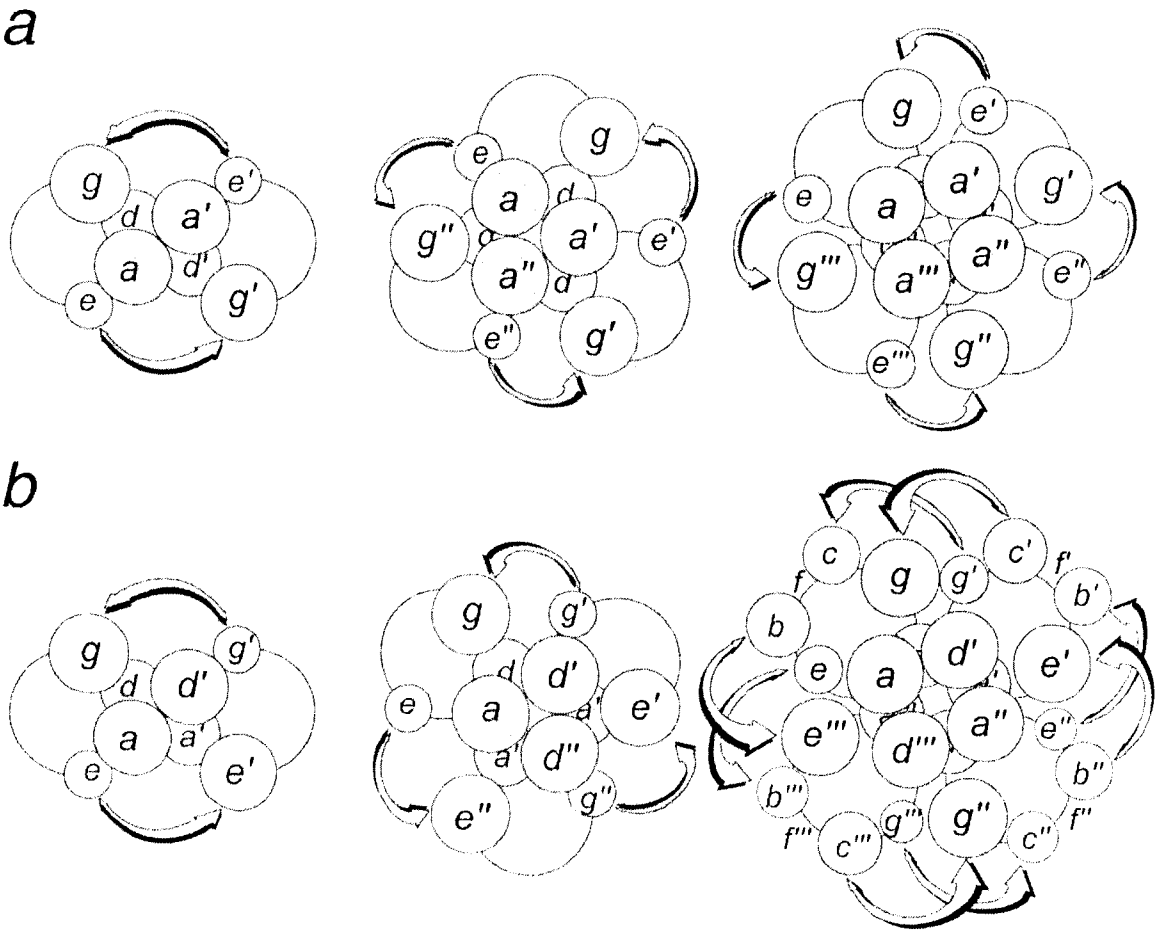


Figure 1.2. Packing comparisons for dimer, trimer and tetramer. (a) Parallel species, where *a* and *d* residues are segregated on the same core layer and *e* and *g* interactions are present at the hydrophilic interfaces. (b) Antiparallel species, where *a* and *d* residues interact in the hydrophobic core and hydrophilic interactions occur between e/e' and g/g'.

In parallel orientations, the hydrophobic core consists of alternating *a* and *d* layers. The hydrophobic *a* residue side chains pack tightly against the same position on an opposing strand in a hole formed by flanking *d*, *e*, and *g* residues (Figure 1.2).[4] In the

antiparallel orientation the core layers are mixed with the *a* residue packing against the *d* residue on the opposing strand. The hydrophobic core configuration provides the foundation for coiled coil design and will be discussed in more detail in the following sections.

The role of the electrostatic interface, in contrast, is not as well-defined or understood as the hydrophobic core. It has been found to contribute to the specificity in both natural and designed systems, although at a lesser extent than that of the hydrophobic core. Particular approaches to self-assembly systems employing this interface will be discussed at length in the following section.

Efforts to control and manipulate coiled coil formation have targeted oligomerization state, relative strand orientation, and specificity for homo- or heteromeric structures. Design of such systems has involved manipulation of the hydrophilic interface and hydrophobic core. The array of this work has applications in protein misfolding models, materials chemistry and biotechnology.[5] The remainder of this chapter will highlight significant contributions to coiled coil design culminating in the current research in this field.

1-3: Coiled Coil Design: Oligomerization

There have been numerous investigations aimed at controlling coiled coil specificity by modification of both the electrostatic and hydrophobic interfaces of coiled coils. Through manipulation of these interfaces, control of the number of helicies, known as the oligomerization state, has been achieved.[6] The leading investigations in this area have focused on known coiled coil sequences, where both nonpolar and polar residues are substituted into the hydrophilic and hydrophobic interfaces. The bZIP transcription

factor, GCN4, has been one of the most predominant templates used in coiled coil investigations. The coiled coil region of this protein, GCN4-p1, exists as a parallel, two-stranded coiled coil.[7] Contacts between the helices include ion pairs and a unique hydrophobic interface that contains a distinctive hydrogen bond between a single pair of buried polar residues.

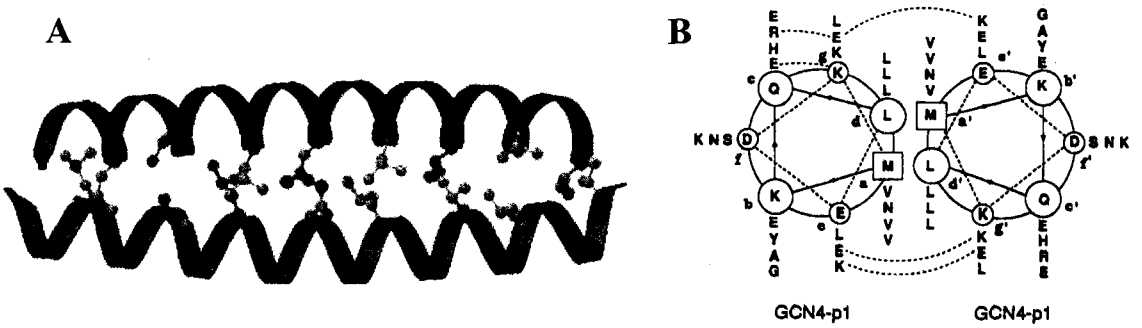


Figure 1.3. GCN4-p1. (A) Central asparagines is shown in color. Exterior residues omitted for clarity. (B) Helical wheel representation. Residues that form ion pairs in the x-ray crystal structure are connected with dashed lines. (Figure B from O'Shea *et al.*, 1991)

Within the hydrophobic core of GCN4-p1 is a central asparagine residue, Asn16 (Figure 1.3). This residue has been found to impart both oligomerization and orientation specificity. Hydrogen bonding between the two Asn16 residues is only possible in the dimeric oligomerization state; altering this would destroy this distinctive bonding, thus destabilizing the higher order oligomer formation. Antiparallel or staggered arrangements would also be destabilizing since the Asn side chains would pack against nonpolar residues in the core. This specificity, however, comes at a cost of stability. The burial of polar residues within the hydrophobic core is thermodynamically unfavorable. Additionally, GCN4-p1 exhibits complementary electrostatic pairings. Positive and negative residues on the leucine zipper are generally alternated along the helices, which

allows for intra- and interhelical ion pairing. These interactions within GCN4-p1 have been the focus of a plethora of studies.

Table 1.1. Hydrophobic Core mutants of GCN4-p1

Position *a*	Position *d*	T_m (°C)	# of Helices
GCN4-p1	GCN4-p1	53	2
I	L	>100	2
I	I	>100	3
L	I	>100	4
V	I	73	-
L	V	81	3
V	L	95	(2,3)
L	L	>100	3

Several of these investigations include core mutations of both polar and nonpolar residues. In one of the first core mutation studies, Kim and coworkers simultaneously changed the *a* and *d* residues of GCN4-p1 to leucine, valine, or isoleucine.[8] They found that these hydrophobic core residue substitutions contributed to a more stable complex but at a loss of oligomerization control (Table 1.1). Interestingly, these two heptad positions exhibit differential behavior when substituted with the same amino acid. Specifically when isoleucine is present in position *a*, and leucine is present in position *d* a dimer forms but when these residues are reversed a tetramer is formed. Analysis of the crystal structure revealed that the apparent difference is due to variations of the "knobs into holes" packing model.

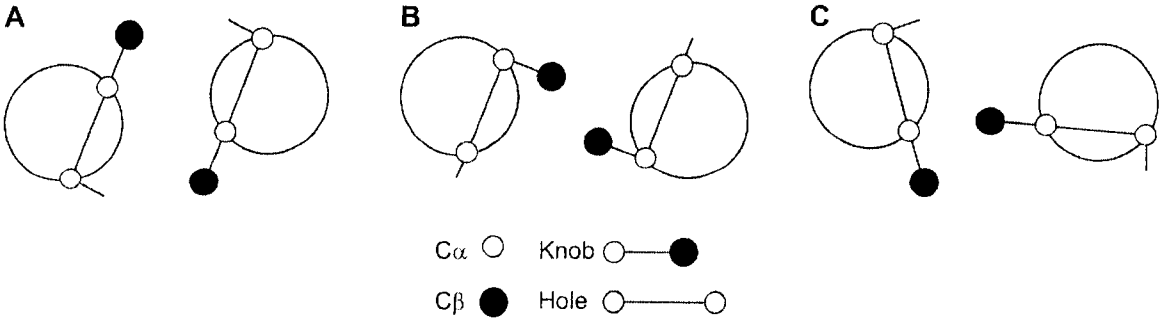


Figure 1.4. Hydrophobic core packing angles. Perpendicular (A), Parallel (B), and Acute (C) are illustrated (not to scale). (Figure adopted from Harbury *et al.*, 1993)

Particularly, they observed that the packing angles of the hydrophobic positions for dimers, trimers, and tetramers differ considerably (Figure 1.4). The packing angle, defined as the angle between the Cα-Cβ bond of one hydrophobic core side chain and the Cα-Cα vector at the bottom of the recipient hole of the opposing strand, is greatly affected by the bulkiness of the β-branched core residues. Perpendicular packing angles were observed at the *d* position of dimers and the *a* position of tetramers, while parallel packing angles were observed at the *a* positions of dimers and *d* positions of tetramers. They concluded that isoleucine could not exist in a perpendicular packing arrangement due to its bulky, β-branched side chain. Additionally, they found that the trimer contains packing angles at both *a* and *d* positions intermediate to those previously described. These angles, termed acute, help explain the preference for the trimeric state when isoleucine is substituted in the *a* and *d* positions. These findings allowed researchers to predict the three-dimensional structure of simple coiled coils based exclusively on hydrophobic packing.

Having found preferential oligomerization for the native polar asparagine residue in the core, the Alber group expanded these previous core mutation studies to include other polar residues such as glutamine, lysine and norleucine.[9] In each case, the central asparagine residue was replaced with glutamine, lysine or norleucine. Similarly to asparagine, when lysine is in the central core position, dimerization is favored. Yet the dimer is less stable than its native GCN4-p1 dimer and its packing differs greatly. Analysis of the crystal structure reveals the ε-amine on the lysine side chain is exposed to solvent and does not pack up against the other lysine residue as is the case with the asparagine residues in the native GCN4-p1. Additionally, the packing difference is

observed with glutamine in the core. Not only did the glutamine specify for both dimeric and trimeric oligomerization states, but direct hydrogen bonding between the two glutamine side chains was not detected. Rather water molecules, one for the dimer and two for the trimer, were observed by crystal structure analysis in the core which seemed to participate in a "hydrogen bonding network." Lastly, when norleucine was present in the core, melting temperatures were over 100 °C, yet this added stability came at a cost of specificity. Norleucine failed to specify a particular oligomerization state much like alanine and valine. The crystal structure of the norleucine dimer showed efficient packing of the aliphatic side chains predominantly within the core. The behavior of buried polar residues in the hydrophobic core is extremely complex as illustrated by these studies.

In addition to single polar core substitutes Hu and coworkers have examined the role of multiple Asn core residues in GCN4-p1.[10] To understand the basis for the specificity of dimer formation, they characterized GCN4 leucine zipper mutants with all sixteen possible permutation and combinations of isoleucines and asparagines at four *a* positions in the dimer interface. They found that homodimeric leucine zippers form with asparagines at different combinations of *a* positions, and that the positioning of buried asparagines is sufficient in determining specificities in which homodimeric and heterodimeric leucine zippers form.

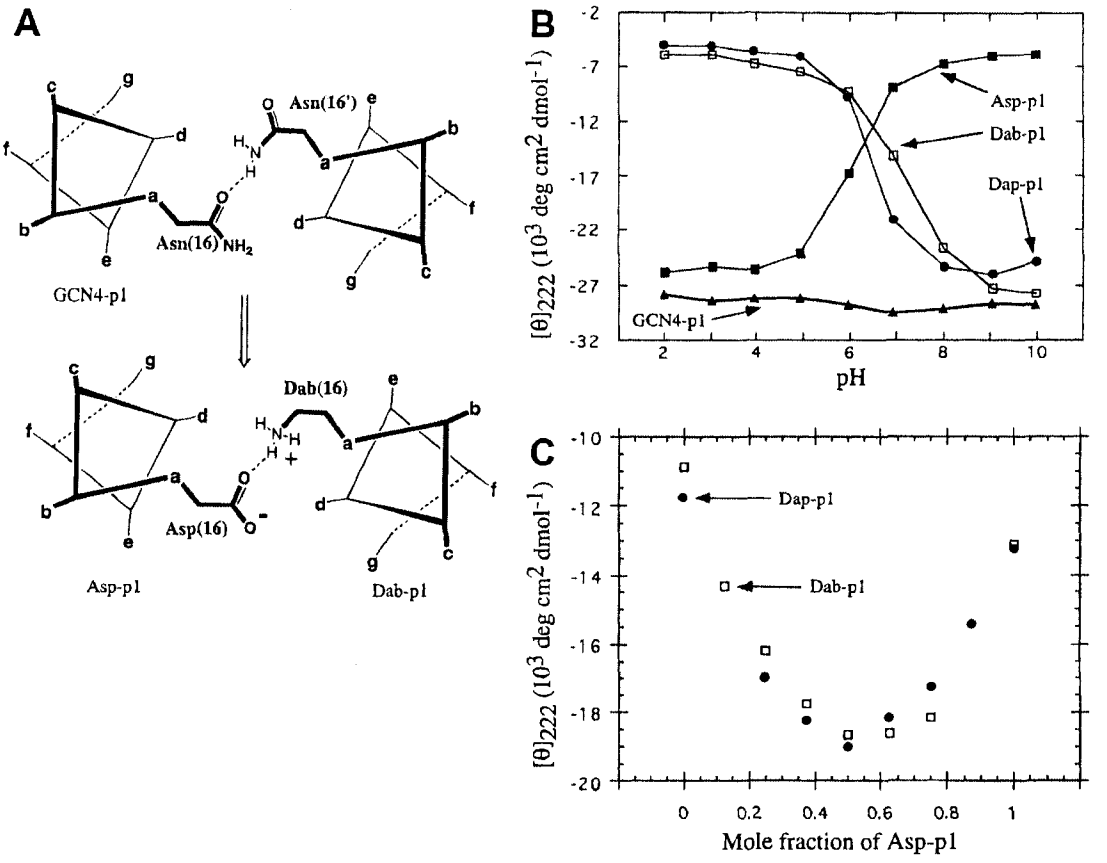


Figure 1.5. (A) Heptad arrangement of GCN4-based peptides with the buried salt bridge shown. (B) Mean residue ellipticity at 222nm monitored as a function of pH. (C) Mean residue ellipticity at 222nm versus the mole fraction of Asp-p1 for Dap-p1 and Dab-p1.

Expanding on these previous studies, DeGrado and coworkers examined the role of buried salt bridges in a heterodimeric GCN4-p1 mimic.[11] They were interested in seeing if a single buried salt bridge would lead to coiled coil specificity (Figure 1.5). For this study the central asparagine residue was substituted with aspartic acid (Asp-p1), diaminobutyric acid (Dab-p1) and diaminopropionic acid (Dap-p1). They found that dimerization is driven by pH, where neutral pH of side chains is required for the formation of homooligomers. Sedimentation equilibrium ultracentrifugation experiments revealed Asp-p1 and Dab-p1 exist as dimers, while Dap-p1 exists as a mixture of dimers and trimers. Additionally, upon mixing equimolar amounts of the Asp-p1 with either

Dab-pl or Dap-pl at a neutral pH, they found formation of heterodimers. This study shows that a single buried salt bridge provides an impressive degree of specificity for forming heterodimers over homodimers. This selectivity does come at a cost of stability; these new complexes are less stable than native GCN4-pl.

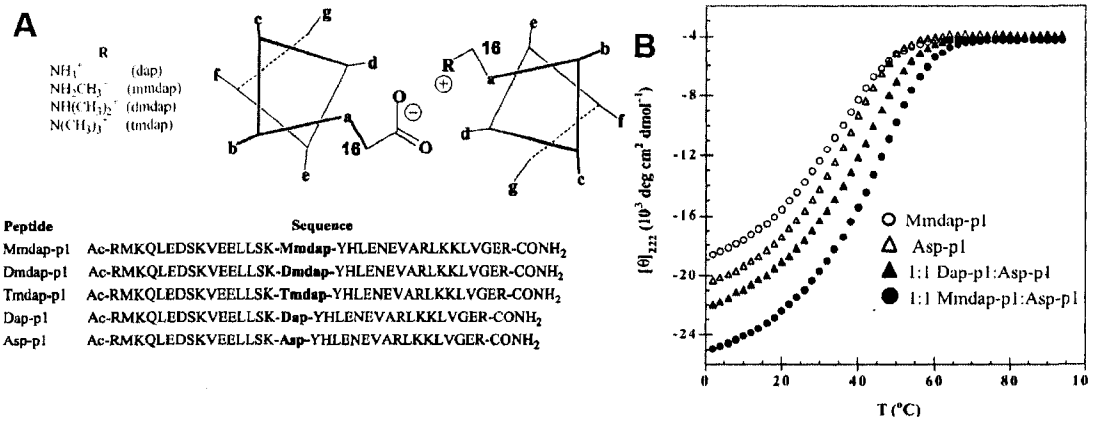


Figure 1.6. Peptides investigated. (A) Helical wheel diagram of a parallel heterdodimeric coiled coil showing positions a-g residues comprising one heptad repeat. Residues of intended buried salt bridge shown between Asp-pl and other designed short basic residues. (B)Thermal denaturation monitored by CD for homo- and heterogeneous solutions of peptide. Total peptide concentration is 100μM in 50mM Mes, 150mM NaCl, pH 6.0. (Figure from Kretisinger *et al.*, 2003)

Schneider and coworkers also examined similar buried salt bridges using varying degrees of hydrophobic basic residues.[12] They substituted mono, di, and tri-methyldiaminopropionic residues in the core of GCN4-pl for the native Asn residue and studied these interactions as homooligomers and as heterodimers with Asp-pl (Figure 1.6A). The monomethyldap peptide (Mmdap-pl) was found to form a more stable homodimer than Dap-pl, while the trimethyldap peptide (Tmdap-pl) was less stable than Dap-pl. To their surprise, the dimethyldap peptide (Dmdap-pl) formed a homotrimer. Heterodimerization occurred with Mmdap-pl:Asp-pl and Tmdap-pl:Asp-pl, although the latter was relatively unstable. Dmdap-pl was incapable of forming a heterodimer with Asp-pl. Interestingly, the incorporation of one methyl group into the amine side

chain resulted in a more stable heterodimer than Dap-p1:Asp-p1, presumably due to a decrease in the desolvation penalty experienced upon folding (Figure 1.6B). This study further illustrates the utility of polar residues in the hydrophobic core directing heterodimer specificity.

Controlling heterodimerization is also possible through manipulations of the electrostatic interface. Kim and coworkers designed a peptide "velcro" system based on studies of two major leucine zippers: the homodimeric GCN4 and the heterodimeric Fos/Jun.[13] The *e* and *g* heptad positions were persubstituted with either lysine (base peptide) or glutamic acid (acid peptide); the central asparagine residue was maintained to favor the parallel orientation of helices. This system favors heterodimer formation by relieving destabilizing electrostatic interactions in the homodimers (Figure 1.7).

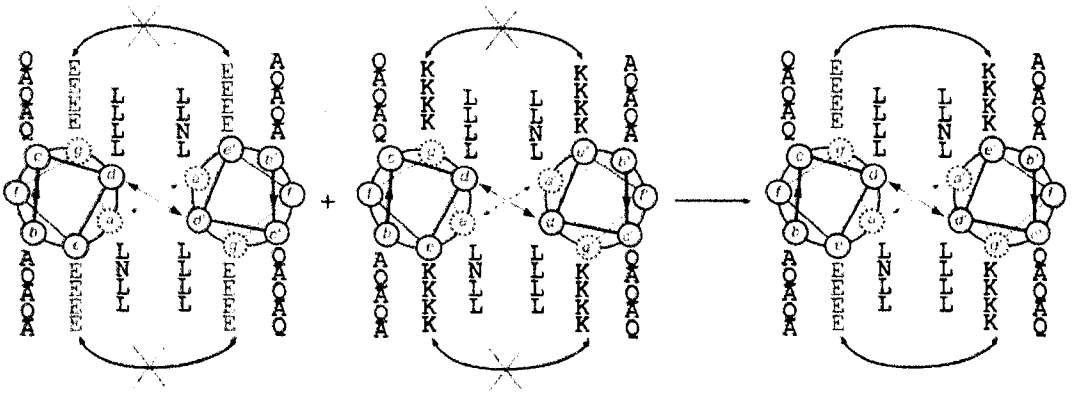
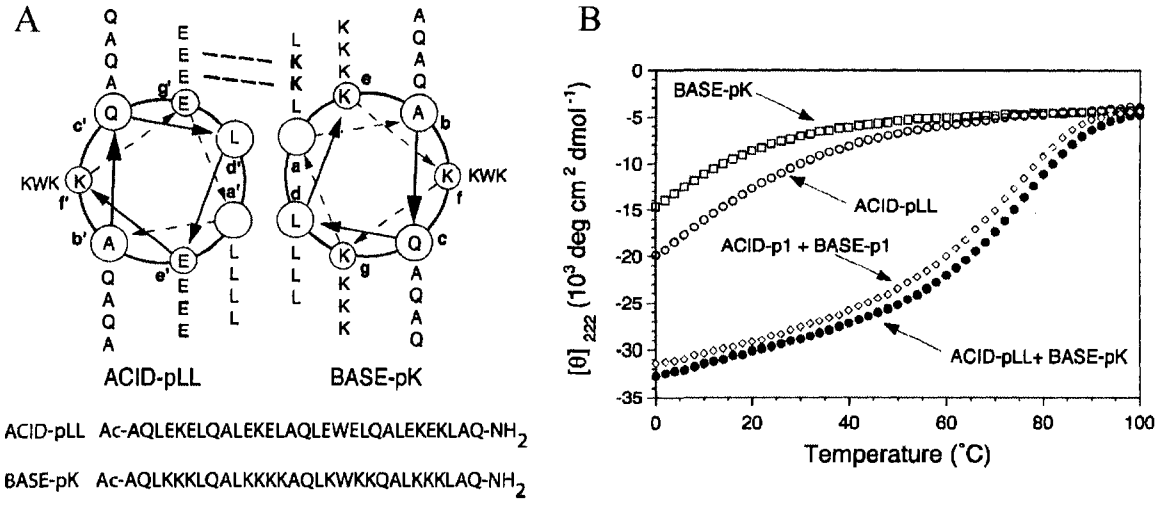


Figure 1.7. Peptide "Velcro." Helical wheel diagram depicting unfavorable electrostatic interactions of the homodimers leading to heterodimer formation.

Kim and coworkers further expanded this study by replacing the central asparagine residue in the Acid-p1 and Base-p1 with a leucine residue.[14] They found that these variants (Acid-pLL and Base-pLL) form stable heterotetramers, which do not fold into a unique structure, that is, the helices lack unique orientation. They concluded that

although the buried polar interactions are destabilizing, they provide an elegant mechanism for introducing structural uniqueness.

Drawing from this study, Lumb and coworkers sought to investigate other core polar interactions by mutating the core of the base peptide to contain two lysine residues, to produce Base-pK.[15] They found that these buried position *a* lysines were capable of polar interactions with the surface *e'* or *g'* Glu residues (Figure 1.8A). This Lys-Glu polar interaction is less destabilizing than the Asn-Asn interaction and is capable of imparting structural uniqueness but not orientation specificity (Figure 1.8B). That is, there was not a preference for parallel or antiparallel configuration.



ACID-pLL Ac-AQLEKELQALEKELAQLEWELQALEKEKLAQ-NH_2

BASE-pK Ac-AQLKKKLQALKKKKAQLKWKKQALKKKLAQ-NH_2

Figure 1.8. (A)Helical-wheel representation of Acid-pLL/Base-pK heterodimer shown as a parallel coiled coil. Polar interactions are denoted by the dashed line. (B) Thermal denaturation of Acid-pLL, Base-pK and Acid-pLL/Base-pK. Total peptide concentration is 10μM in PBS, pH 7.0. (Figure from Campbell *et al.*, 2002)

Utilizing favorable electrostatic interactions has also lead to heterotrimer formation.[16] Alber and coworkers maximized the number of favorable interhelical ion pairs to form three complimentary peptides by substituting complementary charged residues at the *e* and *g* positions (Figure 1.9A). This allowed not only for interhelical

interactions in the heterotrimer, but also for repulsive electrostatic interactions in the homotrimers. Additionally, isoleucine residues and a single glutamine residue were placed at the *a* and *d* positions to favor formation of trimers over dimers and tetramers. When these peptides are combined, there is an increased helicity, and melting temperatures were found to be 30 to 40 °C higher than any of the isolated components (Figure 1.9 B&C). An equimolar mixture of the three peptides forms a helical trimer with high specificity and stability.

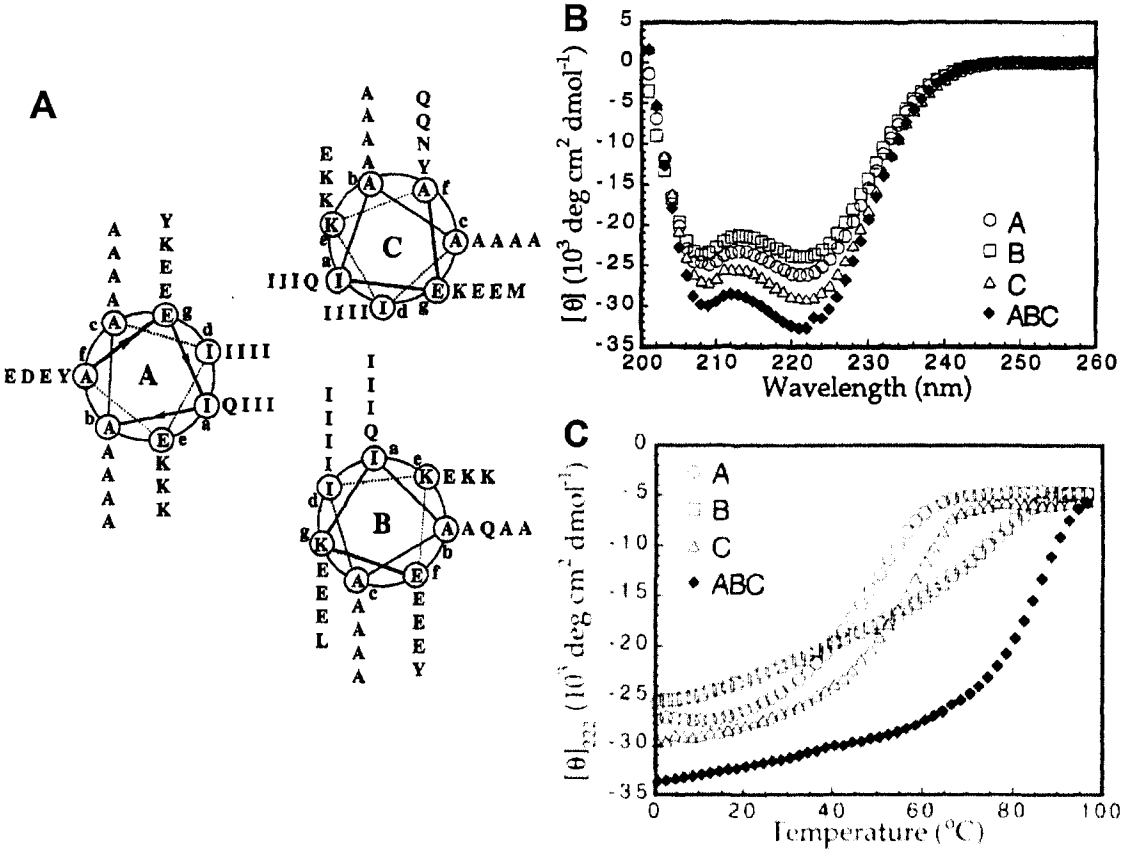


Figure 1.9. Peptides Investigated. (A) Helical wheel representation of the designed heterotrimeric coiled coil. CD spectra of the individual A, B, and C petides and an equimolar mixture of the three peptides, ABC, (B) wavelength scan and (C) thermal denaturation. (Figure from Nautiyal *et al.*, 1995)

These studies have demonstrated the buried polar interactions are necessary for orientation and oligomer specificity. Replacement of these residues shows that although

more stable complexes form, the oligomerization and strand specificity is lost. Additionally manipulating the electrostatic interface contributes to heterodimerization and in some cases assists with polar interactions with core residues. Thus, further manipulation is necessary for control of strand specificity.

1-4: Coiled Coil Design: Orientation

In addition to research on the control of oligomerization, orientation specificity has also garnered much attention. Similar to the previous oligomerization studies, various groups have manipulated of the electrostatic interface and hydrophobic core to direct helix orientation. For instance, in work by DeGrado and coworkers, the electrostatic interface of a known trimeric synthetic coiled coil, coil-Ser, was manipulated to minimize unfavorable interhelical ion pairs to form a stable antiparallel trimer.[17] Additionally, Hodges and coworkers have investigated the relative positions of alanine residues in the hydrophobic core to control the formation of parallel and antiparallel coiled coils.[18]

More recently, Kim and coworkers found that a buried polar interaction directs relative orientations of coiled coils.[19] Utilizing their peptide "velcro" template, they altered the position of the central asparagine residues such that a buried polar interaction would only be possible in an antiparallel configuration. Through the use of covalently linked, disulfide-containing heterodimers, they found that the antiparallel heterodimer is the predominant species at equilibrium when the heterodimers are allowed to undergo thiol—disulfide exchange (Figure 1.10). Remarkably, the potential for a buried polar interaction between Asn residues at the *a* and *d'* positions is sufficient to specify an

antiparallel helix orientation at a preference of 2.3 kcal/mol over the parallel orientation. This work has been the platform for other orientation studies.

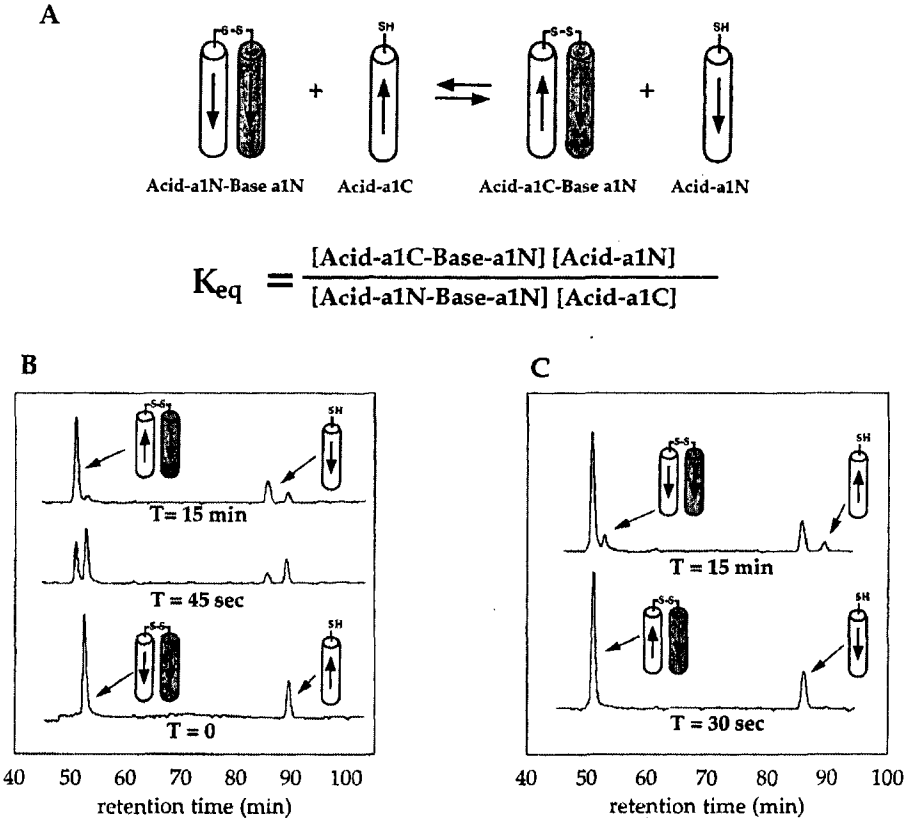


Figure 1.10. (A) Schematic view of the equilibrium thiol—disulfide exchange assay. Arrows indicate the direction of the peptide chain from the N to the C terminus. Rearrangements were carried out in PBS buffer (pH 7) at room temperature. (B) HPLC chromatogram showing the progress of the rearrangement from the parallel heterodimer to the antiparallel heterodimer. (C) HPLC chromatogram showing the essentially unchanged antiparallel heterodimer. (Figure from Oakley and Kim, 1998)

Oakley and coworkers further expanded on this previous work.[20] They kept the Asn residues in the same positions as Kim and coworkers, but they substituted a single residue at the *g* position in each peptide such that in the antiparallel orientation only attractive interactions are expected (Figure 1.11A). Conversely, two potentially repulsive interactions are expected in the parallel confirmation. Once again, to test the helix orientation in this heterodimer, equilibrium thiol—disulfide exchange assays were employed (Figure 1.11B-E). The antiparallel helix orientation was preferred by at least

4.4 kcal/mol over the parallel orientation. In a similar study, this group positioned a single charged residue in each peptide such that favorable interactions can occur only in the parallel orientation while maintaining the buried polar interaction that is only possible in an antiparallel configuration.[21] They found that that interhelical interactions and a buried polar interaction are of approximately equal importance for helix orientation specificity since there was no observable preference for an antiparallel orientation.

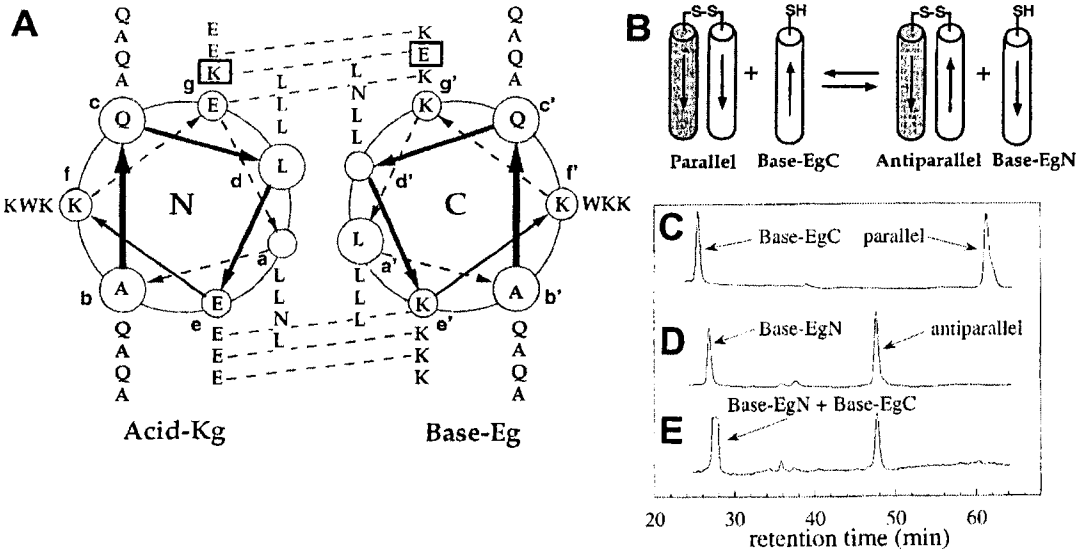


Figure 1.11. Antiparallel Heterodimers. (A) Helical wheel representation. (B) Schematic view of equilibrium thiol—disulfide exchange assay. HPLC chromatograms before (C), after (D) thiol—disulfide rearrangement. (E) HPLC trace showing rearrangement prducts of a 3:1 mixture of Base-EgN and Acid-KgN—Base-EgC.

In addition to manipulating the electrostatic interactions to control orientation, Oakley and coworkers also have modified the hydrophobic core.[22] Using modified acid and base peptides from previous studies, they sought to control orientation by burial of an interhelical salt-bridge involving a *d* position Arg residue on the Acid peptide and a *g'* position Glu residue on the Base peptide. This interaction is only possible in the antiparallel orientation. Additionally, they replaced both Asn core residues in both peptides with a Leu residue. Through thiol-disulfide exchange assays, they found that the potential for forming a partially buried salt-bridge is not sufficient to specify a preference

for orientation. When the peptides are tethered together in a parallel orientation, addition of the corresponding Cys containing peptide to form an antiparallel heterodimer does not lead a dramatic increase in antiparallel heterodimers; rather, it leads to a similar ratio of parallel to antiparallel species. This is also observed with the complementary experiment starting with the tethered antiparallel heterodimer. Interestingly, the single buried Arg residue does specify for a dimer at a lower thermodynamic cost than the two buried Asn residues in the core.

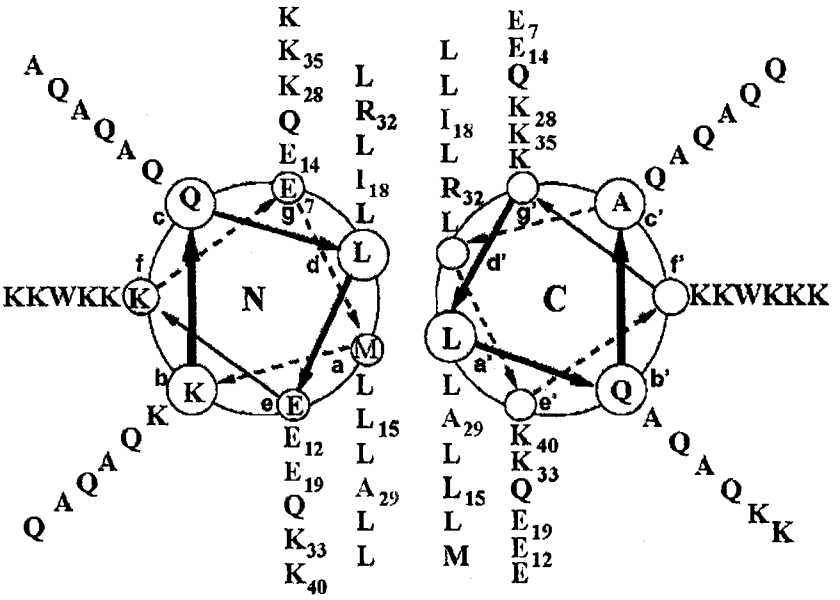


Figure 1.12. Helical wheel representation of the antiparallel homodimer APH. The view is shown in the antiparallel alignment with the basic (blue), acidic (red), and hydrophobic (green) residues expected to contribute to helix orientation highlighted. (Figure from Gurnon *et al.*, 2003)

Together with electrostatic matching and polar substitution, Oakley and coworkers employed nonpolar steric matching to form an antiparallel homodimer, APH.[23] This design includes Glu residues at the N-terminal *e* and *g* positions, Lys residues at the C-terminal *e* and *g* positions, Ile in the *d* position, and Ala in the *a* position (Figure 1.12). In the antiparallel orientation there are eight favorable electrostatic interactions and a favorable Ile-Ala interaction, whereas, in the parallel orientation there are eight

destabilizing interactions and a clashing Ile-Ile and Ala-Ala interactions. Analysis by analytical ultracentrifugation, CD and thiol-exchange experiments showed the design to be successful, forming a stable antiparallel homodimer. Additionally, they were able to successfully express APH in bacteria cells, thus distinguishing it from other model coiled coils and promoting its use in biological settings.

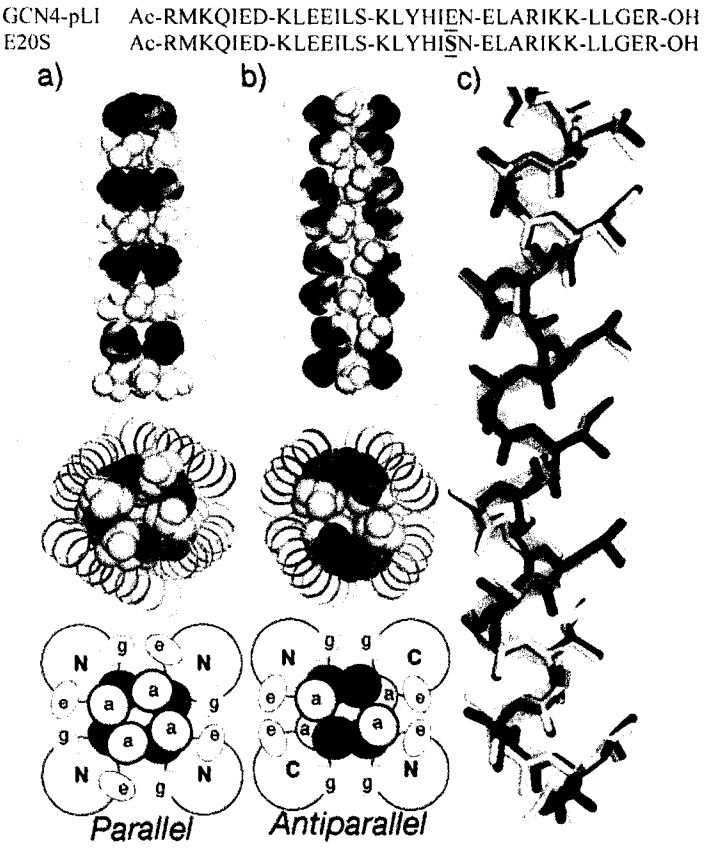


Figure 1.13. Crystal structures of the E20S variant in the parallel and antiparallel tetrameric configurations. Side and top views show the parallel (a) and antiparallel (b) structures. (c) Superposition of single helices from the antiparallel (blue) and parallel (beige) configurations. (Figure from Yadav *et al.*, 2006)

In addition to electrostatic matching and core modification to control orientation, Ghadiri and coworkers found that substituting a single solvent-exposed residue enabled the parallel coiled coil tetramer to populate the antiparallel configuration.[24] Using the GCN4-derived GCN4-pLI as a model, they observed that substituting a serine residue for

the glutamic acid residue at position 20 enabled the switch in crystallographic structure from the commonly observed parallel coiled coil to the antiparallel arrangement. The E20S variant crystallized in both the parallel and antiparallel configurations (Figure 1.13). This suggests that the two configurations are close enough in energy for subtle sequence changes to have important structural consequences. Additionally, they found that the hydrophobic and electrostatic patterning of these coiled coils can accommodate both parallel and antiparallel configurations with minor changes to the helix register and side chain conformations. Their findings emphasize the predisposition of the coiled-coil motif to adopt multiple configurations.

While these previous studies have focused on lateral interactions, Gellman and coworkers have turned their attention to vertical core interactions.[25] As previously discussed, hydrophobic side chains at *a* and *d* pack against each other in a "knobs into hole" motif to create the hydrophobic core. The interactions possible from this packing environment can be grouped into two categories: lateral and vertical interactions. Lateral interactions are defined as interactions between two side chains on a plane that is perpendicular to the helix axis. Vertical interactions involve interactions between two side chains on a plane that is parallel to the helix axis. In this study, the Gellman group investigated interactions between the *a'-a-a'* side chains in antiparallel coiled coils using a combination of experiments and protein data bank analysis. They found experimentally that hetero-triads such as Ile-Leu-Ile and Leu-Ile-Leu form more stable antiparallel complexes then the corresponding homo-triads Ile-Ile-Ile and Leu-Leu-Leu. This trend is found in natural antiparallel coiled coil dimers of known structure, where certain *a'-a-a'* hetero-triads, Ile-Leu-Ile and Leu-Ile-Leu, are favored relative to the corresponding

homo-triads (Table 1.2). This study illustrates that vertical interactions represent an important component to determining coiled coil interaction preferences in complex biological environments.

Table 1.2. Numbers of specific *a'-a-a'* combinations in natural antiparallel coiled coils

	a' = *a'* = Ile			*a'* = *a'* = Leu		
Amino acid at *a*	Observed	Expected	Observed/expected	Observed	Expected	Observed/expected
Ile	6	9.4	0.6	26	16.5	1.6
Leu	19	14.2	1.3	21	24.7	0.9
Sum (all amino acids)	59			103		

(Table from Hadley *et al.*, 2008)

1-5: Previous work in Kennan Group

Research in the Kennan group has focused on molecular assembly mediated by noncovalent forces. Using the α-helical coiled coil as a model for protein-protein association, our group has sought to use both natural and unnatural amino acids in both the hydrophobic and hydrophilic interfaces to develop new specific interactions capable of governing coiled coil formation. Initial work was inspired by findings of Alber and coworkers that trimerization of peptides bearing a core alanine could be induced by addition of a hydrophobic ligand.[26] Using similar alanine mutant peptides in a ratio of two to one with a peptide bearing a core cyclohexlyalanine, heterotrimerization was established (Figure 1.13). The ability of unnatural hydrophobic core side chains to direct specific aggregation has been the stepping stone of further research in this group.[27]

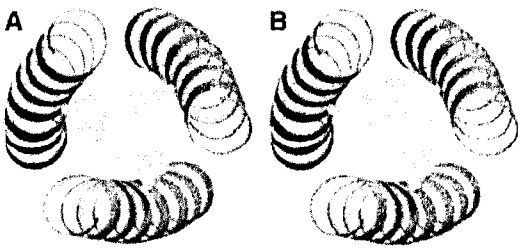


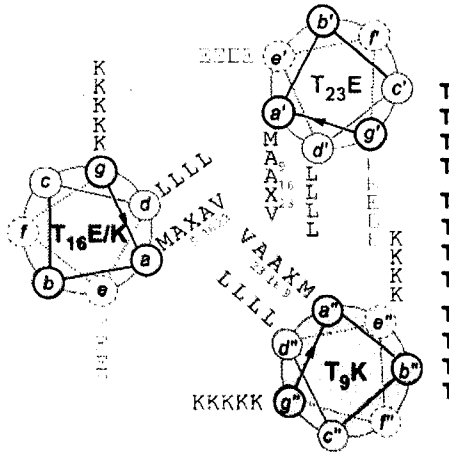
Figure 1.13. Heterotrimerization through steric matching. (A) The entirely Ala core homotrimer creates a pocket in the hydrophobic core. (B) The 2:1 complex of Ala_{16}:Chx_{16} shows the cyclohexlyananine side chain partially filling the void, stabilizing the heterotrimer. (Figure from Schnarr *et al.*, 2001)

Parallel studies of the isolated peptides of the first generation system found that only the alanine bearing peptide was significantly less stable than the heterotrimer. Specificity for the 2:1 complex appeared to arise mainly from poor interactions in the alanine peptide homotrimer. Additionally, the cyclohexylalanine peptides formed complexes even more stable than the designed one. To alleviate this, a second-generation system was developed, in which three sequential *a* residues were replaced, two with alanine and one with cyclohexylalanine.[28] An equimolar mixture of peptides with cyclohexyl sidechains at each possible position formed a parallel 1:1:1 heterotrimer, with each contributing the cyclohexyl group for a different core layer (Figure 1.14). The peptides involved in the heterotrimer were dubbed T_9, T_{16}, and T_{23} (the subscript denotes the position of the cyclohexylalanine residue), since the three-dimensional arrangement somewhat resembles the outcome of a tic-tac-toe game. The remaining sequence is based on the first-generation system, derived in turn from GCN4-p1 and related peptides.

Figure 1.14. Specificity through multiple interactions. Large core *a* sidechains pack against small ones at the same position of other strands. (Figure from Schnarr *et al.*, 2002)

Expanding on the hydrophobic forces governing hetrotrimerization, our group next sought to incorporate the hydrophilic interface to aid in self assembly.[29] To accomplish this, the *e/g* positions of each of the T_9, T_{16}, and T_{23} were substituted with glutamic acid or lysine to allow for favorable electrostatic interactions (Figure 1.15). Investigation of the possible assemblies found that complexes with matched cores but

two or three mismatched hydrophilic interfaces are as unstable as are those containing a single Glu/Glu *e/g* mismatch. However, a single Lys/Lys *e/g* mismatch is tolerated as is a single core alanine layer in the system with matched electrostatics to permit stable complex formation. Exploitation of these tolerated interactions has allowed for the construction of sophisticated assemblies as illustrated by formation of three specific heterotrimers from a mixture of six peptides. Additionally, other more complicated assemblies have been formed by manipulating the pH of the system to allow for less tolerated mismatches as well as sequential and specific replacement of one to all three of the heterotrimer stands.[30]

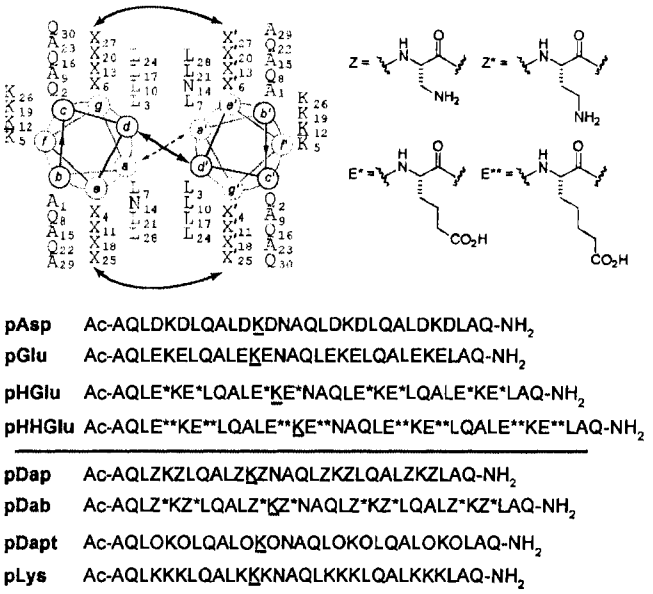


	a d a d a d a d a
T_9	AcNH-RMKQLEKKXEELLSKAQQLEKEAAQLKKLVG-Am
T_9E	AcNH-EMKQLEKEXEELESEAQQLEKEAAQLEKEVG-Am
T_9K	AcNH-KMKQLKKKXEELKSKAQQLKKKAAQLKKKVG-Am
T_9E/K	AcNH-KMKQLEKKXEELESKAQQLEKKAAQLEKKVG-Am
T_{16}	AcNH-RMKQLEKKAEELLSKXQQLEKEAAQLKKLVG-Am
$T_{16}E$	AcNH-EMKQLEKEAEELESEXQQLEKEAAQLEKEVG-Am
$T_{16}K$	AcNH-KMKQLKKKAEELKSKXQQLKKKAAQLKKKVG-Am
$T_{16}E/K$	AcNH-KMKQLEKKAEELESKXQQLEKKAAQLEKKVG-Am
T_{23}	AcNH-RMKQLEKKAEELLSKAQQLEKEXAQLKKLVG-Am
$T_{23}E$	AcNH-EMKQLEKEAEELESEAQQLEKEXAQLEKEVG-Am
$T_{23}K$	AcNH-KMKQLKKKAEELKSKAQQLKKKXAQLKKKVG-Am
$T_{23}E/K$	AcNH-KMKQLEKKAEELESKAQQLEKKXAQLEKKVG-Am
	1 5 10 15 20 25 30

Figure 1.15. Peptides investigated. Helical wheel projection of the totally matched trimer (T_9K:$T_{16}E/K$:$T_{23}E$) to illustrate the interfaces involved (left). Each sequence derives from one of the three parents (T_9, T_{16}, T_{23}) by replacement of all *e/g* residues with Glu (T_nE), Lys (T_nK), or both (T_nE/K) (right). X = cyclohexylalanine. The positions of core modification in the parent peptides are underlined.

Further manipulation of the hydrophobic core has been investigated. By repositioning the *a* substituted residues to the *d* position, antiparallel heterodimerization occurred.[31] Additional utilization of the electrostatic interface with this system allows for switchable strand orientation in peptide assembly by changing the pH.[32] Other hydrophobic core manipulations include the use of phenylalanine as a steric matching

partner in place of cyclohexylalanine.[33] While the phenylalanine derivative was sufficient for trimerization, it was found to be less stable than the parent cyclohexylalanine heterotrimer.



pAsp	Ac-AQLDKDLQALDKDNAQLDKDLQALDKDLAQ-NH_2
pGlu	Ac-AQLEKELQALEKENAQLEKELQALEKELAQ-NH_2
pHGlu	Ac-AQLE*KE*LQALE*KE*NAQLE*KE*LQALE*KE*LAQ-NH_2
pHHGlu	Ac-AQLE**KE**LQALE**KE**NAQLE**KE**LQALE**KE**LAQ-NH_2
pDap	Ac-AQLZKZLQALZKZNAQLZKZLQALZKZLAQ-NH_2
pDab	Ac-AQLZ*KZ*LQALZ*KZ*NAQLZ*KZ*LQALZ*KZ*LAQ-NH_2
pDapt	Ac-AQLOKOLQALOKONAQLOKOLQALOKOLAQ-NH_2
pLys	Ac-AQLKKKLQALKKKNAQLKKKLQALKKKLAQ-NH_2

Figure 1.16. Peptides used. Helical wheel projection demonstrates interhelical *e/g* contatcts by double headed arrows. Sequences of each peptide are given below, with nonostandard amino acids represented by the following code: Z = 2,3-diaminopropionic acid, Z* = 2,4-diaminobutyric acid, E* = homoglutamic acid, E** = homohomoglutamic acid, O = 2,5-diaminopentanoic acid.

In addition to trimer formation, our group has investigated dimerization by manipulating the electrostatic interface.[34] Using the peptide "velcro" system as a template our group has manipulated the electrostatic interface by increasing and decreasing the side chain length of the parent glutamic acid and lysine residues by single methylene units (Figure 1.16). The stability of the dimers comprised of the individual peptides whose e/g residues contain 1-4 methylene units between the backbone and polar terminus were extremely varied. Heterodimers containing 7-8 spacer units exhibited the greatest stability. To establish the stability of these individual dimers, a series of exchange experiments in which one strand of an original heterodimer was specifically

replaced by an added one were preformed. Remarkably, the native-like Glu/Lys pairings were found to undergo exchange with the corresponding derivatized Acid/Base peptide.

This preliminary work in our lab has served as a backbone for further investigations in molecular self-assembly. Starting from a simple concept, complex assemblies requiring both hydrophobic and hydrophilic specificity elements have been targeted. The following chapters describe our progress with polar specificity elements within the hydrophobic core, in both dimer and trimer systems.

1-6: Literature Cited

[1] (a) Mason, J. M.; Arndt, K. M., *Chem. Bio. Chem.* **2004**, 5, 170-176. (b) Engel, D. E.; DeGrado, W. F., *J. Mol. Biol.* **2004**, 337, 1195-1205. (c) Hadley, E. B.; Gellman, S. H., *J. Am. Chem. Soc.* **2006**, 128, 16444-16445. (d) Mason, J. M.; Schmitz, M. A.; Muller, K. M.; Arndt, K. M., *Proc. Nat. Acad. Sci.* **2007**, 103, 8989-8994. (e) Nikolaev, Y.; Pervushin, K., *J. Am. Chem. Soc.* **2007**, 129, 6461-6469.

[2] (a) Woolfson, D. N., *Adv. Protein Chem.* **2005**, 70, 79-112. (b) Parry, D. A. D.; Squire, J. M., *Adv. Protein Chem.* **2005**, 70, 1-10. (c) Lupas, A. N.; Gruber, M., Adv. *Protein Chem.* **2005**, 70, 37-78. (d) Cheng, R. P., *Curr. Opin. Struct. Biol.* **2004**, 14, 512-520. (e) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A., *Annu. Rev. Biochem.* **1999**, 68, 779-819.

[3] O'Neal, K.T.; DeGrado, W.F., *Science* **1990**, 250, 646-651.

[4] Crick, F., *Acta Crystallogr.* **1953**, 6, 689-697.

[5] (a) Liu, J.; Deng, Y.; Cheng, C.; Kallenbach, N. R.; Lu, M., *Proc. Nat. Acad. Sci.* **2007**, 103, 15457-15462. (b) Steinmetz, M. O.; Jelesarov, I.; Matousek, W. M.; Honnappa, S.;

Jahnke, W.; Missimer, J. H.; Frank, S.; Alexandrescu, A. T.; Kammerer, R. A., *Proc. Nat. Acad. Sci.* **2007**, 104, 7062-7067. (c) Tsurkan, M. V.; Ogawa, M. Y., *Inorg. Chem.* **2007**, 46, 6849-6851. (d) Portwich, M.; Keller, S.; Strauss, H. M.; Mahrenholz, C. C.; Kretzshmar, I.; Kramer, A.; Volkmer, R., *Angew. Chem. Int. Ed.* **2007**, 46, 1654-1657. (e) Straussman, R.; Ben-Ya'acov, A.; Woolfson, D. N.; Ravid, S., *J. Mol. Biol.* **2007**, 366, 1232-1242.

[6] (a) Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S., *Protein Science* **1999**, 8, 2312-2329. (b) Arndt, K. M.; Pelletier, J. N.; Muller, K. M.; Alber, T.; Michnick, S. W.; Pluckthun, A., *J. Mol. Biol.* **2002**, 295, 627-639. (c) Akey, D. L.; Malashkevich, V. N.; Kim, P.S., *Biochemistry* **2001**, 40, 6352-6360. (d) Litowski, J. R.; Hodges, R. S., *J. Biol. Chem.* **2002**, 277, 37272-37279. (e) Stockner, T.; Ash, W. L.; MacCallum, J. L.; Tieleman, D. P., *Biophys. J.* **2004**, 87, 1650-1656. (f) Zimenkov, Y.; Dublin, S. N.; Ni, R.; Tu, R. S.; Breedveld, V.; Apkarian, R. P.; Conticello, V. P., *J. Am. Chem. Soc.* **2006**, 128, 6770-6771. (g) Yadav, M. K.; Leman, L. J.; Price, D. J.; Brooks III, C. L.; Stout, C. D.; Ghadiri, M. R., *Biochemistry* **2006**, 45, 4463-4473.

[7] O'Shea, E.K.; Klemm, J.D.; Kim, P.S.; Alber, T., *Science* **1991**, 254, 539-254.

[8] Harbury, P.B.; Zhang, T.; Kim, P. S.; Alber, T., *Science* **1993**, 262, 1401-1407.

[9] Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T., *Nat. Struct. Biol.* **1996**, 3, 1011-1018.

[10] Zeng, X.; Herndon, A. M.; Hu, J. C., *Proc. Nat. Acad. Sci.* **1997**, 94, 3673-3678.

[11] Schneider, J. P.; Lear, J. D.; DeGrado, W. F., *J. Am. Chem. Soc.* **1997**, 119, 5742-5743.

[12] Kretisinger, J. K.; Schneider, J. P., *J. Am. Chem. Soc.* **2003**, 125, 7907-7913.

[13] O'Shea, E. K.; Lumb, K. J.; Kim, P. S., *Current Biology* **1993**, 3, 658-667.

[14] Lumb, K. J.; Kim, P. S., *Biochemistry* **1995**, 34, 8642-8648.

[15] Campbell, K. M.; Lumb, K. J., *Biochemistry* **2002**, 41, 7169-7175.

[16] Nautiyal, S.; Woolfson, D. N.; King, D. S.; Alber, T., *Biochemistry* **1995**, 34, 11645-11651.

[17] Lombardi, A.; Bryson, J. W.; DeGrado, W. F., *Biopolymers* **1997**, 40, 495-504.

[18] Monera, O. D.; Zhou, N. E.; Lavigne, P.; Kay, C. M.; Hodges, R. S., *J. Biol. Chem.* **1996**, 271, 3995-4001.

[19] Oakley, M.G.; Kim, P. S., *Biochemistry* **1998**, 37, 12603-12610.

[20] McClain, D. L.; Woods, H. L.; Oakley, M. G., *J. Am. Chem. Soc.* **2001**, 123, 3151-3152.

[21] McClain, D. L.; Binfet, J. P.; Oakley, M. G., *J. Mol. Biol.* **2001**, 313, 371-381.

[22] McClain, D. L.; Gurnon, D. G.; Oakley, M. G., *J. Mol. Biol.* **2002**, 324, 257-270.

[23] Gurnon, D. G.; Whitaker, J. A.; Oakley, M. G., *J. Am. Chem. Soc.* **2003**, 125, 7518-7519.

[24] Yadav, M. K.; Leman, L. J.; Price, D. J.; Brooks, C. L. III; Stout, C. D.; Ghadiri, M. R., *Biochemistry* **2006**, 45, 4463-4473.

[25] Hadley, E. B.; Testa, O. D.; Woolfson, D. N.; Gellman, S. H., *Proc. Nat. Acad. Sci.* **2008**, 105, 530-535.

[26] Gonzalez Jr., L.; Plecs, J. J.; Alber, T., *Nat. Struct. Biol.* **1996**, 3, 510-515.

[27] Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2001**, 123, 11081-11082.

[28] Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2002**, 124, 9779-9783.

[29] Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2003**, 125, 667-671.

[30] (a) Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2003**, 125, 6364-6365. (b) Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2003**, 125, 13046-13051.

[31] Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2004**, 126, 14447-14451.

[32] Schnarr, N. A.; Kennan, A. J., *Org. Lett.* **2005**, 7, 395-398.

[33] Travisano, P.; Kennan, A. J., *Org. Lett.* **2004**, 6, 4219-4222.

[34] Ryan, S. J.; Kennan, A. J., *J. Am. Chem. Soc.* **2007**, 129, 10255-10260.

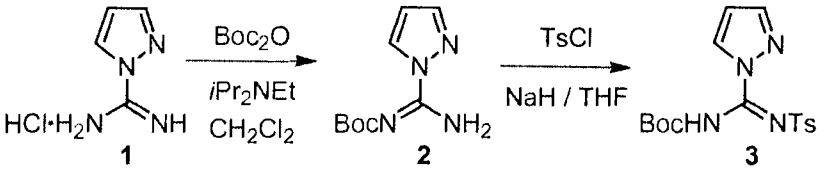
Chapter 2

Unnatural Amino Acid Synthesis and Peptide Incorporation

2-1: Introduction

The most common buried polar interaction found in nature for dimermeric coiled coils involves hydrogen bonding between two asparagine residues. The burial of this residue imparts structural uniqueness with respect to both oligomerization state and strand orientation.[1] Exploration of other residues that can impart this same specificity has garnered much attention.[2] These residues include mainly natural and commercially available unnatural amino acids, although a few of these studies have included synthetically produced, unnatural amino acids.[3] Similarly, work in our group has focused on molecular recognition motifs involving the use of commercially available unnatural amino acids.[4] To dramatically increase the investigation of various polar interactions capable of specific recognition events, we sought to develop a convenient synthetic methodology for functionalizing side chains of commercially available amino acid. Since varying side chain length lysine derivatives are readily available, functional groups involving nitrogen seemed to be a likely target. Thus, the main focus of this research has been on the functionalization of the terminal amine group on lysines and its commercially available derivatives.

2-2: Guanidinylation



Scheme 2.1 Reagent Preparation

One such functional group that has a widespread occurrence in natural products is guanidine. Guanidine groups have inspired considerable development of synthetic methodology.[5] There have been a number of elegant methods for on-resin

functionalization reported involving more than one step.[6] Previous work in our group has eliminated the use of additional steps with the development of pyrazole derivative, **3**, capable of guanidinylating primary amines with impressive speed and efficiency.[7] Synthesis of this reagent is straightforward involving a two-step sequence from commercially available pyrazole, **1**, in a 41% yield (Scheme 2.1).

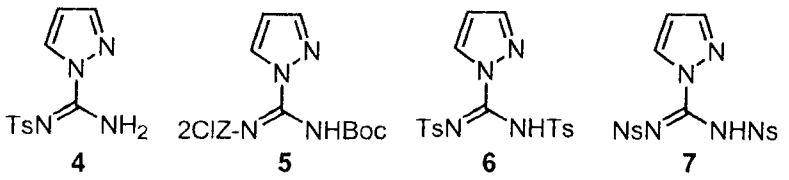
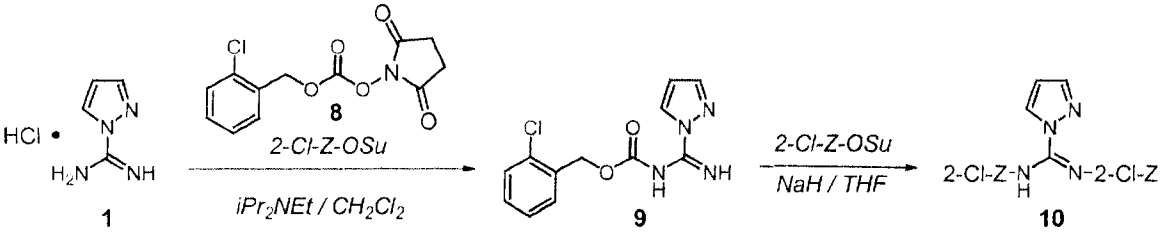


Figure 2.1. Other guanidinylating reagents investigated.

Due to difficulties synthesizing this guanidinylation reagent, **3**, various derivatives of the commercially available pyrazole were synthesized and investigated as guanidinylating reagents (Figure 2.1). These reagents were synthesized under similar conditions to those in Scheme 1. The mono-tosyl protected pyrazole derivative, **4**, was isolated in good yield (90%), but only guanidinylated 35% of hexylamine in solution phase. The Boc and 2-Cl-Z protected pyrazole derivative, **5**, was isolated in poor yield (15%), but the guanidinylation of the primary amine in solution phase was extremely successful with yields near 100%. The di-tosyl protected pyrazole, **6**, and the di-nosyl protected pyrazole, **7**, derivatives were never isolated. The major products of these reactions were the tosyl- and nosyl-protected pyrazole.



Scheme 2.2. Second Generation Reagent Preparation

Previous work had shown that a di-Cbz protected pyrazole derivative was found to functionalize an ornithine side chain on solid phase, although the guanidinylation demanded somewhat elevated temperatures and extended reaction times.[8] To avoid similar reaction conditions, a di-2-Cl-Z protected pyrazole derivative, **10**, was synthesized from commercially available pyrazole carboxamidine, **1**, by sequential reaction with 2-Cl-Z succinimidyl ester, **8**, in a 68% yield (Scheme 2.2). Additionally, the di-2-Cl-Z derivative, **10**, guanidinylated primary amines in the solution phase in excellent yields (~100%). Since this synthesis proceeded in higher yields than the previous guanidinylation reagent and worked in solution phase, a short test peptide was synthesized to examine the guanidinylating capabilities on solid phase (Figure 2.2). The test peptide synthesis was successful so the di-2-Cl-Z derivative, **10**, was used for all of the full length guanidinylated peptides.

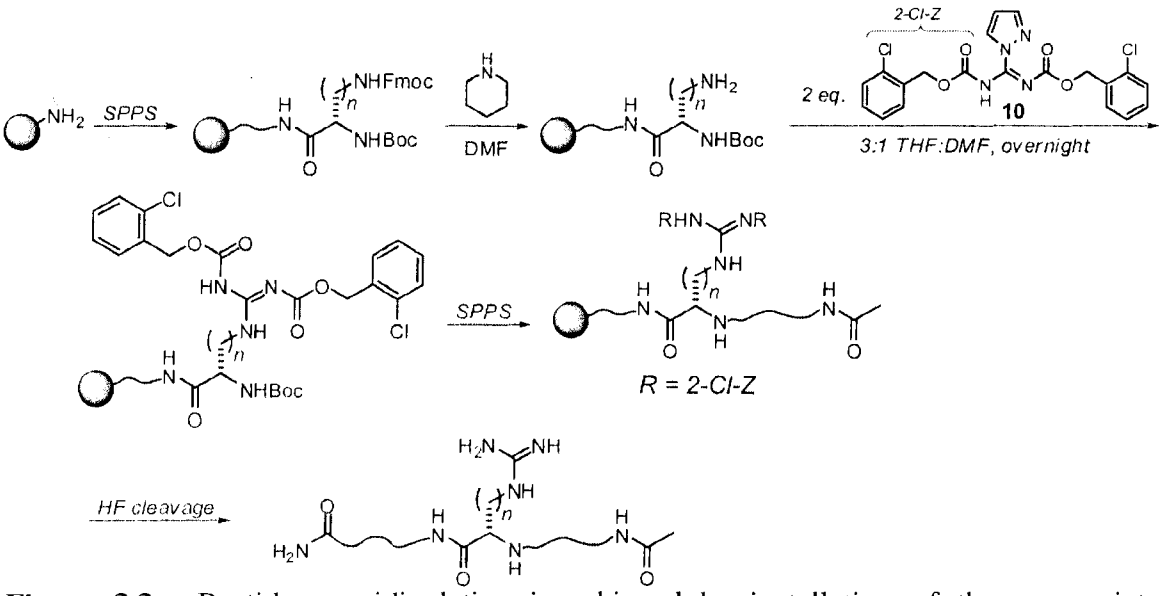
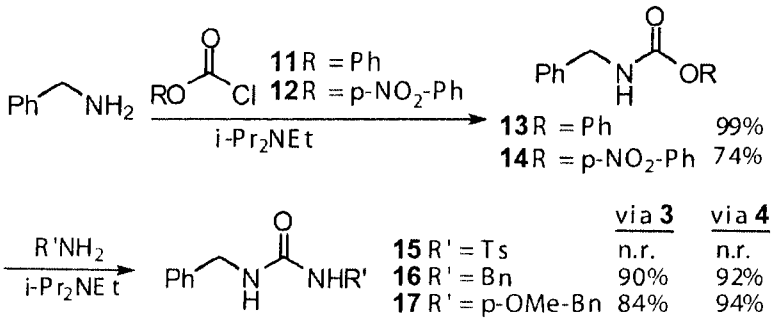


Figure 2.2. Peptide guanidinylation is achieved by installation of the appropriate differentially protected diaminoacid, followed by side chain deprotection (20% piperdine, DMF, 20 minutes) and reaction with the reagent, **10,** (2 eq., 3:1 THF:DMF, room temperature, overnight). Side chain 2-Cl-Z groups are removed during the normal course of HF cleavage from the resin, without any additional steps required.

2-3: Urea Formation

Similar to the guanidine group, urea functional groups also exhibit a widespread abundance in natural products. Ureas have long been employed as crucial recognition elements, principally via hydrogen bonding interactions.[9] Due to this, numerous synthetic methods have been reported in both solution and solid phase synthesis. These syntheses have primarily involved going through an isocyanate[10] or carbamate[11] intermediate. Yet, the majority of these methods include considerable N-substitution of the urea.

In the course of self recognition studies, mono-alkyl ureas (i.e. citrulline analogs) were desired. Therefore, we sought to develop a single method which could convert variable length primary amine side chains to the corresponding ureas, using chemistry and protection schemes compatible with Boc solid phase peptide synthesis (SPPS).[12] Since the carbamate intermediates tend to be more stable, we felt that they would be better suited for SPPS.



Scheme 2.3. Solution phase tests, n.r. no reaction.

Preliminary solution experiments were preformed to ascertain the basic feasibility of the urea formation reaction. Benzyl amine was chosen as a convenient stand-in for the side chain to be functionalized, and carbamate formation by both phenyl, **11,** and *p*-nitrophenyl chloroformate, **12,** were examined. Both carbamates were synthesized in

reasonable yields (Scheme 2.3). However, selecting the amine nucleophile was not as straightforward.

In selecting candidate amine nucleophiles, we focused on those that could reasonably be expected to produce the desired mono-alkyl urea under standard Boc SPPS cleavage conditions. Thus, amine nucleophile candidates were based on those common to protecting groups found in Boc SPPS, since we did not want to involve a separate deprotection step. Initially we examined tosyl amine, since the tosyl group is readily removed under standard Boc cleavage conditions (anhydrous HF). Unfortunately, neither carbamate reacted with the tosyl protected amine under the mild conditions required for convenient transfer to solid-phase applications (*i*-Pr_2NEt, THF, room temperature).

Next, the more electron rich amines, benzyl and *p*-methoxybenzyl amines, were examined due to the widespread application of benzyl derivatives in the protection of acids and alcohols in Boc SPPS. Both amines reacted with either carbamate to give the appropriate ureas in good yield (Scheme 2.3).

Since the solution phase urea formation was successful, efforts were turned toward urea formation in SPPS. To examine whether the reaction could be carried out without significant introduction of additional synthetic side products, a nine residue test peptide containing a single citrulline, Cit, residue was synthesize (**18**, Figure 2.3A).

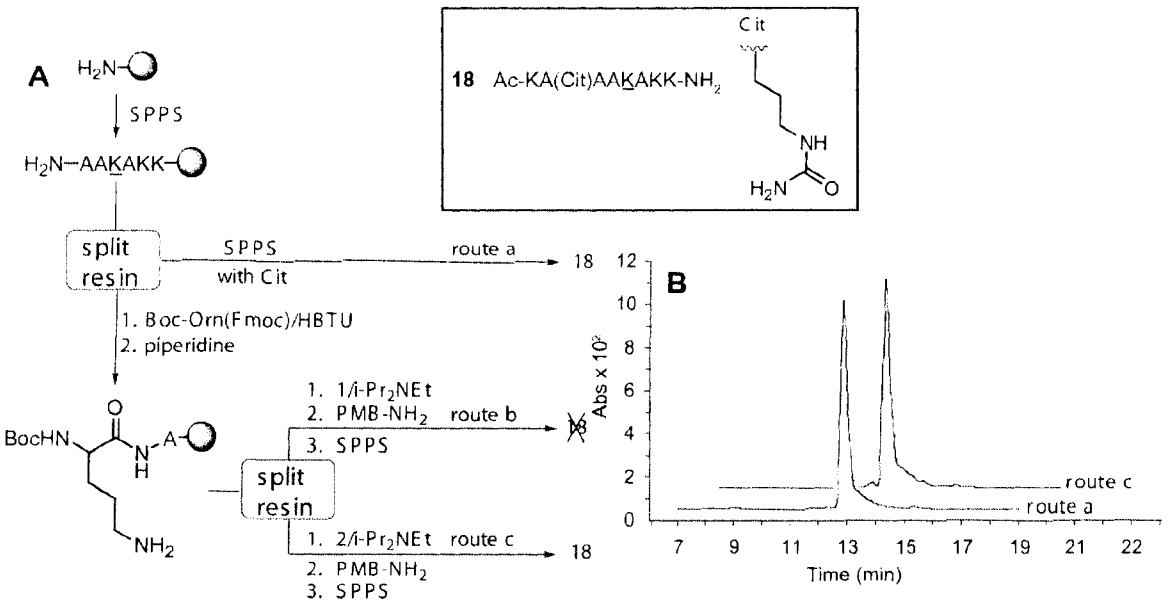


Figure 2.3. Strategy for evaluating the synthesis of peptide **18**. The same sequence was prepared using an authentic Cit monomer, the others formed the urea using chloroformates **11** or **12** and *p*-methoxybenzyl amine. Underlined lysine side chains are capped with acetamidobenzoate groups as spectroscopic labels.

In an effort to isolate side reactions due to urea formation from those incurred in the normal course of SPPS, we prepared the sequence via three different routes (Figure 2.3A). The first involved standard SPPS using an authentic Cit monomer, while the other two were derived from the same initial synthesis of the first seven residues. Following Fmoc removal on the side chain amine of α-Boc-protected ornithine residue, the resin was split equally and subjected to urea formation using either carbamate **11** or **12** and *p*-methoxybenzyl amine (route b or c, Figure 2.3A).

The three routes to **18** afforded clean products, although only those prepared via routes a and c gave the appropriate mass by ESI analysis. Comparing the HPLC traces of the crude peptides, one prepared using authentic Cit and the other using the synthetic sequence with chloroformate **12**, reveals that they are essentially identical (Figure 2.3B).

Additionally, this similarity suggests that urea formation reactions do not contribute to significant formation of additional side products.

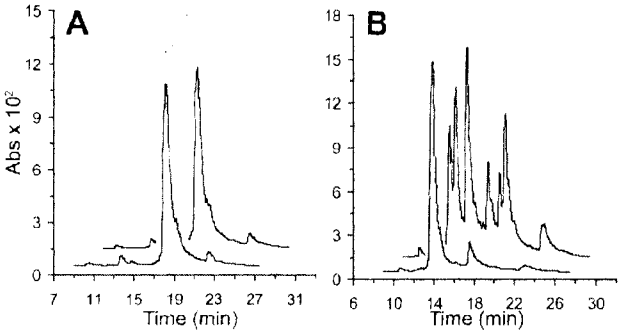


Figure 2.4. Comparison of on-resin carbamate reactivity. (A) Crude HPLC traces from attempted synthesis of peptide analogous to **18** using **11**, in which the second step converting carbamate to urea was included (front trace) or omitted (rear trace). Similarity of the traces suggests that the phenyl carbamate is unreactive. (B) The same experiment except using **12** to form the *p*-nitrophenyl carbamate. When the amine step is present, the synthesis is largely successful (front trace), but when it is omitted the more reactive carbamate leads to numerous side reactions (rear trace).

The peptide prepared via route b, using chloroformate **11**, afforded a product that was 77 mass units too high, which would fit the structure bearing an unreacted phenyl carbamate. Although the analogous test structures were successful in solution phase synthesis, this was not the case in solid phase syntheses. The apparent heterogeneity of the solid phase environment impeded the reaction. To test this, the synthesis of an analogous peptide to **18** was prepared using route b except for half the resin was removed following carbamate formation. The other half was reacted with benzyl amine in an attempt at urea formation, and then the last three residues were attached to each portion of the resin using standard SPPS methods. The crude HPLC traces of both peptides are essentially identical, supporting the idea that reaction via **11** stops at the carbamate stage (Figure 2.4A). The analogous experiment using chloroformate **12** reveals considerable side reactions in the peptide whose carbamate was not converted to the urea, consistent with the expectation of the *p*-nitro derivative being more reactive (Figure 2.4B).

Additionally, this HPLC trace reveals the formation of a side product, which upon ESI analysis was found to be the N-benzyl urea peptide. To avoid this, *p*-methoxybenzyl amine was used as the amine nucleophile in the urea formation reactions.

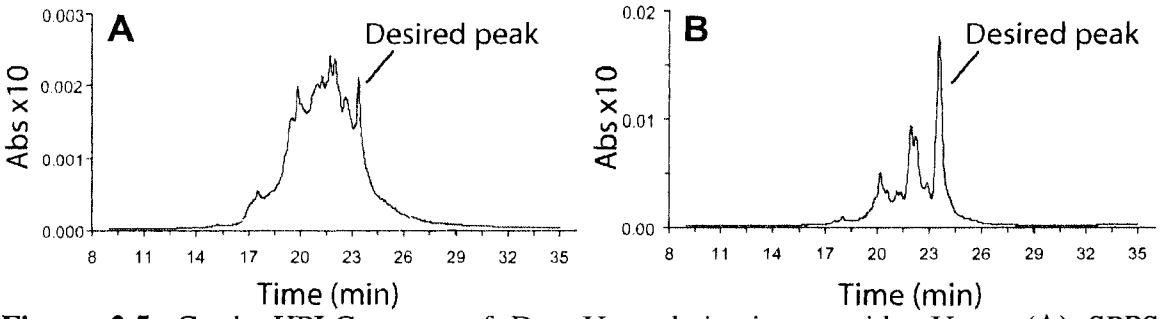


Figure 2.5. Crude HPLC traces of Dap Urea derivative peptide, Ur_K. (**A**) SPPS completed via first generation urea formation: carbamate formation then urea formation with the addition of *p*-methoxybenzylamine. (**B**) SPPS completed via second generation urea formation: addition of the *p*-methoxybenzyl-*p*-nitrophenyl carbamate to the unprotected side chain amine.

After success with the test peptide urea formation on both the ornithine and lysine terminal amines, shorter sidechains were investigated. Both diaminobuteric acid (Dab) and diaminopropionic acid (Dap) urea derivatives were synthesized using the *p*-nitrophenyl chloroformate followed by addition of the *p*-methoxybenzyl amine. The Dab urea derivative synthesis was successful but the Dap urea derivative proved to be more difficult (Figure 2.5A).

Therefore, a second generation urea formation method was developed. Instead of the previous two step, on resin urea formation reaction, a single step urea formation method was developed. This new method involves the addition of four equivalents of *p*-methoxybenzyl-*p*-nitrophenyl carbamate, **19,** to the resin bound unprotected side chain amine (Figure2.6). Use of the pre-formed carbamate, **19,** dramatically reduces side product formation (Figure 2.5B).

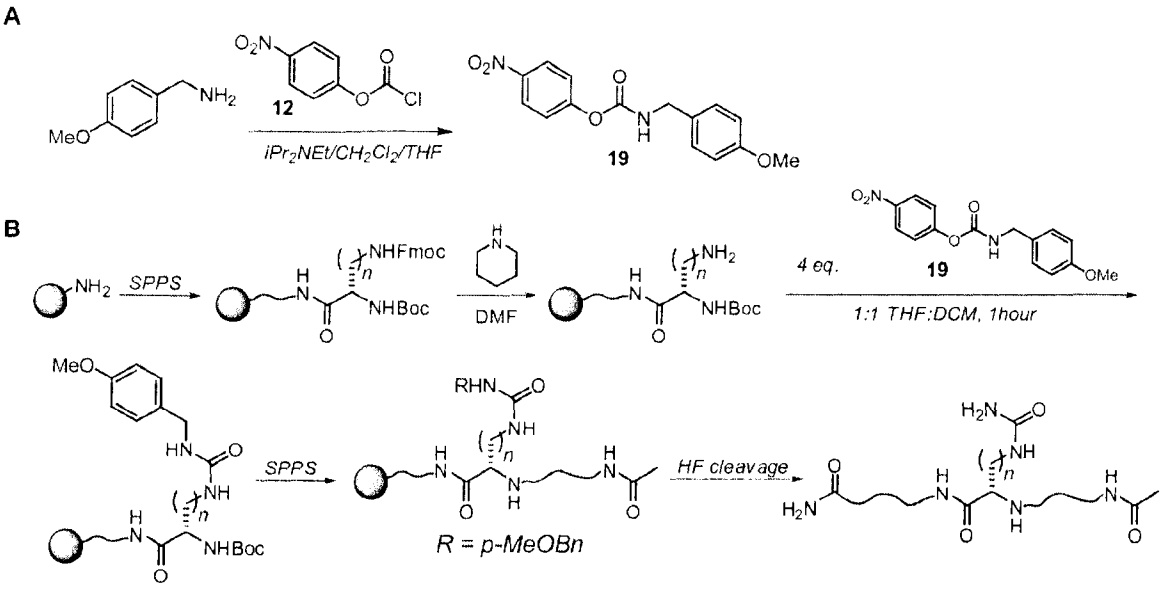


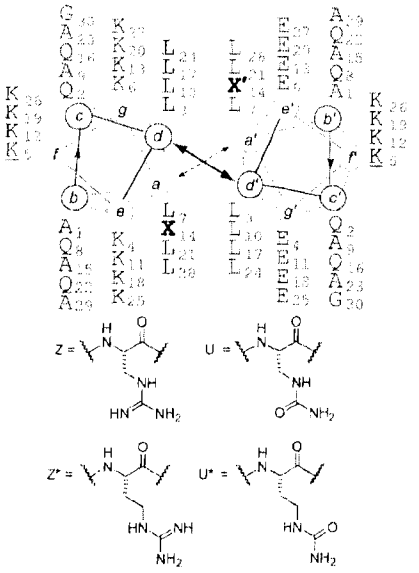
Figure 2.6. On-resin urea formation. (**A**) *p*-methoxybenzyl-*p*-nitrophenyl Carbamate, **19**, was prepared from commercially available *p*-MeObenzylamine and *p*-nitrophenylchloroformate, **12**. (**B**) Peptide urea formation is achieved by installation of the appropriate differentially protected diaminoacid, followed by side chain deprotection (20% piperidine, DMF, 20 minutes) and reaction with the reagent (4 eq., 1:1 THF:DCM, room temperature, 1 hour). The side chain *p*-methoxybenzyl group comes off during the normal course of HF cleavage from the resin, without any additional steps required.

2-4: Full Length Peptides

To examine the self-assembling capabilities of peptides containing these functional groups, longer peptides were synthesized (30 residues) with the polar residue at position 14. These sequences for probing polar core interactions are derived from the Acid-p1/Base-p1 heterodimer reported by Kim and coworkers, itself derived from the coiled coil domain of the leucine zipper, GCN4 (Figure 2.7).[13] Each peptide contains either Glu or Lys at all *e/g* positions, to discourage homodimer formation. The seven basic and nine acidic peptides differ only by identity of the core polar residue (position 14). Six contain core guanidine groups, spaced from the backbone of acidic or basic sequences by one ($pGd_{K/E}$), two ($pGd^*_{K/E}$) or three ($pArg_{K/E}$) methylenes. Similarly, six contain core urea groups, spaced from the backbone of acidic or basic sequences by one

($pUr_{K/E}$), two ($pUr^*_{K/E}$) or three ($pCit_{K/E}$) methylenes. Two contain asparagine in the core ($pAsn_{K/E}$), and two contain an acidic residue in the core, aspartic acid ($pAsp_E$) and glutamic acid ($pGlu_E$).

For the incorporation of unnatural amino acids, methods as described in the previous sections were used to functionalize the terminal amine of diaminopropionic acid (Dap) for the Gd and the Ur peptides and diaminobuteric acid (Dab) for the Gd* and the Ur* peptides. The second generation urea formation was used for the Ur peptides, while the first generation method was used for the Ur* peptides.



pUr_K	Ac-AQLKKKLQALKKK**U**AQLKKKLQALKKKLAG-NH_2
pUr^*_K	Ac-AQLKKKLQALKKK**U***AQLKKKLQALKKKLAG-NH_2
$pCit_K$	Ac-AQLKKKLQALKKK**Cit**AQLKKKLQALKKKLAG-NH_2
pGd_K	Ac-AQLKKKLQALKKK**Z**AQLKKKLQALKKKLAG-NH_2
pGd^*_K	Ac-AQLKKKLQALKKK**Z***AQLKKKLQALKKKLAG-NH_2
$pArg_K$	Ac-AQLKKKLQALKKK**R**AQLKKKLQALKKKLAG-NH_2
$pAsn_K$	Ac-AQLKKKLQALKKK**N**AQLKKKLQALKKKLAG-NH_2
pUr_E	Ac-AQLEKELQALEKE**U**AQLEKELQALEKELAG-NH_2
pUr^*_E	Ac-AQLEKELQALEKE**U***AQLEKELQALEKELAG-NH_2
$pCit_E$	Ac-AQLEKELQALEKE**Cit**AQLEKELQALEKELAG-NH_2
pGd_E	Ac-AQLEKELQALEKE**Z**AQLEKELQALEKELAG-NH_2
pGd^*_E	Ac-AQLEKELQALEKE**Z***AQLEKELQALEKELAG-NH_2
$pArg_E$	Ac-AQLEKELQALEKE**R**AQLEKELQALEKELAG-NH_2
$pAsn_E$	Ac-AQLEKELQALEKE**N**AQLEKELQALEKELAG-NH_2
$pAsp_E$	Ac-AQLEKELQALEKE**D**AQLEKELQALEKELAG-NH_2
$pGlu_E$	Ac-AQLEKELQALEKE**E**AQLEKELQALEKELAG-NH_2

Figure 2.7. Peptides Synthesized. Helical wheel projection (above left) demonstrates interactions in heterodimers. Buried polar residues are represented by X or X' (see position 14). Sequences of each peptide are given below, with non-standard amino acids represented by the following code (structures of each given bottom left): Z/U = guanidine/urea functionalized diaminopropionic acid, Z*/U* = guanidine/urea functionalized diaminobutyric aid, Cit = citrulline. Underlined lysine side chains are capped with acetamidobenzoate groups as spectroscopic labels.

2-5: Experimental Information

General Procedures:

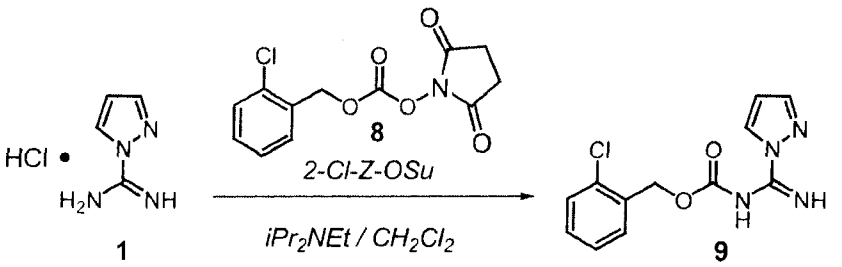
Tetrahydrofuran was freshly distilled from sodium-benzophenone ketyl. Methylene chloride and diisopropylethyl amine (DIEA) were distilled from calcium hydride. Column chromatography was performed on EM Science silica gel 60 (230 – 400 Mesh). Thin layer chromatography was performed on EM Science 0.25 mm silica gel 60-F plates. Visualization was accomplished with UV light or $KMnO_4$ followed by heating. All chemicals except amino acids were obtained from Sigma-Aldrich (Milwaukee, WI) unless noted. Side chain Fmoc protected amino acids were obtained from Bachem (CA), all other amino acids were obtained from NovaBiochem (CA).

^1H NMR and ^{13}C spectra were obtained on a Varian Innova spectrometer at 300 MHz (^1H) or 75 MHz (^{13}C) and ambient temperature. Data are reported as follows: chemical shift in parts per million (δ, ppm) from an internal standard (residual $CHCl_3$ taken as 7.26 ppm for ^1H, $CDCl_3$ taken as 77.0 ppm for ^{13}C), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, and m = multiplet), coupling constant (Hz), and integration. Mass spectra were obtained on a Fisons VG Autospec.

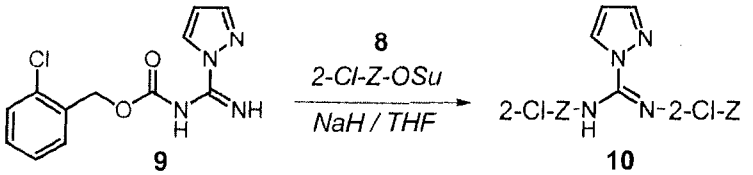
Peptide Synthesis: Amino acids were obtained from NovaBiochem (San Diego) or Bachem (Torrance, CA), including Boc-Dap(Fmoc)-OH and Boc-Dab(Fmoc)-OH (both from Bachem). Peptides were prepared according to the *in situ* neutralization protocol for solid-phase peptide synthesis developed by Kent.[14] Each peptide was purified by reverse-phase HPLC (C-18 column, solvent A: 1% CH_3CN in H_2O, 0.1% (v/v) CF_3CO_2H; solvent B: 10% H_2O in CH_3CN, 0.07% (v/v) CF_3CO_2H), and the identity of purified samples was confirmed by electrospray mass spectrometry (Finnegan LCQ-

Duo). All peptides are C-terminally amidated and N-terminally acetylated; each contains an acetamidobenzoate group on the sidechain nitrogen of a solvent-exposed lysine as a spectroscopic label.

*Reagent synthesis **10**:*



To a stirred solution of 1H-pyrazole-1-carboxamidine hydrochloride (4.79g, 32.67 mmol) and DIEA (11 mL, 65.35 mmol) in 30 mL of dry CH_2Cl_2 was added a solution of 2-Cl-Z-Su, **8**, (9.3g, 32.67 mmol) in 10 mL of CH_2Cl_2. After stirring at room temperature for three and half hours, 30 mL of CH_2Cl_2 was added and the mixture was washed with $NaHCO_3$ (sat.), water, and brine and dried over Na_2SO_4. After solvent removal under reduced pressure, 8.59g of pure product, **9**, was obtained (94 %). **^{1}H NMR** (300 MHz, $CDCl_3$): δ 9.80 (s, 1H), 8.46 (d, J = 2.6 Hz, 1H), 7.70 (d, J = 1.5 Hz, 1H), 7.48 (m, 1H), 7.36 (m, 1H), 7.27 (m, 1H), 6.43 (dd, J = 1.5 Hz, J = 1.1 Hz, 1H), 5.33 (s, 2H). **^{13}C NMR** (75 MHz, $CDCl_3$): δ 163.9, 155.8, 144.0, 134.3, 133.5, 129.7, 129.6, 129.5, 129.2, 127.1, 109.6, 64.9. **HRMS**: m/z calculated for $C_{12}H_{11}ClN_4O_2$ = 278.0571; found = 279.0649 [M+H]

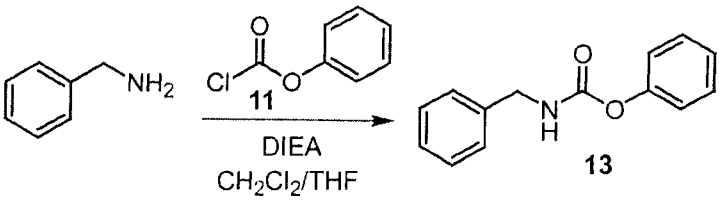


To a suspension of NaH (1.7g, 42.70 mmol), in 10 mL of THF was added drop-wise a solution of the mono 2-Cl-Z derivative, **9**, (3.4 g, 12.20 mmol) in 5 mL of THF at 0°C. After the resulting mixture was stirred for half an hour at room temperature, a solution of 2-Cl-Z-Su, **8**, in 20 mL of THF was added. After further stirring for 5 hours and 15 minutes at 0°C, water and CH_2Cl_2 were added and the organic layer was washed with $NaHCO_3$ (sat.), water, and brine, then dried over Na_2SO_4. After solvent removal under reduced pressure, the residue was chromatographed (5:1 hexane:ethyl acetate) to yield 4.38 g of pure product,**10**, (80 %). **^{1}H NMR** (300 MHz, $CDCl_3$): δ 9.36 (s, 1H), 8.32 (d, *J* = 2.6 Hz, 1H), 7.64 (d, *J* = 1.5 Hz, 1H), 7.53 (m, 1H), 7.39 (m, 3H), 7.27 (m, 4H), 6.46 (dd, J = 2.6 Hz, J = 1.8 Hz, 1H), 5.36 (s, 2H), 5.32 (s, 2H). **^{13}C NMR** (75 MHz, $CDCl_3$): δ 158.2, 150.9, 143.5, 138.8, 133.9, 133.1, 132.9, 132.7, 130.4, 130.3, 130.1, 129.7, 129.1, 128.5, 127.5, 127.2, 126.8, 110.7, 65.9, 65.6. **HRMS**: m/z calculated for $C_{20}H_{16}N_4O_4Cl_2$ = 446.0549; found = 447.0627 [M+H]

Solution Phase Reactions:

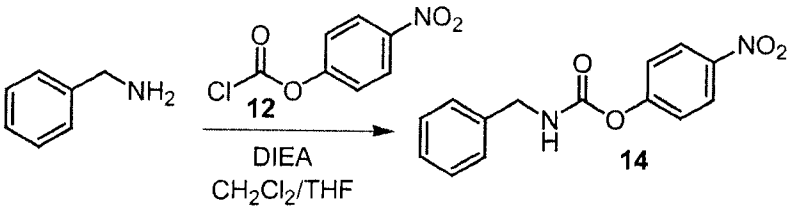
A mixture of di-2-Cl-Z guanidinylating reagent, **10**, (0.06mmol) and amine (0.055mmol) in 0.5mL of THF was stirred at room temperature for 5 minutes. Water and ethyl acetate were added, and the organic layer was washed with 0.5M citric acid, water and brine and dried over Na_2SO_4. After removal of solvent *in vacuo*, the residue was chromatographed in 3:1 ethyl acetate: hexane to afford pure product.

*Synthesis of **13***



To a stirred solution of benzylamine (200 μL, 1.83 mmol) and DIEA (638 μL, 3.66 mmol) in 4 mL of 1:1 CH_2Cl_2/THF was added a solution of phenylchloroformate, **11**, (230 μL, 1.83 mmol) in 2 mL of 1:1 CH_2Cl_2/THF. After stirring at room temperature for 24 hours, 5 mL of CH_2Cl_2 was added and the mixture was washed with $NaHCO_3$ (sat), water, and brine, then dried over Na_2SO_4. After solvent removal under reduced pressure, the residue was chromatographed (CH_2Cl_2) to afford 414 mg of the pure product, **13**, (99%). **^{1}H NMR** (300 MHz, $CDCl_3$): δ 7.39 (m, 7H), 7.23 (m, 3H), 5.36 (s, 1H), 4.46 (d, J = 5.9 Hz, 2H) **^{13}C NMR** (100 MHz, $CDCl_3$): δ 154.96, 151.24, 138.28, 129.51, 128.97, 127.93, 127.89, 125.56, 121.81, 45.481 **IR** (cm^{-1}) 3287, 1702, 1537,1484, 1454, 1250, 1206 **HRMS** m/z calculated for $C_{14}H_{13}NO_2$ [M+H] = 228.1025; found 228.1026

*Synthesis of **14**:*



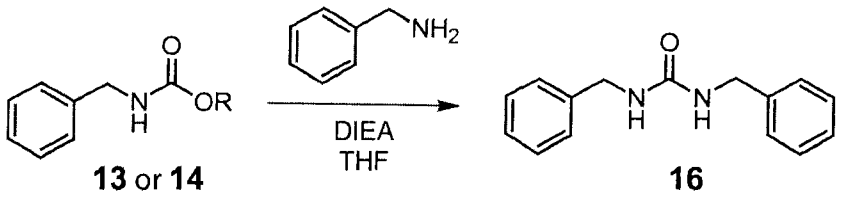
To a stirred solution of benzylamine (500 μL, 4.58 mmol) and DIEA (1.6 mL, 9.16 mmol) in 4 mL of 1:1 CH_2Cl_2/THF was added a solution of p-nitrophenylchloroformate, **12**, (9.2 g, 45.8 mmol) in 2 mL of 1:1 CH_2Cl_2/THF. After stirring at room temperature for 24 hours, 5 mL of CH_2Cl_2 was added and the mixture was washed with saturated $NaHCO_3$, water, and brine then dried over Na_2SO_4. After solvent removal under reduced pressure, the residue was chromatographed (CH_2Cl_2) to afford 895 mg of the pure product, **14**, (74%). **^{1}H NMR** (300 MHz, $CDCl_3$): δ 8.26 (d, J = 9.1 Hz, 2H), 7.39 (m, 7H), 5.45 (s, 1H), 4.48 (d, J = 6.2 Hz, 2H) **^{13}C NMR** (100 MHz,

$CDCl_3$): δ 156.07, 153.34, 144.96, 137.56, 129.10, 128,17, 127.93, 125.33, 122.16, 45.61 **IR** (cm^{-1}) 3288, 1708, 1521 **HRMS** m/z calculated for $C_{14}H_{12}N_2O_4$ [M+H] = 273.0875; found 273.0884

General Procedure for Solution Phase Urea Formation:

To a stirred solution of amine (10 eq.) and DIEA (10 eq.) in 2 mL of THF was added a solution of carbamate (1 eq.) in 3 mL of THF. After stirring at room temperature for 72 hours, 5 mL of CH_2Cl_2 was added and the organic layer was washed with saturated $NaHCO_3$, water and brine then dried over Na_2SO_4. The residue was chromatographed (CH_2Cl_2) to afford pure product.

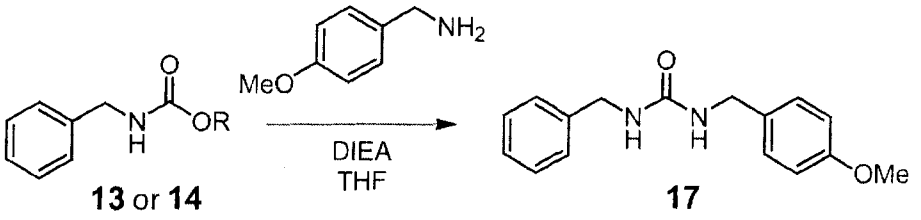
*Specific results for synthesis of **16**:*



13 or **14** **16**

Using 472 μL (4.3 mmol) of benzyl amine and 100 mg (0.44 mmol) phenyl carbamate **13**, 95 mg (90%) pure **16** was obtained. Using 165 μL (1.5 mmol) benzyl amine and 42 mg (0.17 mmol) *p*-nitrophenyl carbamate **14**, 34 mg (92%) of pure **16** was obtained. **^{1}H NMR** (300 MHz, $CDCl_3$): δ 7.32 (m, 10H), 4.83 (s, 2H), 4.32 (s, 4H) **^{13}C NMR** (100 MHz, $CDCl_3$): δ 158.199, 139.22, 128.84, 127.62, 127.53, 44.77 **IR** (cm^{-1}) 3320, 1625 **HRMS** m/z calculated for $C_{15}H_{16}N_2O$ [M+H] = 241.1341; found 241.1349

*Specific results for synthesis of **17**:*



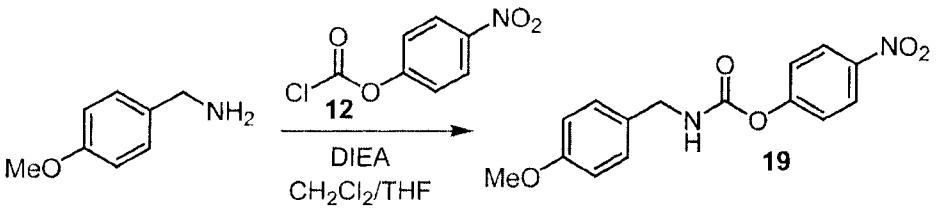
13 or **14** **17**

Using 575 μL (4.4 mmol) *p*-methoxybenzyl amine and 100 mg (0.44 mmol) phenyl carbamate **13**, 112 mg (94%) pure **17** was obtained. Using 238 μL (1.8 mmol) benzyl amine and 50 mg (0.18 mmol) *p*-nitrophenyl carbamate **14**, 47 mg (92%) of pure **17** was obtained. **^{1}H NMR** (300 MHz, $CDCl_3$): δ 7.32 (m, 5H), 7.18 (d, J = 8.6 Hz, 2H), 6.84 (d, J = 8.6 Hz, 2H), 4.55 (s, 2H), 4.34 (s, 2H), 4.27 (s, 2H), 3.78(s, 3H) **^{13}C NMR** (100 MHz, $CDCl_3$): δ 158.98, 158.40, 139.39, 131.37, 128.93, 128.75, 127.57, 127.40, 114.13, 55.43, 44.59, 44.13 **IR** (cm^{-1}) 3287, 1702, 1537,1484, 1454, 1250, 1206 **HRMS** m/z calculated for $C_{16}H_{18}N_2O_2$ [M+H] = 271.1446; found 271.1443

General procedure for solid phase urea formation first generation:

General procedure: After removal of the Fmoc protecting group with 20% piperdine in DMF, the resin was rinsed twice with DMF, twice with THF, and twice with 1:1 CH_2Cl_2:THF. After rinsing, 10 eq. of *p*-nitrophenylchloroformate, **11**, or phenylchloroformate, **12**, was added to the resin with 10 eq. of DIEA in 1:1 CH_2Cl_2:THF. After reacting for 30 minutes, the resin was rinsed with CH_2Cl_2:THF, then DMF. 10 eq. of *p*-methoxybenzyl amine was then added with 10 eq. of DIEA in DMF and was allowed to react for 15 to 30 minutes.

*Second Generation Urea Formation Reagent Synthesis **19**:*



To a stirred solution of *p*-methoxybenzyl amine (500 μl, 5.49 mmol) and DIEA (1.9 mL, 10.99 mmol) in 2 mL of dry 1:1 THF:CH_2Cl_2 was added a solution of *p*-nitrophenylchloroformate (1.2g, 6.04 mmol) in 4 mL of THF:CH_2Cl_2 at 0°C. An

additional 10 ml of THF:CH_2Cl_2 was then added and the reaction was brought to room temperature. After stirring for 42 hours, 10 mL of CH_2Cl_2 was added and the mixture was washed with saturated $NaHCO_3$, water, and brine and dried over Na_2SO_4. After solvent removal under reduced pressure, the residue was chromatographed (3:1 hexane:ethyl acetate) to afford 816 mg of pure product was obtained (48%, 57.8%BSM). **^{1}H NMR** (400 MHz, $CDCl_3$): δ 8.25 (d, *J* =2.1 Hz, 2H), 7.33 (d, *J* = 2.1 Hz, 2H), 7.28 (d, *J* = 8.7 Hz, 2H), 6.91 (d, *J* = 8.5, 2H), 5.42 (s, 1H), 4.40 (s, *J* = 6.0, 2H), 3.81 (s, 3H). **^{13}C NMR** (100 MHz, $CDCl_3$): δ 159.57, 156.14, 153.28, 144.96, 129.42, 128.95, 125.344, 122.61, 122.19, 114.45, 55.55, 45.135. **HRMS**: m/z calculated for $C_{15}H_{14}N_2O_5$ = 302.0903; found = 302.0902 [M+H]

General procedure for solid phase urea formation second generation:

General procedure: After removal of the Fmoc protecting group with 20% piperdine in DMF, the resin was rinsed twice with DMF, then twice with THF, then twice with 1:1 CH_2Cl_2:THF. After rinsing, 4 eq. of *p*-methoxybenzyl-*p*-nitrophenyl carbamate, **19,** was added to the resin in 1:1 CH_2Cl_2:THF. After reacting for 1 hour, the resin was rinsed with CH_2Cl_2:THF, then DMF, and SPPS was resumed.

2-6: Literature Cited

[1] (a) Woolfson, D. N., *Adv. Protein Chem.* **2005**, 70, 79-112. (b) Parry, D. A. D.; Squire, J. M., *Adv. Protein Chem.* **2005**, 70, 1-10. (c) Lupas, A. N.; Gruber, M., Adv. *Protein Chem.* **2005**, 70, 37-78. (d) Cheng, R. P., *Curr. Opin. Struct. Biol.* **2004**, 14, 512-520. (e) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A., *Annu. Rev. Biochem.* **1999**, 68, 779-819.

[2] (a) Harbury, P.B.; Zhang, T.; Kim, P. S.; Alber, T., *Science* **1993**, 262, 1401-1407. (b) Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T., *Nat. Struct. Biol.* **1996**, 3, 1011-1018. (c) Campbell, K. M.; Lumb, K. J., *Biochemistry* **2002**, 41, 7169-7175. (d) Nautiyal, S.; Woolfson, D. N.; King, D. S.; Alber, T., *Biochemistry* **1995**, 34, 11645-11651.

[3] (a) Schneider, J. P.; Lear, J. D.; DeGrado, W. F., *J. Am. Chem. Soc.* **1997**, 119, 5742-5743. (b) Kretisinger, J. K.; Schneider, J. P., *J. Am. Chem. Soc.* **2003**, 125, 7907-7913.

[4] (a) Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2001**, 123, 11081-11082. (b) Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2002**, 124, 9779-9783. (c) Ryan, S. J.; Kennan, A. J., *J. Am. Chem. Soc.* **2007**, 129, 10255-10260.

[5] (a) Moroni, M.; Koksch, B.; Osipov, S. N.; Crucianelli, M.; Frigerio, M.; Bravo, P.; Burger, K., *J. Org. Chem.* **2001**, 66, 130-133. (b) Brewer, M.; Rich, D. H., *Org. Lett.* **2001**, 3, 945-948. (c) Katrizky, A.R.; Rogovoy, B.V.; Chassaing, C.; Vvedensky, V., J. *Org. Chem.* **2000**, 65, 8080-8082.

[6] (a) Mamai, A.; Madalengoitia, J. S., *Org. Lett.* **2001**, 3, 561-564. (b) Ghosh, A.K.; Hol, W. G. J.; Fan, E., *J. Org. Chem.* **2001**, 66, 2161-2164. (c) Aapf, C. W.; Creighton, C.J.; Tomioka, M.; Goodman, M., *Org. Lett.* **2001**, 3, 1133-1136. (d) Dahmen, S.; Braese, S., *Org. Lett.* **2000**, 2, 3563-3565.

[7] Zhang, Y.; Kennan, A.J. *Org Lett* **2001**, 3, 2341-2344.

[8] Benatowicz, M.S.; Wu, Y.; Matsueda, G.R. *Tetrahedron Lett* **1993**, 34, 3389-3392.

[9] (a) Norwick, J. S., Org Biomol Chem 2006, 4, 3869-3885. (b) Rebek, J., Jr. Angew Chem Int Ed 2005, 44, 2068-2078. (c) Takemoto, Y. Org Biomol Chem 2005, 3, 4299-4306.

[10] (a) Guichard, G.; Semetey, V.; Didierjean, C.; Aubry, A.; Briand, J.; Rodriguez, M. *J Org Chem* **1999**, 40, 2749-2752. (b) Limal,D.; Semetey, V.; Dalbon,P.; Jolievet, M.; Briand, J. *Tetrahedron Lett* **1999**, 40, 2749-2752. (c) Chong, P.Y.; Petillo, P.A. Tetrahedron Lett 1999, 40, 4501-4504. (d) Burgess, K.; Ibarzo, J.; Linthicum, D.S.; Russell, D.H.; Shin, H.; Shitangkoon, A.; Totani, R.; Zhang, A.J. *J Am Chem Soc* **1997**, 119, 1556-1564. (e) Nowick, J.S.; Powell, N.A.; Nguyen, T.M.; Noronha, G. *J Org Chem* **1992**, 57, 7364-7366.

[11] (a) Gallou, I.; Ericksson, M.; Zeng, X.; Senanayake, C.; Faina, V. *J Org Chem* **2005**, 70, 6960-6963. (b) Lee, S.; Matsushita, H.; Koch, G.; Zimmermann, J.; Clapham, B.; Janda, K.D. *J Comb Chem* **2004**, 6, 822-827. (c) Wiszniewska, A.; Kunce, D.; Chung, N.N.; Schiller, P.W.; Izdebski, J. *Lett Pept Sci* **2003**, 10, 33-39. (d) Boeijen, A.; Liskamp, R.M. *Eur J Org Chem* **1999**, 2127-2135. (e) Xiao, X.; Ngu, K.; Chao, C.; Patel, D.V. *J Org Chem* **1997**, 62, 6968-6973. (f) Kim, J.; Yingzhi, B.; Paikoff, S.J.; Schultz, P.A. *Tetrahedron Lett* **1996**, 37, 5305-5308. (g) Hutchins, S.M.; Chapman, K.T. *Tetrahedron Lett* **1995**, 36, 2583-2586.

[12] Diss, M. L.; Kennan, A. J. *Biopolymers* **2007**, 86, 276-281.

[13] O'Shea, E. K.; Lumb, K.; Kim, P. S. *Curr Biol* **1993**, 3, 658-667.

[14] Schnöelzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S.B.H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

Chapter 3

Investigation of Guanidine Interactions

3-1: Introduction

As discussed in the previous chapters, our group and others have been interested in the design and investigation of self-assembling peptide systems derived from α-helical coiled coils.[1] Efforts to control and manipulate coiled coil formation have targeted oligomerization state, relative strand orientation, and specificity for homo- or heteromeric structures. Our group has developed steric matching of hydrophobic core side chains as a route to specific trimer formation, while other groups have used polar side chains in the core.[2] Expanding on these previous studies we sought to establish similar levels of control in dimeric systems. New buried polar residues, such as arginine and its derivatives have been investigated.[3]

3-2: Guanidine-Guanidine Interactions

With the successful synthesis of guanidine containing peptides, we turned our attention to investigating guanidine—guanidine interactions in the core. To begin this study, heterodimers bearing identical guanidine side chains in place of the parent asparagines were investigated (Figure 3.1). Anticipating that the burial of like charges might substantially destabilize potential complexes, circular dichroism (CD) was employed to assay this. Surprisingly, spectra of 1:1 pGd^*_K/pGd^*_E and $pArg_K/pArg_E$ mixtures exhibited dramatic helicity increases compared to the average component signals (Figure 3.1C/E). Thermal unfolding experiments revealed cooperative transitions only for the mixtures that also have reasonably high stability (Figure 3.1D/F, T_m = 59 °C in each case), although these were somewhat diminished from the parent $pAsn_K/pAsn_E$ complex (T_m = 77°C). These observations indicate a specific and significant interaction between the component peptides, consistent with heterodimer

formation. Similarly, the pGd_K/pGd_E mixture exhibits a slightly more subtle effect that is less stable than the other two heterodimer mixture (T_m = 49 °C). This indicates an overall trend of increased heterodimer stability with increases in side chain length (Figure 3.1A/B).

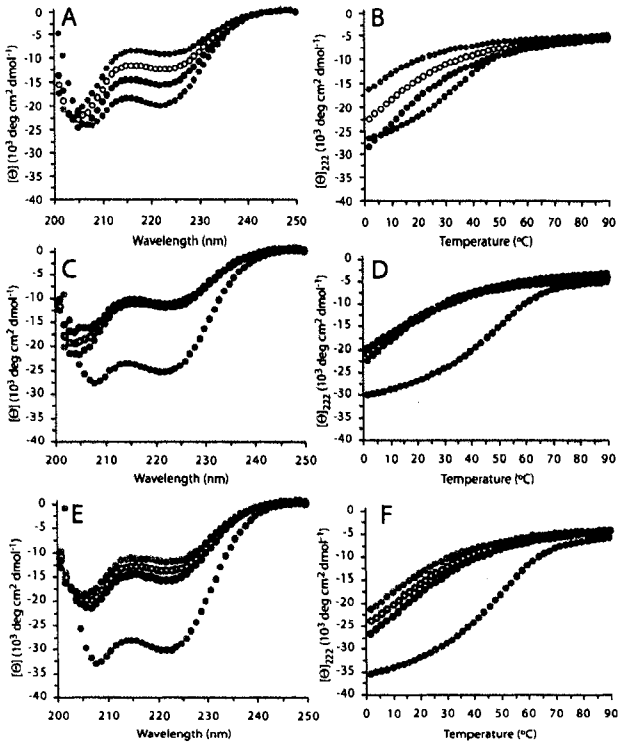


Figure 3.1. Buried guanidine-guanidine interactions. Wavelength (**A**, **C**, **E**) and thermal denaturation (**B**, **D**, **F**) CD spectra for pure solutions of $pXaa_K$ (red) and $pXaa_E$ (blue), where Xaa = Gd (**A**, **B**), Gd* (**C**, **D**), and Arg (**E**, **F**). In each case, traces for an equimolar mixture (green) and the calculated weighted average signal (open circles) are also given.

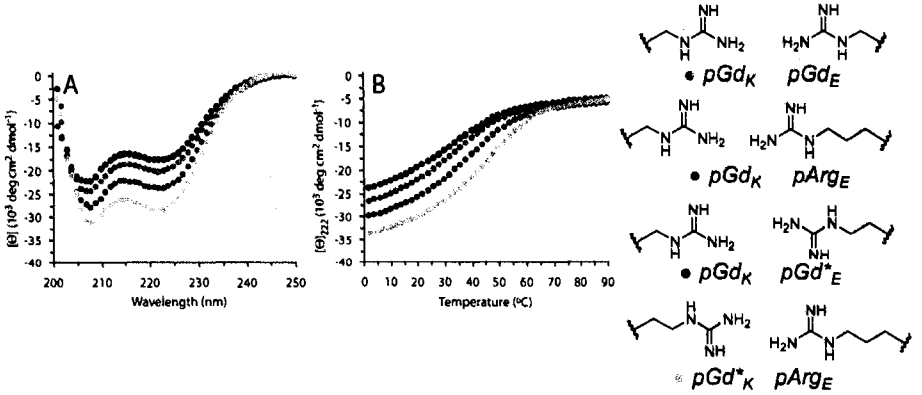


Figure 3.2. Mixed guanidine—guanidine interactions. Wavelength (**A**) and thermal denaturation (**B**) CD spectra for the indicated equimolar mixtures.

The reasonable stability in complexes with identical side chain length guanidine functionality lead us to examine complexes with mixed guanidine side chain length. Spectra of equimolar pGd_K/pGd^*_E, $pGd_K/pArg_E$, $pGd^*_K/pArg_E$ mixtures also exhibit reasonable room-temperature helicities and cooperative thermal unfolding profiles (Figure 3.2, Table 3.1). The same stability trend, favoring complexes with more total methylenes in the buried polar side chains is also observed within the mixed side chain length complexes.

Table 3.1. Observed T_m Values for Guanidine Core Pairs

Sample	T_m°C	Sample	T_m°C
pGd_K/pGd_E	49	$pGd_K/pArg_E$	47
pGd^*_K/pGd^*_E	59	$pGd^*_K/pArg_E$	55
$pArg_K/pArg_E$	59	$pArg_K/pGd_E$	47
pGd_K/pGd^*_E	43	$pArg_K/pGd^*_E$	51
pGd^*_K/pGd_E	49		

Examining the mixed-core complexes gives rise to another possible complication: the exchange of the polar core residues on the acidic and basic strands produces a distinct complex (e.g., pGd_K/pGd^*_E is not the same as pGd^*_K/pGd_E). Although the expected difference between these two distinct complexes was minimal, we nonetheless examined the reverse mixed-core mixtures pGd^*_K/pGd_E, $pArg_K/pGd_E$, $pArg_K/pGd^*_E$ mixtures (Figure 3.3, Table 3.1). Surprisingly, a modest compression in stability range was observed. The two mixed pGd_E complexes are very similar to each other and to the original pGd_K/pGd_E mixture. Even the $pArg_K/pGd^*_E$ signal is closer to the others than its counterpart.

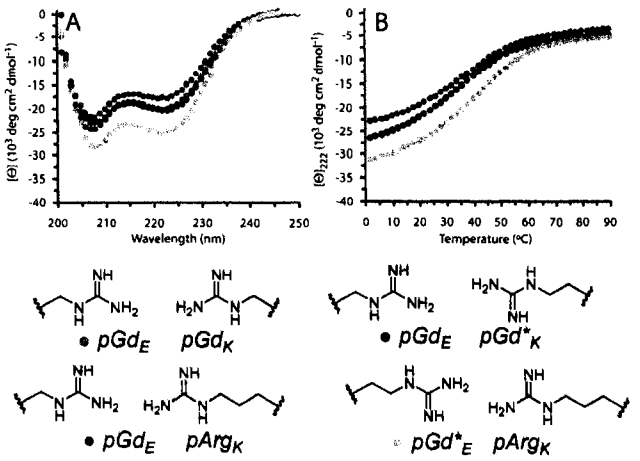


Figure 3.3. Core-swapped mixed guanidine—guanidine interactions. Wavelength (**A**) and thermal denaturation (**B**) CD spectra for the indicated equimolar mixtures.

3-3: Guanidine-Amide Interactions

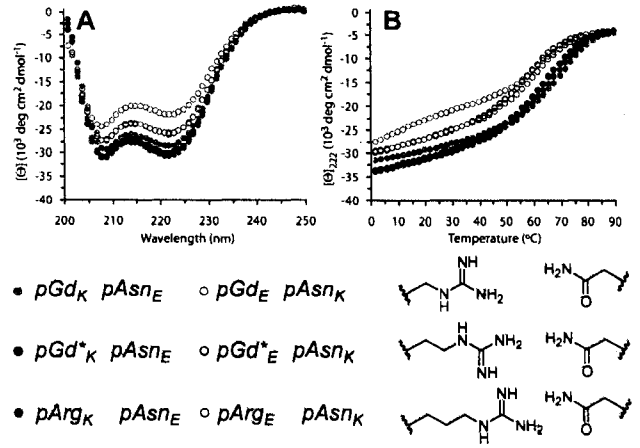


Figure 3.4. Guanidine/Asn core interactions. Wavelength (**A**) and thermal denaturation (**B**) CD spectra for the indicated equimolar mixtures. Closed and open circles of the same color indicate data from complexes that differ by exchange of the indicated buried polar side chains (from the acidic or basic peptide to the other).

Since the electrostatically repulsive guanidine/guanidine pairings gave surprisingly stable complexes, we next examined the interaction of each guanidinylated side chain with the parent neutral Asn residue. A further increase in overall complex stability was expected. Inspection of the wavelength and thermal unfolding CD spectra supports this expectation (Figure 3.4, Table 3.2). Equimolar mixtures of pGd_K, pGd^*_K,

$pArg_K$ with $pAsn_E$ exhibited cooperative unfolding transitions and substantial melting temperatures (T_m = 71-73 °C). Reverse complexes formed by mixing pGd_E, pGd^*_E, $pArg_E$ with $pAsn_K$ were also investigated to determine the impact of local strand interactions (Figure 3.4, Table 3.2). Similarly, as with the mixed guanidine complexes, a small but definite distinction in stability exists, favoring those with guanidinylated side chains on the basic peptide (e.g., T_m = 71 °C for $pArg_K$/ $pAsn_E$, versus 65 °C for $pAsn_K$/ $pArg_E$). The difference is particularly pronounced in the room-temperature helicity of the two Gd complexes, although the gap narrows at the melting transitions. Overall, a reduced stability variance is observed within each side chain length set. The length of the guanidine side chain seems largely irrelevant when paired opposite Asn, as opposed to another guanidine.

Table 3.2. Observed T_m Values for Guanidine Amide Core Pairs

sample	T_m°C	sample	T_m°C
$pGd_K/pAsn_E$	73	$pAsn_K/pGd_E$	63
$pGd^*_K/pAsn_E$	71	$pAsn_K/pGd^*_E$	63
$pArg_K/pAsn_E$	71	$pAsn_K/pArg_E$	65

3-4: Guanidine-Acid Interactions

Table 3.3. Observed T_m Values for Guanidine Acid Core Pairs

sample	T_m°C	sample	T_m°C
$pGd_K/pAsp_E$	61	$pGd_K/pGlu_E$	59
$pGd^*_K/pAsp_E$	67	$pGd^*_K/pGlu_E$	63
$pArg_K/pAsp_E$	65	$pArg_K/pGlu_E$	63

Since the complexes with the guanidine side chain present on the basic peptide (i.e., pGd_K, pGd^*_K, or $pArg_K$) exhibited greater stability then their counter complexes, (i.e., those with pGd_E, pGd^*_E, or $pArg_E$) only the acidic forms of the acid core peptides were synthesized (see Chapter 2 for synthesis details). Interactions between the

residues containing guanidine side chains with Asp or Glu residues were of particular interest since the overall neutral complex may benefit from core electrostatic attraction.

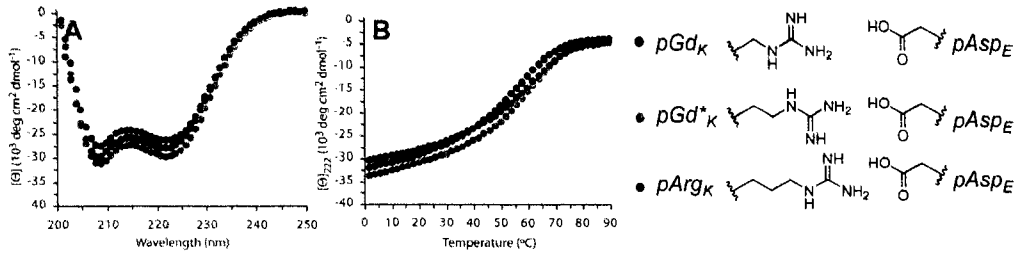


Figure 3.5. Guanidine/Asp core interactions. Wavelength (**A**) and thermal denaturation (**B**) CD spectra for the indicated equimolar mixtures.

As with the other core interactions previously discussed, equimolar mixtures of pGd_K, pGd^*_K, and $pArg_K$ with either $pAsp_E$ or $pGlu_E$ exhibit good helicities and thermal stabilities (Figure 3.5-3.6, Table 3.3). The range of melting temperatures (59 to 67 °C) is comparable to those of the guanidine/amide pairs, albeit slightly lower. Similarly, the guanidine/acid pairs display an insensitivity to polar residue chain length, with only a small change between the $pGd_K/pAsp_E$ and $pArg_K/pAsp_E$ complexes (Figure 3.5). This is also the case with the $pGd_K/pGlu_E$ and $pArg_K/pGlu_E$ complexes (Figure 3.6).

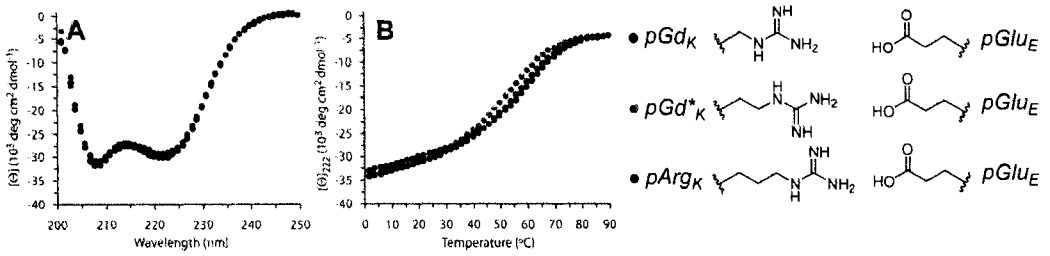


Figure 3.6. Guanidine/Glu core interactions. Wavelength (**A**) and thermal denaturation (**B**) CD spectra for the indicated equimolar mixtures.

3-5: Summary of Guanidine Interactions

Wavelength scans and thermal denaturations of all possible dimers were investigated (see Appendix 3 for details). Additionally, analytical ultracentrifugation

experiments were carried out on the $pGdn_K/pAsp_E$, $pGdn_K/pGlu_E$, and the same side chain length guanidine peptides $pGdn_K/pGdn_E$. Each of these experiments indicated dimer formation (Table 3.4).

Table 3.4. Molecular Weights from Anlytical Ultracentrifugation

Sample	MW_{calc}	MW_{obs}	% difference
$pGd_K/pAsp_E$	7200	6950	3.47
$pGd^*_K/pAsp_E$	7214	7082	1.83
$pArg_K/pAsp_E$	7228	7237	-0.125
$pGd^*_K/pGlu_E$	7229	6350	12.5
$pArg_K/pGlu_E$	7243	7128	1.59
pGd_K/pGd_E	7214	6256	13.3
$pArg_K/pArg_E$	7270	7060	2.89

A comparison of CD spectra for each guanidine side chain paired with each of its possible dimer partners reveals some interesting features (Figure 3.7). In each case, solutions with $pAsn_E$ as the partner core residue exhibit the most thermally stable profiles, followed by the $pAsp_E$ and $pGlu_E$ complexes, and finally guanidine/guanidine mixtures. The degree of separation between these classes varies quite significantly with the length of the side chain. For example, for interactions involving pGd_K (Figure 3.7A/B), there is a sharp distinction between the most stable pairing with $pAsn_E$ and the least stable pairing with pGd^*_E. With the longer side chains there is a more gradual distinction between the stabilities of the dimers. The trace for the $pArg_K$ with $pArg_E$ mixture differs slightly from those of the $pAsp_E$ and $pGlu_E$ mixtures (Figure3.7 E/F).

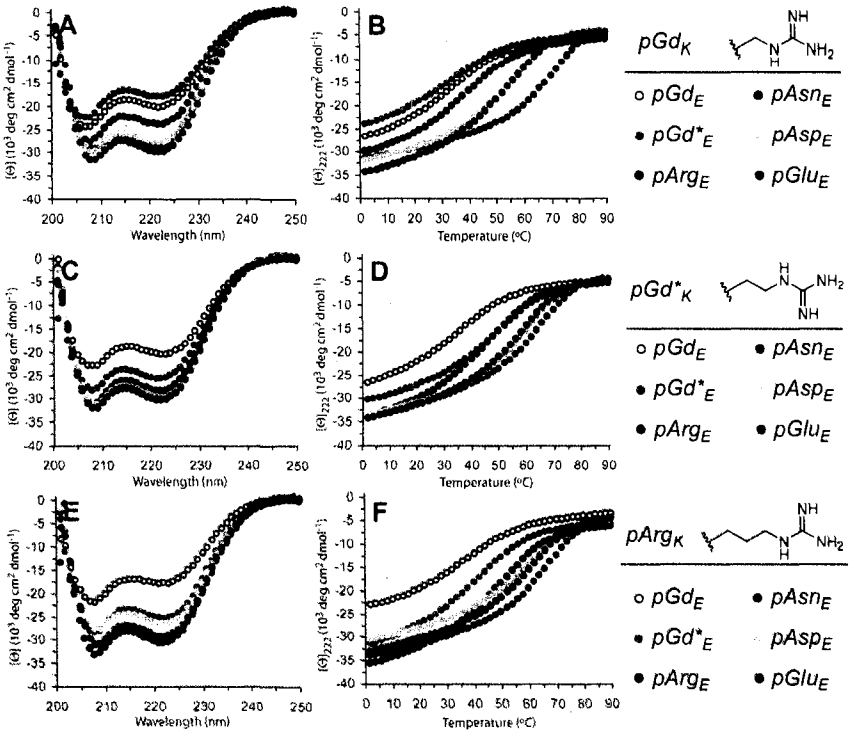


Figure3.7. Wavelength and melting curve comparisons. Wavelength (**A**, **C**, **E**) and thermal denaturation (**B**, **D**, **F**). CD spectra for pGd_K (**A**, **B**), pGd^*_K (**C**, **D**), and $pArg_K$ (**E**, **F**) with each of the indicated acidic peptides replotted for comparison.

3-6: Investigation of Self Assembly Capabilities

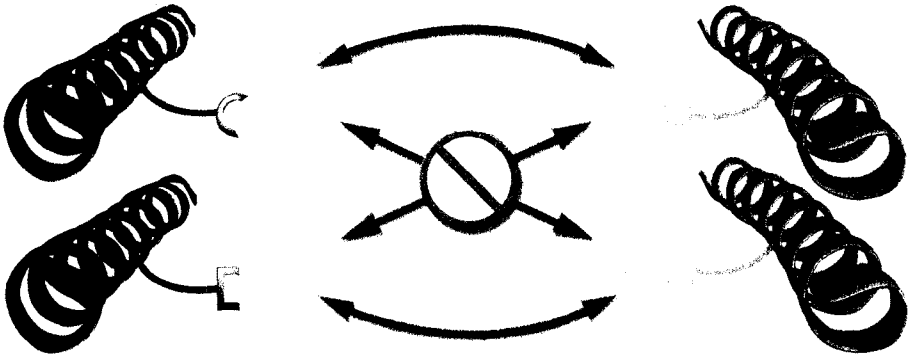


Figure 3.8. Self assembly schematic. For self assembly to occur, only two distinct heterodimers will form, when two different basic peptides (blue) are mixed with two different acidic peptides (red).

Since the new peptide interactions exhibited reasonable stability, we wanted to examine the interaction's ability to be preserved in the presence of another heterodimer (Figure 3.8). To study this we employed Ni-affinity experiments.[2] An N-terminal

GlyGly$(His)_6$ tag confers affinity for Ni-NTA agarose beads upon both the tagged peptides and any species which binds to it. After exposure to beads, removal of the supernatant, washing to eliminate low-affinity peptides, and elution with imidazole buffer to recover bound material, each fraction is analyzed by HPLC (Figure 3.9). In the present case, if one of the four peptides is tagged and orthogonal recognition occurs, the elution fraction is expected to be dominated by the tagged peptide and its partner, while the two matched untagged peptides reside in the supernatant.

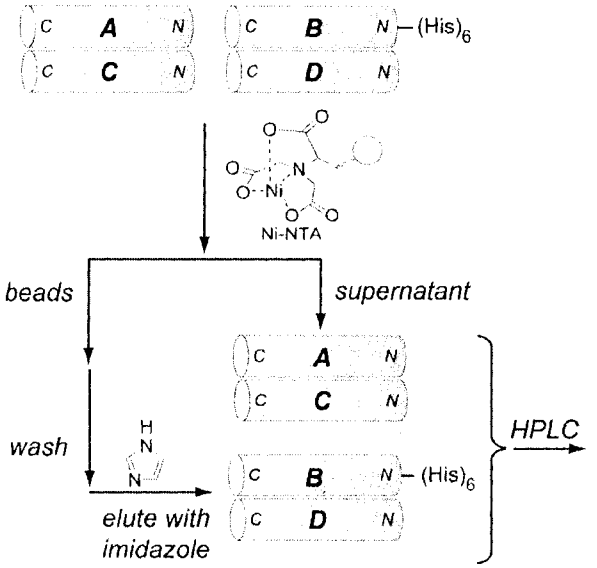


Figure 3.9. Ni-NTA affinity tag analysis scheme. Initially, peptides A, C, D and B, which bears an N-terminal Gly-Gly-$(His)_6$ affinity tag, were mixed together. Upon exposure to Ni-NTA agarose beads, B is bound through the His-tag, and D is bound through interaction with B. A and C, which do not interact with the beads or the tagged peptide, remain unbound. After supernatant removal and washing with blank buffer, bound material is eluted by treatment with imidazole. HPLC analysis reveals the identity of unbound (supernatant fraction) and bound (elution fraction) material.

The reasonable stability of the Arg/Arg core pair as well as the Gd*/Asp and Gd*/Glu pairs lead us to investigate the pair's abilities to be preserved within the presence of another pair. It was hoped that equimolar mixtures of $pArg_K/pArg_E/pGd^*_K/pAsp_E$ and $pArg_K/pArg_E/ pGd^*_K/pGlu_E$ would each afford two

separate dimers, rather than a mixture of the four energetically reasonable assemblies (discounting electrostatically incompatible binding of two acidic or basic peptides). For these two mixtures, tagged derivatives of the Asp ($pAsp_{E\ His}$) and Glu ($pGlu_{E\ His}$) peptides were synthesized using standard methods. The analysis of each mixture demonstrates essentially nonspecific recognition (Figure 3.10). Both elution fractions contain each of the basic peptides, suggesting that $pArg_K/pAsp_E$ and $pArg_K/pGlu_E$ complexes form in addition to the desired ones. The supernatants also contain both basic species, consistent with formation of the unintended $pGd^*_K/pArg_E$ complex in each case.

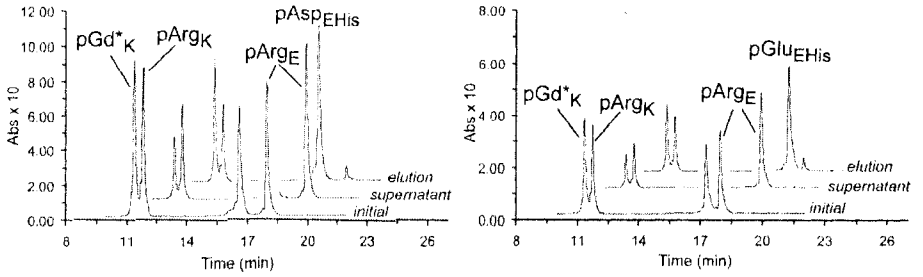


Figure 3.10. Ni-NTA analysis of $pArg_K/pArg_E$ pGd^*_K Asp/Glu selective dimer formation. Equimolar initial mixtures of $pArg_K/pArg_E/pGd^*_K/pAsp_{E\ His}$ (left) and $pArg_K/pArg_E/pGd^*_K/pGlu_{E\ His}$ (right) give rise to supernatant and elution fractions containing all reasonable complexes. Initial solutions are 20μM total peptide, pH 7.0, 150mM NaCl, 10mM phosphate.

Following this lack of successful self assembly using two new core—core matchups, we next investigated whether guanidine/guanidine peptides could recognize each other in the context of the known Asn/Asn interaction, which would require identifying only one new viable core—core interaction. Thus, equimolar solutions of $pAsn_{KHis}/pAsn_E/pGd_K/pGd_E$, $pAsn_{KHis}/pAsn_E/pGd^*_K/pGd^*_E$, and $pAsn_{KHis}/pAsn_E/pArg_K/pArg_E$, were analyzed by Ni-NTA affinity (Figure 3.11). Unfortunately, once again, essentially no specificity was observed as both supernatant and elution fractions contained significant quantities of each acidic peptide in all three experiments.

These unsuccessful selectivity experiments lead us to consider guidelines for self-assembly experiments. Selection of two reasonable complexes is not a sufficient criterion to ensure specific recognition. In particular, examination of the stability of the alternative complexes as an additional consideration to predict likely candidates is necessary. Our initial efforts seemed less promising by this measure. Using melting temperatures as a crude measure of stability, we recognized that the reasonable $pArg_K/pArg_E$ interaction could be balanced by the comparable $pGd^*_K/pArg_E$ affinity. Similarly, the very stable $pAsn_K/pAsn_E$ contact (T_m = 77 °C) is nearly balanced by the uniformly high guanidine/amide T_m values (more than balanced for the relatively weak pGd^*_K/pGd^*_E interactions).

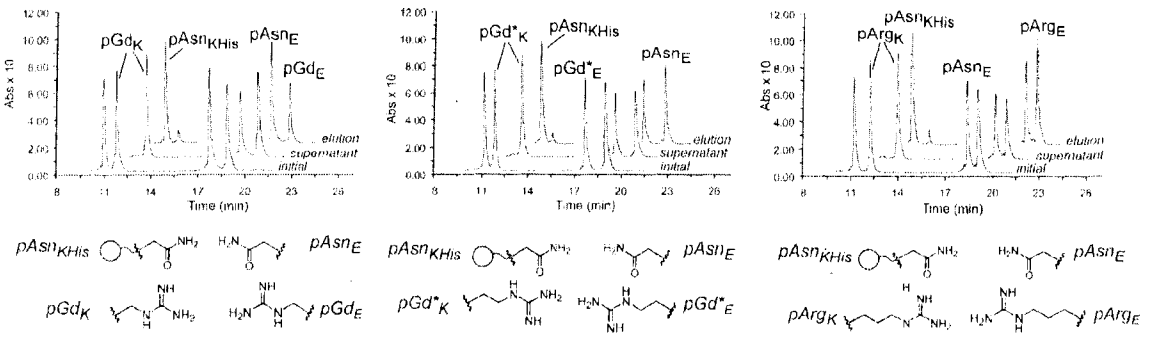


Figure 3.11. Ni-NTA analysis of four component Asn/Asn Gdn/Gdn mixtures. Equimolar ratios of the indicated peptides were subjected to the Ni-NTA protocol, revealing in all cases relatively non-specific binding. Solutions are as in Figure 3.10.

Table 3.5. Observed T_m Values for Asn/Acid Core Pairs

Sample	T_m(°C)	Sample	T_m(°C)
$pAsn_K/pAsp_E$	51	$pAsn_K/pAsp_E$	77
$pAsn_K/pGlu_E$	65		

Keeping this in mind, we next targeted all-guanidine systems. Mixtures of pGd_K with pGd_E or $pArg_E$ exhibited similar stabilities (T_m = 49 and 47 °C, respectively), while $pArg_K$ demonstrated a preference for $pArg_E$ over pGd_E (T_m = 59 vs. 47 °C). Similar arguments apply to pGd^*_K, with slightly more favorable numbers (T_m values of

59 vs. 51°C and 59 vs. 55°C). Despite this determination, analysis of equimolar $pArg_K/pArg_E/pGd_K/pGd_E$ and $pArg_K/pArg_E/pGd^*_K/pGd^*_E$ mixtures (using $pArg_{K\ His}$) again failed to support high levels of specificity (Figure 3.12). However, the slightly unequal ratios of the acidic peptides in the supernatant and elution fractions indicated some relative preferential binding.

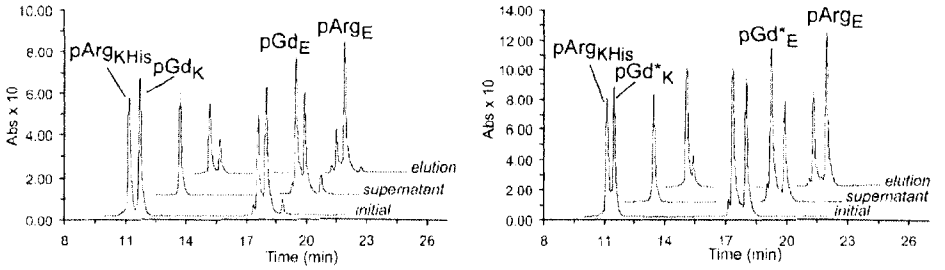


Figure 3.12. Ni-NTA analysis of four component Arg/Arg, Gd/Gd and Gd*/Gd* complexes. Equimolar initial mixtures of $pArg_{K\ His}/pArg_E/pGd_K/pGd_E$ (left) and $pArg_{K\ His}/pArg_E/pGd^*_K/pGd^*_E$ (right) give rise to supernatant and elution fractions containing all reasonable complexes. Solutions are as in Figure 3.10.

With the above reasoning, selective recognition seems most likely when at least one of the unintended complexes is destabilized. Thus, to exploit the stable Asn/Asn contact, we sought another pair in which at least one partner is mismatched with Asn. Although the guanidine/Asn contacts were largely favorable, we wondered if the same would apply to Asp or Glu matched with Asn. Accordingly, we measured wavelength and thermal unfolding CD profiles for equimolar solutions of $pAsn_K/pAsp_E$ and $pAsn_K/pGlu_E$, and compared them with the parent $pAsn_K/pAsn_E$ complex (Figure 3.13, Table 3.5).

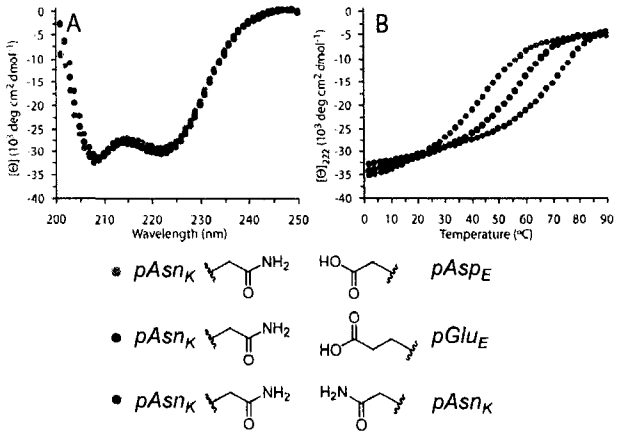


Figure 3.13. Asn/acid interactions. Wavelength (A) and thermal denaturation (B) CD spectra for equimolar mixtures of $pAsn_K/pAsn_E$ (green), $pAsn_K/pAsp_E$ (blue), and $pAsn_K/pAsp_E$ (red).

The Asp residue appears to be relatively mismatched with Asn (T_m = 51 °C), unlike with the Glu residue which has a relatively strong interaction with Asn (T_m = 65°C). Thus we predicted that guanidine/Asp peptides should be able to specifically recognize each other in the presence of the Asn/Asn contact, while the corresponding guanidine/Glu complexes would be nonspecific. To examine both of these possibilities, equimolar solutions of $pAsn_K/pAsn_E$ with $pGd_K/pAsp_E$, $pGd^*_K/pAsp_E$, $pArg_K/pAsp_E$, $pGd_K/pGlu_E$, $pGd^*_K/pGlu_E$, and $pArg_K/pGlu_E$ were investigated by Ni-NTA methods ($pAsn_{K\ His}$ in all but the latter case, which used $pGlu_{E\ His}$).

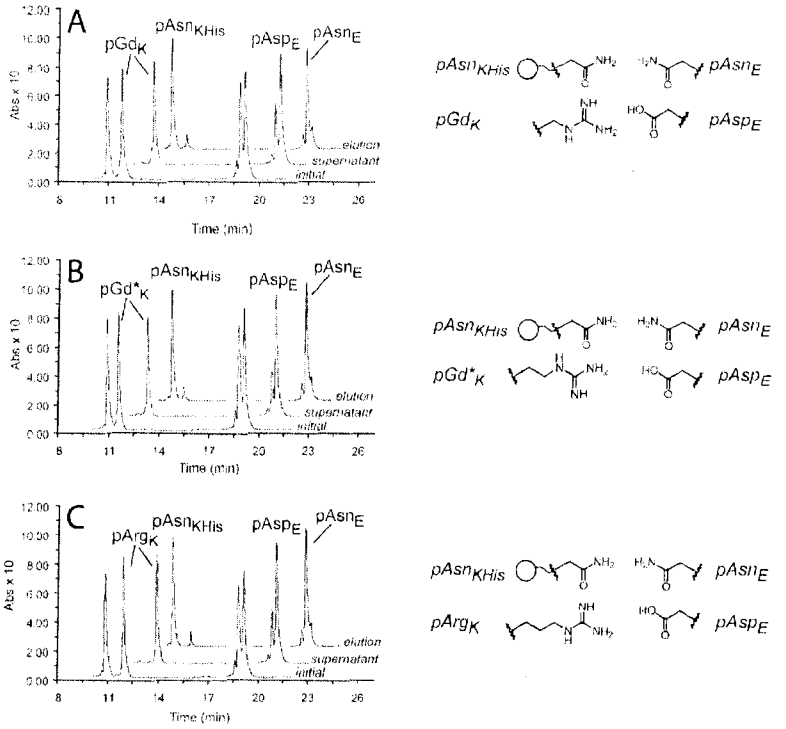


Figure 3.14. Ni-NTA analysis of four component Asn/Asn and Gdn/Asp mixtures. Equimolar ratios of (A) $pAsn_{KHis}/pAsn_E/pGd_K/pAsp_E$, (B) $pAsn_{KHis}/pAsn_E/pGd^*_K/pAsp_E$, and (C) $pAsn_{KHis}/pAsn_E/pArg_K/pAsp_E$ all demonstrate specific recognition. Elution fractions are dominated by the Asn peptides, indicating preferential binding of $pAsp_E$. Solutions are as in Figure 3.10.

Analysis of the Asp complexes reveals that indeed these pairs are capable of orthogonal recognition in the presence of an Asn/Asn interaction (Figure 3.14). In each case, the supernatant fraction largely contains the guanidine/Asp peptides, while the elution fraction is dominated by $pAsn_{K\ His}$ and $pAsn_E$. To our knowledge, this is the first example of simultaneous self-selection of independent heterodimeric coiled coils driven only by a single residue buried polar-residue substitution.

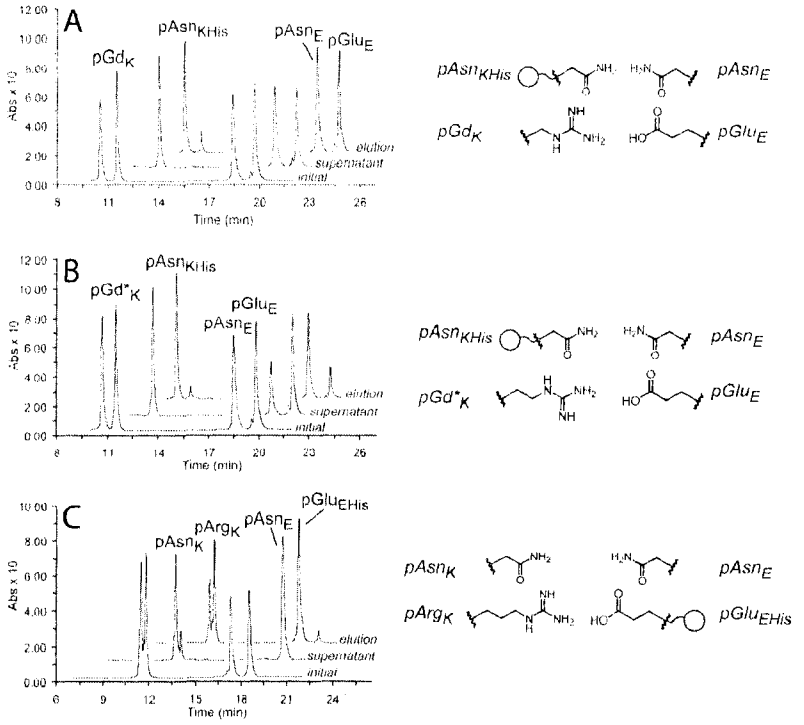


Figure 3.15. Ni-NTA analysis of four component Asn/Asn and Gdn/Glu mixtures. Equimolar ratios of (A) $pAsn_{KHis}/pAsn_E/pGd_K/pGlu_E$, (B) $pAsn_{KHis}/pAsn_E/pGd^*_K/pGlu_E$, and (C) $pAsn_K/pAsn_E/pArg_K/pGlu_{EHis}$ all demonstrate nonspecific recognition. Supernatant and elution fractions are consistant with formation of all possible dimmers in significant quantities. Solutions are as in Figure 3.10.

Further supporting the exquisite specificity of this interaction, the corresponding guanidine/Glu complexes, differing only by the addition of a single methylene unit in the core, do not display correspondingly specific binding (Figure 3.15). In each case, as with the early experiments, significant quantities of each possible complex are indicated by the presence of additional peptides in both supernatant and elution fractions.

Regardless of the overall relative tolerance of varying core residue chain length and charge states, only a select subset of these new heterodimers were capable of recognition based only on polar core residue identity. The outcomes of these experiments suggest that specific formation of two independent heterodimers from

initial four-peptide mixtures requires at least one mismatched alternative complex. The poor stability of Asn/Asp cores is capable of directing impressive recognition levels.

3-7: Experimental Information

Ni-NTA Affinity Tag Experiments: A 0.5 mL sample of a 50% slurry of Ni-NTA agarose (Qiagen) in an Eppendorf tube was centrifuged for 30 s, followed by removal of the supernatant. Peptide solution (1 mL of 20 μM or 10 μM total peptide, 5 μM each peptide) was added, and the tube was repeatedly inverted for 5 minutes. The sample was centrifuged (30 s) and the supernatant (flow-through fraction) was removed. The procedure was then repeated with 1 mL of buffer (wash fraction) and 1 mL of buffer containing 250 mM imidazole (elution fraction), except that the wash fraction was only agitated for 30 seconds.[4] Solutions were analyzed by RP-HPLC.

CD Spectroscopy: All experiments were performed on an Aviv model 202 circular dichroism spectrometer. Sample concentrations were measured by UV absorbance of the acetamidobenzoate label at 270 nm. Wavelength data are the average of three scans from 250 to 200 nm in 1 nm steps. Thermal denaturation experiments at 222 nm were run from 0° to 90°C in two-degree steps, at a two-degrees/minute rate of increase with one-minute equilibration and data averaging at each temperature. T_m values were obtained from minima of the first derivative of θ vs. 1/T plots.[5]

Analytical Ultracentrifugation: Sedimentation equilibrium experiments were performed using a Beckman XL-A analytical ultracentrifuge equipped with an An60-Ti or An50-Ti rotor. Data were collected using 12 mm path length six-sector centerpieces at 270 nm. Samples were dialyzed against the reference buffer at 4°C overnight. Data were collected at 38,000 and 48,000 r.p.m. at concentrations spanning 17-55 μM.

Samples were judged equilibrated when three consecutive scans taken 1h apart were indistinguishable (equilibration was complete in 12h in all cases). Solvent densities and partial molar volumes were calculated in the manner prescribed by Durchschlag and Zipper,[6] using one of three v_bar values for Gd and Gd* (Table 3.6). Data were analyzed using Ultrascan[7] and fit globally to ideal single species. For fits see Appendix 4.

Table 3.6. v_bar values for unnatural amino acids and peptides containing them

Residue	v_bar	Peptide	v_bar
Gd	0.702	pGd_K	0.782
Gd*	0.746	pGd_E	0.733
		pGd^*_K	0.783
		pGd^*_E	0.734

3-8: Literature Cited

[1] (a) Tsurkan, M. V.; Ogawa, M. Y., *Inorg. Chem.* **2007**, *46*, 6849-6851.(b) Steinmetz, M. O.; Jelesarov, I.; Matousek, W. M.; Honnappa, S.; Jahnke, W.; Missimer, J. H.; Frank, S.; Alexandrescu, A. T.; Kammerer, R. A., *Proc. Nat. Acad. Sci. U. S. A.* **2007**, *104*, 7062-7067. (c) Ryan, S. J.; Kennan, A. J., *J. Am. Chem. Soc.* **2007**, *129*, 10255-10260. (d) Portwich, M.; Keller, S.; Strauss, H. M.; Mahrenholz, C. C.; Kretzschmar, I.; Kramer, A.; Volkmer, R., *Angew. Chem. Int. Ed.* **2007**, *46*, 1654-1657. (e) Nikolaev, Y.; Pervushin, K., *J. Am. Chem. Soc.* **2007**, *129*, 6461-6469. (f) Yoder, N. C.; Kumar, K., *J. Am. Chem. Soc.* **2006**, *128*, 188-191. (g) Son, S.; Tanrikulu, C.; Tirrell, D. A., *ChemBioChem* **2006**, *7*, 1251-1257.

[2] (a)Schnarr, N. A.; Kennan, A. J., *Org. Lett.* **2005**, 7, 395-398. (b) Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2004**, 126, 14447-14451. (c) Campbell, K. M.; Lumb, K. J., *Biochemistry* **2002**, 41, 7169-7175. (d) McClain, D. L.; Gurnon, D. G.;

Oakley, M. G., *J. Mol. Biol.* **2002**, 324, 257-270. (e) Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S., *Protein Science* **1999**, 8, 2312-2329. (f) Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T., *Nat. Struct. Biol.* **1996**, 3, 1011-1018.

[3] Diss, M. L.; Kennan, A. J., *J. Am. Chem. Soc.* **2008**, 130, 1321-1327.

[4] Method patterned after that in: Brown, M.B.; Saucer, R.T. Pr*oc. Nat. Acad. Sci. U.S.A.* **1999**, 96, 1983-1988.

[5] Cantor, C.R.; Schimmel, P.R. Biophyscial *Chemistry of Macromolecules, Pt.3: The Behavior of Biological Macromolecules*; W.H. Freeman: New York, NY, 1980, p. 1132.

[6] Durchshlag, H.; Zipper, P. *Progr Collaoid Polym Sci* **1994**, 94, 20-39.

[7] Ultrascan (version 7.1). Developed by Borries, Demeler, The University of Texas Health Science Center at San Antonio.

Chapter 4

Investigation of Urea Interactions

4-1: Introduction

Like guanidine, urea is very abundant in nature. Ureas have long been employed as a crucial recognition element, principally via hydrogen bonding interactions.[1] Thus, the development of mono-substituted urea side chains in solid phase peptide synthesis (SPPS) was crucial for this programmed self-assembly investigation of α-helical coiled coils.[2]

Since the variable-length guanidinylated core residues supported well-defined dimeric coiled coils with a variety of binding partners and were capable of selective recognition, we sought to study the impact of variable-length urea containing core residues on the same system.[3] Neutral urea/urea binding, driven by complementary and efficient hydrogen-bond donor/acceptor interactions, has a vast array of applications in synthetic self-assembly systems. Thus, urea-terminated side chains were studied to see if they could function as directing elements in dimeric coiled coil formation.

4-2: Urea-Urea Interactions

Table 4.1. Observed T_m Values for Urea Core Pairs.

Sample	T_m (°C)	Sample	T_m (°C)
pUr_K/pUr_E	63	$pUr_K/pCit_E$	63
pUr^*_K/pUr^*_E	63	$pUr^*_K/pCit_E$	67
$pCit_K/pCit_E$	79	$pCit_K/pUr_E$	71
pUr_K/pUr^*_E	63	$pCit_K/pUr^*_E$	75
pUr^*_K/pUr_E	65		

Six peptides were synthesized containing a core urea group, spaced from the backbone of acidic or basic sequences by one ($pUr_{K/E}$), two ($pUr^*_{K/E}$) or three ($pCit_{K/E}$) methylenes (see Chapter 2 for more details). With these three urea-substituted sequences of each electrostatic type (Ur/Ur*/Cit) there are nine possible urea-containing mixtures. By design, reasonable complexes require one acidic and one basic peptide (having Glu or

Lys, respectively, in all *e/g* positions). Of these nine possible mixtures, three pairings will have identical side chain length polar core groups. Examination of the equimolar solutions of these identical side chain length urea groups by circular dichrosim (CD) spectroscopy shows that the mixtures exhibit wavelength and thermal denaturation profiles consistent with coiled coil formation (Figure 4.1). The Cit/Cit interaction exhibits a thermal stability that slightly exceeds that of the parent Asn/Asn system (T_m = 79 vs. 77 °C), while the other urea interactions exhibit similar helicities and identical melting temperatures (T_m = 63 °C, Table 4.1).

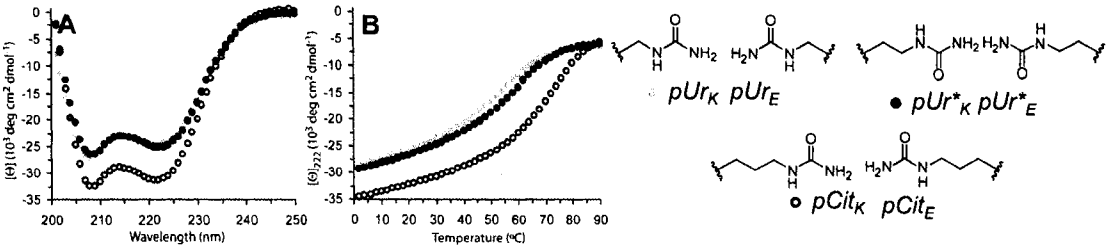


Figure 4.1. Buried Urea—Urea interactions. Wavelength (A) and thermal denaturation (B) CD spectra for the indicated equimolar mixtures.

Examination of the mixed urea side chain lengths by CD spectroscopy shows that the 1:1 mixtures exhibit room-temperature helicities and thermal stabilities well in excess of calculated component weighted average signals, which is consistent with complex formation (Figure 4.2.). In general, use of a neutral polar group contributed to heterodimer stability (T_m = 63-79 °C, Table 4.1). Any influence of polar residue side chain length on the overall stability is subtle. There is a general trend favoring longer chains that is less pronounced than within the guanidine system (see Chapter 3 for details). In some contexts, chain length can be significant. Both the Ur/Cit and Ur*/Cit complexes are much more stable when the shorter chain is attached to the acidic parent

sequence (ΔT_m = 8 °C in each case, Table 4.1). Analogous experiments with buried guanidine pairs revealed a much lower preference for one configuration over the other.

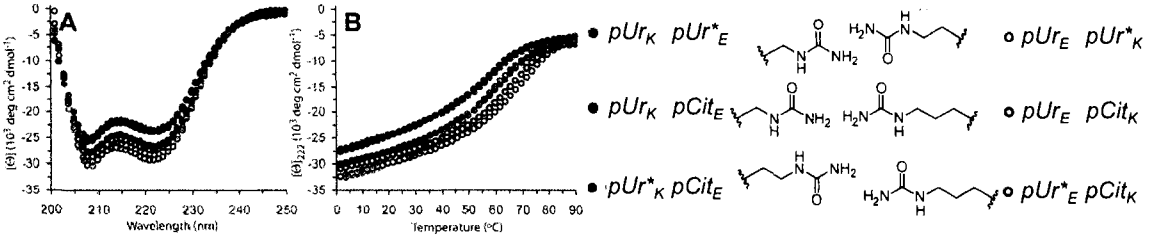


Figure 4.2. Mixed Urea—Urea interactions. Wavelength (A) and thermal denaturation (B) CD spectra for the indicated equimolar mixtures. Open/closed circles of the same color denote complexes that differ by exchange of polar residue between acidic and basic sequences.

4-3: Urea-Amide Interactions

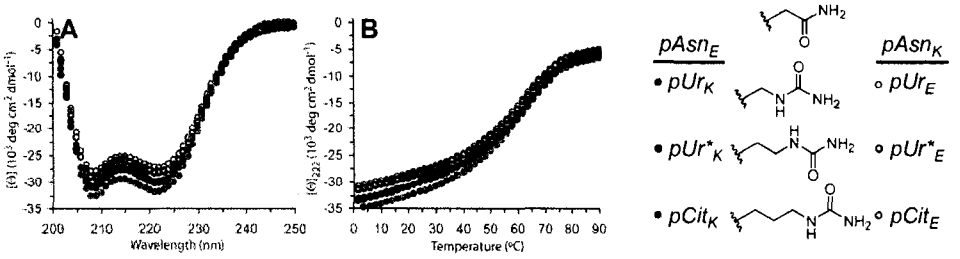


Figure 4.3. Buried Amide—Urea contacts. Wavelength (A) and thermal denaturation (B) CD spectra for the indicated equimolar mixtures. Open/closed circles of the same color denote complexes that differ by exchange of polar residue between acidic and basic sequences.

Given the selectivity observed in Gdn/Asn pairs, we next investigated heterodimers that positioned each urea derivative opposite an asparagine residue (Figure 4.3). In contrast to the interaction of either group with guanidines, the amide/urea pairs were essentially insensitive to the parent strand sequence. Mixing of the amide/urea pairs produced very stable complexes in each case (T_m = 67 to 71 °C, Table 4.2)

Table 4.2. Observed T_m Values for Amide/Urea Core Pairs.

Sample	T_m (°C)	Sample	T_m (°C)
$pUr_K/pAsn_E$	67	$pAsn_K/pUr_E$	67
$pUr^*_K/pAsn_E$	67	$pAsn_K/pUr^*_E$	67
$pCit_K/pAsn_E$	67	$pAsn_K/pCit_E$	71

4-4: Urea-Acid Interactions

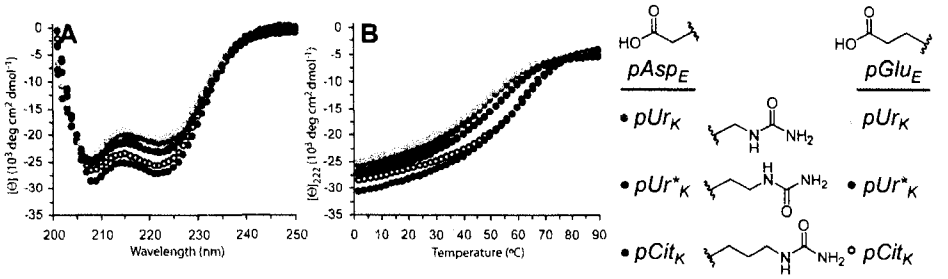


Figure 4.4. Buried Acid—Urea contacts. Wavelength (A) and thermal denaturation (B) CD spectra for the indicated equimolar mixtures.

Additionally, we evaluated mixtures of each basic urea peptide with either $pAsp_E$ or $pGlu_E$ (Figure 4.4). Similar to the previous mixtures, stable structures were observed (T_m = 55 to 67 °C), albeit somewhat less stable than the amide/urea ones (Table 4.3). As with the guanidine core residues, little to no preference was exhibited for binding to either Asp or Glu partners. Although a subtle trend is observed, complexes containing the longer the side chain length of the urea tend to form a more stable dimer.

Table 4.3. Observed T_m Values for Acid/Urea Core Pairs.

Sample	T_m (°C)	Sample	T_m (°C)
$pUr_K/pAsp_E$	57	$pUr_K/pGlu_E$	57
$pUr^*_K/pAsp_E$	55	$pUr^*_K/pGlu_E$	61
$pCit_K/pAsp_E$	67	$pCit_K/pGlu_E$	67

4-5: Urea-Guanidine Interactions

Finally, we began to explore interactions between each of the three urea and the three guanidine side chains (Figure 4.5). To begin, we paired each urea with one of the three guanidinylated residues. Wavelength and thermal unfolding CD curves for the eighteen possible combinations exhibit reasonably cooperative transitions, and a broad range of melting temperatures (T_m = 51 to 75 °C, Table 4.4). Once again, complexes

with the polar group on the shortest side chain length were the least stable, while the more stable complexes contained residues with polar groups on longer side chains.

Table 4.4. Observed T_m Values for Guanidine/Urea Core Pairs.

Sample	T_m (°C)	Sample	T_m (°C)
pUr_K/pGd_E	51	pGd_K/pUr_E	61
pUr^*_K/pGd_E	51	pGd_K/pUr^*_E	57
$pCit_K/pGd_E$	55	$pGd_K/pCit_E$	68
pUr_K/pGd^*_E	59	pGd^*_K/pUr_E	61
pUr^*_K/pGd^*_E	51	pGd^*_K/pUr^*_E	61
$pCit_K/pGd^*_E$	60	$pGd^*_K/pCit_E$	73
$pUr_K/pArg_E$	55	$pArg_K/pUr_E$	67
$pUr_K/pArg_E$	61	$pArg_K/pUr^*_E$	67
$pCit_K/pArg_E$	69	$pArg_K/pCit_E$	75

These guanidine/urea pairs exhibit an interesting preference for parent sequence identity, with higher T_m values observed universally when the guanidine residue is on the basic rather than acidic peptide. The strength of this effect is rather varied, with ΔT_m ranging from 2 to 13 degrees. It is most noticeable in structures that pair citrulline with one of the short chain guanidines (e.g. $pCit_K/pGd_E$ vs. $pGd_K/pCit_E$). This trend is consistent with previously observed Gdn/Asn pairs, although in that case all three guanidine side chains displayed an intermediate preference for the basic peptide (ΔT_m = 6 to 10 °C).

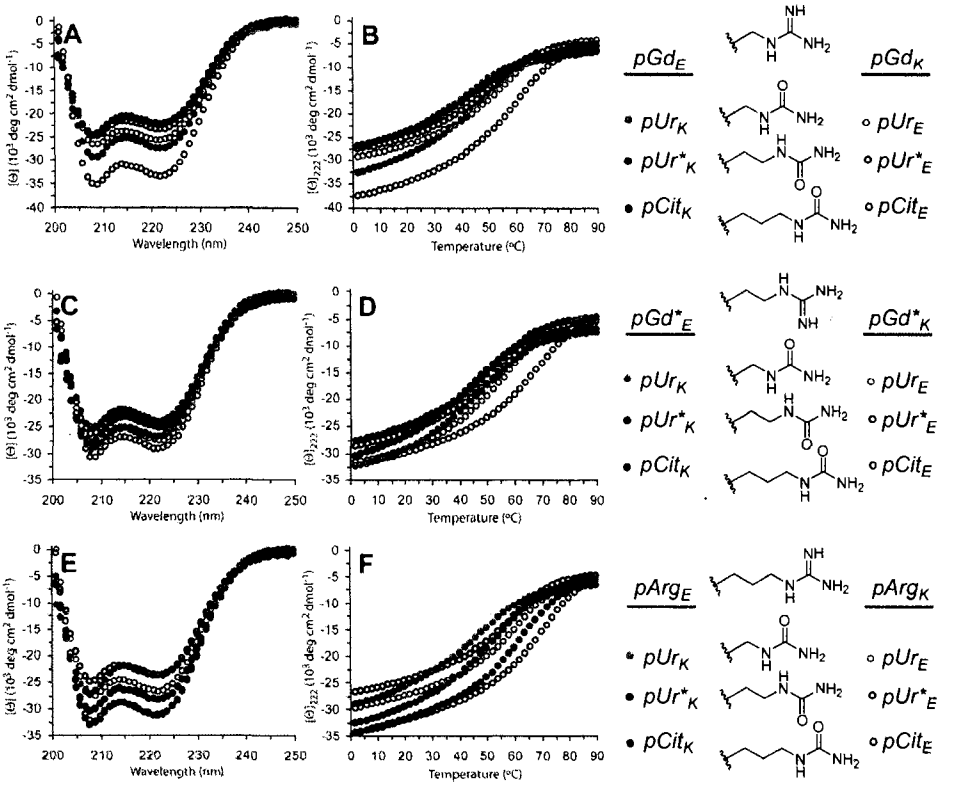


Figure 4.5. Buried Guanidine—Urea contacts. Wavelength (A, C, E) and thermal denaturation (B, D, F) CD spectra for the indicated equimolar mixtures. Open circles indicate that the acidic peptide bears the urea, filled circles indicate that the guanidine is on the acidic peptide.

4-6: Summary of All Core Interactions

Wavelength scans and thermal denaturations for all possible dimers were investigated (Appendix 3). Additionally, analytical ultracentrifugation experiments were carried out on several urea containing mixtures. Each of these experiments supports dimer formation (Table 4.5).

Table 4.5. Molecular Weights from Anlytical Ultracentrifugation

Sample	MW_{calc}	MW_{obs}	% difference
pUr_K/pUr_E	7217		
pUr^*_K/pUr^*_E	7244	7237	-0.0966
$pCit_K/pCit_E$	7272	6958	-4.32
$pUr^*_K/pAsn_E$	7215	7533	4.41
$pUr_K/pCit_E$	7244	7486	3.34
$pCit_K/pArg_E$	7272	7613	4.69

Comparison of the wavelength and melting curves for each urea derivative with its non-urea partners indicates some general trends (Figure 4.6). The short-chain guanidines were generally the weakest affinity partners, but similar length Asp and Glu residues were only slightly better. The neutral Asn contact consistently gives the most helical and stable complexes, but overall discrimination between groups tends to compress as the urea side chain lengthens, as evident by the similar affinity of $pCit_K$ for either $pArg_E$ or $pAsn_E$. pGd_E and pGd^*_E are even reasonable partners for $pCit_K$, suggesting that subtle structural changes can materially alter recognition profiles.

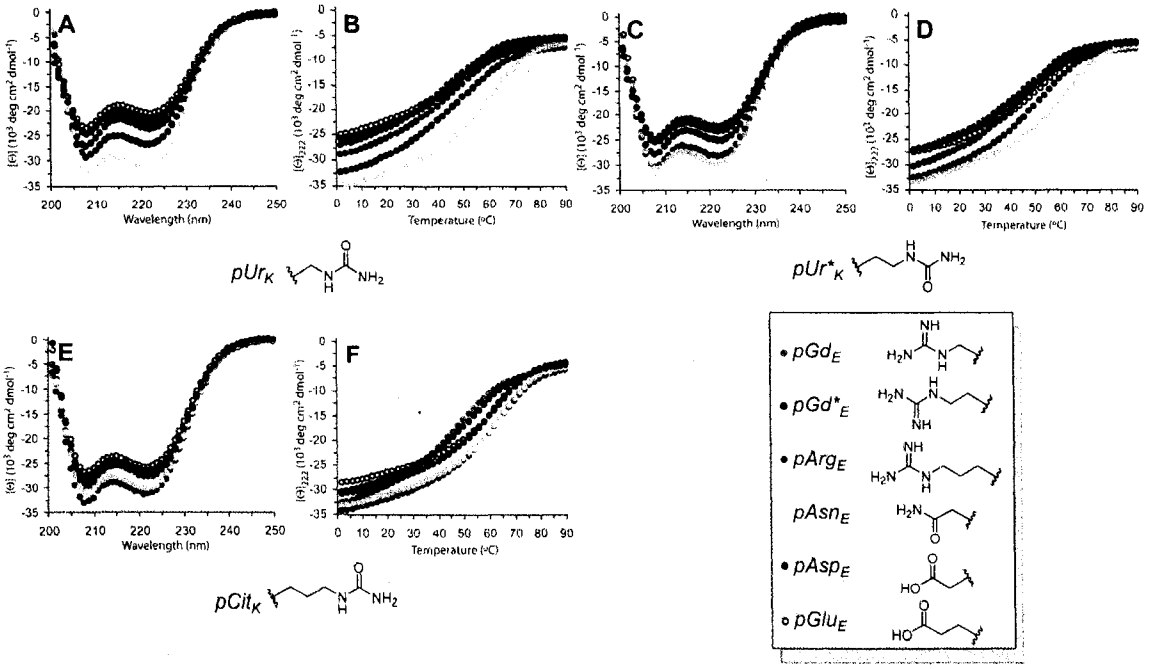


Figure 4.6. Wavelength and melting curve comparisons. Wavelength (**A**, **C**, **E**) and thermal denaturation (**B**, **D**, **F**). CD spectra for pUr_K (**A**, **B**), pUr^*_K (**C**, **D**), and $pCit_K$ (**E**, **F**) with each of the other polar groups investigated.

4-7: Urea Directed Interactions Capable of Specific Dimerization (4-component)

With these new buried polar group interactions catalogued, we turned our attention to their implementation as directing elements in multicomponent mixtures. As discussed in the previous chapter, $pAsn_K$ was found to selectively bind to $pAsn_E$ in the presence of $pAsp_E$ and the $pGdn_K$ peptides. Using the same experimental protocol,

various combinations of urea containing peptides were investigated. Since the $pCit_K/pCit_E$ and $pAsn_K/pAsn_E$ dimers are each more stable than the mixed pairs (T_m = 79 and 77 vs 71 and 67 °C), we expected that an equimolar mixture of all four peptides would sort into the like-group heterodimers. As expected, the citrulline and asparagines cores specifically recognize each other (Figure 4.7A). Essentially none of the citrulline peptides show up in the elution fraction when the $pAsn_E$ is tagged.

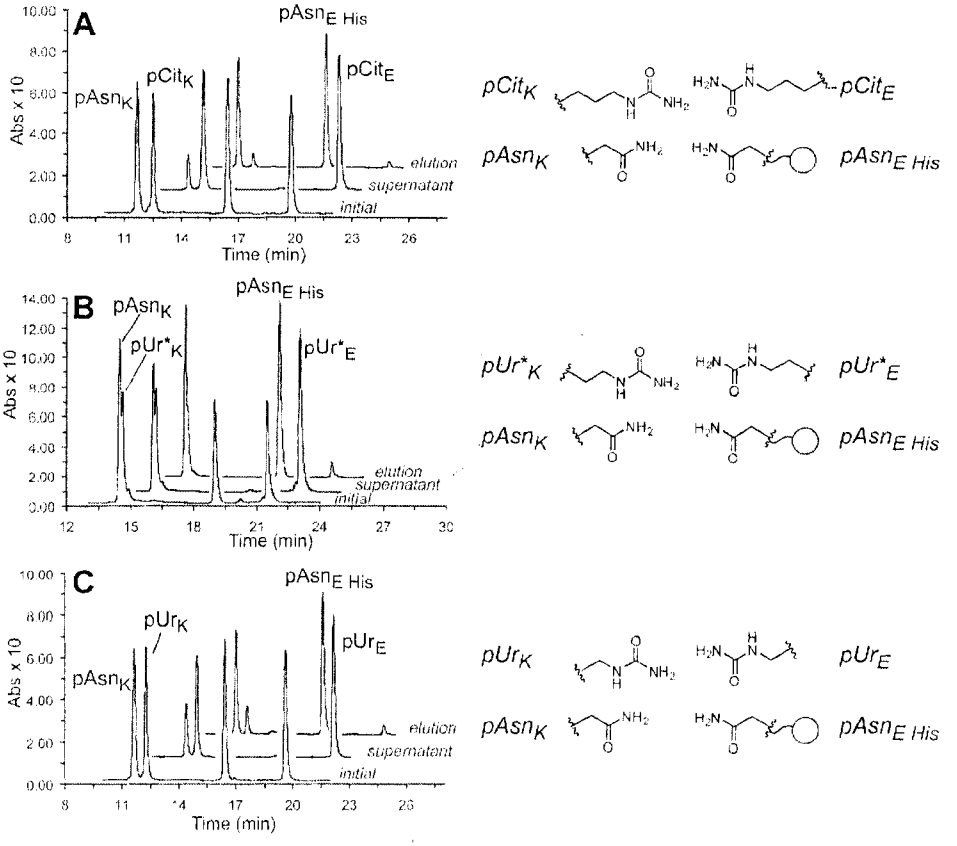


Figure 4.7. Ni-NTA analysis of four component Asn/Asn and urea/urea mixtures. Equimolar rations of: (**A**) $pAsn_K/pAsn_{EHis}/pCit_K/pCit_E$, (**B**) $pAsn_K/pAsn_{EHis}/pUr^*_K/pUr^*_E$, and (**C**) $pAsn_K/pAsn_{EHis}/pUr_K/pUr_E$ all demonstrate specific recognition. Elution fractions are dominated by the Asn peptides, indicating preferential self-recognition of each polar group. Initial solutions are 20 μM total peptide, pH 7.0, 150 mM NaCl, 10 mM phosphate.

Next, we investigated the shorter-chain ureas. Their like-core dimers are a bit less stable (T_m = 63 °C for both pUr_K/pUr_E and pUr^*_K/pUr^*_E), and the urea/amide affinities

are similar (all have T_m = 67 °C). As discussed previously, these competing T_m values were found to be a rough guide to self assembly in the four-component selective systems. Since these particular competing T_m values are near the edge of what had worked previously, we expected to observe results consistent with borderline cases. In examining the elution fractions of these shorter-chain ureas, we see that this is the case. When $pAsn_{E\ His}$ is used, the elution fractions are dominated by the asparagine peptides, but with slightly more leakage to the undesired pairs than is observed in the citrulline case (Figure 4.7B, C).

Due to the significant number of possible four-component selective systems, we preformed a detailed analysis of competing T_m values to identify other systems that would be capable of selective recognition. Fortunately, a large number of possible combinations seemed favorable. Thus, we selected a few to test (Figure 4.8). A mixture of $pGd^*_K/pAsp_E$ (T_m = 67 °C) and pUr^*_K/pUr^*_E (T_m = 63 °C) is a good candidate for selective recognition, since the $pUr^*_K/pAsp_E$ complex alternative is weak (T_m = 55 °C) and there is no gain from the corresponding pGd^*_K/pUr^*_E alternative (T_m = 61 °C). In keeping with this reasoning, Ni-NTA analysis of an equimolar four-component mixture is consistent with specific recognition (Figure 4.8A). When the $pAsp_E$ peptide is tagged, the elution fraction contains mainly the pGd^*_K and $pAsp_E$ peptides. Similarly, when $pAsn_K/pGd_E$ (T_m = 63 °C) and $pCit_K/pCit_E$ (T_m = 79 °C) are combined, selective recognition is observed since the alternative complexes are weaker, $pAsn_K/pCit_E$ (T_m = 71 °C) and $pCit_K/pGd_E$ (T_m = 55 °C). The elution fraction contains mainly the $pAsn_{K\ His}$ and pGd_E peptides, while the supernatant contains mostly the other two peptides (Figure 4.8B). To demonstrate that two urea pairs are not required for selective dimerization, we

examined a mixture of $pGd_K/pAsp_E$ and $pUr^*_K/pArg_E$ (T_m = 61 °C for both pairs, Figure 4.8C). Both competing complexes are less stable (T_m = 47 or 55 °C), which leads to reasonable results.

Previous four-component selective systems found selectivity based on the relatively unstable $pAsn_K/pAsp_E$ interaction (T_m = 51 °C). Utilizing this and possible interactions involving urea, additional selective dimerization was found. A mixture of $pGd^*_K/pAsp_E$ and $pAsn_K/pUr^*_E$ (T_m = 67 °C for both dimers) is a good candidate for selective recognition since it takes advantage of the weak $pAsn_K/pAsp_E$ complex alternative (T_m = 51 °C), and there is no gain from the corresponding pGd^*_K/pUr^*_E alternative (T_m = 61 °C, Figure 4.8D). Selectivity is also observed when the pUr^*_E peptide of the previous mixture is replaced by pUr_E, which does not change any of the melting termperatures (Figure 4.8E). Thus, when $pAsp_E$ is tagged, the supernatant contains mainly the $pAsn_K$ and pUr_E peptides, and the elution fraction contains predominantly the other two peptides. Additionally, urea/acid interactions are preserved in the presence of Asn/Asn interactions. When $pAsn_K/pAsn_E$ is combined with $pCit_K/pAsp_E$ (T_m = 77 and 67 °C respectively), selective dimerization is observed (Figure 4.8F). Once again, the unfavorable interaction between the alternative $pAsn_K/pAsp_E$ (T_m = 51 °C) and the slightly more favorable alternative $pCit_K/pAsn_E$ (T_m = 67 °C) drives the dimerization. With $pAsn_K$ tagged, the elution fraction contains the two asparagine peptides, while the supernatant contains the other two peptides.

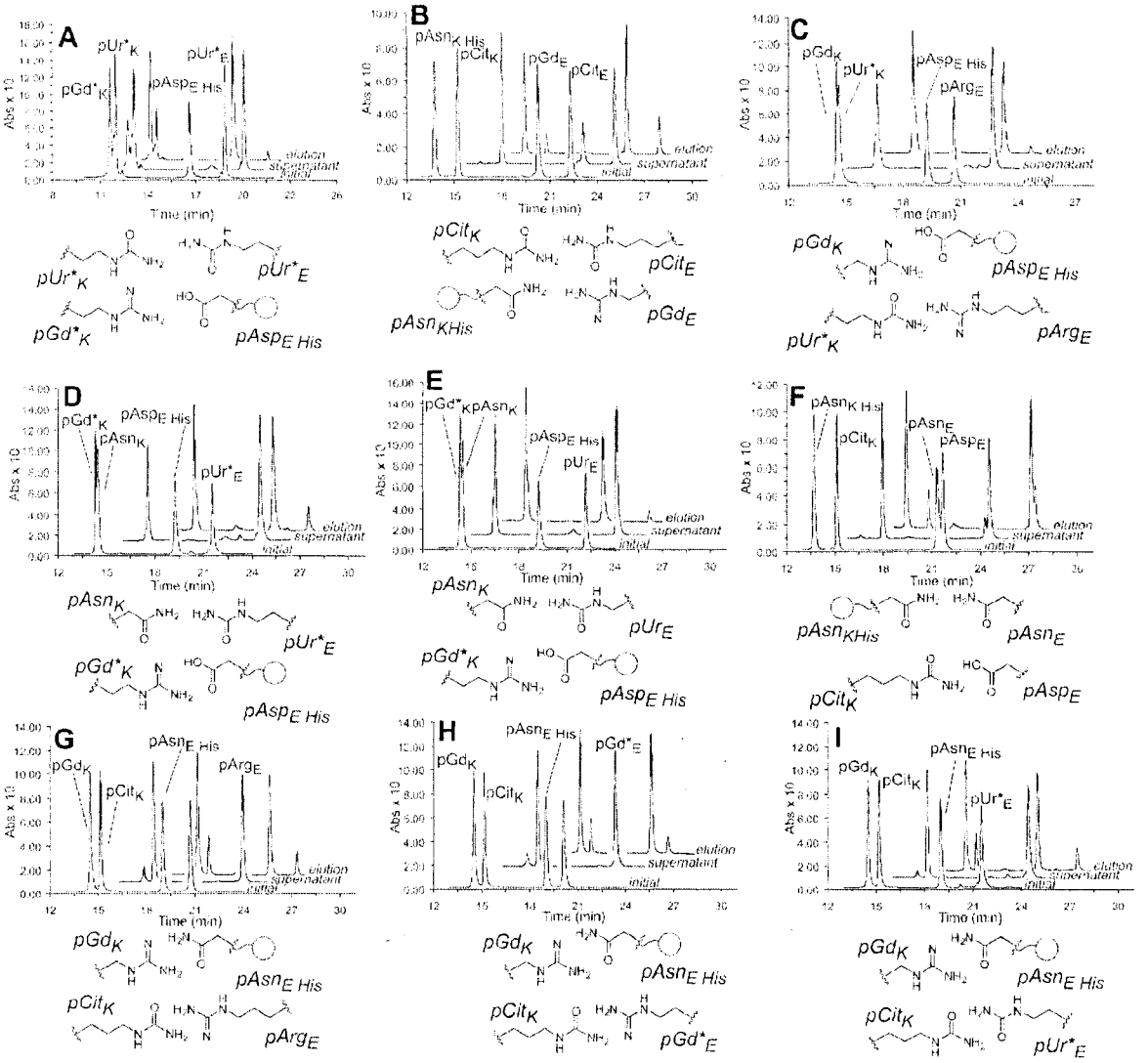


Figure 4.8. Ni-NTA analysis of other four component mixtures. Equimolar ratios of the indicated peptides all demonstrate specific recognition. Elution fractions are dominated by the two distinct peptides, indicating preferential self-recognition of each polar group. Solutions are as in Figure 4.7.

Though guanidine-guanidine core contacts are tolerated, they are generally less stable than guanidine/neutral ones. Thus a mixture of $pGd_K/pAsn_E$ (T_m = 73 °C) and $pCit_K/pArg_E$ (T_m = 69 °C) is a good candidate for selective recognition, since the $pGd_K/pArg_E$ alternative is weak (T_m = 47 °C) and there is no gain from the corresponding $pCit_K/pAsn_E$ complex (T_m = 67 °C, Figure 4.8G). When the $pAsn_E$ peptide is tagged, the supernatant contains essentially only the $pCit_K/pArg_E$ pair, with the remaining peptides

dominating the elution fraction. Similarly, when the $pArg_E$ peptide is replaced by the short pGd^*_E analogue, selectivity is observed (Figure 4.8H). Although the $pCit_K/pGd^*_E$ is slightly less stable than the $pCit_K/pArg_E$ (T_m = 60 vs. 69 °C), this is offset by the decreased stability of the pGd_K/pGd^*_E (T_m = 43 °C). Additionally, when the acidic guanidine peptide is replaced with pUr^*_E, similar selectivity is observed (Figure 4.8I). The increased stability of the undesired pGd_K/pUr^*_E complex (T_m = 57 °C) is somewhat offset by an increase in the desired $pCit_K/pUr^*_E$ dimer (T_m = 75 °C).

In addition to finding systems that selectively dimerize, general trends were observed in cases where self-recognition did not occur. When $pCit_K/pCit_E$ (T_m = 77 °C) and $pAsn_{K\ His}/pGdn_E$ (T_m ranging from 61 °C for pGd_E to 67 °C for pGd^*_E) are combined, the resulting supernatant and elution fractions contain varying degrees of the acid guanidine core peptide, $pGdn_E$, which can be attributed to the side chain length stability trends previously discussed in this chapter (Figure 4.9). When pGd_E is present, the supernatant contains only a small amount of pGd_E, and consequently the amount of $pCit_K/pCit_E$ complex present in the elution fraction is small (Figure 4.9a). When the side chain length is lengthened and pGd^*_E is present, there is significantly more leakage present in both fractions. That is, in the supernatant there is more pGd^*_E, and in the elution fraction, there is more of the undesired complex $pCit_K/pCit_E$ (Figure 4.9b). Upon addition of the $pArg_E$ peptide, there is a significant increase of the undesired complex $pCit_K/pCit_E$ in the elution fraction, while the amount of $pArg_E$ in the supernatant is similar to that of the $pCit_E$ (Figure 4.9c). These findings are consistent with the stability trends of guanidine citrulline interactions. The shorter the side chain length of the guanidine, the less stable the complex. This is illustrated in the melting temperatures, where the

higher the temperature the more stable the complex. The melting temperatures of the $pCit_K/pGd_E$ complex is the least stable (T_m = 55 °C), the $pCit_K/pGd^*_E$ complex is slightly more stable (T_m = 60 °C), and the $pCit_K/pArg_E$ complex is the most stable (T_m = 69 °C), which makes $pArg_E$ a more competitive binding partner for $pCit_K$ than the corresponding shorter side chain length acidic guanidine peptides.

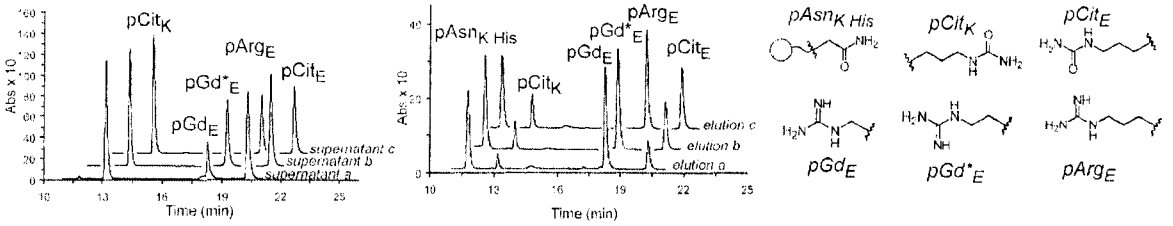


Figure 4.9. Ni-NTA analysis of trends in the supernatant and elution fractions. Mixtures with $pCit_K/pCit_E/pAsn_{KHis}$ with either pGd_E (a), pGd^*_E (b) or $pArg_E$(c). Analysis of supernatant (left) shows side chain length of $pGdn_E$ peptide increases as does the amount the $pGdn_E$ peptide. Analysis of the elution fraction (right) shows an increase in the undesired dimer $pCit_K/pCit_E$ with the increase in side chain length of the $pGdn_E$ peptide. Solutions were prepared as in Figure 4.7.

Similarly, this same trend is observed with the reverse peptides (i.e., $pGdn_K/pCit_E$). When $pCit_K/pCit_{E\ His}$ (T_m = 77 °C) and $pGdn_K/pAsp_E$ (T_m ranging from 63 °C for pGd_K to 65 °C for $pArg_E$) are combined, varying degrees of selective dimerization are observed (Figure 4.10). Since the $pGd_K/pCit_E$ alternative complex is the least stable (T_m = 68 °C), the degree of unselective recognition is the lowest of the three mixtures (Figure 4.10a), meaning that the amount of the undesired dimers are minimal amounts in both the supernatant and elution fractions. The $pArg_K/pCit_E$ alternative complex is the most stable (T_m = 75 °C), and the degree of unselective recognition is the highest of the three mixtures (Figure 4.10c). The amount of $pCit_K$ in the supernatant fraction and the amount of $pArg_K$ in the elution fraction are the highest of three mixtures. The stability and melting temperature of the $pGd^*_K/pCit_E$ alternative complex (T_m = 73 °C) is between

the other two possible $pGdn_K/pCit_E$ complexes, so the resulting selectivity of this mixture is between that of the other two mixtures (Figure 4.10b).

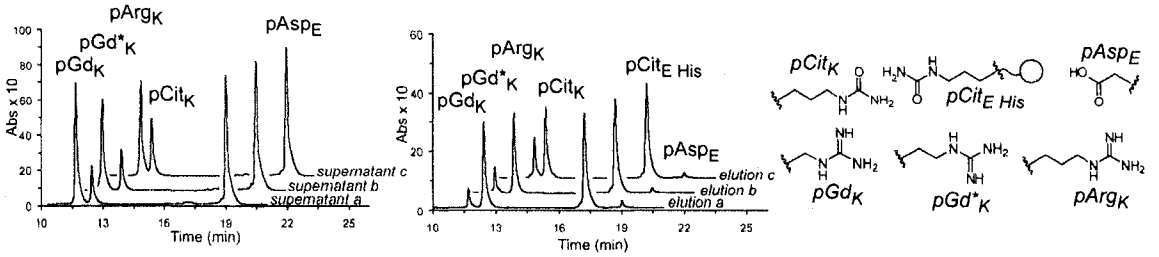


Figure 4.10. Ni-NTA analysis of trends in the supernatant and elution fractions. Mixtures with $pCit_K/pCit_{EHis}/pAspn_E$ with either pGd_K (a), pGd^*_K (b) or $pArg_K$(c). Analysis of supernatant (left) shows side chain length of $pGdn_K$ peptide increases as does the amount the $pCit_K$ peptide. Analysis of the elution fraction (right) shows an increase in the undesired dimer $pGdn_K$ with the increase in side chain length of the $pGdn_K$ peptide. Solutions were prepared as in Figure 4.7.

4-8: Urea Directed Interactions Capable of Specific Dimerization (6-component)

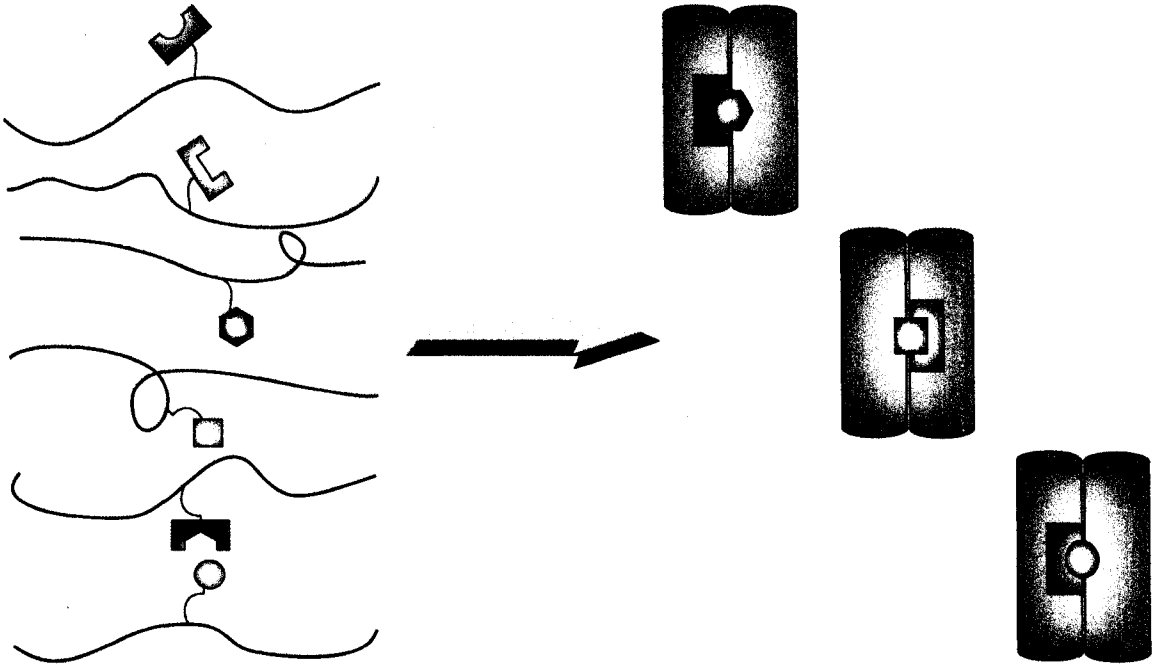


Figure 4.11. Six component self assembly scheme. If six different peptides are combined together, 3 basic and 3 acidic, will three distinct heterodimers form?

With the significant increase in available orthogonal recognition interactions demonstrated, we began to wonder if more complicated systems were feasible (Figure 4.11). In particular, the preference for self-recognition in the urea-functionalized

peptides suggested that a urea/urea pair might be added to already successful four-component systems. Impressively, such six-component systems do indeed form three independent heterodimeric coiled-coils, with little or no contamination from competing complexes (Figure 4.12). Mixtures of $pAsn_K/pAsn_E$, $pCit_K/pCit_E$, and $pGdn_K/pAsp_E$ were investigated. The overall system displayed a remarkable level of recognition specificity, with the identity of the guanidine peptide seemingly relatively unimportant.

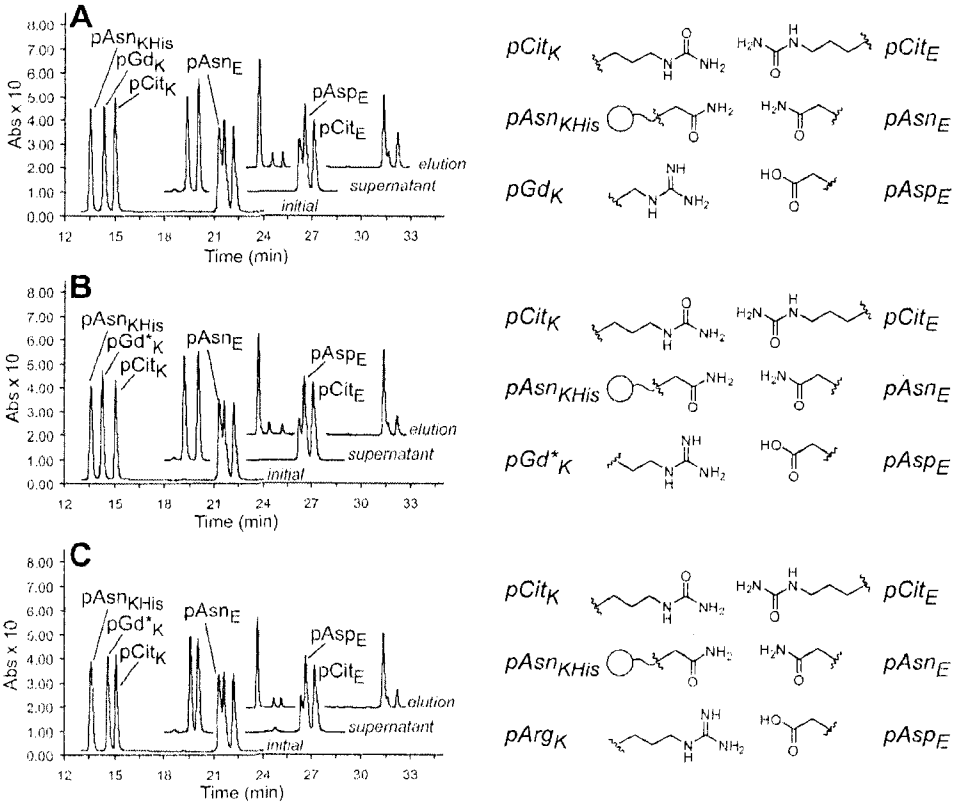


Figure 4.12. A Ni-NTA analysis of six component Asn/Asn, Gdn/Asp, urea/urea mixtures. The given HPLC traces result from experiments with equimolar mixtures of: (A) $pAsn_{KHis}/pAsn_E/pCit_K/pCit_E/pGd_K/pAsp_E$, (B) $pAsn_{KHis}/pAsn_E/pCit_K/pCit_E/pGd^*_K/pAsp_E$, and (C) $pAsn_{KHis}/pAsn_E/pCit_K/pCit_E/pGd_K/pAsp_E$. Each demonstrates specific recognition, as the elution fractions are dominated by Asn peptide. Initial solution were prepared as in Figure 4.7.

The difficulty of constructing a viable six-component recognition system is underscored by replacing the $pAsp_E$ peptide in the successful mixture with $pGlu_E$ (Figure 4.13). Despite the subtle structural change, essentially all specificity is removed. Each of

the two other possible binding partners for each component of the tagged dimer are observed in both supernatant and elution fraction.

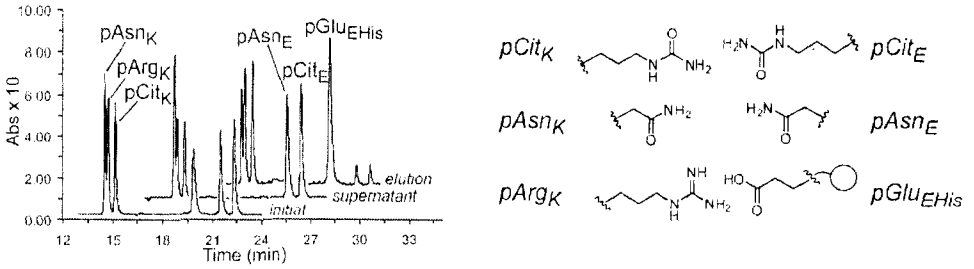


Figure 4.13. An unsuccessful six component recognition experiment. An equimolar mixtures of $pAsn_K/pAsn_E/pCit_K/pCit_E/pGd_K/pGlu_{EHis}$, fails to exhibit specific recognition. The elution fraction contains substantial amounts of all thee basic peptide, demonstrating that Glu is not an appropriate binding partner for Arg in this context, despite the success of Asp. Initial solutions were prepared as in Figure 4.7.

4-9: Experimental Information

Ni-NTA Affinity Tag Experiments: A 0.5 mL sample of a 50% slurry of Ni-NTA agarose (Qiagen) in an Eppendorf tube was centrifuged for 30 s, followed by removal of the supernatant. Peptide solution (1 mL of 20 μM or 10 μM total peptide, 5 μM each peptide) was added, and the tube was repeatedly inverted for 5 minutes. The sample was centrifuged (30 s) and the supernatant (flow-through fraction) was removed. The procedure was then repeated with 1 mL of buffer (wash fraction) and 1 mL of buffer containing 250 mM imidazole (elution fraction), except that the wash fraction was only agitated for 30 seconds.[4] Solutions were analyzed by RP-HPLC.

CD Spectroscopy: All experiments were performed on an Aviv model 202 circular dichroism spectrometer. Sample concentrations were measured by UV absorbance of the acetamidobenzoate label at 270 nm. Wavelength data are the average of three scans from 250 to 200 nm in 1 nm steps. Thermal denaturation experiments at 222 nm were run from 0° to 90°C in two-degree steps, at a two-degrees/minute rate of

increase with one-minute equilibration and data averaging at each temperature. T_m values were obtained from minima of the first derivative of θ vs. 1/T plots.[5]

Analytical Ultracentrifugation: Sedimentation equilibrium experiments were performed using a Beckman XL-A analytical ultracentrifuge equipped with an An60-Ti or An50-Ti rotor. Data were collected using 12 mm path length six-sector centerpieces at 270 nm. Samples were dialyzed against the reference buffer at 4°C overnight. Data were collected at 38,000 and 48,000 r.p.m. at concentrations spanning 17-55 μM. Samples were judged equilibrated when three consecutive scans taken 1h apart were indistinguishable (equilibration was complete in 12h in all cases). Solvent densities and partial molar volumes were calculated in the manner prescribed by Durchschlag and Zipper,[6] using one of three v_bar values for Gd and Gd* (Table 4.6). Data were analyzed using Ultrascan[7] and fit globally to ideal single species. For fits see Appendix 4.

Table 4.6. v_bar values for unnatural amino acids and peptides containing them

Residue	v_bar	Peptide	v_bar
Ur	0.637	pUr_K	0.781
Ur*	0.687	pUr_E	0.732
		pUr^*_K	0.783
		pUr^*_E	0.734

4-10: Literature Cited

[1] (a) Nowick, J.S. *Org Biomol Chem* **2006**, 4, 3869-3885. (b) Rebek, J., Jr. *Angew Chem Int Ed* **2005**, 44, 2068-2078. (c) Takemoto, Y. *Org Biol Chem* **2005**, 3, 4299-4306.

[2] Diss, M.L.; Kennan, A.J. *Biopolymers* **2007**, 86, 276-281.

[3] Diss, M.L.; Kennan, A.J. *J. Am. Chem. Soc.* **2008**, 130, 1321-1327.

[4] Method patterned after that in: Brown, M.B.; Saucer, R.T. Pr*oc. Nat. Acad. Sci. U.S.A.* **1999**, 96, 1983-1988.

[5] Cantor, C.R.; Schimmel, P.R. Biophyscial *Chemistry of Macromolecules, Pt.3: The Behavior of Biological Macromolecules*; W.H. Freeman: New York, NY, 1980, p. 1132.

[6] Durchshlag, H.; Zipper, P. *Progr Collaoid Polym Sci* **1994**, 94, 20-39.

[7] Ultrascan (version 7.1). Developed by Borries, Demeler, The University of Texas Health Science Center at San Antonio.

Chapter 5

Buried Polar Interactions within a Trimeric System

5-1: Introduction

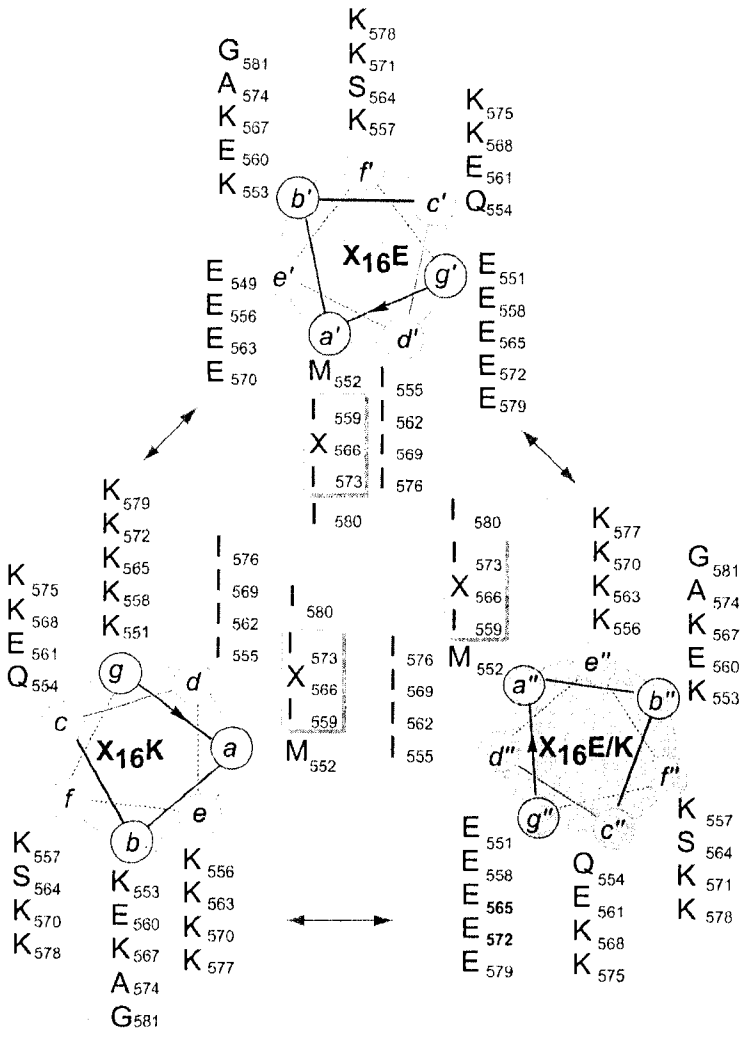


Figure 5.1. Helical wheel projection. The sequences are derived from the original T_{16} system, with the steric matching residue replaced with a polar one, as indicated by the X. Underlined lysine side chains are capped with acetamidobenzoyl groups as spectroscopic labels.

Since the variable-length guanidinylated and urea core residues support formation of well-defined dimeric coiled coils with a variety of binding partners, we sought to examine these same buried polar interactions within a trimeric system. Previous work in

our group developed steric matching of hydrophobic core side chains as a route to specific trimer formation.[1] Expanding on this, we sought to use a similar sequence with a single buried polar residue in the core of the trimeric peptide in hopes of directing trimer formation (Figure 5.1). The buried polar interactions investigated are similar to those investigated within the dimeric system involving polar interactions between guanidine, urea, acid and amide functional groups.

5-2: Peptides Synthesized

Utilizing the previous on resin methodology for guanidinylation and urea formation, twelve peptides were synthesized. These peptides consist of four basic peptides ($pXxx_K$), with lysine in the *e/g* positions, four acidic peptides ($pXxx_E$), with glutamic acid in the *e/g* positions, and four mixed peptides ($pXxx_{E/K}$), with lysine in the *e* position and glutamic acid in the *g* position. Each of these peptides has a buried polar residue in the central *a* position (i.e. residue 16, Figure 5.2).

```
                    a  d  a  d  a  d  a  d  a
pCit_K      AcNH-KMKQIKKKIEEIKSKUKKIKKKIAAIKKKIG-Am
pArg_K      AcNH-KMKQIKKKIEEIKSKRKKIKKKIAAIKKKIG-Am
pHCit_K     AcNH-KMKQIKKKIEEIKSKU*KKIKKKIAAIKKKIG-Am
pHArg_K     AcNH-KMKQIKKKIEEIKSKR*KKIKKKIAAIKKKIG-Am
pCit_E/K    AcNH-EMKQIKKEIEEIKSEUKKIKKEIAAIKKEIG-Am
pGlu_E/K    AcNH-EMKQIKKEIEEIKSEEKKIKKEIAAIKKEIG-Am
pGln_E/K    AcNH-EMKQIKKEIEEIKSEQKKIKKEIAAIKKEIG-Am
pHCit_E/K   AcNH-EMKQIKKEIEEIKSE  *KKIKKEIAAIKKEIG-Am
pCit_E      AcNH-EMKQIEKEIEEIESEUKKIEKEIAAIEKEIG-Am
pArg_E      AcNH-EMKQIEKEIEEIESERKKIEKEIAAIEKEIG-Am
pHCit_E     AcNH-EMKQIEKEIEEIESEU*KKIEKEIAAIEKEIG-Am
pHArg_E     AcNH-EMKQIEKEIEEIESER*KKIEKEIAAIEKEIG-Am
                551         560         570         580
```

Figure 5.2. Peptides Used. Sequences of each peptide are given, with nonstandard amino acids represented by the following code: U = Citrulline, U* = Homocitrulline, and R* = Homoarginine. Underlined lysine side chains are capped with acetamidobenzoyl groups as spectroscopic labels.

5-3: Various Interactions Investigated

Previous work in the dimer systems showed that the *e/g* positions influenced the overall stability of the dimer. To investigate this within the trimer system, two identical core interactions were examined in which the basic peptide and the acidic peptide core residues were swapped (Figure 5.3). The observed wavelength and thermal melt spectra show nearly identical plots for each trimer, indicating that the *e/g* positions really do not impact the overall helicity or stability of the trimer. Therefore, only one possible interaction of the core polar groups was investigated. Meaning, since there are multiple combinations of peptides that would give the same polar core interactions (i.e. $pCit_K/pCit_{E/K}/pHArg_E$ and $pHArg_K/pCit_{E/K}/pCit_E$), only one was investigated (in this case $pCit_K/pCit_{E/K}/pHArg_E$). All of the possible core interactions between the varying side chain length urea and guanidine groups with glutamic acid and glutamine residues were investigated (Appendix 3).

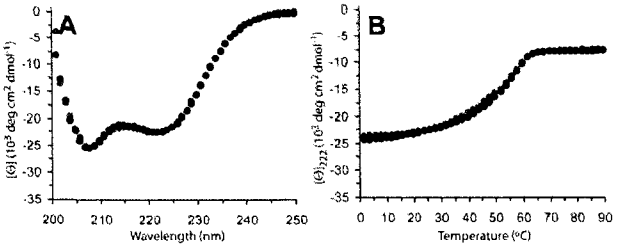


Figure 5.3. Core-swapped mixed trimer interactions. Wavelength (A) and thermal denaturation (B) CD spectra for $pArg_K/pGln_{E/K}/pHCit_E$ (red) and $pHCit_K/pGln_{E/K}/pArg_E$ (blue).

5-4: Stability Trends

Although we anticipated that burial of these polar residues might substantially destabilize potential complexes, only mixtures containing two guanidine groups exhibited a CD signal consistent with that of the average component signals, which is expected for

no complex formation (Figure 5.4). In each case the weighted average calculation and the wavelength scan are nearly identical, indicating a destabilizing interaction.

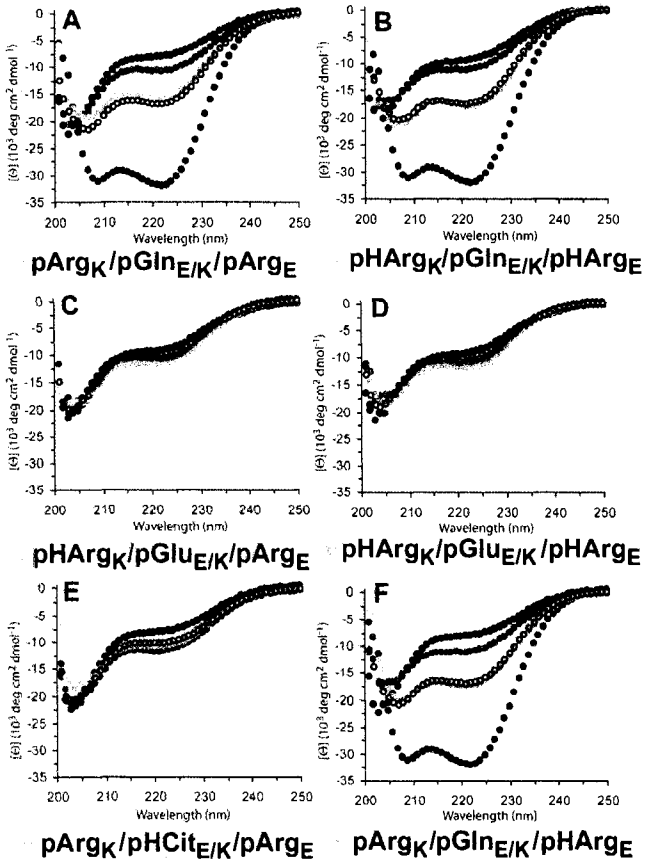


Figure 5.4. Core interactions with two guanidines. Wavelength CD spectra for pure solutions of pXxx$_K$ (blue), pXxx$_E$ (red), pXxx$_{E/K}$ (green). In each case, the calculated weighted-average signal (open circles) and the traces for an equimolar mixture (orange) are as indicated.

The weak complexes containing two guanidines do exhibit a stability trend in regards to the other residue present in the core. The most stable of these mixtures contains a glutamine residue in the core followed by a citrulline residue (Figure 5.5). Homocitrulline and glutamic acid residues lead to similar stability with the homocitrulline interaction being slightly more stable than that of glutamic acid. The length of the side chain on the guanidine group does contribute to the overall stability of the complex. Mixtures with arginine as the only source of the guanidine are the least

stable (Figure 5.5A). The stability increases slightly when homoarginine is the only source of guanidine in the core (Figure 5.5B). Yet, mixtures with both arginine and homoarginine residues in the core are the most stable (Figure 5.5C).

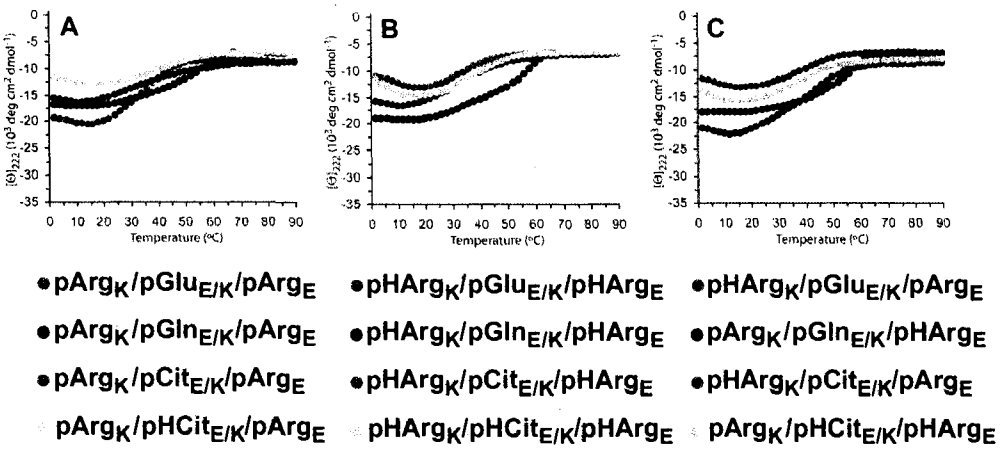


Figure 5.5. Buried interactions with two guanidines. Thermal unfolding data for the indicated equimolar solutions with (A) two pArg peptides (B) two pHArg peptides and (C) one pArg and one pHArg peptide.

Interactions that contain only one guanidine core residue do not exhibit these same destabilizing interactions as is the case with two guanidine core residues (Figure 5.6). The length of the side chain of the guanidine group does not seem to impact the overall stability and helicity as much as the presence of two guanidine groups in the core. This finding is consistent with previous guanidine—guanidine interactions in the core since the burial of like charges tends to be destabilizing.

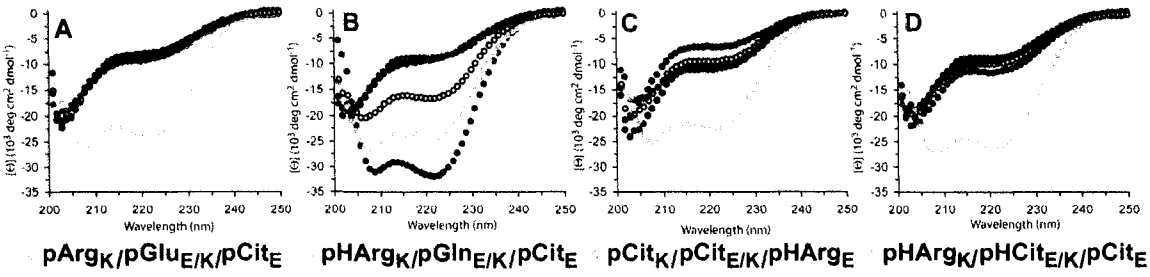


Figure 5.6. Core interactions with only one guanidine. Wavelength CD spectra for pure solutions of pXxx$_K$ (blue), pXxx$_E$ (red), pXxx$_{E/K}$ (green). In each case, the calculated weighted-average signal (open circles) and the traces for an equimolar mixture (orange) are as indicated.

With the role of guanidine interactions established, we next began to investigate the role of urea interactions in the core. Interactions involving at least one citrulline were investigated (Figure 5.7). Examination of the thermal denaturation reveals an expected stability preference for an entirely neutral urea core. While all three mixtures of an entirely urea core exhibit similar stabilities, the complex with two homocitrulline residues is the most stable, followed closely by the other two all urea containing mixtures. The mixed urea cores containing one guanidine group are far more stable then when two guanidine groups are present, yet are slightly less stable than when no guanidine groups are present. The most stable of these complexes is the $pHArg_K/pHCit_{E/K}/pCit_E$ complex, with the $pCit_K/pCit_{E/K}/pHArg_E$ complex closely following.

Table 5.1. Melting Temperatures of trimers containing at least one pCit.

Sample	T_m(°C)	Sample	T_m(°C)
$pArg_K/pCit_{E/K}/pCit_E$	47	$pCit_K/pCit_{E/K}/pCit_E$	51
$pCit_K/pCit_{E/K}/pHArg_E$	51	$pCit_K/pHCit_{E/K}/pCit_E$	57
$pArg_K/pHCit_{E/K}/pCit_E$	45	$pHCit_K/pCit_{E/K}/pHCit_E$	57
$pHArg_K/pHCit_{E/K}/pCit_E$	51	$pHCit_K/pHCit_{E/K}/pHCit_E$	59
$pHArg_K/pHCit_{E/K}/pHCit_E$	49	$pArg_K/pHCit_{E/K}/pHCit_E$	41

The length of the side chain does not significantly impact the overall stability of the complex as much as the polar interactions in the core do. The increased side chain length does increase the melting temperature of the mixture, going from an entirely citrulline core to one with two homocitrullines and one citrulline, the melting temperature is increased by six degrees (Table 5.1). Additionally replacing the $pArg_K$ with $pHArg_E$ increases the melting temperature by at least four degrees.

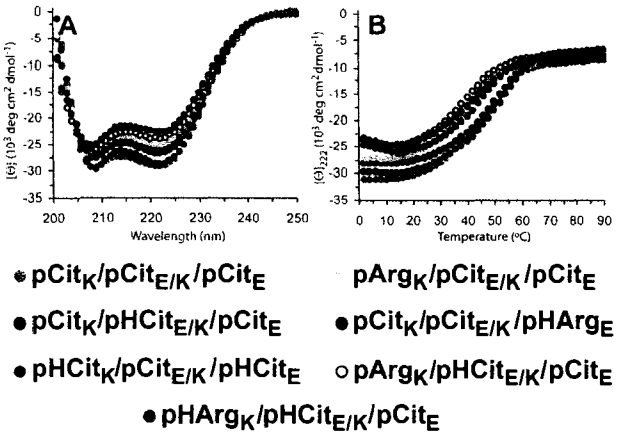


Figure 5.7. Buried interactions with one citrulline residue. Wavelength (A) and thermal unfolding data (B) for the indicated equimolar mixtures.

Core interactions that contain urea from a homocitrulline are very similar to interactions with a urea from a citrulline (Figure 5.8). The entirely homocitrulline comprised core is the most stable with an increased melting temperature of eight degrees over the entirely citrulline comprised core (Table 5.1). Additionally, when the $pHCit_K$ in the $pHCit_K/pHCit_{E/K}/pHCit_E$ complex is replaced with either pArgK or pHArgK there is a significant decrease in melting temperature (18 °C and 10 °C, respectively).

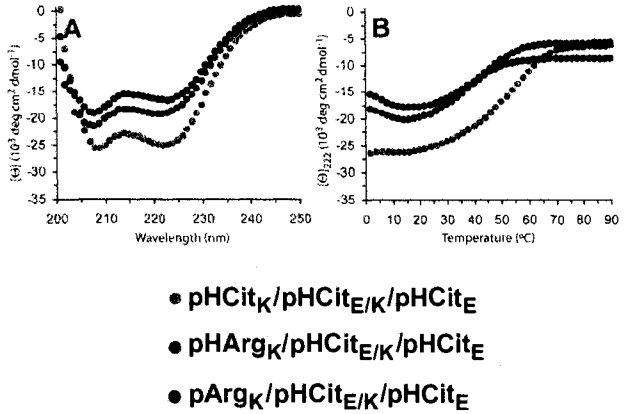


Figure 5.8. Buried interactions with pHCit. Wavelength (A) and thermal unfolding data (B) for the indicated equimolar mixtures.

Next we turned our attention to other polar core interactions. Since the aspartic acid residue in the dimeric system bound nicely with the guanidinylated residue, we

sought to establish a similar interaction in the trimer system with glutamic acid. It was thought that the additional methylene was necessary for the core interactions within the trimeric geometry. Similarly to the previous core interactions with urea, interactions between glutamic acid and two ureas were the most stable (Figure 5.9).

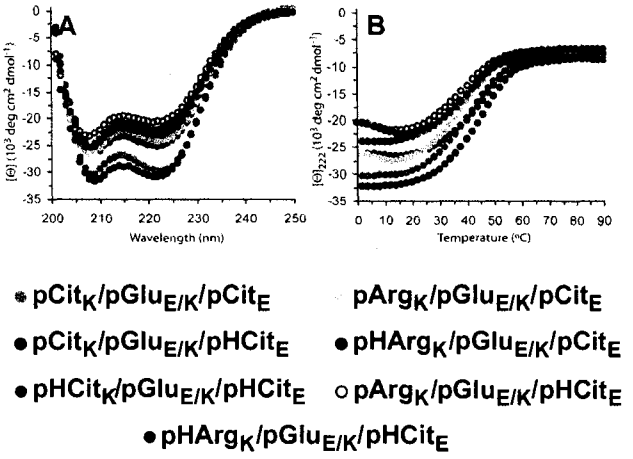


Figure 5.9. Buried interactions with $pGlu_{E/K}$. Wavelength (A) and thermal unfolding data (B) for the indicated equimolar mixtures.

The $pCit_K/pGlu_{E/K}/pHCit_E$ complex was the most stable followed closely with both the $pHCit_K/pGlu_{E/K}/pHCit_E$ and $pCit_K/pGlu_{E/K}/pCit_E$ complexes. The mixed core interactions between urea, guanidine and the acid were slightly less stable than the previous complexes. The melting temperature range is less significant with the $pGlu_{E/K}$ peptide complexes than with the urea complexes (Table 5.2). The range for the $pGlu_{E/K}$ peptide complexes is eight degrees, whereas with the urea peptide complexes, the range is eighteen degrees. Overall, complexes with the $pGlu_{E/K}$ peptide are less stable than with any other $pXxx_{E/K}$ peptide. This can be attributed to the destabilizing charge of the acidic residue in the neutral hydrophobic core of the trimer.

Table 5.2. Melting Temperatures of trimers containing $pGlu_{E/K}$.

Sample	T_m(°C)	Sample	T_m(°C)
$pArg_K/pGlu_{E/K}/pCit_E$	45	$pCit_K/pGlu_{E/K}/pCit_E$	45
$pHArg_K/pGlu_{E/K}/pCit_E$	45	$pCit_K/pGlu_{E/K}/pCit_E$	51
$pArg_K/pGlu_{E/K}/pHCit_E$	43	$pHCit_K/pGlu_{E/K}/pHCit_E$	51
$pHArg_K/pGlu_{E/K}/pHCit_E$	43		

Having had previous success with amide interactions within the dimer, we sought to examine the role of amide interactions within the trimer system. Previous core investigations had shown that a single core glutamine residue prefers the geometry of a trimer, whereas a core asparagine residue prefers dimeric geometry.[2] Thus, we chose to use Gln as our amide source in the trimer system. Wavelength and thermal denaturation experiments were run on mixtures containing $pGln_{E/K}$ (Figure 5.10). Evaluation of this data showed surprisingly similar complex stabilities. Additionally, all of the melting temperatures for complexes containing $pGln_{E/K}$ are 59 °C. Mixtures containing $pGln_{E/K}$ tended to be more stable than mixtures containing $pGlu_{E/K}$, but less stable then mixtures with either $pCit_{E/K}$ or $pHCit_{E/K}$.

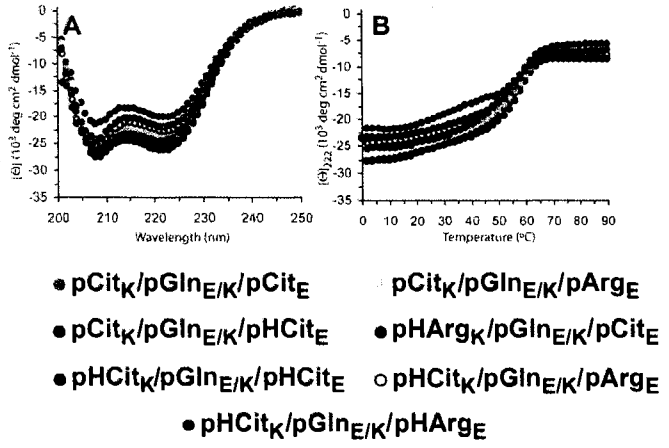
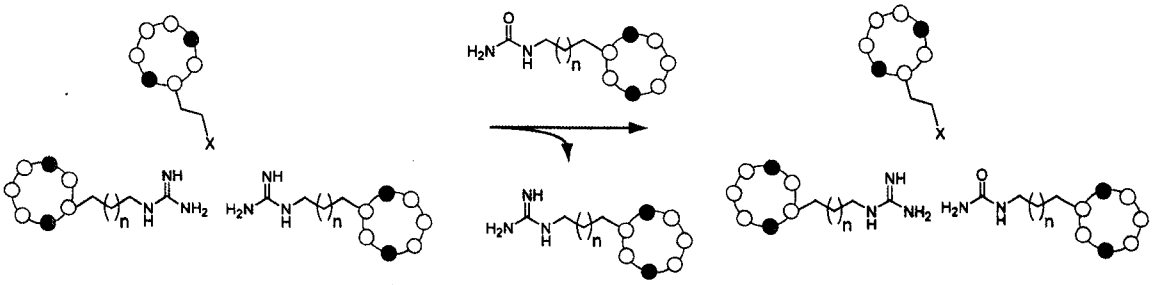


Figure 5.10. Buried interactions with $pGln_{E/K}$. Wavelength (A) and thermal unfolding data (B) for the indicated equimolar mixtures.

5-5: Exchange Experiments



Scheme 5.1. General exchange experiment scheme.

After establishing stability information for each possible trimer core interaction, we sought to manipulate the weak interactions to yield stronger ones. It was hoped that we could exchange a single peptide of a weak complex for one that would yield a stronger complex (Scheme 5.1). To assay this, we employed Ni-affinity tagging experiments (Figure 5.11).

Since trimer interactions with two guanidine groups tend to be relatively weak and unstable, we sought to exchange out one of the guanidinylated residues with a residue containing urea. This was thought to be a favorable exchange since trimer interactions with two urea groups and one guanidine are far more stable than interactions between two guandine groups and one urea. An equimolar mixture of $pArg_{KHis}/pCit_{E/K}/pArg_{E}$ was prepared for such an investigation (Figure 5.12). After supernatant removal and washing with blank buffer, $pCit_{E}$ was added, followed by a second supernatant removal and washing, as well as elution of bound materials from the agarose beads. Unfortunately, after examination of each fraction by HPLC, the second supernatant did not contain the desired $pArg_{E}$ peptide. Rather it contained mainly $pCit_{E}$. Additionally, the elution fraction did not contain the desired more stable trimer of

$pArg_{KHis}/pCit_{E/K}/pCit_E$; rather it contained a mixture of the acidic peptides in addition to the $pArg_{KHis}$ and $pCit_{E/K}$ peptides.

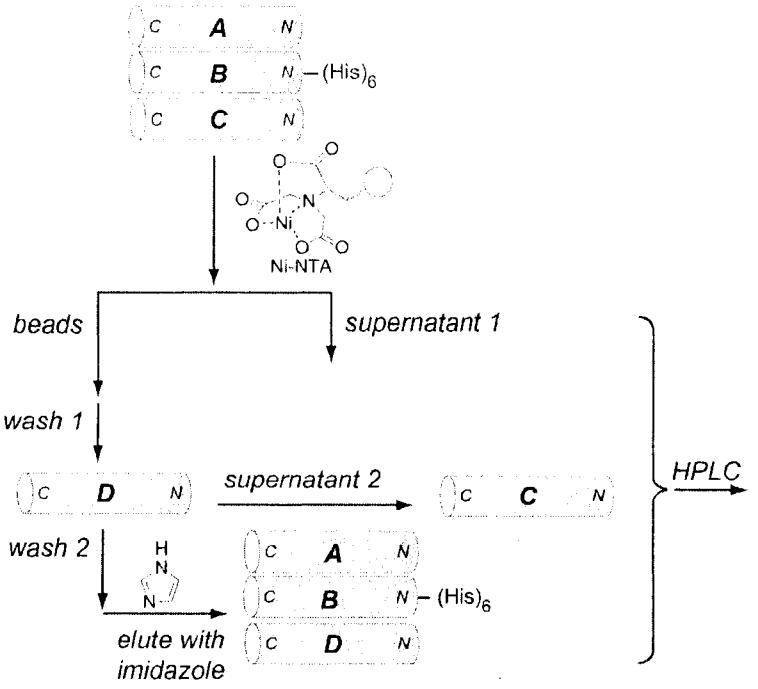


Figure 5.11. Ni-NTA affinity tag exchange experiment scheme. Initially a weak trimer composed of equimolar amounts of **A**, **B**, which bears an N-terminal Gly-Gly-$(His)_6$ affinity tag, and **C** is exposed to Ni-NTA agarose beads. Upon exposure **B** is bound through the His tag, along with **A** and **C** which are bound to **B** through core polar interactions. After supernatant removal and washing with blank buffer, an equimolar amount of peptide **D** is added which forces peptide **C** out of the trimer and the bound materials. After a second supernatant removal and washing with blank buffer, bound material is eluted by treatment with a guanidine-imidazole buffer. HPLC analysis reveals the material in each fraction.

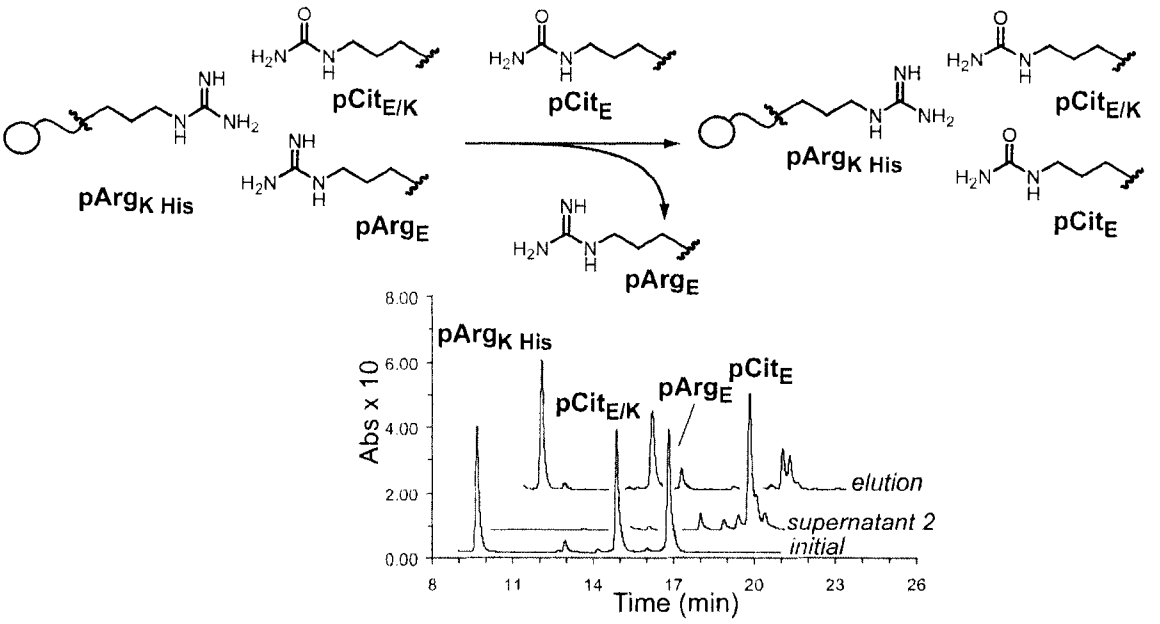


Figure 5.12. Exchange Experiment. The front trace is the initial fraction containing equimolar amounts of each peptide, $pArg_{KHis}/pCit_{E/K}/pArg_E$, the middle trace is the second supernatant fraction containing mainly $pCit_E$, and the back trace is the elution fraction containing of $pArg_{KHis}$, $pCit_{E/K}$, with $pArg_E$ and $pCit_E$.

A second set of Ni-affinity exchange conditions were next examined involving an equimolar solution of $pHArg_K/pHCit_{E/K\ His}/pHCit_E$ (Figure 5.13). It was hoped that the addition the $pCit_E$ peptide would replace $pHCit_E$ since the $pHArg_K/pHCit_{E/K}/pCit_E$ complex is more stable than the $pHArg_K/pHCit_{E/K}/pHCit_E$ complex (ΔT_m = 3 °C, Table 5.1). Unfortunately, this stability difference was not significant enough for the Ni-affinity experiment to be successful. Evaluation of the second supernatant fraction revealed a 2:1 ratio of $pCit_E$ to $pHCit_E$. Additionally, the elution fraction contains a similar ratio of the acidic peptides as well as the $pHArg_K$ and $pHCit_{E/KHis}$ peptide. This result is not overly surprising since the difference in melting temperatures required for selective dimerization is greater than three degrees.

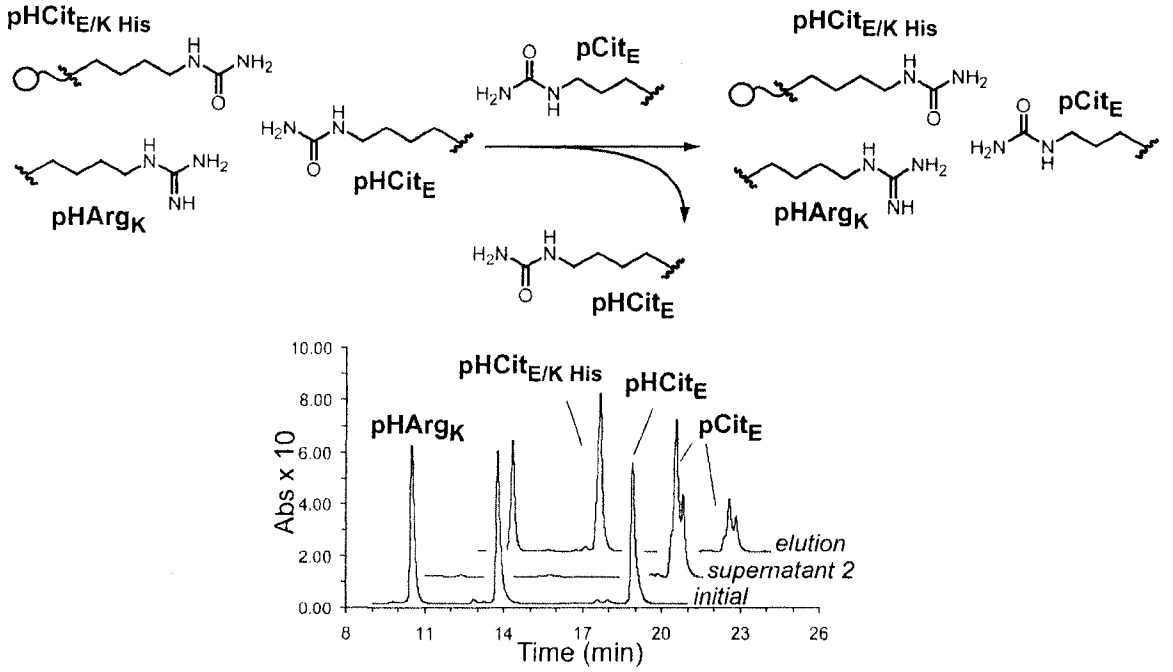


Figure 5.13. Exchange Experiment. The front trace is the initial fraction containing equimolar amounts of each peptide, pHArg$_K$/pHCit$_{E/K\ His}$/pHCit$_E$, the middle trace is the second supernatant fraction containing mainly pCit$_E$ and pHCit$_E$, and the back trace is the elution fraction containing of pHArg$_K$, pHCit$_{E/K\ His}$, with pCit$_E$ and pHCit$_E$.

Further utilizing the weak trimer interactions involving guanidine, an equimolar solution of pHArg$_K$/pHCit$_{E/K\ His}$/pHCit$_E$ was prepared to examine exchange possibilities going from a mixed guanidine/urea core to an all urea core (Figure 5.14). It was hoped that the exchange pHArg$_K$ with pHCit$_K$ would be successful, since the difference in melting temperatures between the pHArg$_K$/pHCit$_{E/K}$/pHCit$_E$ and HCit$_K$/pHCit$_{E/K}$/pHCit$_E$ is more significant (ΔT_m = 10 °C, Table 5.1). Following the same Ni-affinity exchange protocol from above, pHCit$_K$ was added after the wash fraction was removed. Investigation of the second supernatant fraction reveals a 1:2 ratio of pHArg$_K$ to pHCit$_K$. Likewise, evaluation of the elution fraction shows a near 1:1 ratio of the two basic peptides in addition to the pHCit$_{E/K\ His}$ and pHCit$_E$ peptides. Evaluation of the wavelength scan and thermal unfolding of these two trimers reveals a substantial gain in stability going from the mixed core to an entirely urea core, with the stability of

$pArg_K/pHCit_{E/K}/pHCit_E$ being between the two other peptides (Figure 5.8). Since the weaker of the two mixed core trimers containing two pHCit peptides failed the exchange experiment, it is unlikely that the intermediate trimer $pArg_K/pHCit_{E/K}/pHCit_E$ complex would work.

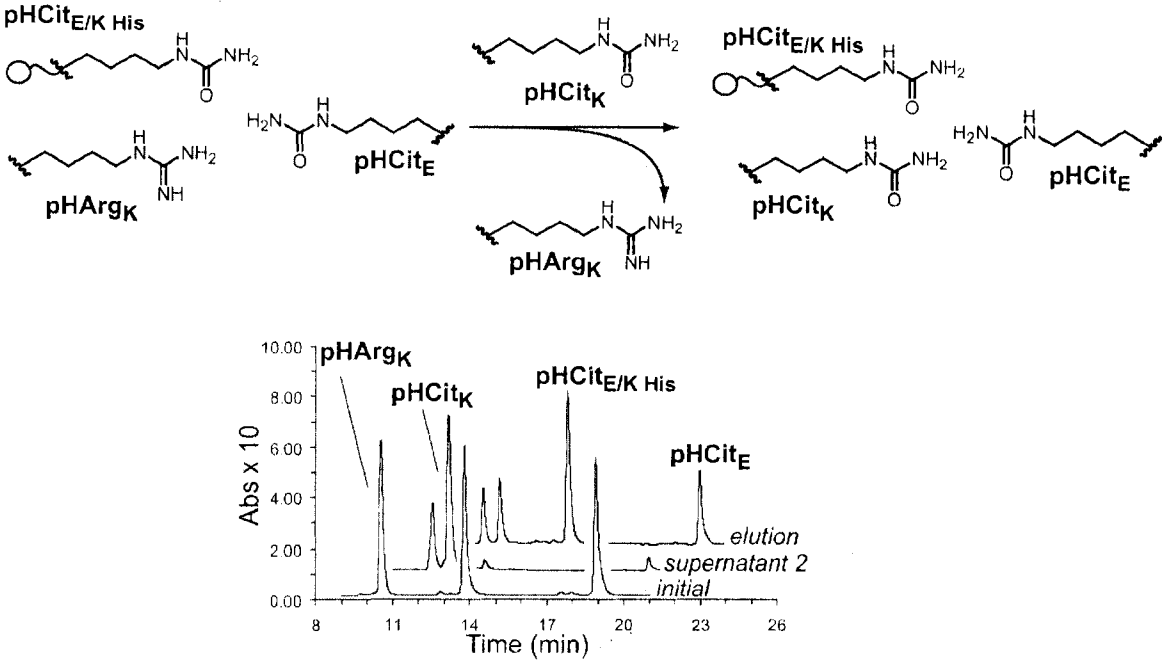


Figure 5.14. Exchange Experiment. The front trace is the initial fraction containing equimolar amounts of each peptide, $pHArg_K/pHCit_{E/K\ His}/pHCit_E$, the middle trace is the second supernatant fraction containing $pHArg_K$ and $pHCit_K$, and the back trace is the elution fraction containing of nearly equal amounts of $pHArg_K$ and $pHCit_K$ with $pHCit_{E/K\ His}$ and $pHCit_E$.

5-6: Future Directions

Since the stability of the possible trimer complexes were fairly similar, exchange experiments were unsuccessful. Preparing other peptides with different polar functional groups in the core would be a likely future direction for this project. With new polar interactions, the difference in the stability between the trimer complexes will hopefully increase, shedding light on the difference in T_m values necessary for successful exchange

experiments and possibly selective trimerization. This project is only limited by the amount of additional, plausible, polar functional groups.

5-7: Experimental Information

Circular Dichroism experiments were performed as explained in Chapter 3 and 4. All peptides were prepared and characterized as described in the previous chapters. Ni-NTA agarose beads were purchased as a 50% slurry from Qiagen.

Ni-NTA Affinity Analysis of Exchange Experiments. Ni-NTA agarose slurry (1 mL) is added to a 1.5 mL Eppendorf tube and centrifuged for 1 min, followed by supernatant removal (discarded). Initial peptide solution (1mL, 10μM total peptide concentration) is added to the beads. The Eppendorf tube is repeatedly inverted for 5 min and centrifuged for 30s, followed by supernatant removal (initial supernatant fraction). The beads are then washed with 1mL of PBS buffer (10mM phospahate, 150 mM NaCl, pH 7) repeated inversion for 30s, centrifugation for 30s, and supernatant removal (wash fraction). The exchange peptide (1mL, 3.3μM total peptide concentration) is then added to the beads. The Eppendorf tube is repeatedly inverted for 5 min and centrifuged for 30 s, followed by supernatant removal (second supernatant fraction). The beads are then washed with 1mL of PBS buffer (10mM phospahate, 150 mM NaCl, pH 7) repeated inversion for 30s, centrifugation for 30s, and supernatant removal (second wash fraction). All remaining bound peptide are eluted by addition of 6M Guanidine HCl with 250 mM imidazole in PBS buffer (10mM phospahate, 150 mM NaCl, pH 7) and further repeated inverting of the Eppendorf for 5 min and centrifuged for 30s, followed by supernatant removal (elution fraction). Each fraction is analyzed by reverse-phase HPLC: analytical

C18 column, linear gradients of solvent A (1% acetonitrile in water, 0.1% v/v CF_3CO_2H) and solvent B (10% water in acetonitrile, 0.07% v/v CF_3CO_2H).

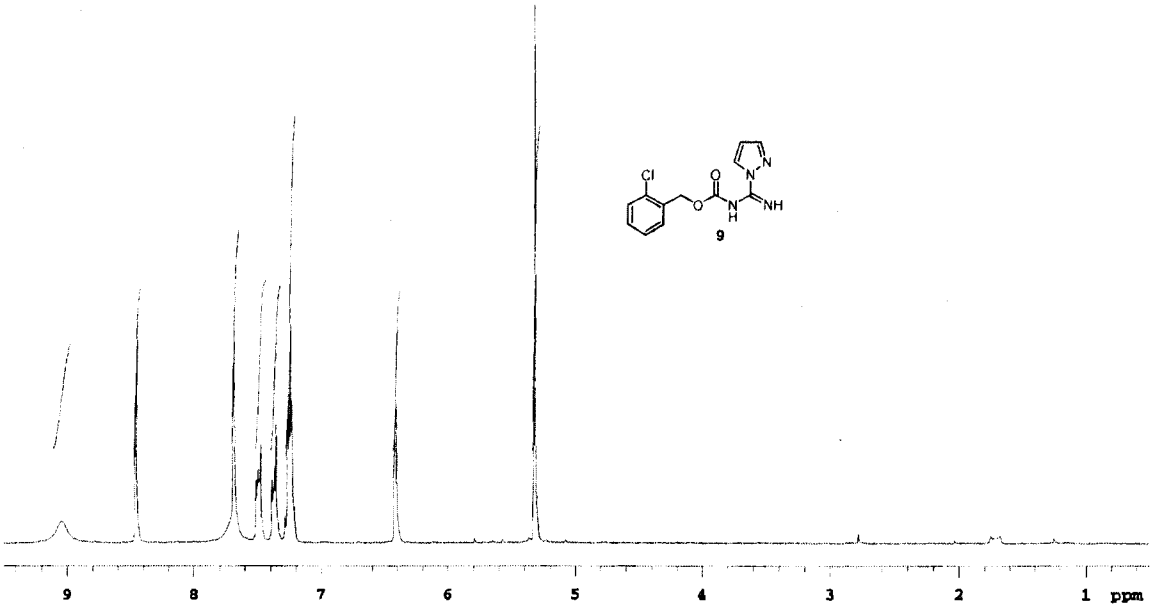
5-8: Literature Cited

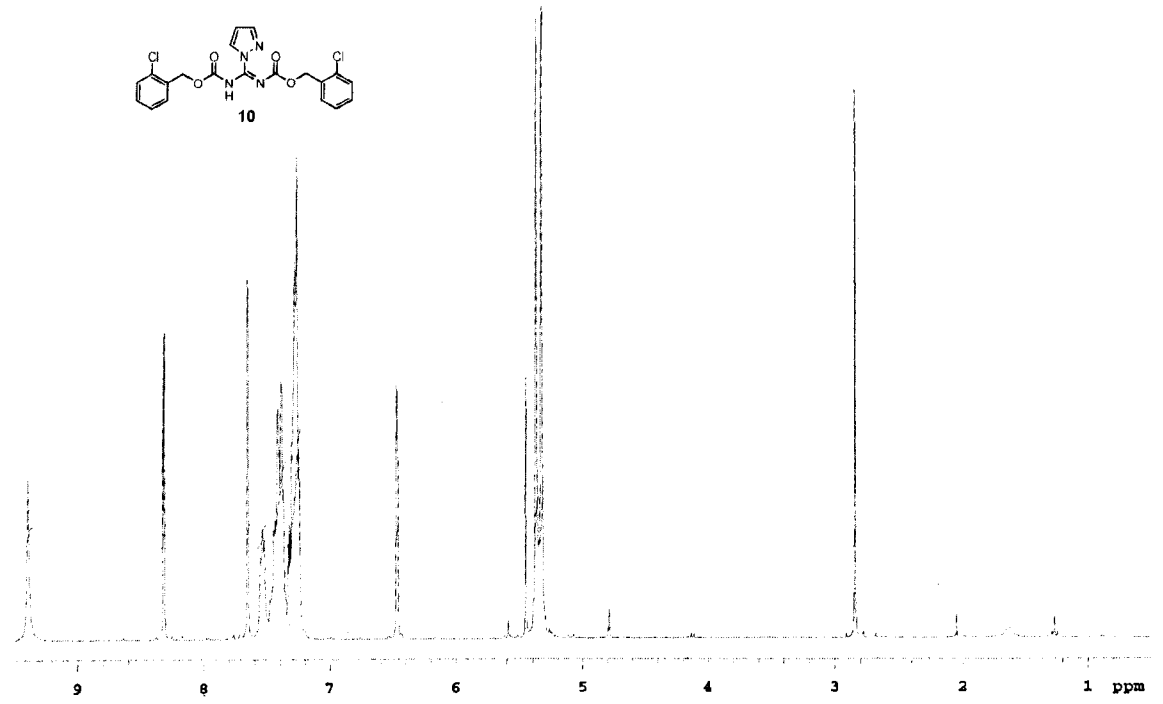
[1] (a) Schnarr, N.A.; Kennan, A.J. *Org Lett* **2005**, 7, 395-398. (b) Schnarr, N.A.; Kennan, A.J. *J. Am. Chem. Soc.* **2004**, 126, 14447-14451.

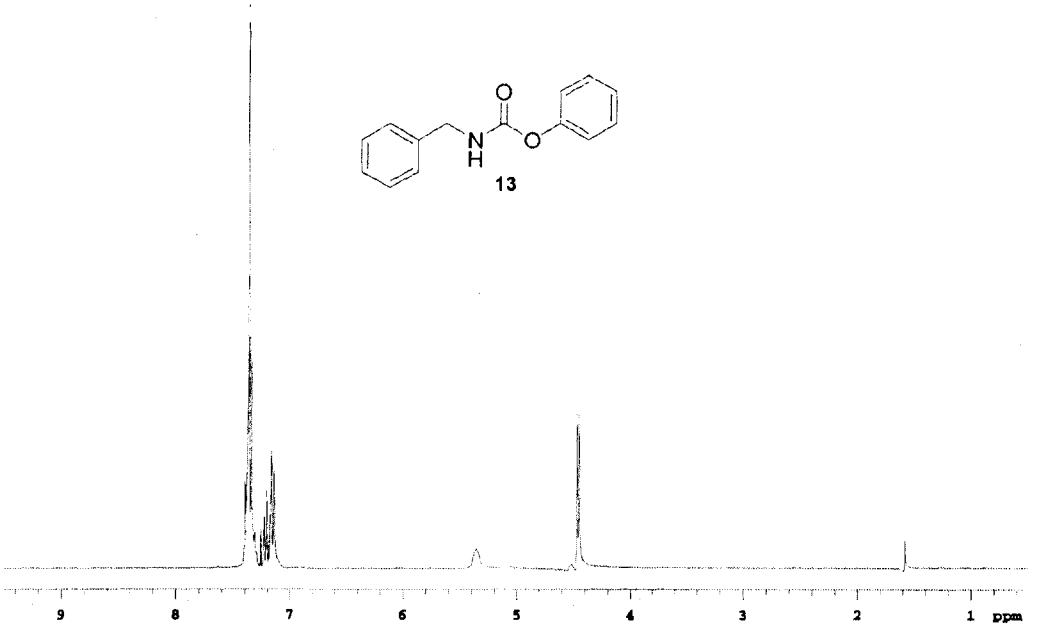
[2] Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T., *Nat. Struct. Bio.* **1996**, 3, 1011-1018.

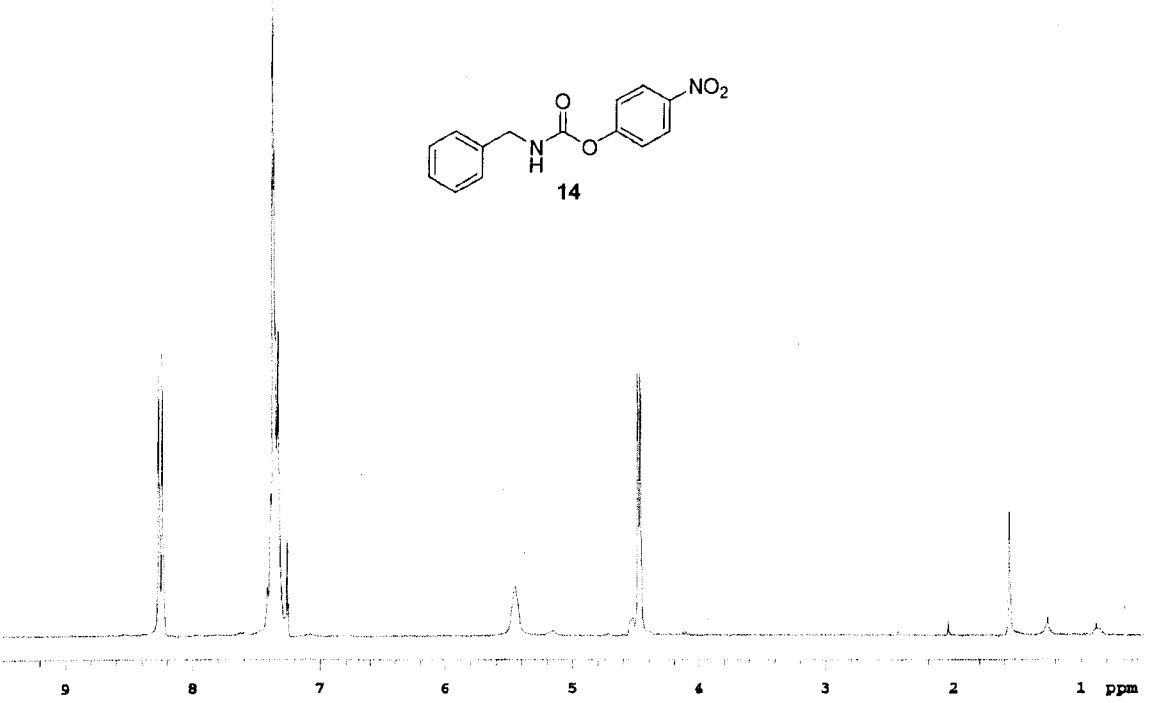
Appendix I

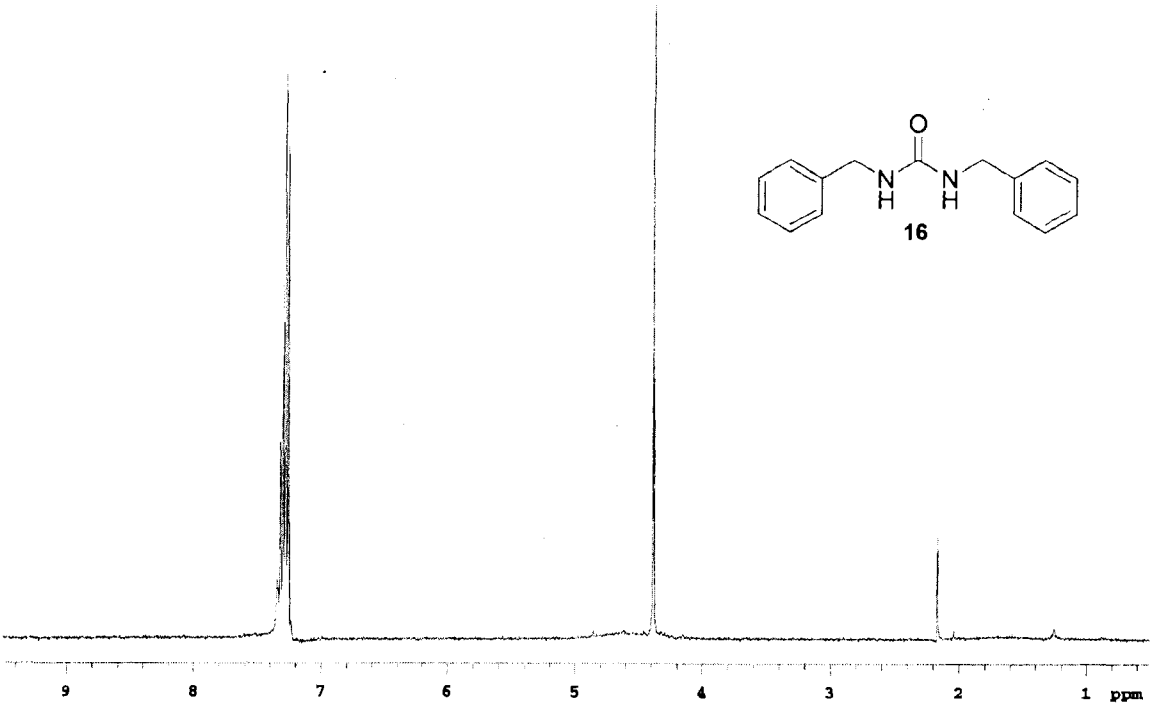
NMR Spectra

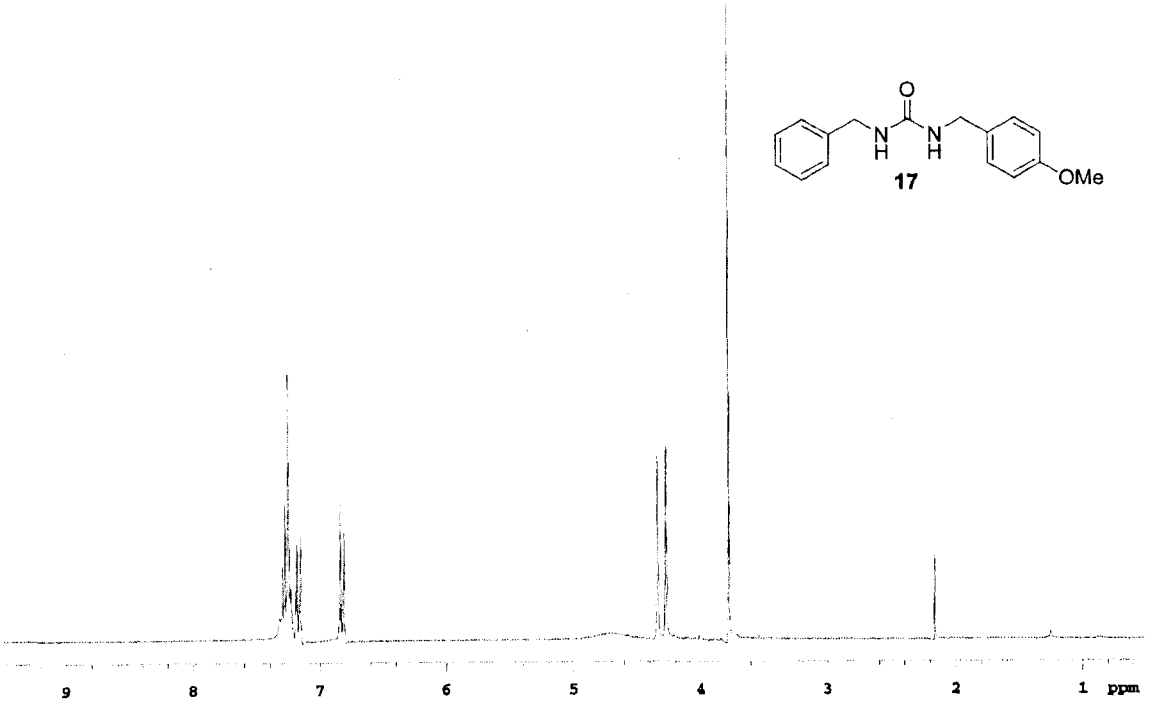

Cl
O
N
H
NH
N
9
9
8
7
6
5
4
3
2
1
ppm

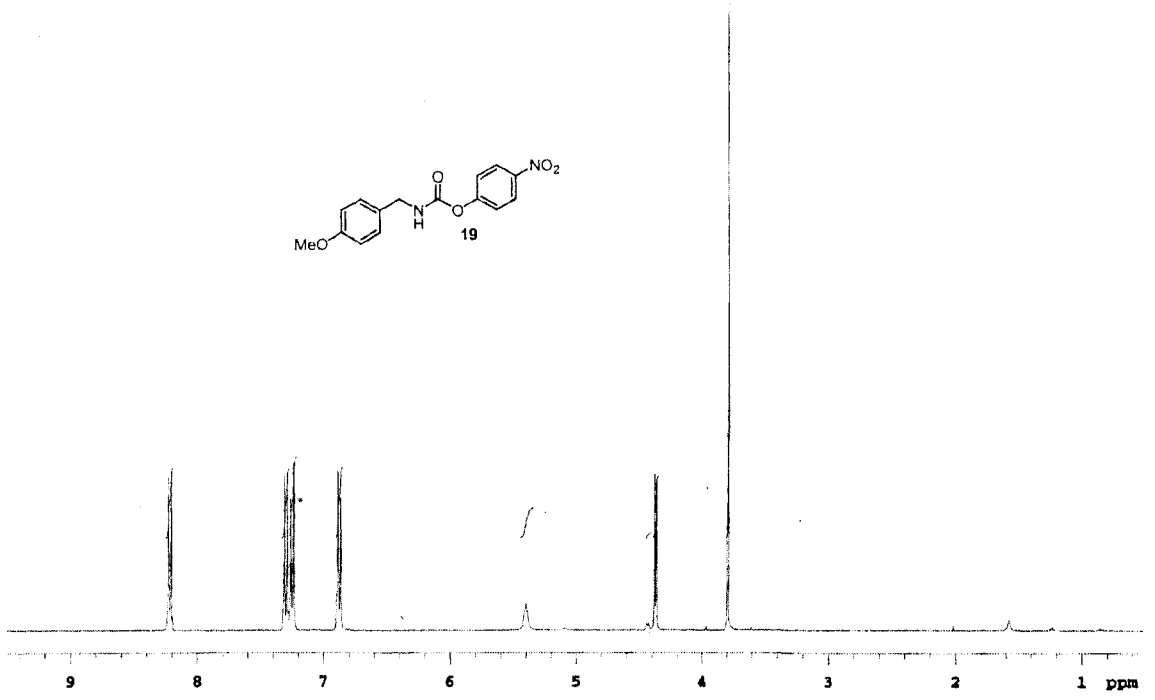

Cl
O
N
H
N
N
O
O
Cl
10
9
8
7
6
5
4
3
2
1
ppm


O
N
H
O
13
9
8
7
6
5
4
3
2
1
ppm


NO2
O
N
H
O
14
9
8
7
6
5
4
3
2
1
ppm


O
N
H
N
H
16
9
8
7
6
5
4
3
2
1
ppm

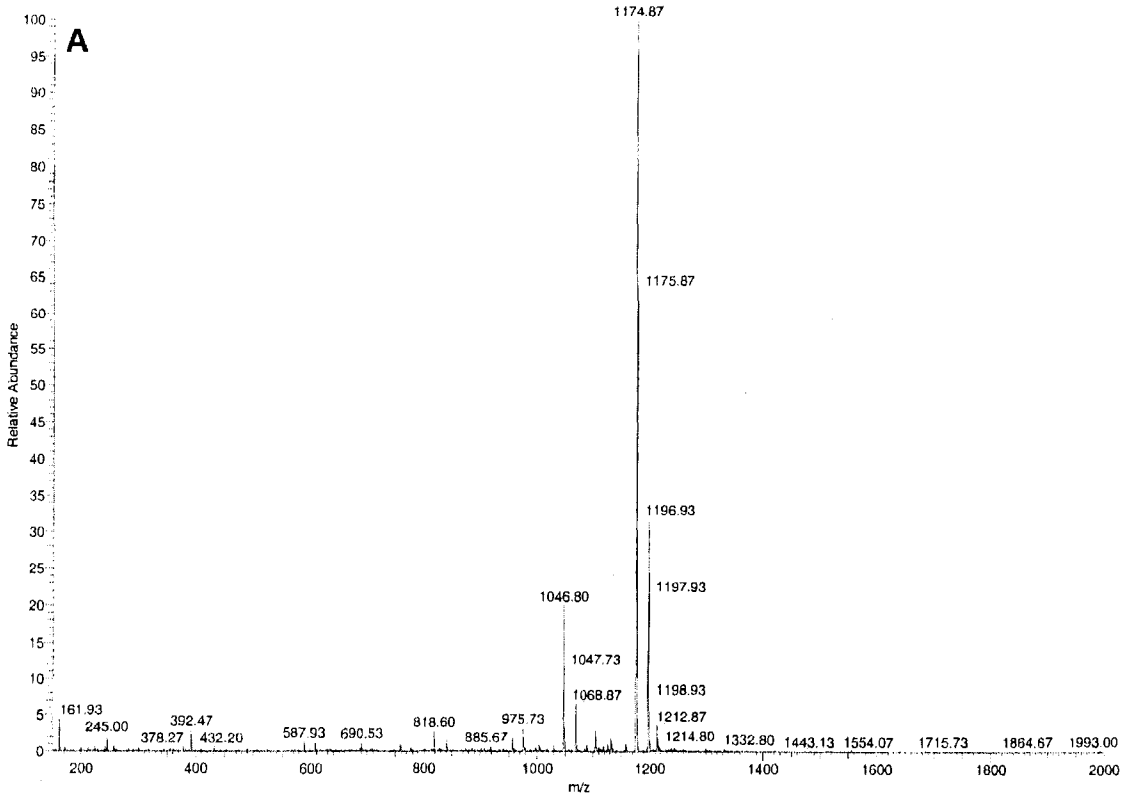

O
N
H
N
H
17
OMe
9
8
7
6
5
4
3
2
1
ppm

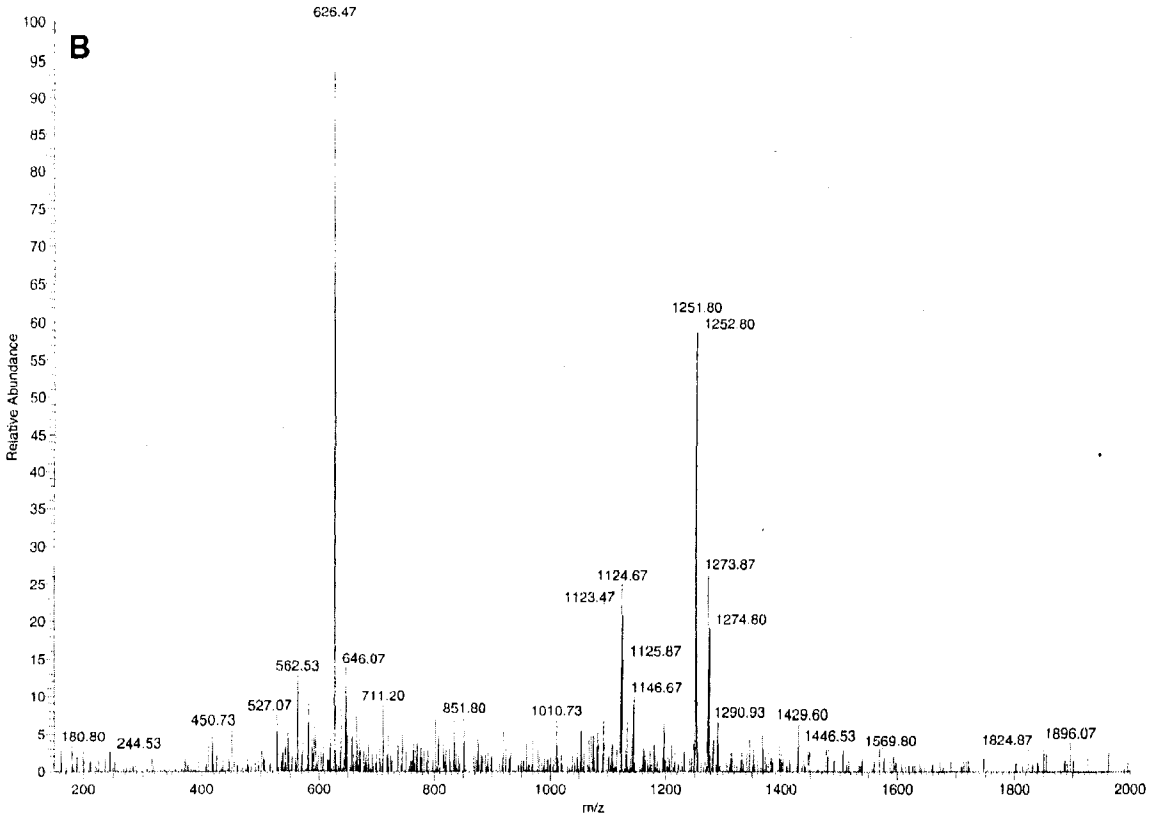

MeO
NO2
19
9
8
7
6
5
4
3
2
1
ppm

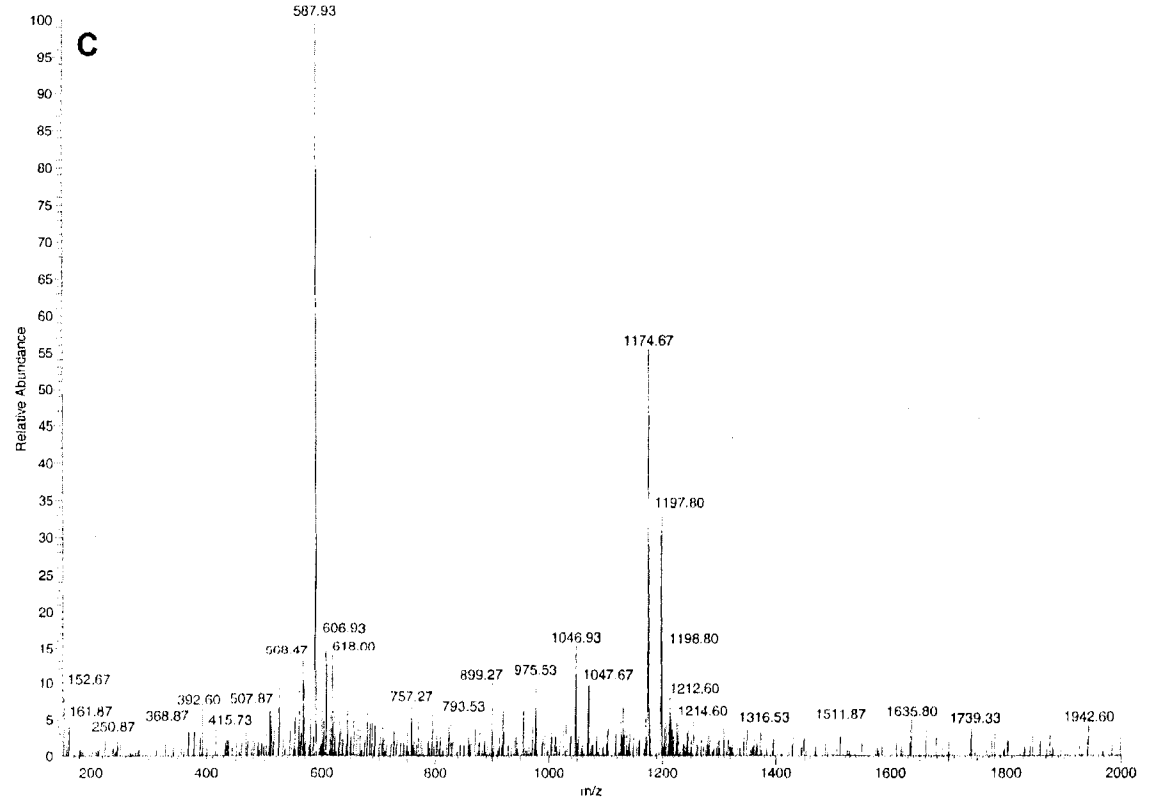
Appendix II

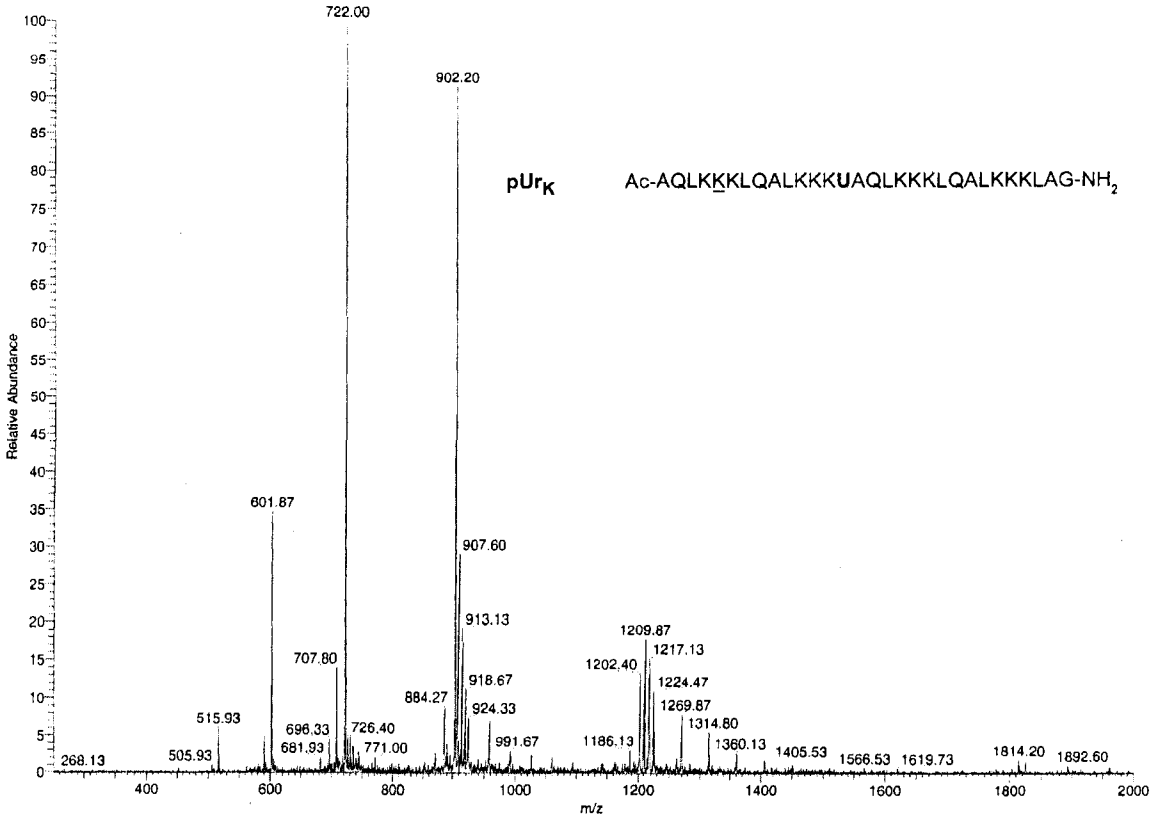
Peptide Characterization

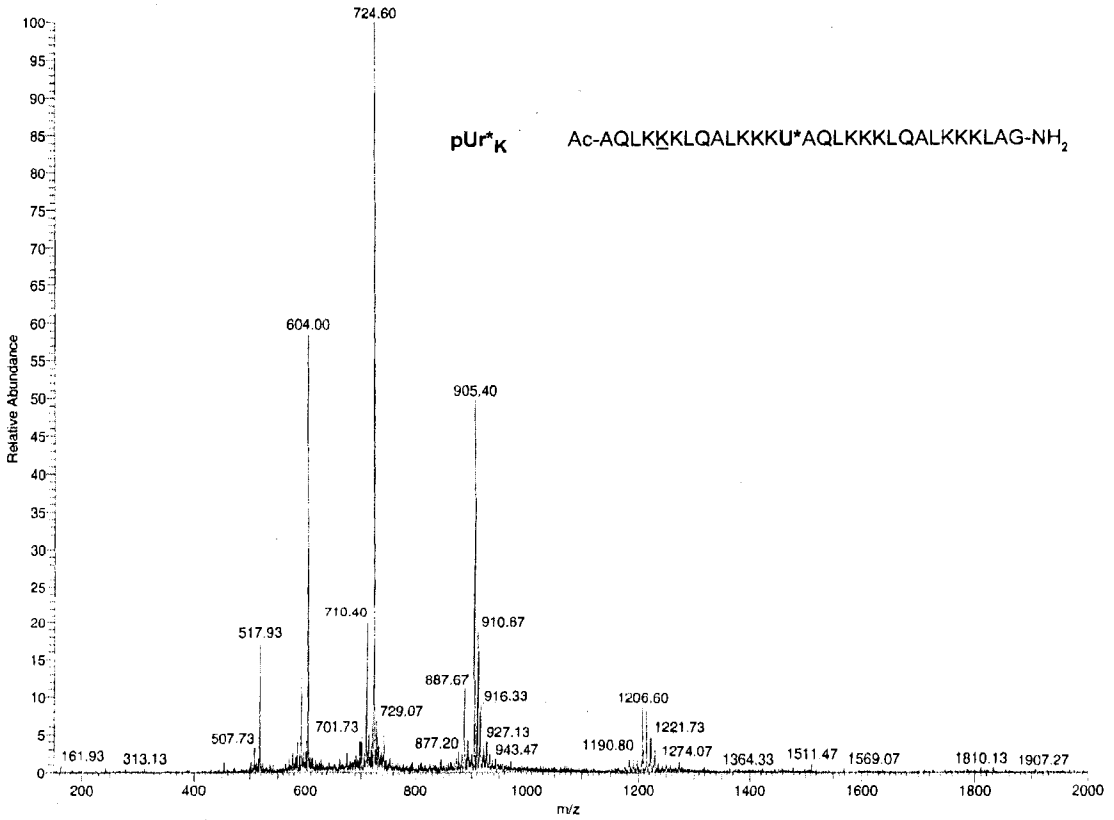
Figure A-1: Electrospray mass spec data for various preparation of crude **18**: (A) via route a (authentic Cit); (B) via route b (phenyl carbamate); (C) via route c (*p*-nitrophenyl carbamate). In each case the expected mass is 1174.4, and some [M+Na$^+$] peaks are also observed. In (B) the observed mass (1251.8) is 77 mass units higher than expected.

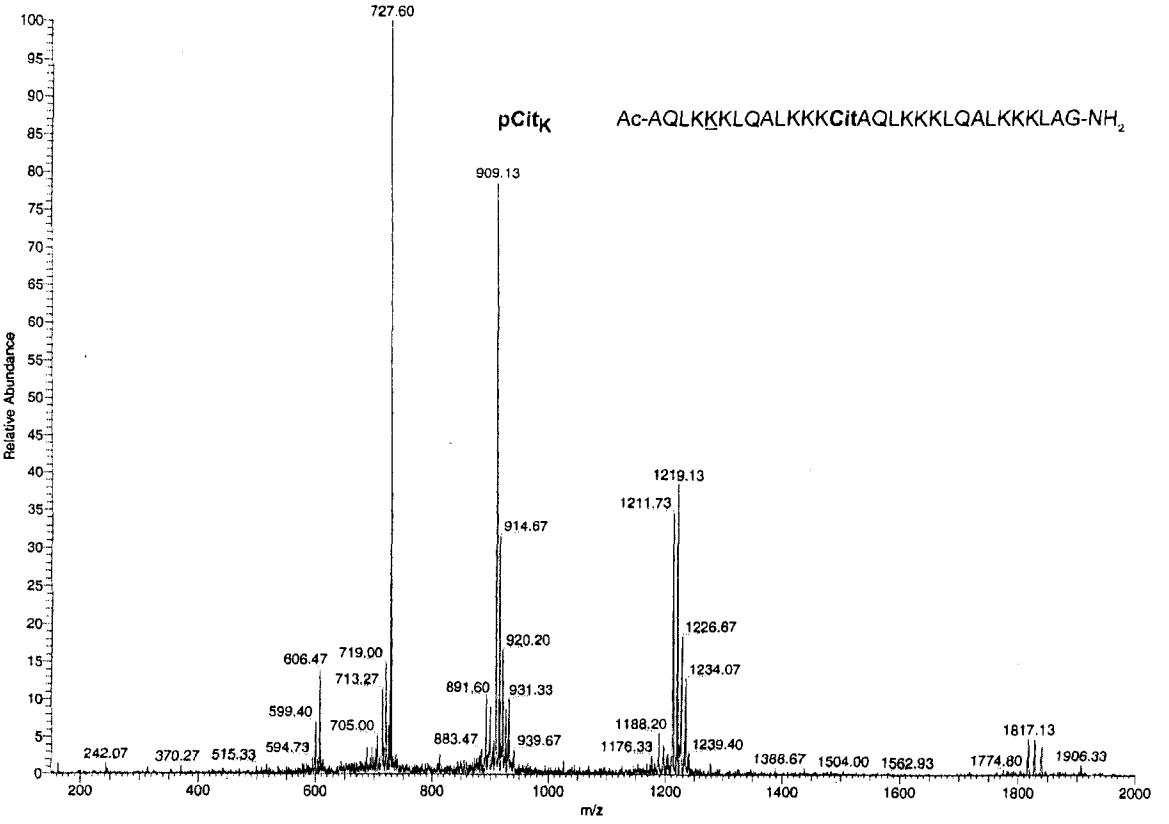


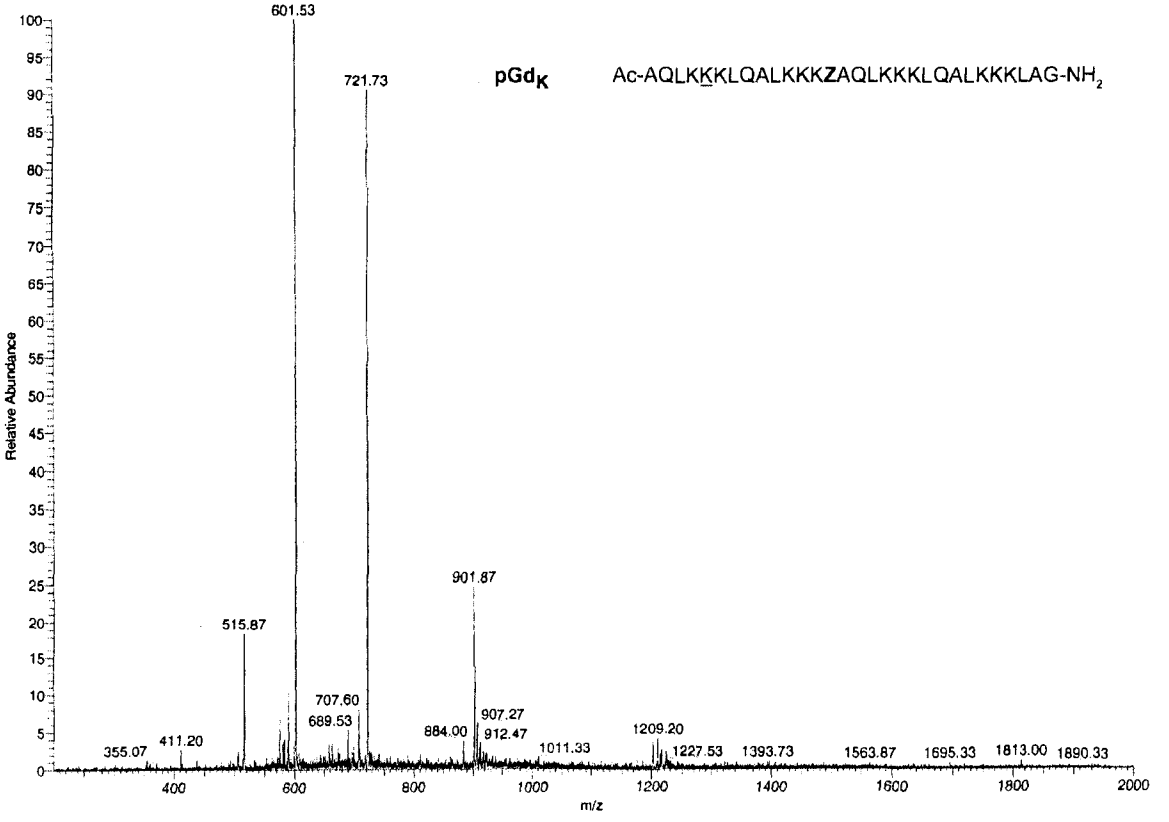

B
626.47
1251.80
1252.80
1124.67
1123.47
1273.87
1274.80
1125.87
1146.67
562.53
646.07
711.20
527.07
450.73
851.80
1010.73
1290.93
1429.60
1446.53
1569.80
1824.87
1896.07
180.80
244.53
Relative Abundance
m/z

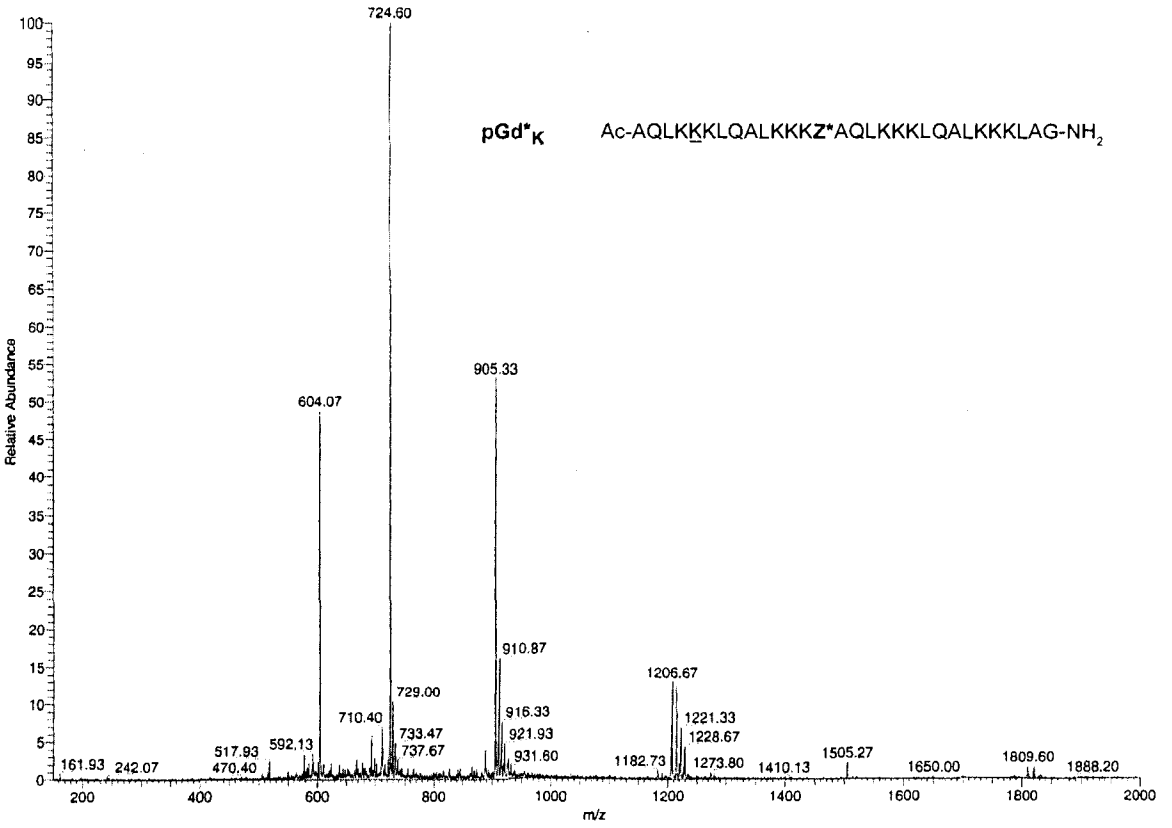

C
587.93
1174.67
1197.80
606.93
618.00
568.47
1046.93
1198.80
1047.67
899.27
975.53
757.27
793.53
1212.60
1214.60
152.67
161.87
250.87
368.87
392.60
507.87
415.73
1316.53
1511.87
1635.80
1739.33
1942.60
Relative Abundance
m/z

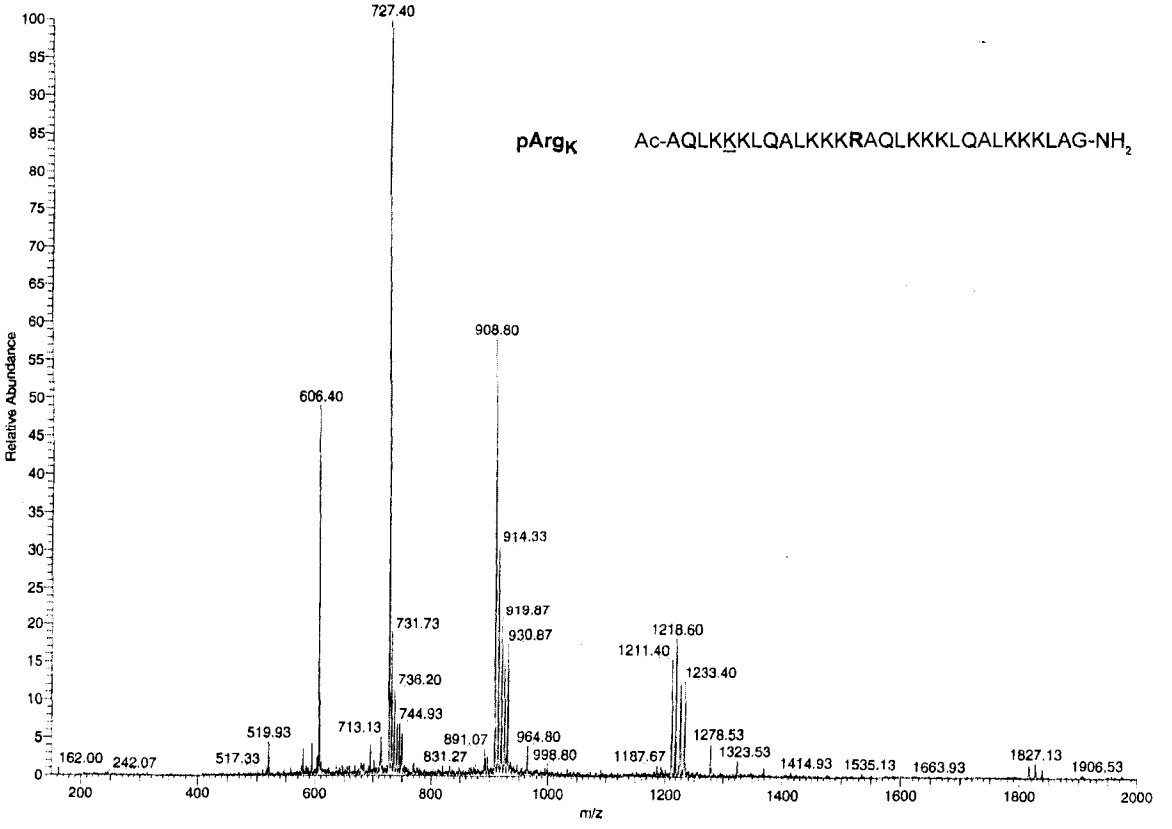

722.00
902.20
pUr$_K$ Ac-AQLKKLQALKKKUAQLKKKLQALKKKLAG-NH$_2$
Relative Abundance
601.87
907.60
913.13
1209.87
1217.13
707.80
1202.40
1224.47
918.67
884.27
1269.87
924.33
515.93
1314.80
696.33
726.40
681.93
771.00
991.67
1186.13
1360.13
1405.53
268.13
505.93
1566.53
1619.73
1814.20
1892.60
m/z

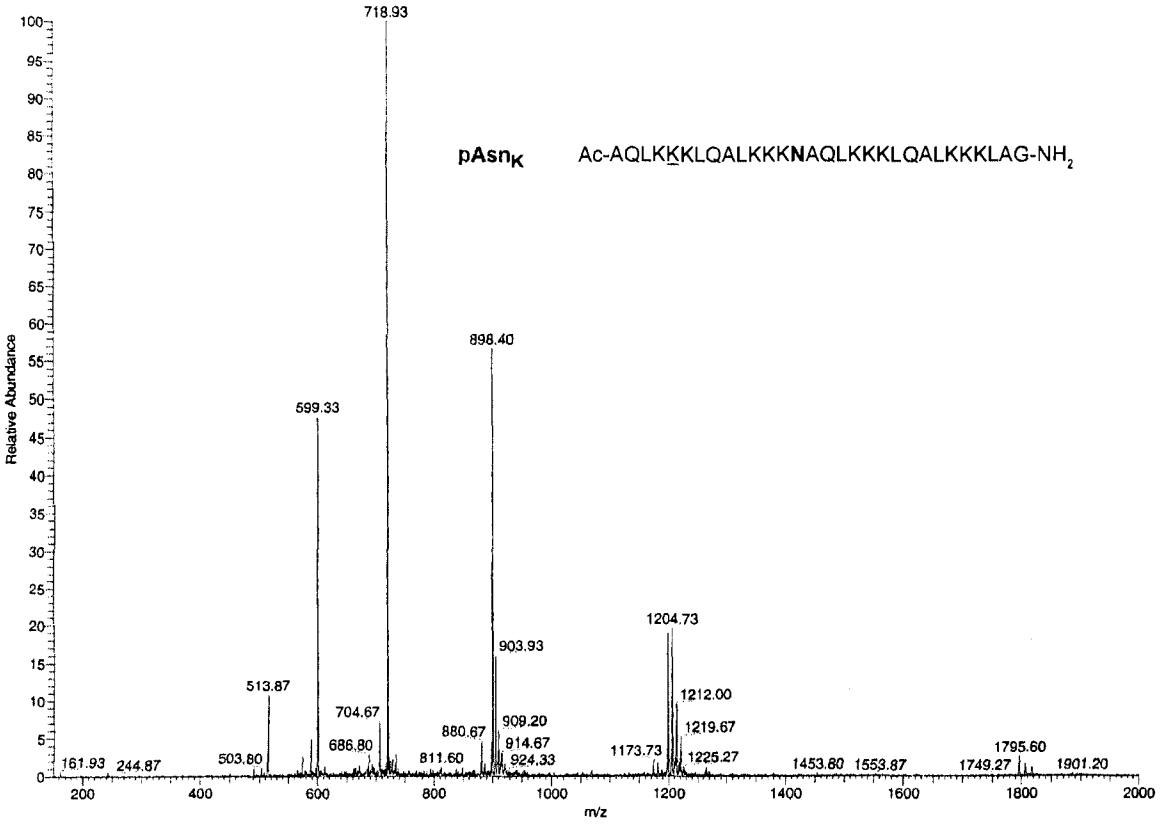

724.60
pUr*$_K$ Ac-AQLKKLQALKKKU*AQLKKKLQALKKKLAG-NH$_2$
Relative Abundance
604.00
905.40
710.40
910.67
517.93
887.67
916.33
729.07
1206.60
701.73
927.13
1221.73
507.73
877.20
943.47
1190.80
1274.07
161.93
313.13
1364.33
1511.47
1569.07
1810.13
1907.27
m/z

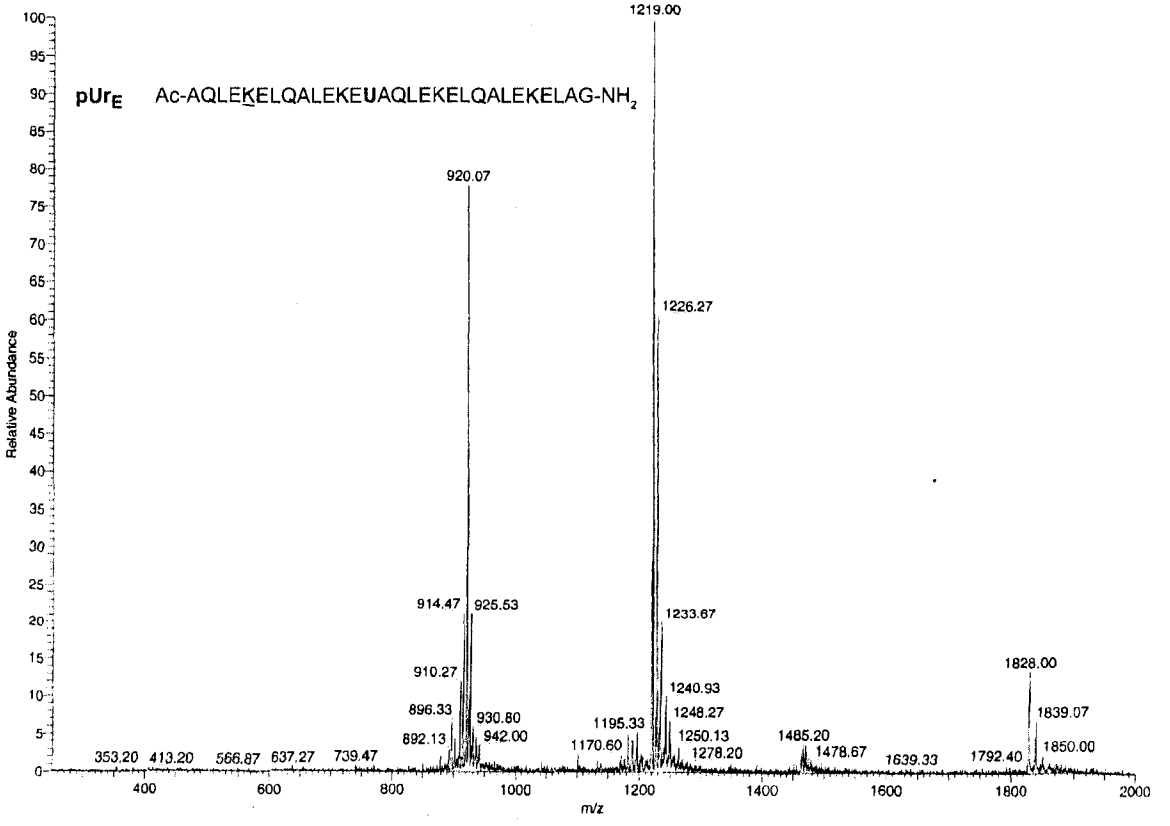

pCit$_K$
Ac-AQLKKKLQALKKKCitAQLKKKLQALKKKLAG-NH$_2$
Relative Abundance
m/z
727.60
909.13
1219.13
1211.73
914.67
1226.67
920.20
606.47
719.00
1234.07
713.27
891.60
931.33
599.40
705.00
1188.20
1817.13
883.47
939.67
1176.33
1239.40
242.07
370.27
515.33
594.73
1366.67
1504.00
1562.93
1774.80
1906.33

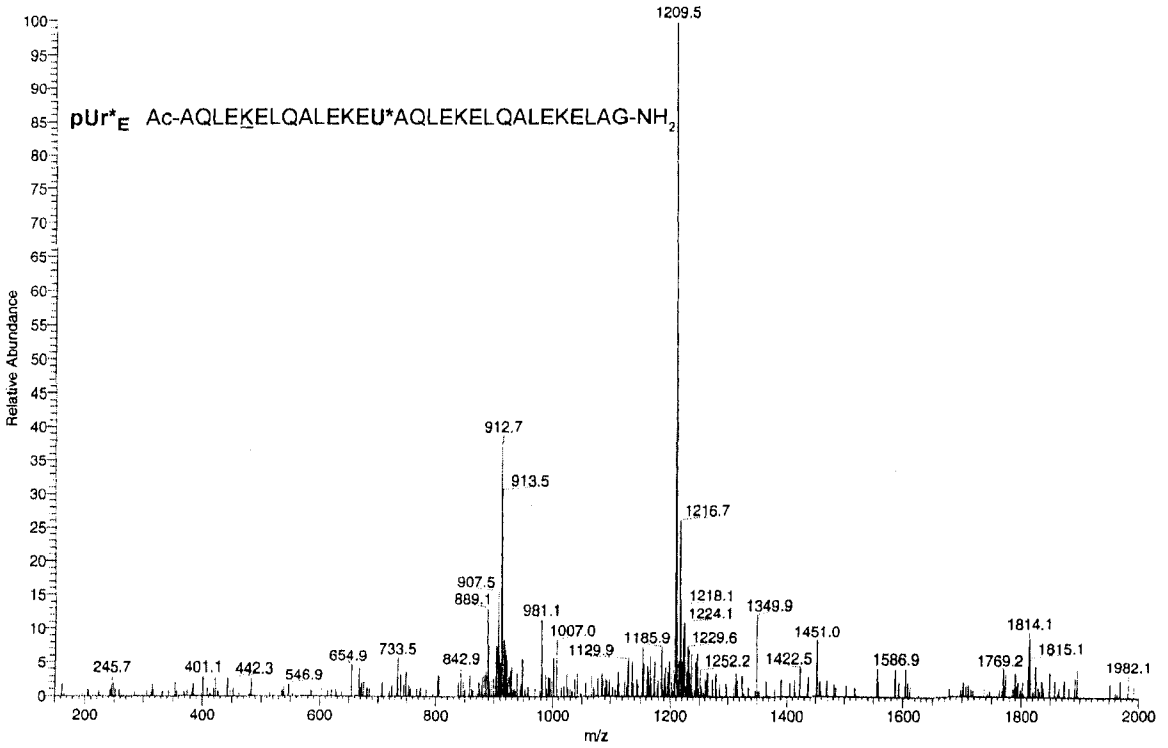

pGd$_K$
Ac-AQLKKKLQALKKKZAQLKKKLQALKKKLAG-NH$_2$
Relative Abundance
m/z
601.53
721.73
901.87
515.87
707.60
689.53
907.27
884.00
912.47
1209.20
355.07
411.20
1011.33
1227.53
1393.73
1563.87
1695.33
1813.00
1890.33

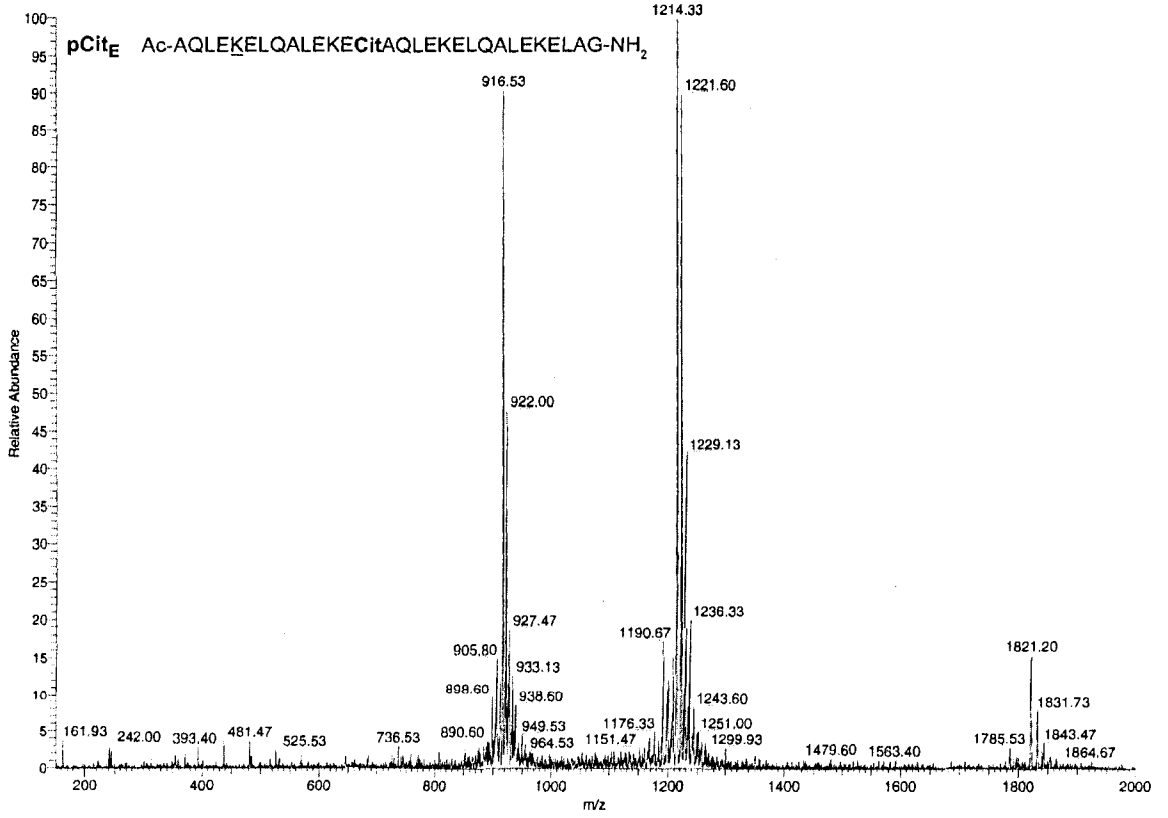

pGd*K
Ac-AQLKKKLQALKKKZ*AQLKKKLQALKKKLAG-NH2
Relative Abundance
m/z
724.60
905.33
604.07
910.87
1206.67
729.00
710.40
916.33
1221.33
733.47
921.93
1228.67
737.67
931.60
517.93
592.13
470.40
161.93
242.07
1182.73
1273.80
1410.13
1505.27
1650.00
1809.60
1888.20

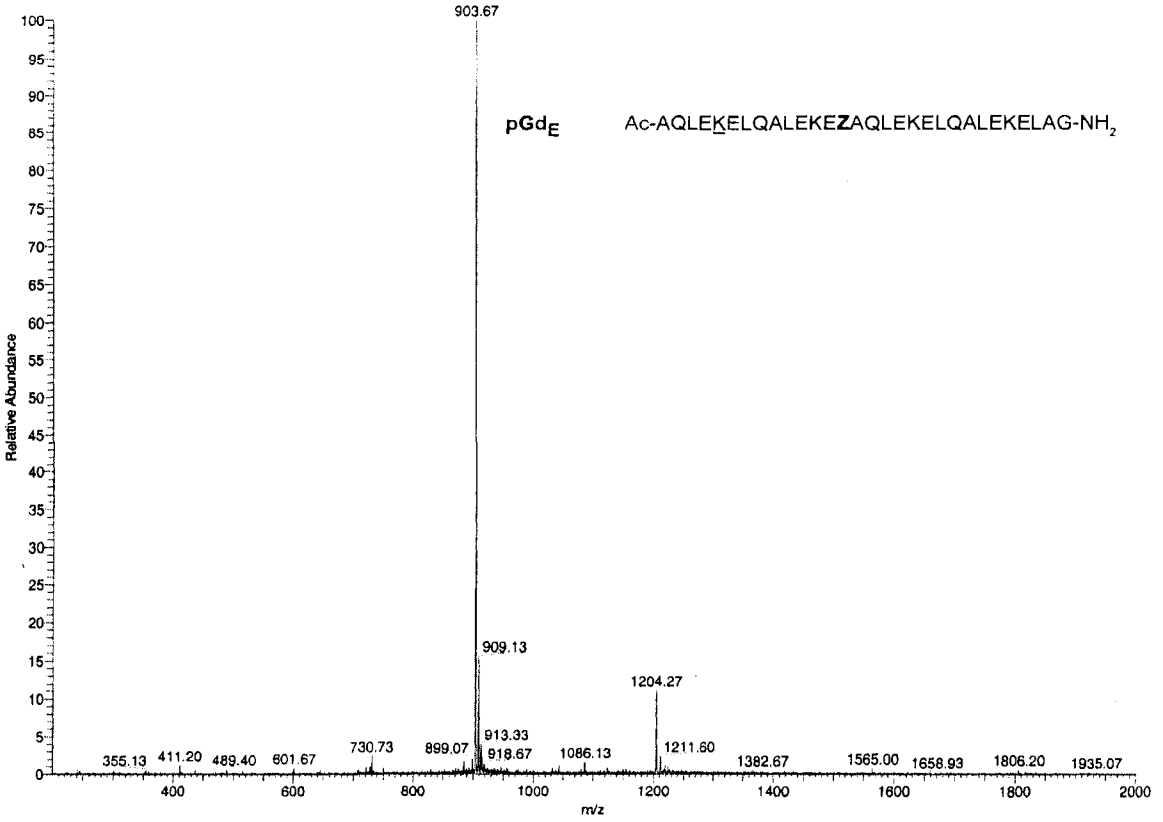

pArgK
Ac-AQLKKKLQALKKKRAQLKKKLQALKKKLAG-NH2
Relative Abundance
m/z
727.40
908.80
606.40
914.33
919.87
731.73
930.87
1218.60
1211.40
736.20
1233.40
744.93
519.93
713.13
891.07
964.80
1278.53
831.27
998.80
1187.67
1323.53
517.33
162.00
242.07
1414.93
1535.13
1663.93
1827.13
1906.53

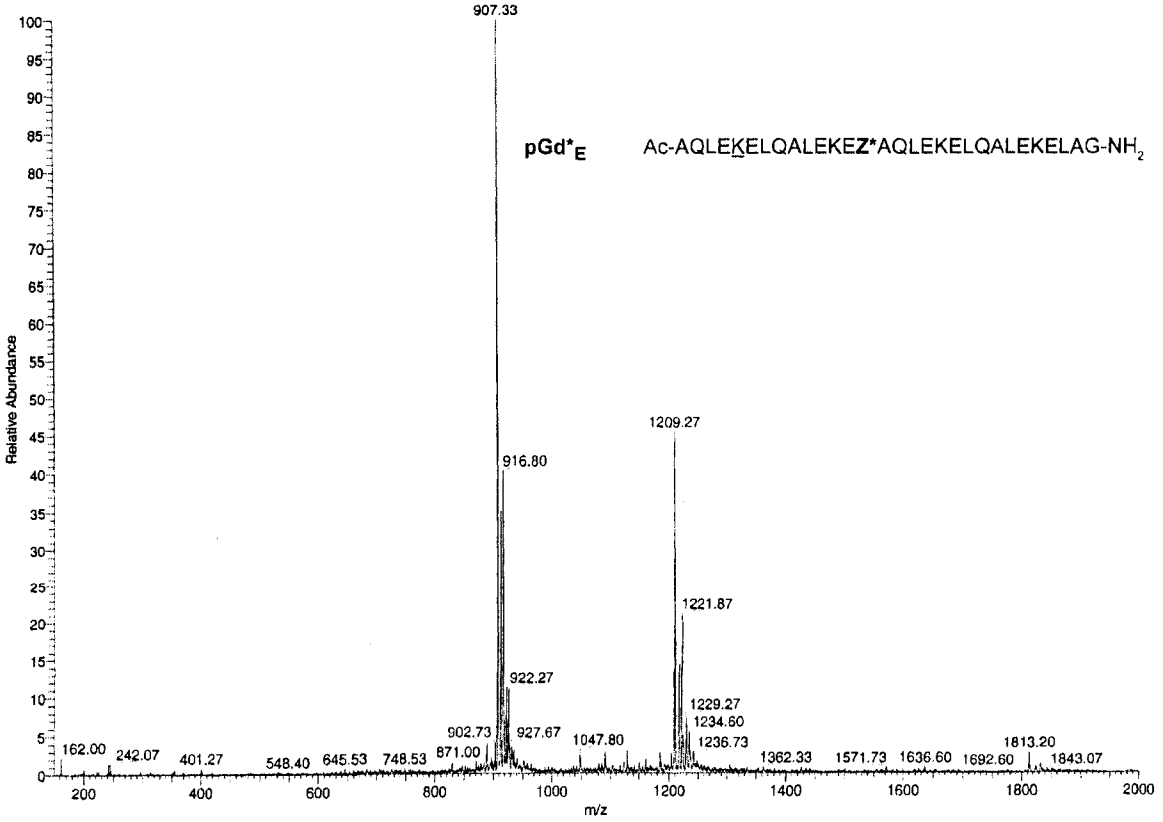

pAsn$_K$ Ac-AQLKKLQALKKKNAQLKKKLQALKKKLAG-NH$_2$
Relative Abundance
m/z
718.93
898.40
599.33
1204.73
903.93
513.87
1212.00
704.67
909.20
880.67
1219.67
686.80
914.67
1795.60
161.93
244.87
503.80
811.60
924.33
1173.73
1225.27
1453.60
1553.87
1749.27
1901.20

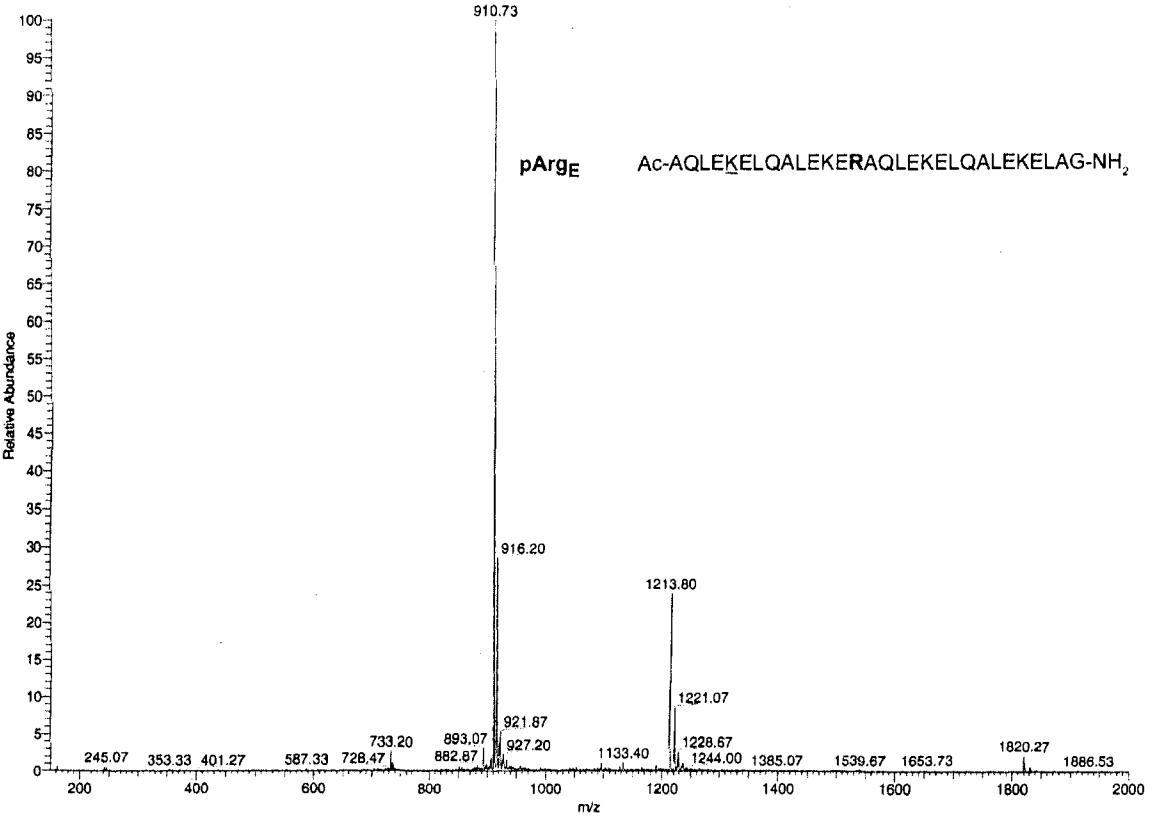

pUr$_E$ Ac-AQLEKELQALEKEUAQLEKELQALEKELAG-NH$_2$
Relative Abundance
m/z
1219.00
920.07
1226.27
914.47
925.53
1233.67
910.27
1828.00
1240.93
896.33
930.80
1248.27
1839.07
892.13
942.00
1195.33
1250.13
1485.20
1850.00
353.20
413.20
566.87
637.27
739.47
1170.60
1278.20
1478.67
1639.33
1792.40

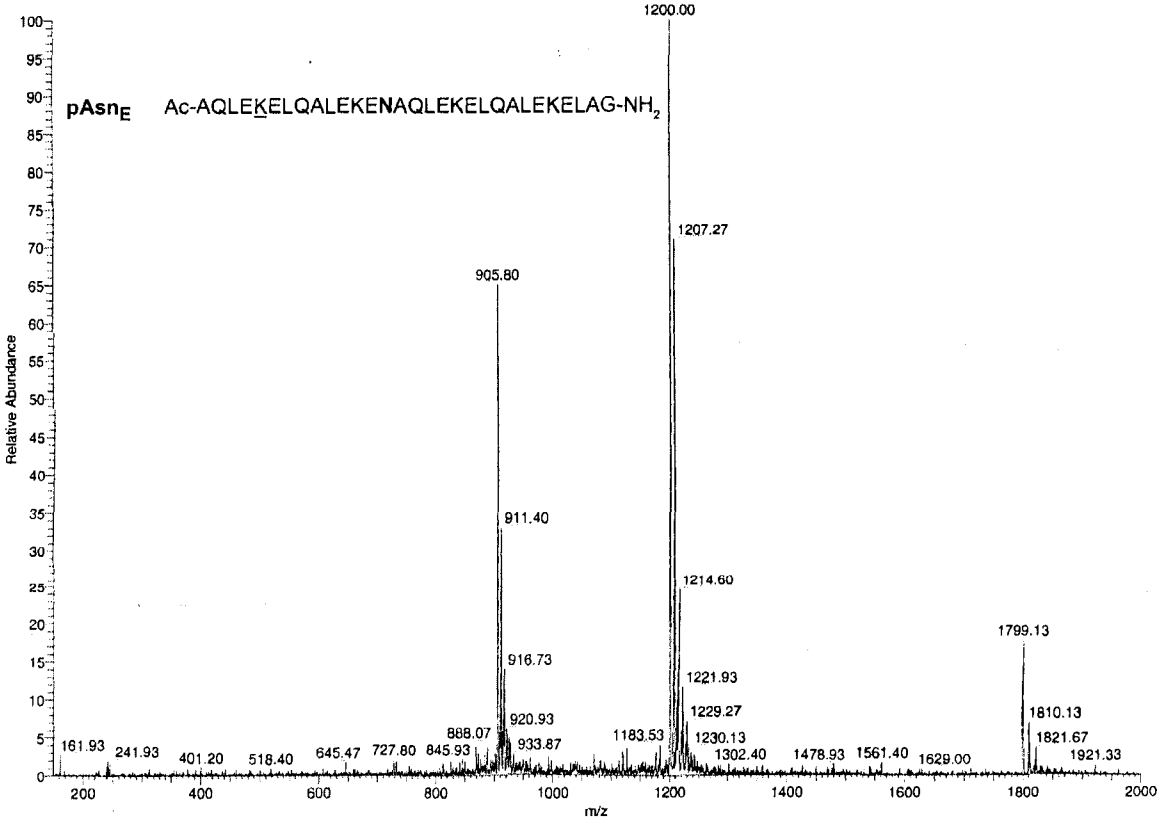

pUr*E Ac-AQLEKELQALEKEU*AQLEKELQALEKELAG-NH2
Relative Abundance
m/z
1209.5
912.7
913.5
1216.7
907.5
889.1
981.1
1007.0
1129.9
1185.9
1218.1
1224.1
1229.6
1252.2
1349.9
1422.5
1451.0
1586.9
1814.1
1815.1
1769.2
1982.1
245.7
401.1
442.3
546.9
654.9
733.5
842.9

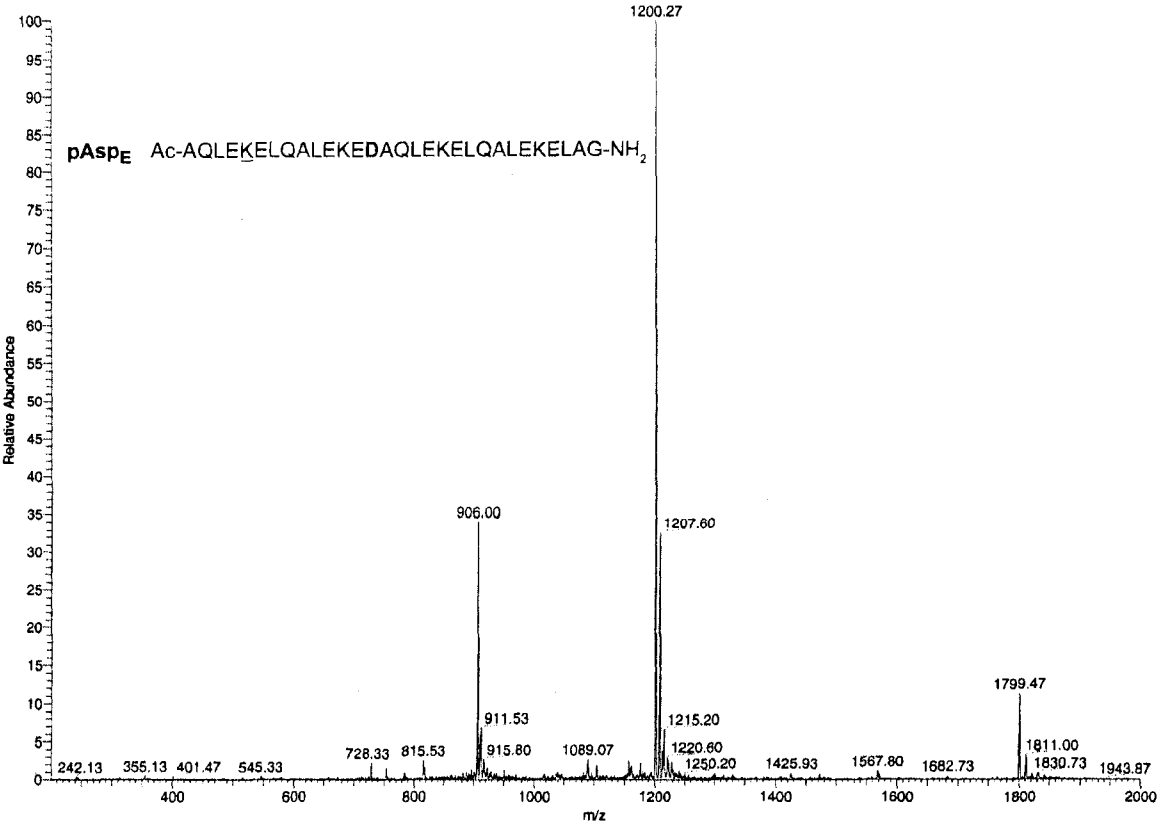

pCitE Ac-AQLEKELQALEKECitAQLEKELQALEKELAG-NH2
Relative Abundance
m/z
1214.33
916.53
1221.60
922.00
1229.13
927.47
1236.33
1190.67
905.80
933.13
898.60
938.60
1243.60
1251.00
1299.93
1821.20
1831.73
1843.47
1864.67
1785.53
1479.60
1563.40
1176.33
1151.47
949.53
964.53
890.60
736.53
525.53
481.47
393.40
242.00
161.93

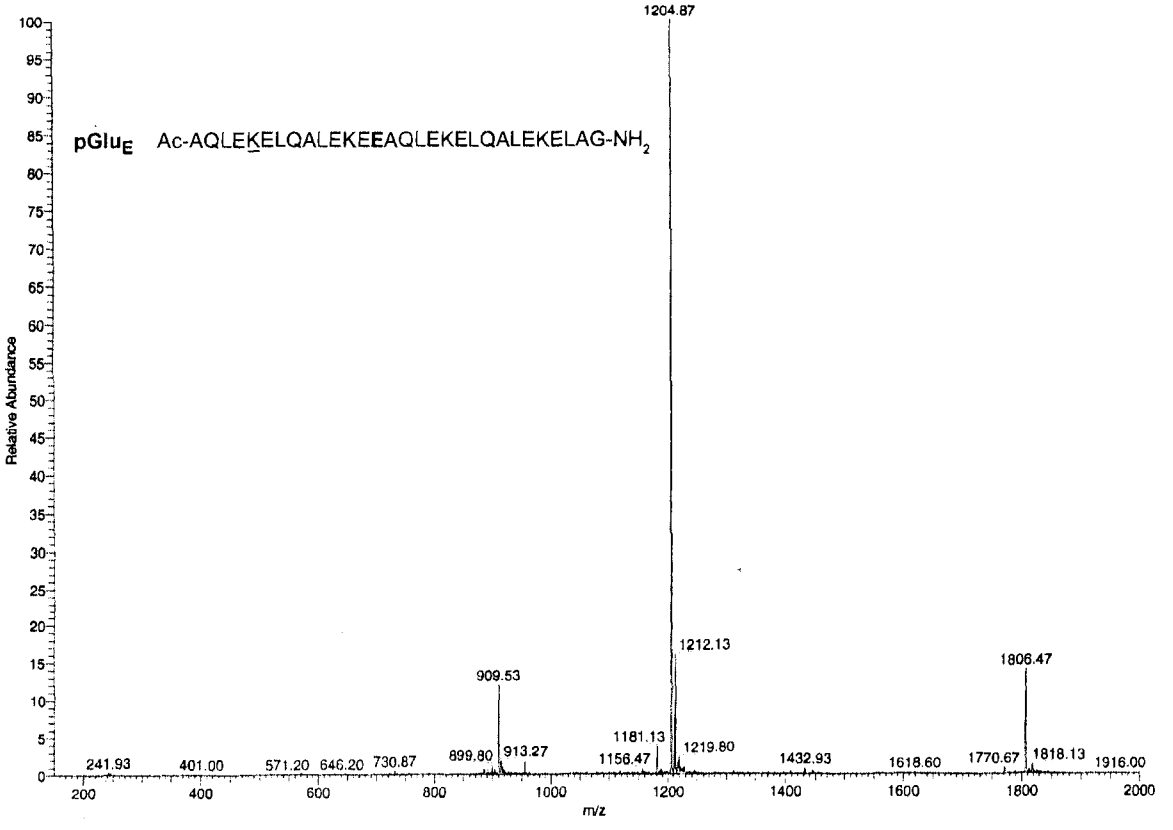

903.67
pGd$_E$
Ac-AQLEKELQALEKEZAQLEKELQALEKELAG-NH$_2$
Relative Abundance
909.13
1204.27
913.33
355.13
411.20
489.40
601.67
730.73
899.07
918.67
1086.13
1211.60
1382.67
1565.00
1658.93
1806.20
1935.07
m/z


907.33
pGd*$_E$
Ac-AQLEKELQALEKEZ*AQLEKELQALEKELAG-NH$_2$
Relative Abundance
1209.27
916.80
1221.87
922.27
1229.27
1234.60
1236.73
902.73
927.67
1047.80
871.00
162.00
242.07
401.27
548.40
645.53
748.53
1362.33
1571.73
1636.60
1692.60
1813.20
1843.07
m/z


pArg$_E$ Ac-AQLE$\underline{K}$ELQALEKE**R**AQLEKELQALEKELAG-NH$_2$
Relative Abundance
m/z
910.73
916.20
1213.80
1221.07
921.87
927.20
733.20
893.07
882.87
728.47
587.33
401.27
353.33
245.07
1133.40
1228.67
1244.00
1385.07
1539.67
1653.73
1820.27
1886.53


pAsn$_E$ Ac-AQLE$\underline{K}$ELQALEKE**N**AQLEKELQALEKELAG-NH$_2$
Relative Abundance
m/z
1200.00
1207.27
905.80
911.40
1214.60
916.73
1221.93
1229.27
920.93
888.07
933.87
1183.53
1230.13
1302.40
1478.93
1561.40
1629.00
1799.13
1810.13
1821.67
1921.33
161.93
241.93
401.20
518.40
645.47
727.80
845.93

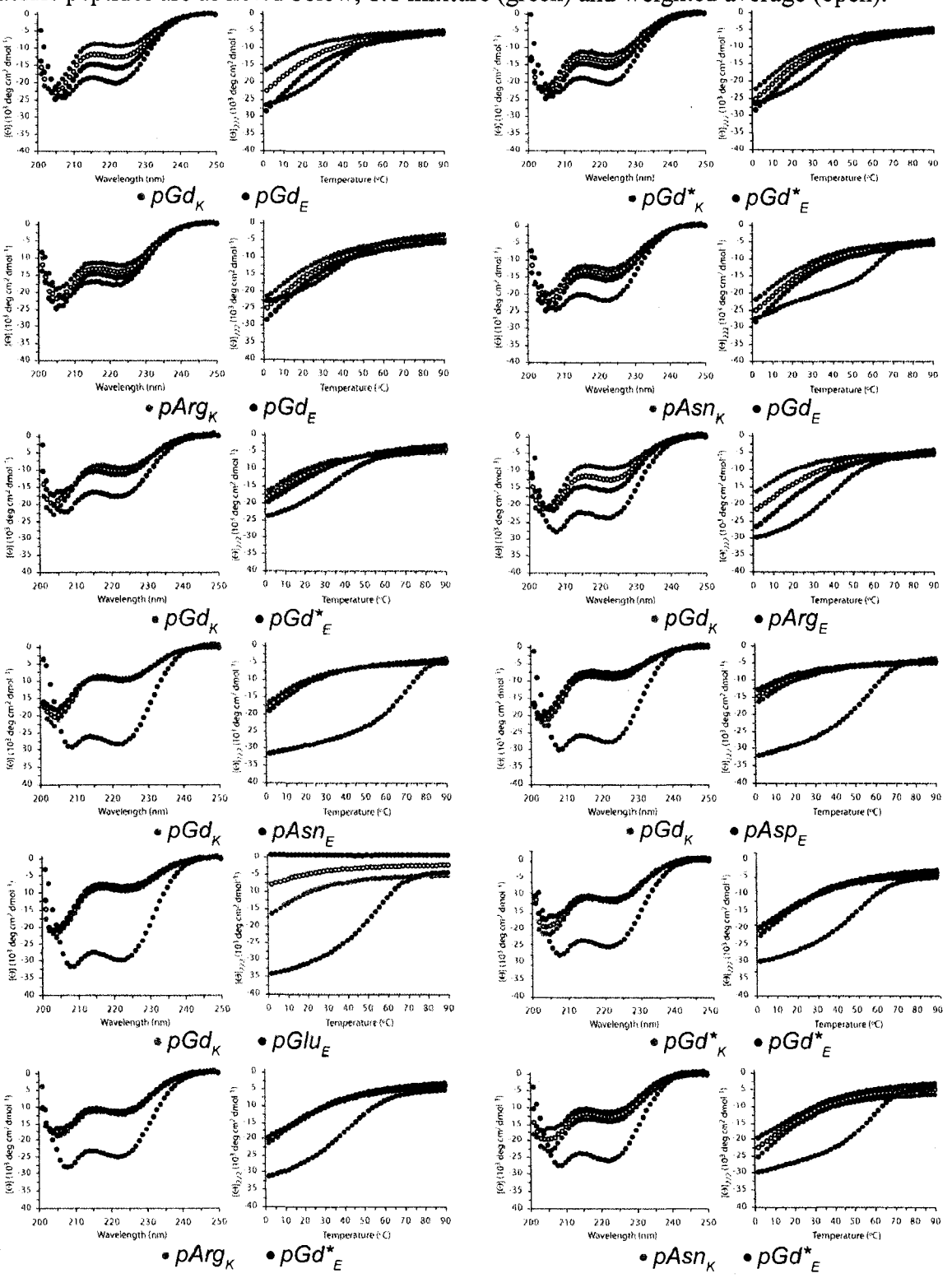

pAsp$_E$ Ac-AQLEKELQALEKEDAQLEKELQALEKELAG-NH$_2$
Relative Abundance
m/z
1200.27
906.00
1207.60
1799.47
911.53
1215.20
1220.60
1811.00
1830.73
1943.87
242.13
355.13
401.47
545.33
728.33
815.53
915.80
1089.07
1250.20
1425.93
1567.80
1682.73

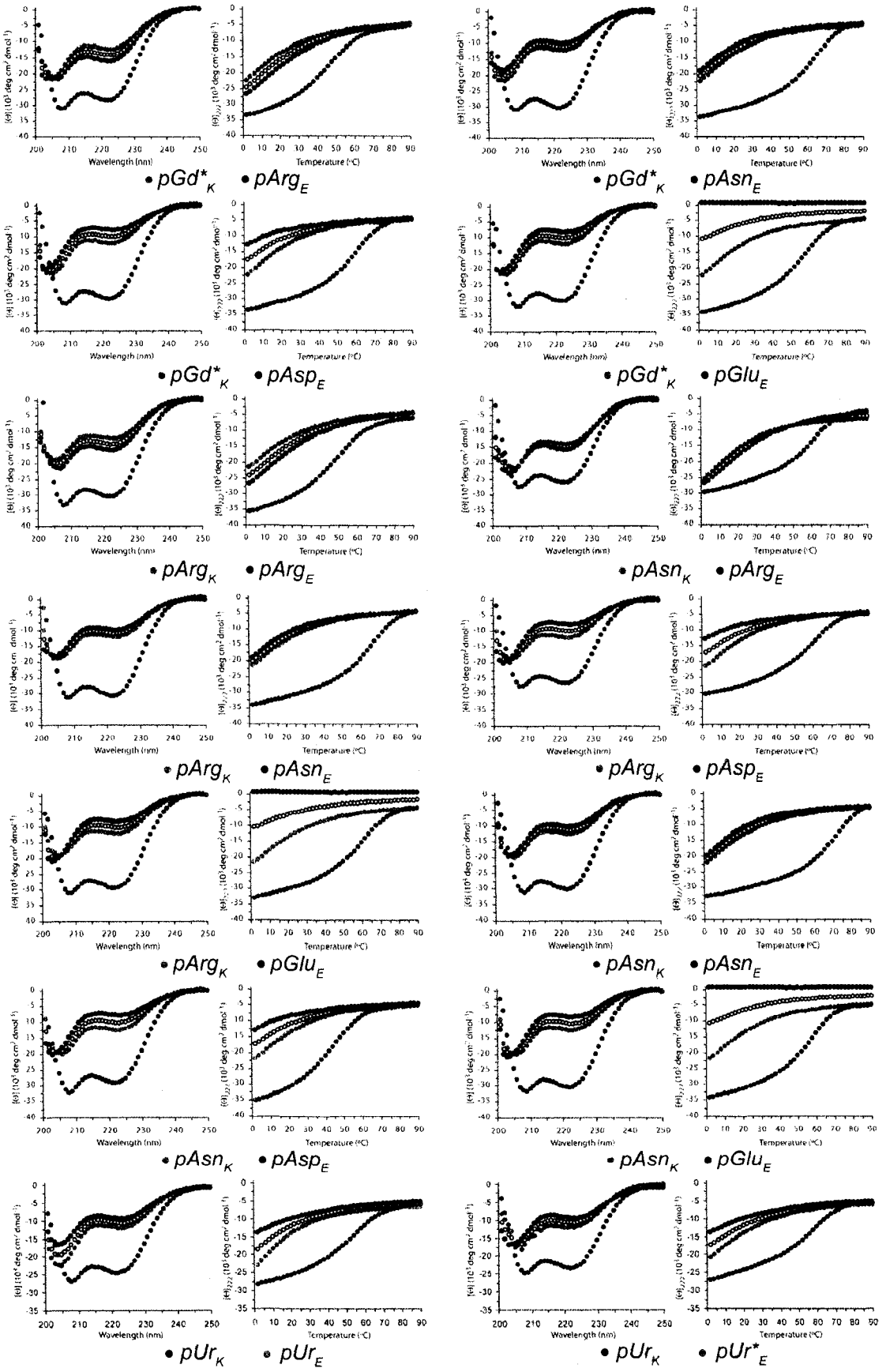

pGlu$_E$ Ac-AQLEKELQALEKEEAQLEKELQALEKELAG-NH$_2$
Relative Abundance
m/z
1204.87
909.53
1212.13
1806.47
1181.13
1219.80
1818.13
1916.00
241.93
401.00
571.20
646.20
730.87
899.80
913.27
1156.47
1432.93
1618.60
1770.67

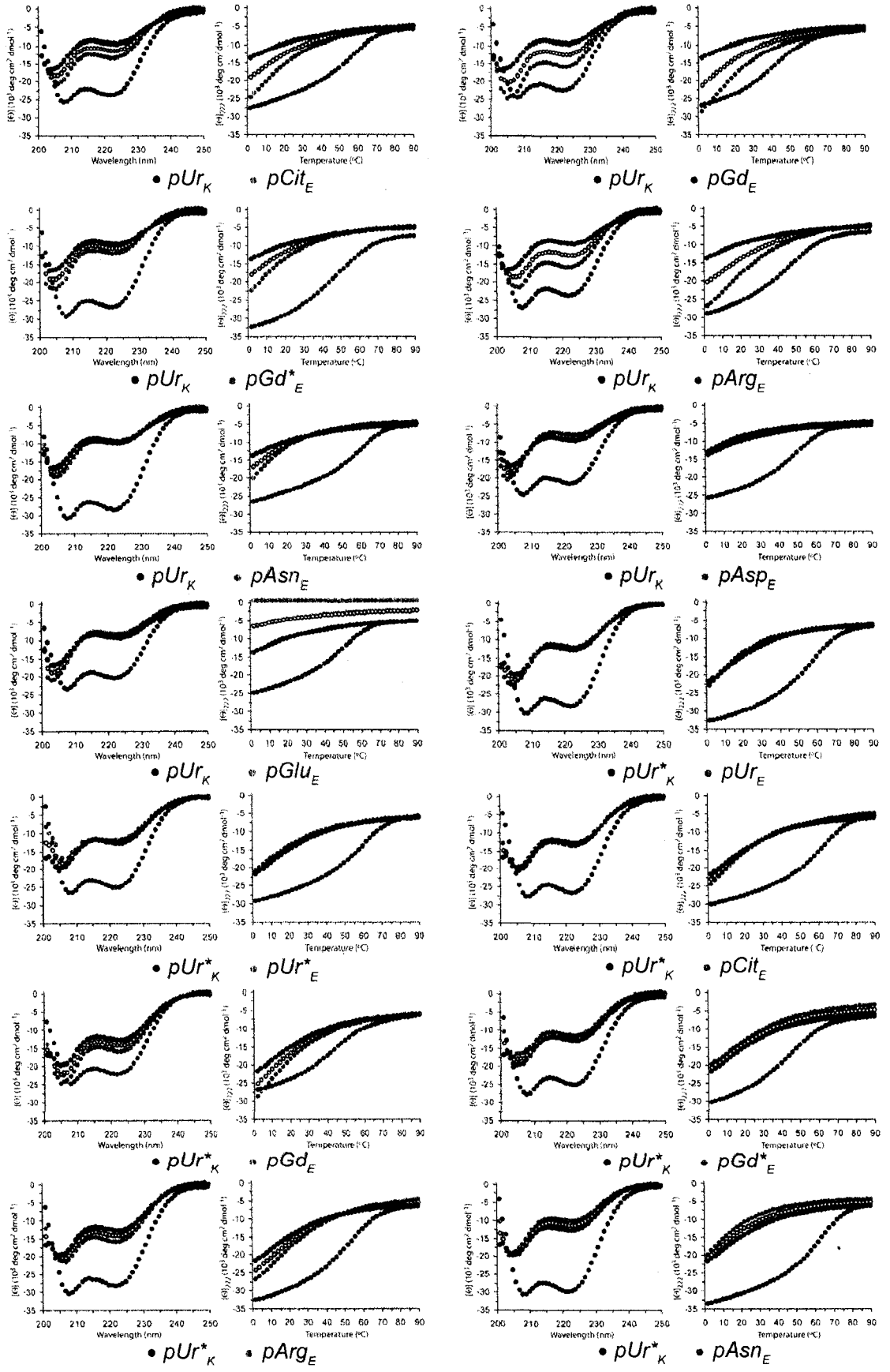
Appendix III

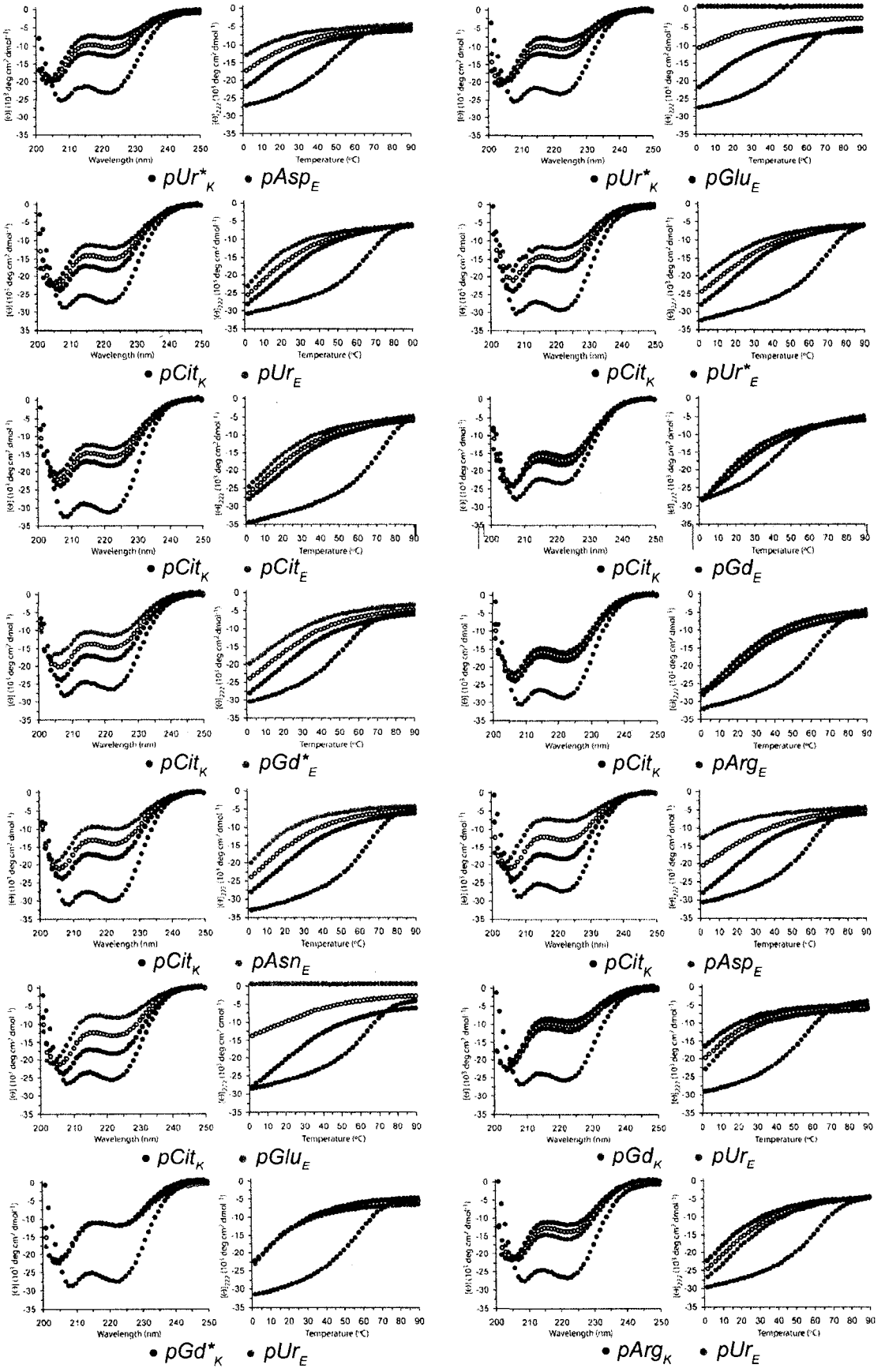
CD Data

Figure A-1: Wavelength and thermal denaturation of all possible dimers. Basic and acidic peptides are as noted below, 1:1 mixture (green) and weighted average (open).




Wavelength (nm)
Temperature (°C)
• pGd*K • pArgE
• pGd*K • pAsnE
• pGd*K • pAspE
• pGd*K • pGluE
• pArgK • pArgE
• pAsnK • pArgE
• pArgK • pAsnE
• pArgK • pAspE
• pArgK • pGluE
• pAsnK • pAsnE
• pAsnK • pAspE
• pAsnK • pGluE
• pUrK • pUrE
• pUrK • pUr*E


Wavelength (nm)
Temperature (°C)
• pUr_K • $pCit_E$
• pUr_K • pGd_E
• pUr_K • pGd^*_E
• pUr_K • $pArg_E$
• pUr_K • $pAsn_E$
• pUr_K • $pAsp_E$
• pUr_K • $pGlu_E$
• pUr^*_K • pUr_E
• pUr^*_K • pUr^*_E
• pUr^*_K • $pCit_E$
• pUr^*_K • pGd_E
• pUr^*_K • pGd^*_E
• pUr^*_K • $pArg_E$
• pUr^*_K • $pAsn_E$


pUr*K
pAspE
pUr*K
pGluE
pCitK
pUrE
pCitK
pUr*E
pCitK
pCitE
pCitK
pGdE
pCitK
pGd*E
pCitK
pArgE
pCitK
pAsnE
pCitK
pAspE
pCitK
pGluE
pGdK
pUrE
pGd*K
pUrE
pArgK
pUrE

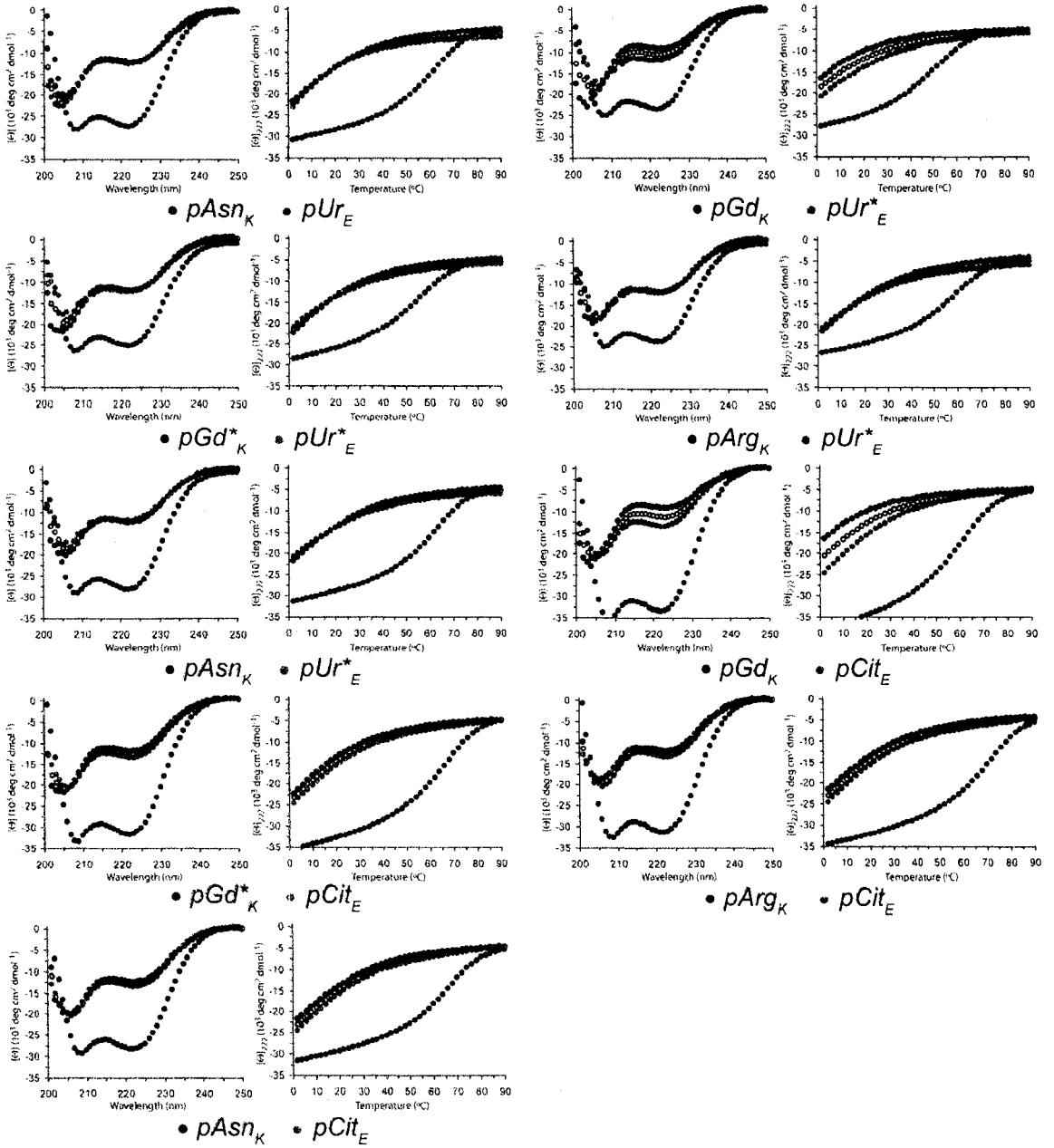
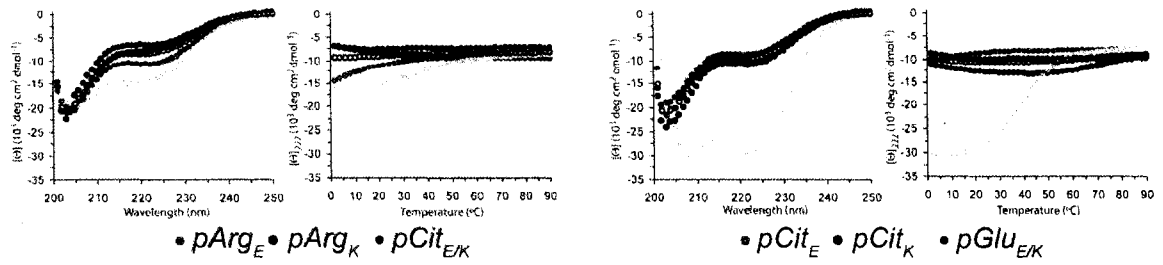
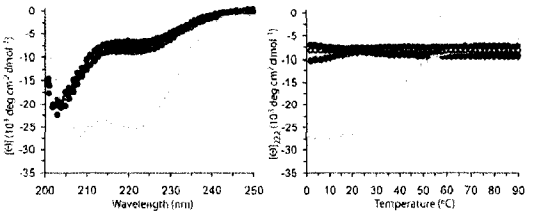
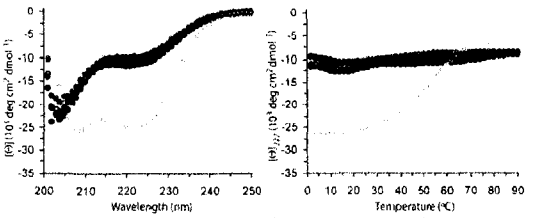


Figure A-2: Wavelength and thermal denaturation of all possible trimers. Basic (blue), acidic (red) and mixed (green) peptides are as noted below, 1:1:1 mixture (orange) and weighted average (open).

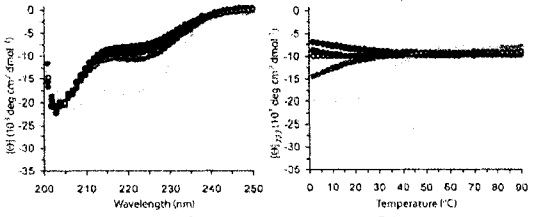




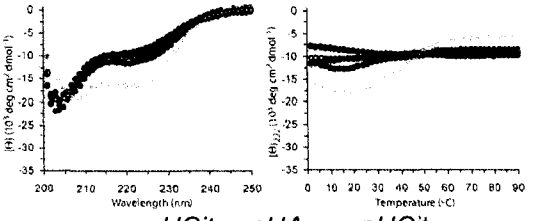
• $pCit_E$ • $pArg_K$ • $pCit_{E/K}$



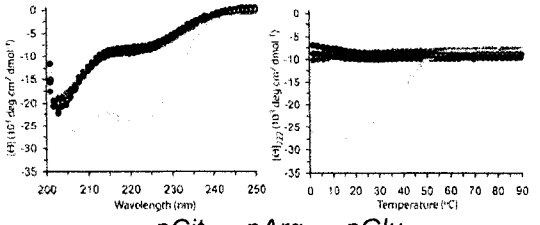
• $pHCit_E$ • $pHCit_K$ • $pHCit_{E/K}$



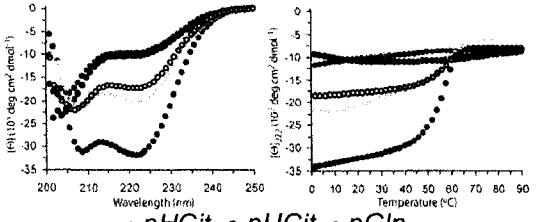
• $pArg_E$ • $pArg_K$ • $pGlu_{E/K}$



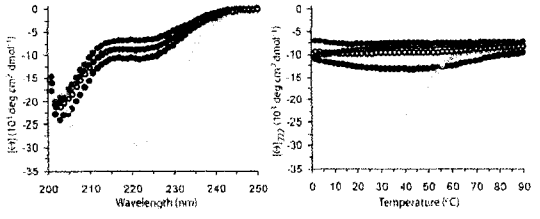
• $pHCit_E$ • $pHArg_K$ • $pHCit_{E/K}$



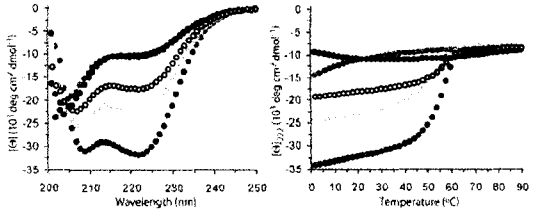
• $pCit_E$ • $pArg_K$ • $pGlu_{E/K}$



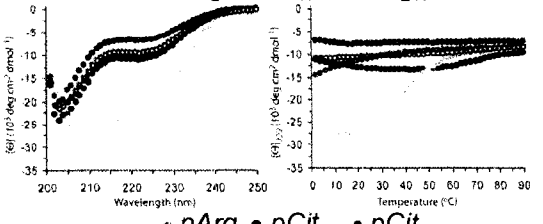
• $pHCit_E$ • $pHCit_K$ • $pGln_{E/K}$



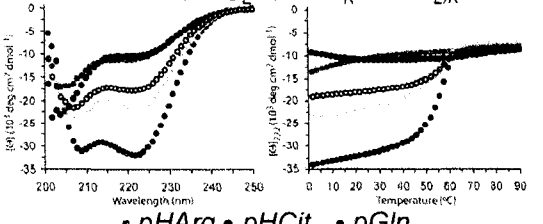
• $pCit_E$ • $pCit_K$ • $pCit_{E/K}$



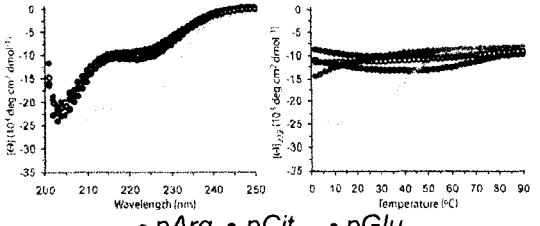
• $pArg_E$ • $pHCit_K$ • $pGln_{E/K}$



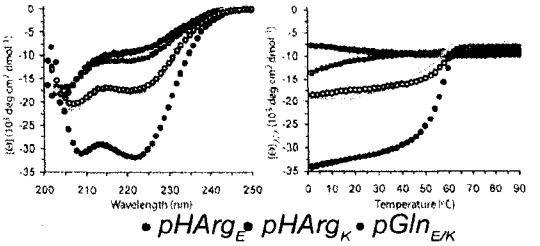
• $pArg_E$ • $pCit_K$ • $pCit_{E/K}$



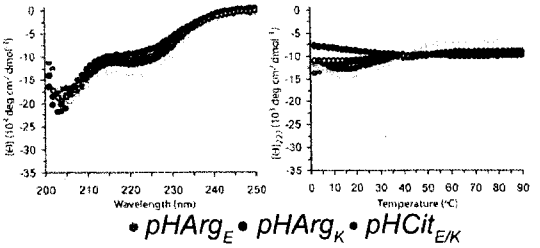
• $pHArg_E$ • $pHCit_K$ • $pGln_{E/K}$



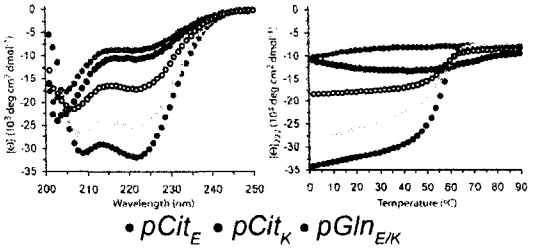
• $pArg_E$ • $pCit_K$ • $pGlu_{E/K}$



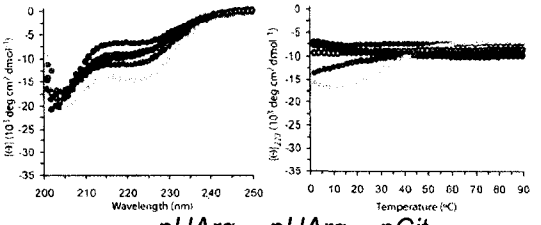
• $pHArg_E$ • $pHArg_K$ • $pGln_{E/K}$



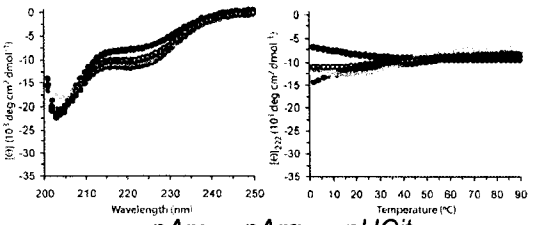
• $pHArg_E$ • $pHArg_K$ • $pHCit_{E/K}$



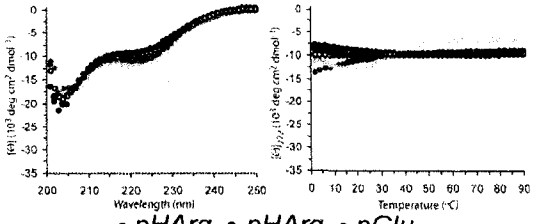
• $pCit_E$ • $pCit_K$ • $pGln_{E/K}$



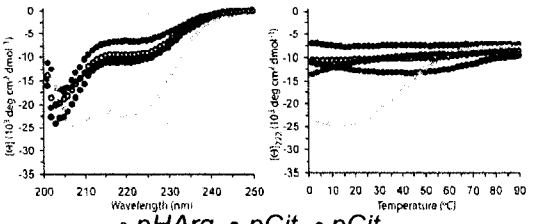
• $pHArg_E$ • $pHArg_K$ • $pCit_{E/K}$



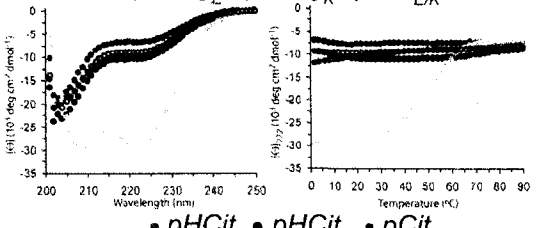
• $pArg_E$ • $pArg_K$ • $pHCit_{E/K}$



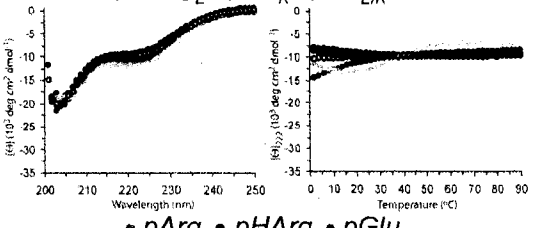
• $pHArg_E$ • $pHArg_K$ • $pGlu_{E/K}$



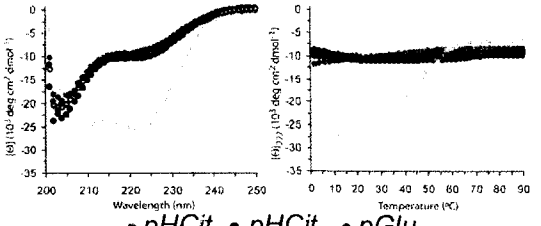
• $pHArg_E$ • $pCit_K$ • $pCit_{E/K}$



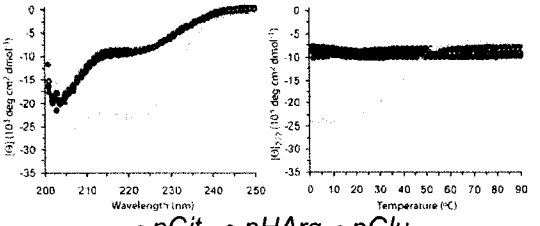
• $pHCit_E$ • $pHCit_K$ • $pCit_{E/K}$



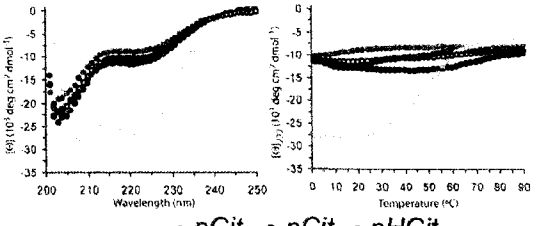
• $pArg_E$ • $pHArg_K$ • $pGlu_{E/K}$



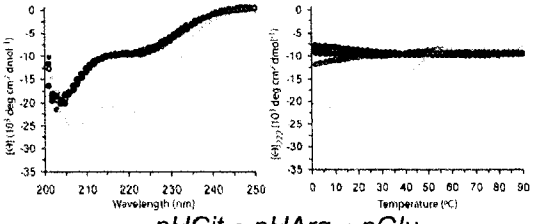
• $pHCit_E$ • $pHCit_K$ • $pGlu_{E/K}$



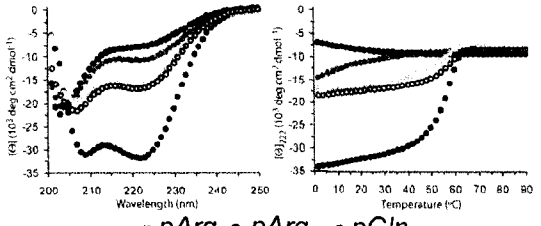
• $pCit_E$ • $pHArg_K$ • $pGlu_{E/K}$



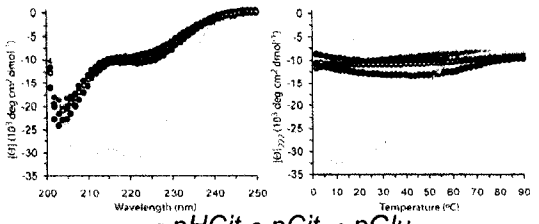
• $pCit_E$ • $pCit_K$ • $pHCit_{E/K}$



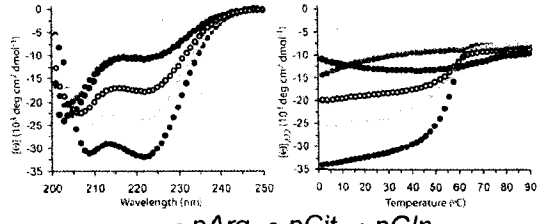
• $pHCit_E$ • $pHArg_K$ • $pGlu_{E/K}$



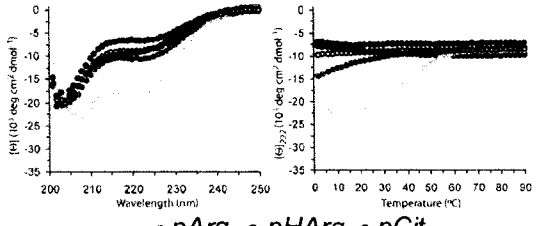
• $pArg_E$ • $pArg_K$ • $pGln_{E/K}$



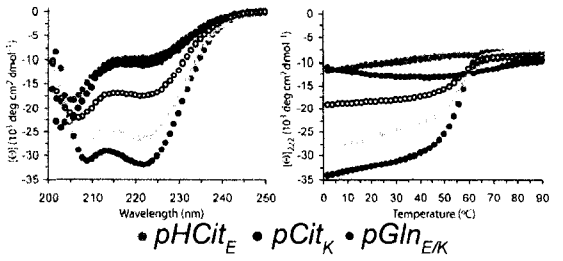
• $pHCit_E$ • $pCit_K$ • $pGlu_{E/K}$



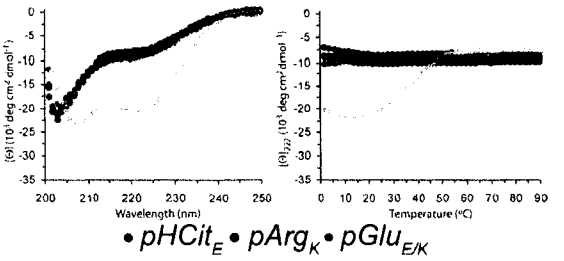
• $pArg_E$ • $pCit_K$ • $pGln_{E/K}$



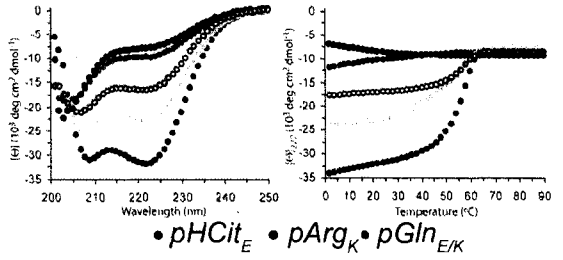
• $pArg_E$ • $pHArg_K$ • $pCit_{E/K}$



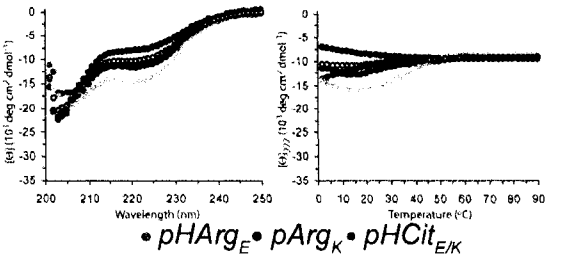
• $pHCit_E$ • $pCit_K$ • $pGln_{E/K}$



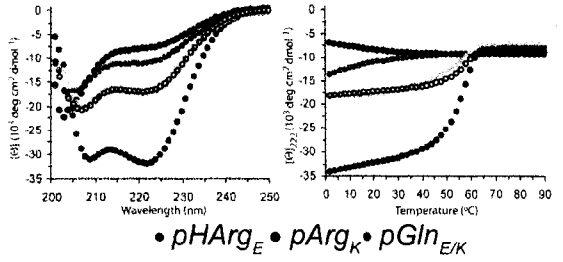
• $pHCit_E$ • $pArg_K$ • $pGlu_{E/K}$



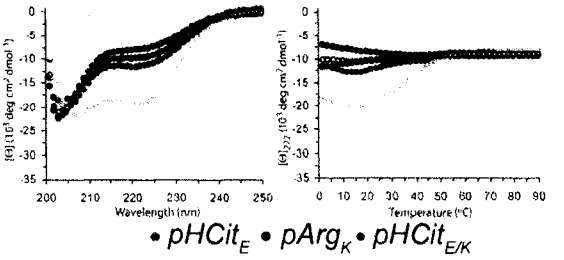
• $pHCit_E$ • $pArg_K$ • $pGln_{E/K}$



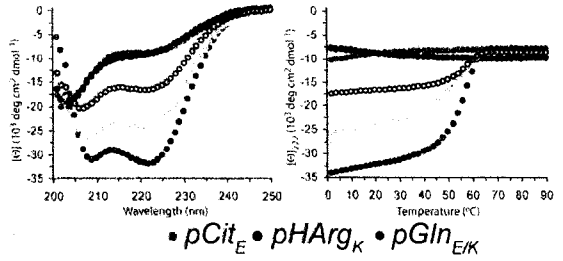
• $pHArg_E$ • $pArg_K$ • $pHCit_{E/K}$



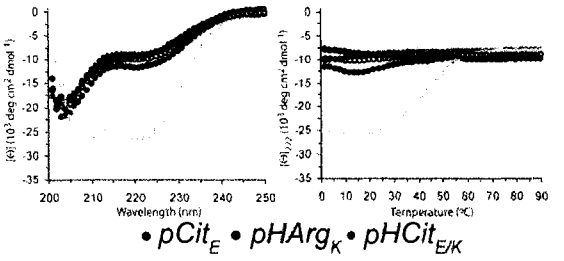
• $pHArg_E$ • $pArg_K$ • $pGln_{E/K}$



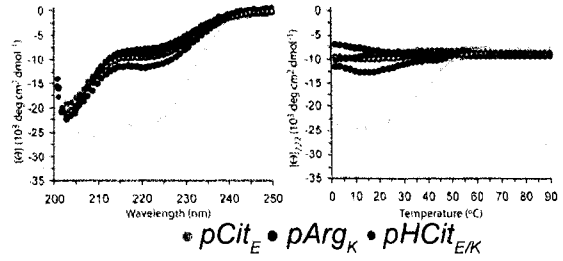
• $pHCit_E$ • $pArg_K$ • $pHCit_{E/K}$



• $pCit_E$ • $pHArg_K$ • $pGln_{E/K}$



• $pCit_E$ • $pHArg_K$ • $pHCit_{E/K}$

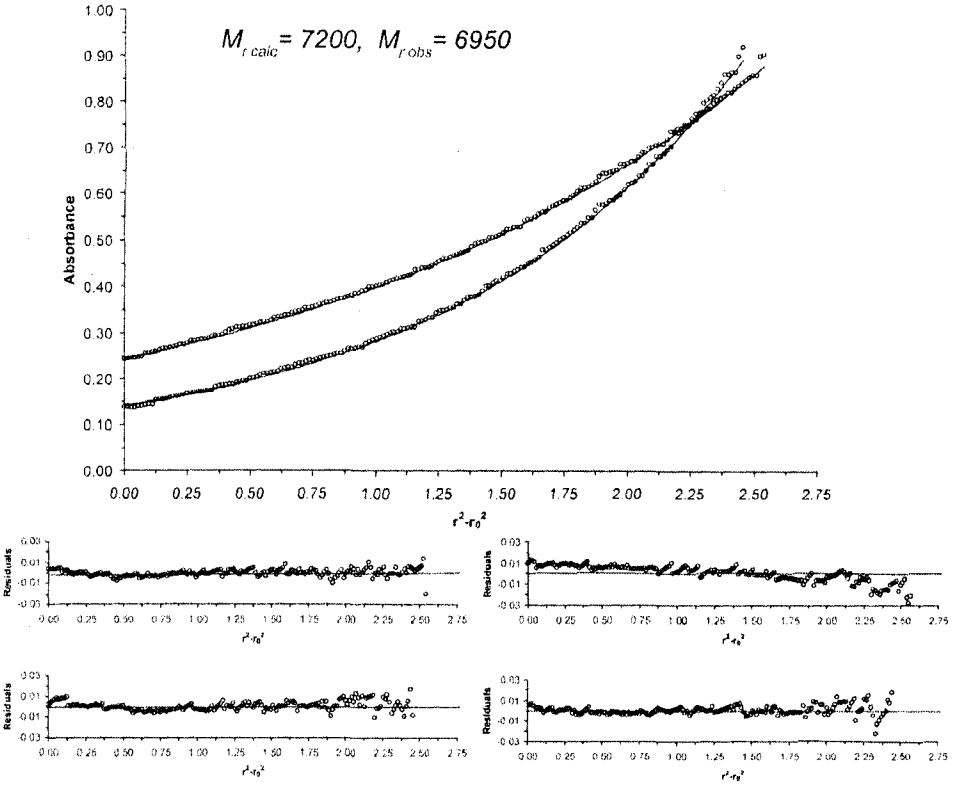


• $pCit_E$ • $pArg_K$ • $pHCit_{E/K}$

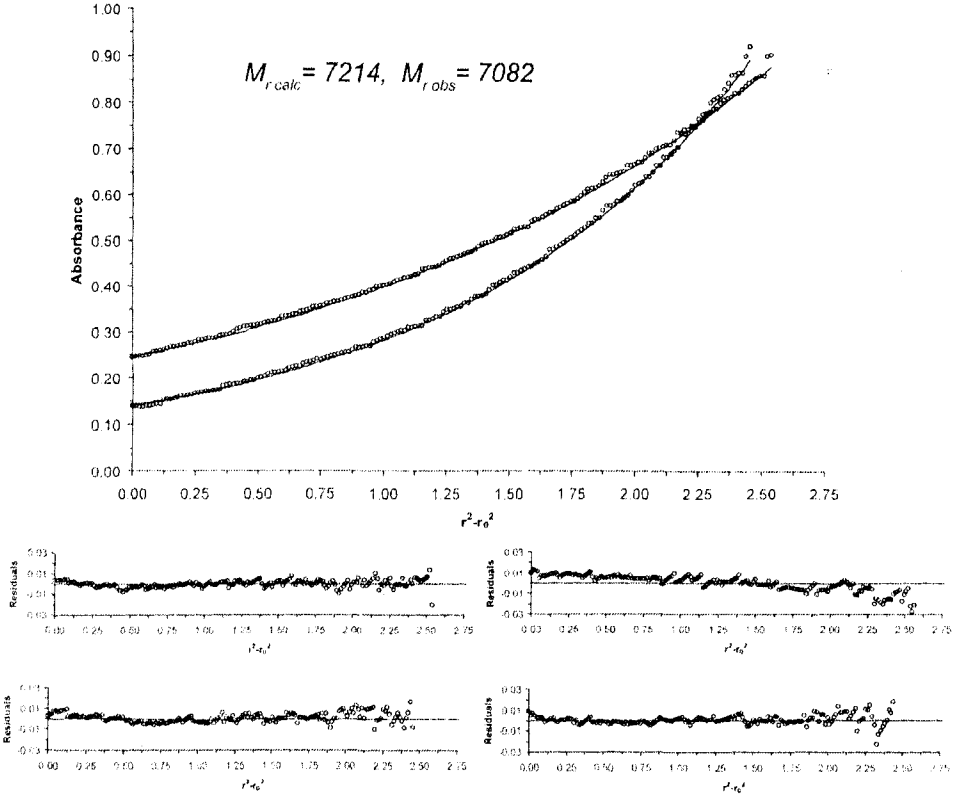
Appendix IV

Sedimentation Data

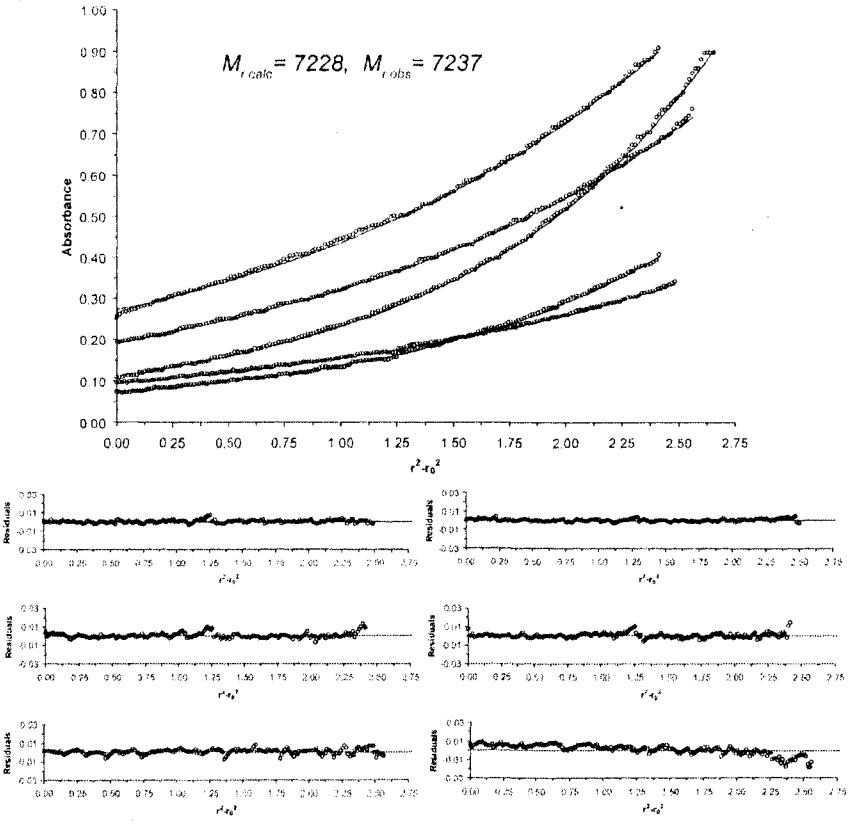
$pGd_K/pAsp_E$



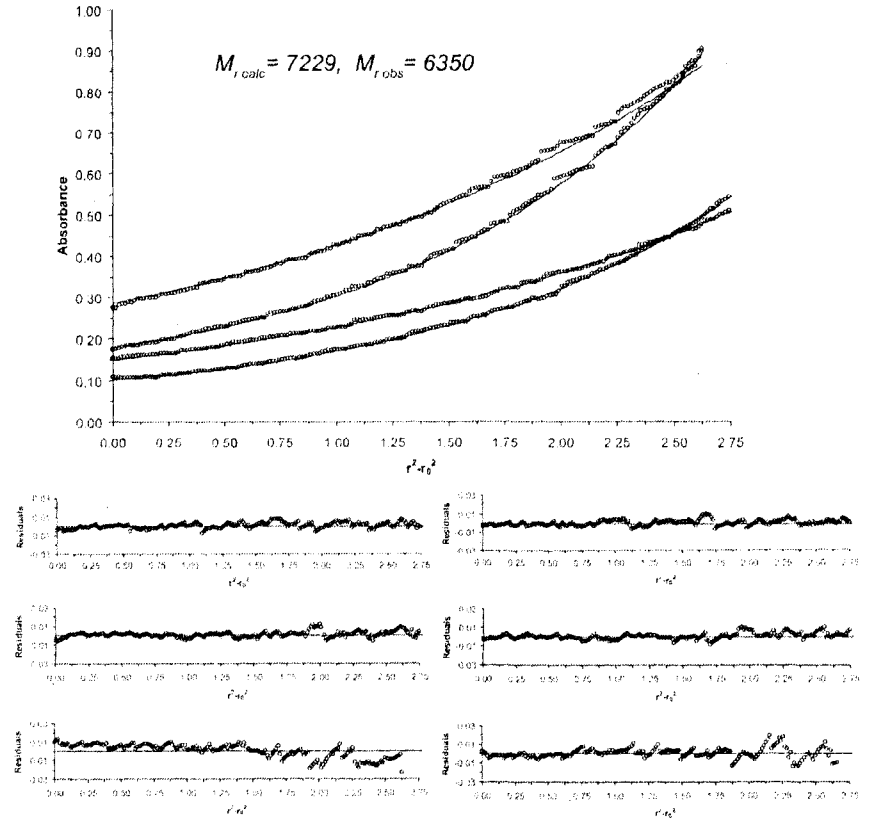
$pGd^*_K/pAsp_E$



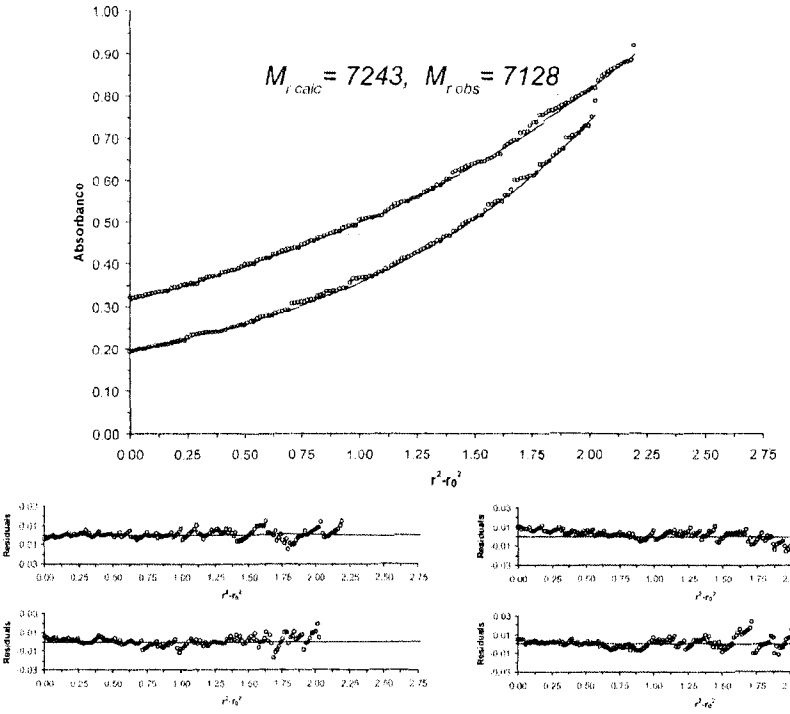
$pArg_K/pAsp_E$



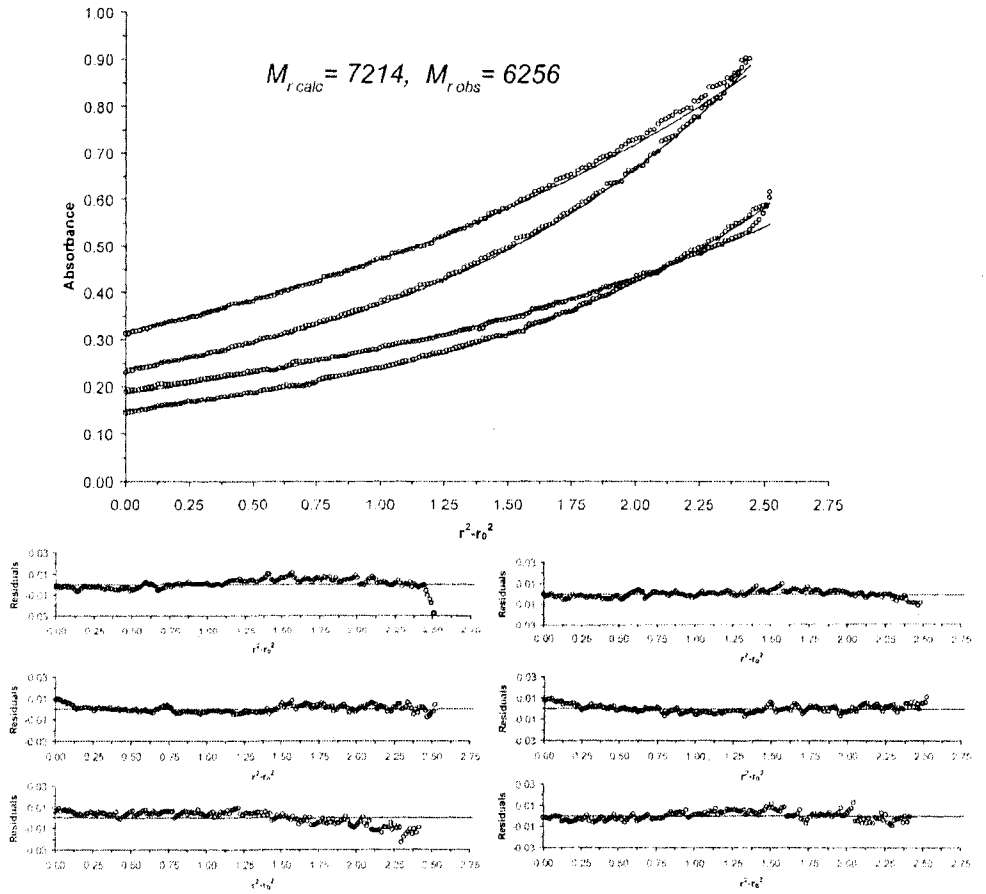
$pGd^*_K/pGlu_E$



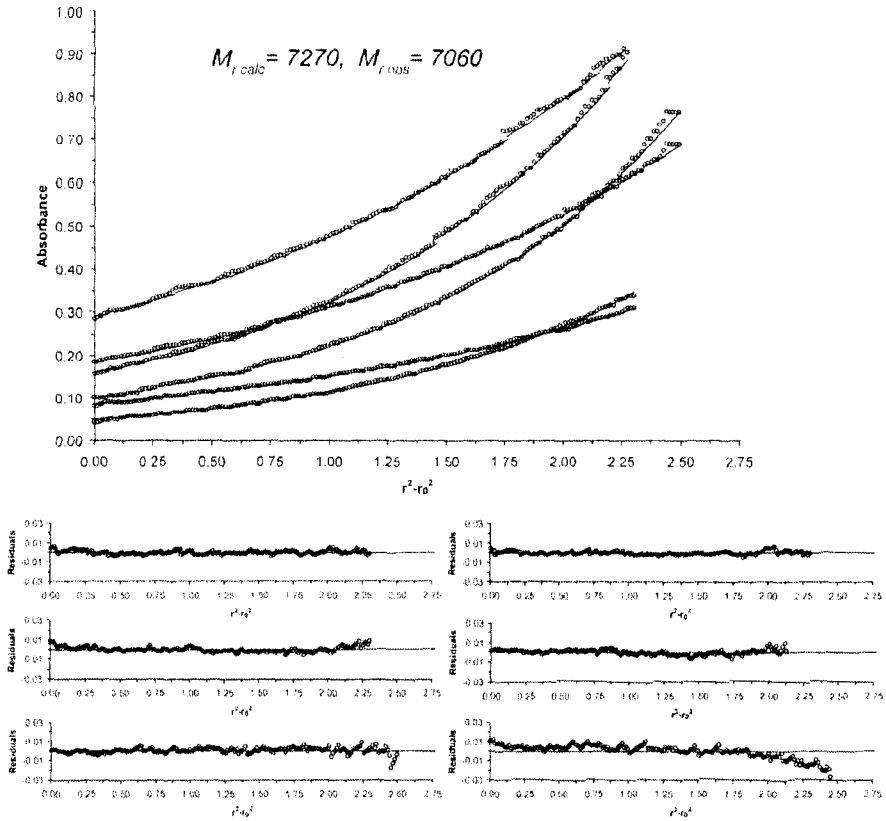
$pArg_K/pGlu_E$



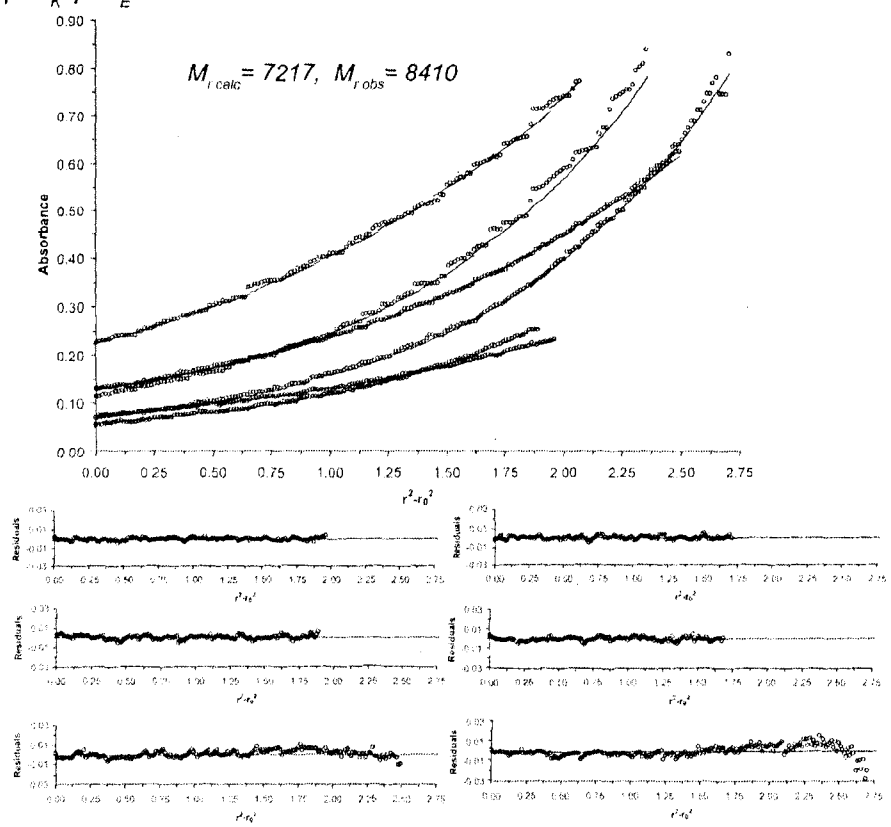
pGd_K/pGd_E



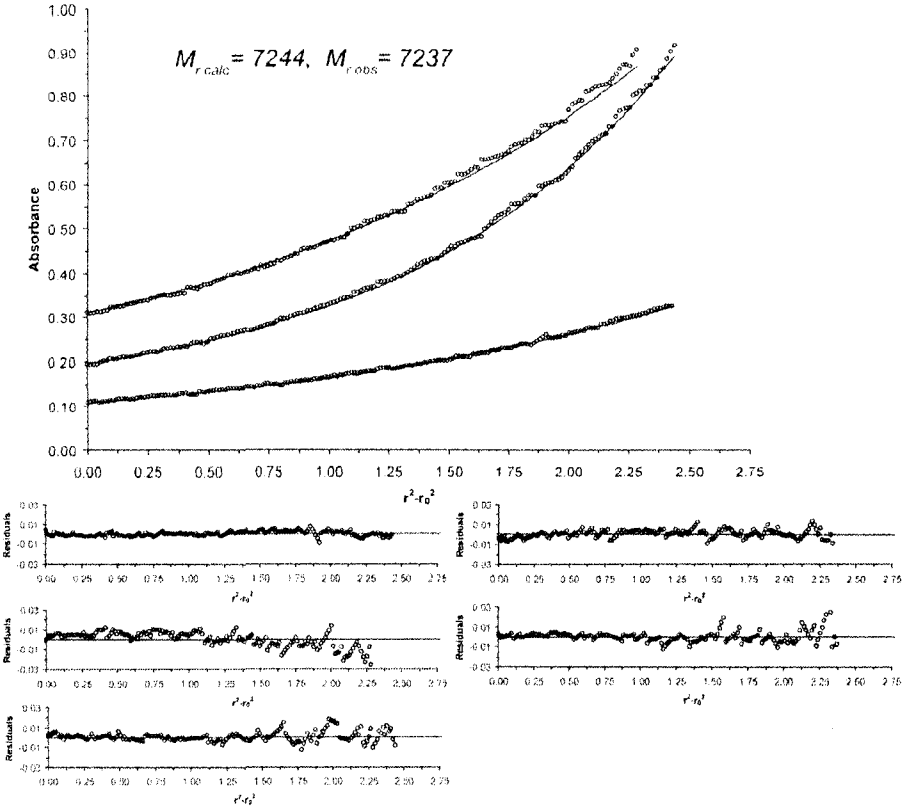
$pArg_K/pArg_E$



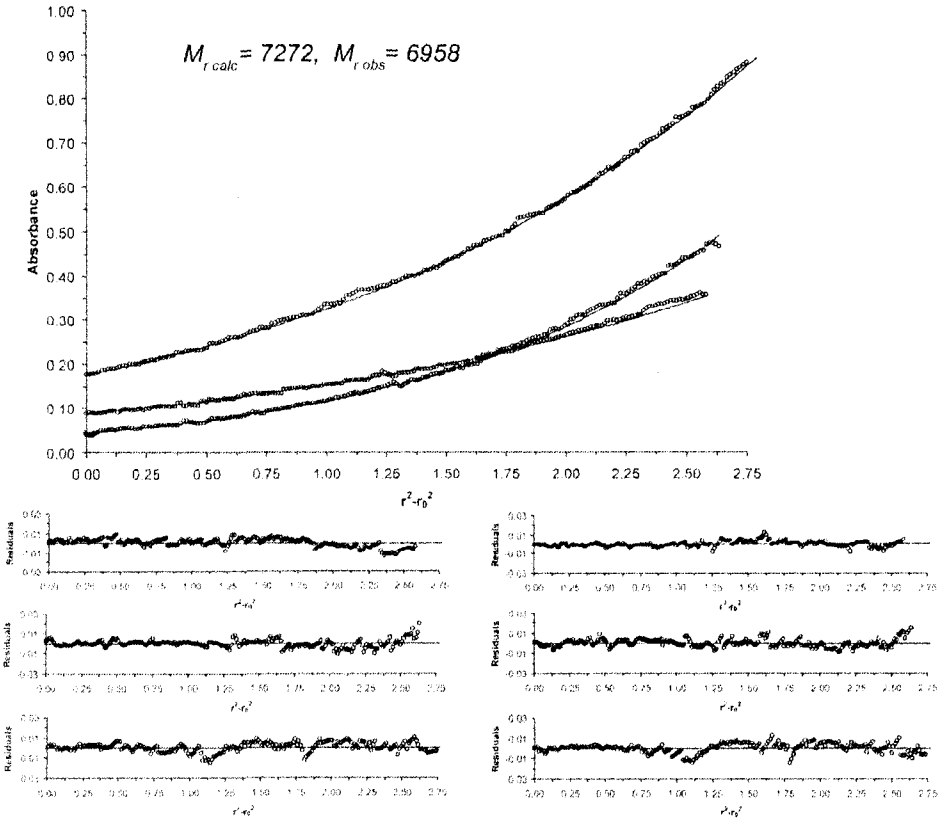
pUr_K/pUr_E



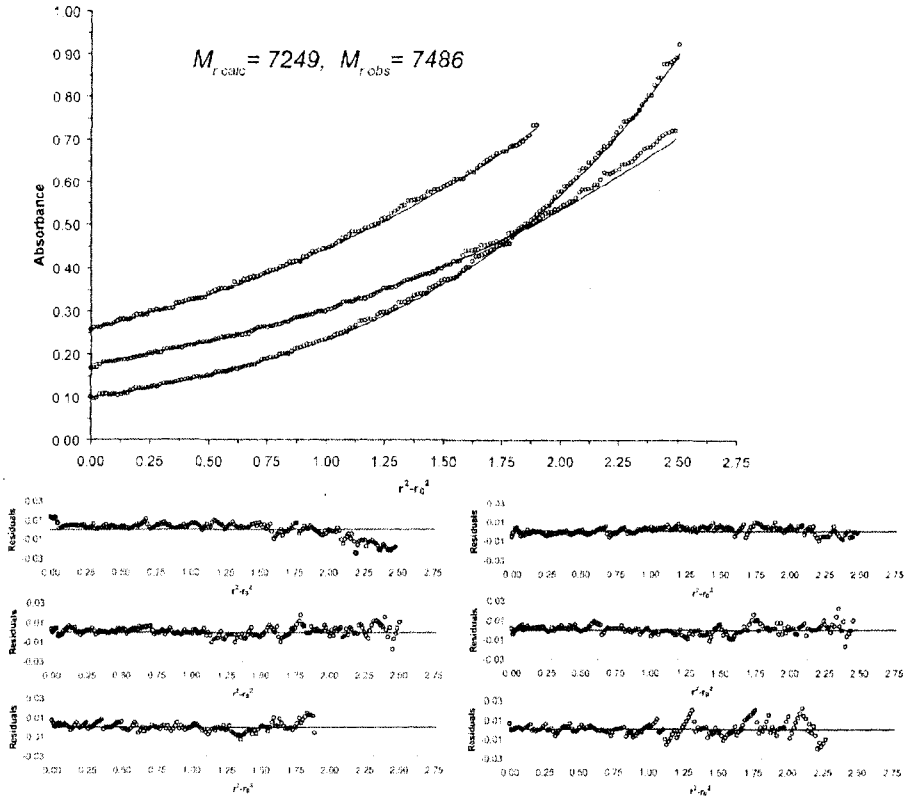
pUr^*_K / pUr^*_E



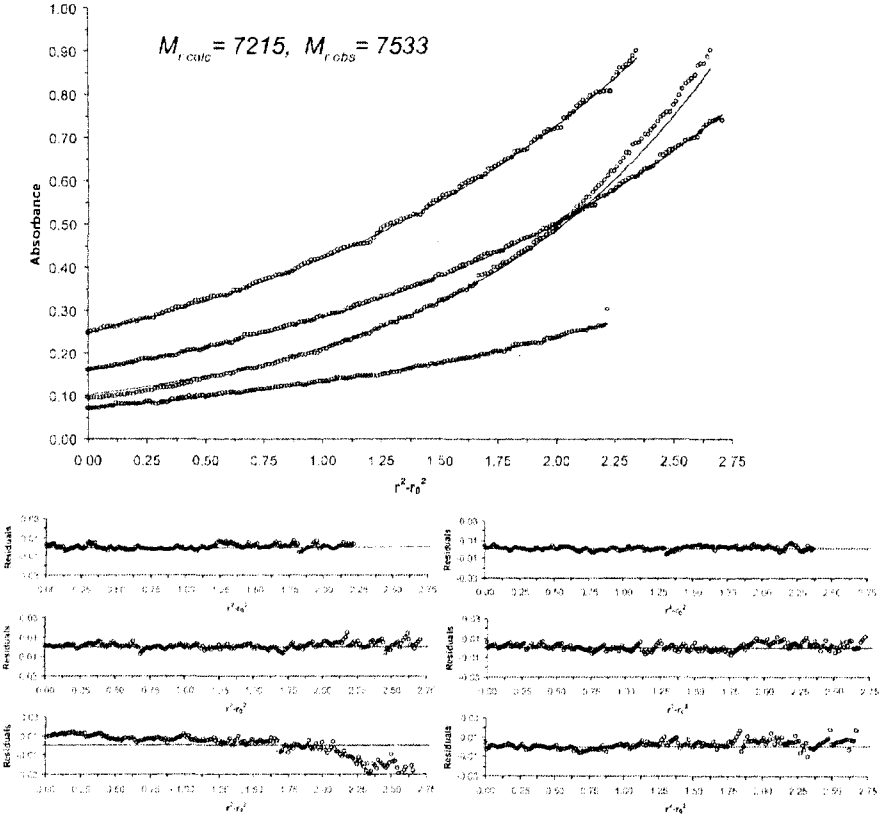
$pCit_K / pCit_E$



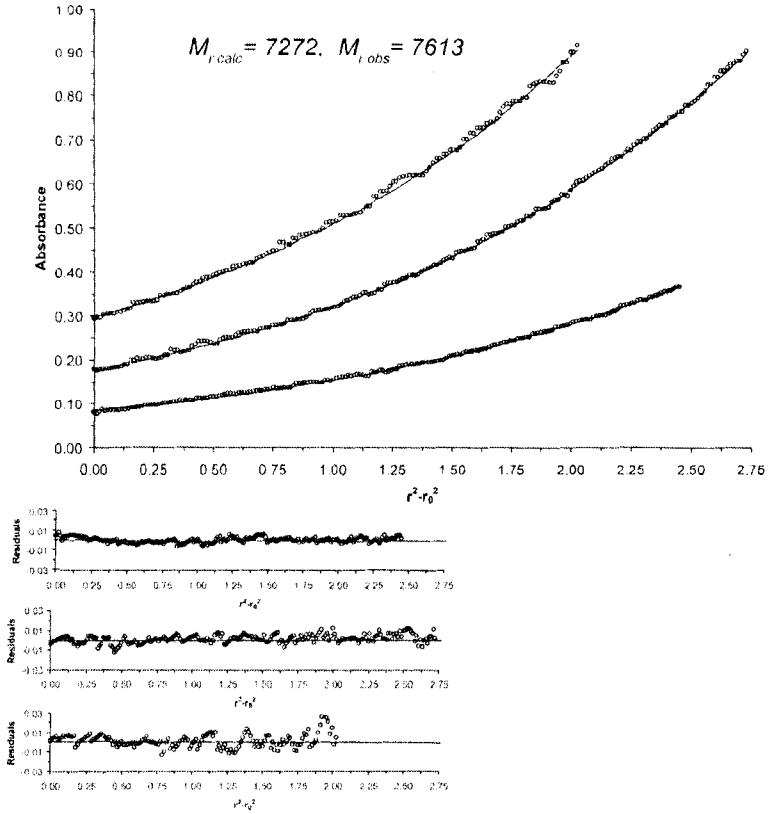
$pUr_K/pCit_E$



$pUr^*_K/pAsn_E$



$pCit_K/pArg_E$


$M_{r\,calc}$ = 7272. $M_{r\,obs}$ = 7613
Absorbance
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00 2.25 2.50 2.75
$r^2-r_0^2$
Residuals