β-Hairpin Peptides and WW Domains Designed for Selective Recognition of Oligonucleotides

Amanda Lynn Stewart

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2009

Approved by:

Professor Marcey Waters

Professor Dorothy Erie

Professor Gary Pielak

Professor David Lawrence

Professor Linda Spremulli

Abstract

Amanda L. Stewart: β-Hairpin Peptides and WW Domains Designed for Selective Recognition of Oligonucleotides
(Under the direction of Professor Marcey L. Waters)

Protein-nucleic acid interactions are crucial in a variety of biological processes. Protein interactions with single-stranded DNA are particularly important in DNA replication, repair, and telomere regulation. The interactions involved in the binding of a designed β-hairpin dimer, $(WKWK)_2$, to ssDNA and dsDNA were previously explored, and the peptide was found to bind ssDNA with a K_d of 3 μM via a combination of aromatic and electrostatic interactions, whereas binding to duplex DNA was driven primarily by electrostatic interactions. In this work, the effects of folding and chirality were studied to determine factors that contribute to affinity and selectivity for ssDNA versus dsDNA. Binding studies showed that (1) folding is crucial for binding to both ss- and dsDNA and (2) chirality affects binding for duplex DNA but not for ssDNA. Taken together, these studies reveal different modes of binding for ss- and duplex DNA, with different driving forces, but in each case peptide structure contributes significantly to binding.

In another study, a β-sheet peptide based on a WW domain sequence was redesigned for the molecular recognition of ssDNA. A previous report showed that $(WKWK)_2$ binds ssDNA with low micromolar affinity but with little selectivity over dsDNA. This work extends those studies to a three-stranded β-sheet designed to mimic the OB-fold. The new peptide binds ssDNA with low micromolar affinity and shows enhanced selectivity for

ssDNA. The redesigned peptide no longer binds its native ligand, the polyproline helix. This indicates that the peptide has been redesigned for the function of binding ssDNA.

Structural studies indicate that this peptide consists of a structured β-hairpin made of Strands 2&3 with a less structured strand 1, which provides affinity for ssDNA but does not improve the stability of the full peptide. Both function and stability are gained by incorporating a novel binding pocket into the peptide, and the redesigned peptide successfully mimics the OB-fold domain. Further mutations were made to design a mutant with increased structure, affinity, and selectivity for ssDNA. Knowledge gained from these binding and structural studies may lead to better designs of β-sheet peptides designed to target nucleotides, damaged DNA, and ssDNA.

Acknowledgements

I must first acknowledge that without my family and friends, I could not have accomplished the great task of writing this dissertation or completing the other requirements for this doctoral degree. I would especially like to thank my parents, David and Julie Stewart, for their support, guidance, encouragement, and unwavering love. I am so thankful to have parents who instilled in me the love of God and the constant reminder of where my strength comes from each day. As I think about all the blessings I have been given, I can not thank you enough for the many things you have done for me throughout my life. If I attempted to thank you in writing for even a portion of what you've done, the size of this volume would not compare to it. I also must thank my brother, Ric, for his endless love, compassion, and humor. You kept me going in the good days and even some of the tougher days throughout my tenure at UNC. Your expressions of good cheer and wit helped me through many difficult times, even when you didn't know it, and I am very grateful.

I am truly blessed to have very loving grandparents who have supported me throughout my life. Thank you both for the great examples you have been for me. I could never thank you enough for everything you have done to help me get to this point.

I have been blessed with so many friends and family who have all given themselves in some way to support my efforts in college and in graduate school. Thank you all so much for what you have added to my life. While many of you have stated that you could never complete a Ph.D. in chemistry, I must confess to you all that I could not have finished this journey without you. You all had a part in this in some special way.

My fourth grade teacher, Mrs. Cindy Jones, was the first to teach me the most basic principles of chemistry and helped to instill in me a love for science. She has been a constant supporter of me on this long journey from fourth grade to the completion of this degree, and I am very thankful for her many expressions of encouragement along the way. I owe much gratitude to all of my teachers and professors. You have all played a special role in helping me achieve this degree.

I would like to thank Marcey Waters for allowing me to work on some very exciting projects at UNC. Marcey, you have been a great advisor and you taught me so much about mentoring. Thanks for challenging me to investigate, learn, and teach at a higher level than I ever imagined.

Thanks to my committee members for all of the advice that you have given me and for having the right solution at the right time.

Waters lab members, you have been great colleagues. Thank you especially for putting up with me during the past year. I thank you all for helping me accomplish this goal. I could not have done this without you.

Table of Contents

List of Tables

Table Page

List of Figures

List of Abbreviations

A	Adenine
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	Adenosine triphosphate
C	Cysteine
C	Cytosine
CD	Circular Dichroism
Cit	Citrulline
CspA; B	Cold shock protein A; cold shock protein B
Cys	Cysteine
D	Aspartic acid
Dab	Diaminobutyric acid
DIPEA	Diisopropylethylamine
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA; duplex DNA
E	Glutamic acid
F	Phenylalanine
G	Glycine
G	Guanine
Gln	Glutamine

Glu	Glutamic acid
Gly	Glycine
H	Histidine
HBTU	O-benzotriazole-N,N,N’,N’,-tetramethyluronium hexafluorophosphate
His	Histidine
HOBT	N-hydroxybenzotriazole
I	Isoleucine
Ile	Isoleucine
K	Lysine
L	Leucine
Leu	Leucine
Lys	Lysine
N	Asparagine
NBT/BCIP	Nitro Blue Tetrazolium chloride/5-Bromo-4-Chloro-3’-Indolyl Phosphate p-toluidine salt
NMR	Nuclear magnetic resonance
O	Ornithine
OB-fold	Oligonucleotide/oligosaccharide binding domain
Orn	Ornithine
P	Proline
Phe	Phenylalanine
PhG	Phenylglycine
Pro	Proline
Q	Glutamine

R	Arginine
RNA	Ribonucleic acid
RPA	Replication protein A
S	Serine
SAAP	Streptavidin/alkaline phosphatase complex
Ser	Serine
SSB	Single-stranded DNA binding protein
ssDNA	Single-stranded DNA
T	Threonine
T	Thymidine
TAMRA	5-(and -6)-Carboxytetramethylrhodamine, mixed isomers
TFA	Trifluoroacetic acid
Thr	Threonine
TIPS	Triisopropylsilane
Trp	Tryptophan
Tyr	Tyrosine
V	Valine
Val	Valine
W	Tryptophan
Y	Tyrosine

Chapter I

Introduction

A. Background and Significance

i. Protein-Nucleotide Interactions

Proteins and nucleotides are essential elements to life and the life cycle. They are both a part of and contribute to the central dogma of biochemistry. From single nucleotide triphosphates to individual strands of both ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) to complex structures of duplex DNA and RNA, nucleotides function in a variety of ways on a number of levels. There are specific proteins to recognize each of these different nucleotide structures. This work focuses on the specific interactions that proteins have with nucleic acids. There are many levels in which this study could be pursued. This volume will attempt to look at the specific interactions and the roles that proteins play in recognizing nucleotides using structured peptides as model systems. More specifically, a major theme and focus of this work is to understand the specific driving forces involved in the recognition of single-stranded DNA (ssDNA) by proteins.

ii. Single-stranded DNA Binding Proteins and Their Significance

The reasons behind the focus on ssDNA recognition are that this recognition event is important in so many different processes and has vast influence on so many of the crucial processes in nature. The primary class of single-stranded DNA binding proteins (SSBs) contains what is termed the oligonucleotide/oligosaccharide binding domain (OB-fold

domain) which specifically recognizes single-stranded DNA and RNA sequences.[1] Many of the classic SSB proteins are known simply for their ability to discriminate between ssDNA and double-stranded DNA (dsDNA). Their roles in metabolism, however, range from recognizing ssDNA sequences during replication, stabilizing ssDNA in its functional single-stranded state, protecting ssDNA from nucleases, melting secondary DNA structures, assisting in homologous recombination, and recognizing DNA damage.[2] OB-fold domains are found in a variety of different proteins and are known to specifically recognize single strands of DNA and RNA as well as single bases that may emerge during replication, recombination, and at damaged DNA sites.[1] Most OB-fold domains are composed of several basic amino acids that bind to the phosphates in the DNA backbone, and they utilize aromatic residues which usually form a cleft which recognizes the nucleotide bases.[3] A study of proteins containing OB-fold domains or other ssDNA binding domains reveals that their mode of recognition and their roles in recognizing bases as well as in stabilizing protein-DNA complexes during binding and DNA repair vary greatly depending on the protein and its function.

---

[1] (a) Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115-133. (b) Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004**, *14*, 36-42. (c) Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

[2] (a) Pestryakov, P. E.; Lavrik, O. I. *Biochemistry (Moscow)* **2008**, *73*, 1388-1404. (b) Shamoo, Y.; Friedman, A. M.; Parsons, M. R.; Konigsberg, W. H.; Steitz, T. A. *Nature* **1995**, *376*, 362-366. (c) Mer, G.; Bochkarev, A.; Gupta, R.; Bochkareva, E.; Frappier, L.; Ingles, J. C.; Edwards, A. M.; Chazin, W. J. *Cell*, **2000**, *103*, 449-456.

[3] (a) Kloks, C. P.; Spronk, C. A.; Lasonder, E.; Hoffmann, A.; Vuister, G. W.; Grzesiek, S.; Hilbers, C. W. *J. Mol. Biol.* **2002**, *316*, 317-326. (b) Schroder, K.; Graumann, P.; Schnuchel, A.; Holak, T. A.; Marahiel, M. A. *Mol. Microbiol.* **1995**, *16*, 699-708.

Replication protein A (Figure 1.1), for example, is a classic ssDNA binding protein known for its six OB-fold domains, four of which are involved in DNA binding.[1, 4] Its functions range from binding ssDNA during replication to facilitating the coordination of DNA repair machinery by specifically interacting with DNA repair proteins.[2c] It is also associated with nucleotide excision repair, base excision repair, and recombination.[2c]

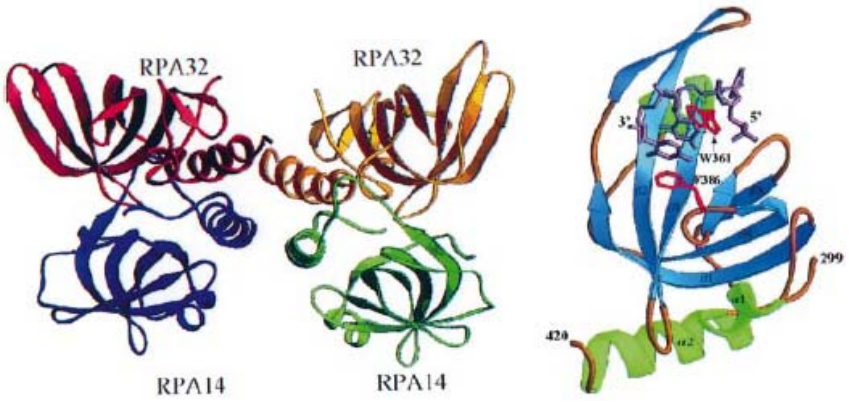


Figure 1.1. Structure of crystallographic dimer of RPA14/32$_{45–170}$. The RPA14 subunits are shown in blue and green and the RPA32 subunits in red and yellow. (b) Structure of RPA70$_{299–420}$ subunit in complex with oligodeoxycytosine; two Trp residues involved in binding are shown in red and the crystallographically determined structure of three bases is highlighted in purple.[4c]

[4] (a) Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. *Nature* **1997**, *385*, 176-181. (b) Wold, M. S. *Annu. Rev. Biochem.* **1997**, *66*, 61-92. (c) Bochkarev, A.; Bochkareva, E.; Frappier, L.; Edwards, A. M. *EMBO J.* **1999**, *18*, 4498-4504.

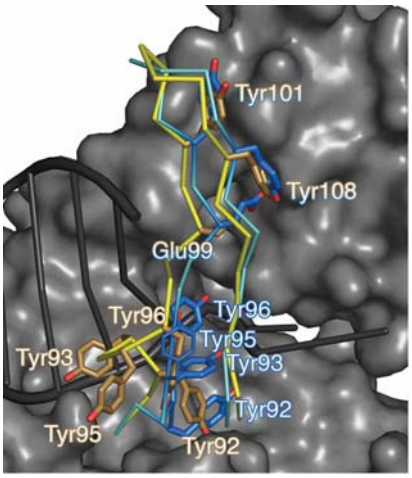
Other examples of β-hairpins or β-sheet proteins involved in DNA repair include UvrB,[5] Chlorella virus DNA ligase,[6] Rad4,[7] and BRCA2 protein.[1b] UvrB is the primary prokaryotic nucleotide excision repair (NER) protein, and its structure reveals that it interacts with DNA by inserting a β-hairpin between the strands of the duplex DNA, locking one of the strands between the hairpin and another domain within the protein (Figure 1.2). Upon DNA binding, UvrB ATPase is activated which leads to translocation. In the DNA-helicase complex, one of the DNA strands threads behind the β-hairpin, and the hairpin forms strong interactions with this 3' overhang. Three crucial Tyr residues, Tyr 92, 93, and 96 form an aromatic pocket that stabilizes the complex. Tyr 96 is also known to bind the C-G base pair which is adjacent to the 3' overhang.[5]

[5] (a) Malta, E.; Moolenaar, G. F.; Goosen, N. *J. Biol. Chem.* **2006**, *281*, 2184-2194. (b) Truglio, J. J.; Karakas, E.; Rhau, B.; Wang, H.; DellaVecchia, M. J.; Van Houten, B.; Kisker, C. *Nat. Struct. Mol. Biol.* **2006**, *13*, 360-364. (c) Skorvaga, M.; DellaVecchia, M. J.; Croteau, D. L.; Theis, K.; Truglio, J. J.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2004**, *279*, 51574-51580. (d) Skorvaga, M.; Theis, K.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2002**, *277*, 1553-1559.

[6] Nair, P. A.; Nandakumar, J.; Smith, P.; Odell, M.; Lima, C. D.; Shuman, S. *Nat. Struct. Mol. Biol.* **2007**, *14*, 770-778.

[7] Min, J. H.; Pavletich, N. P. *Nature* **2007**, *449*, 570-576.

(a)



(b)

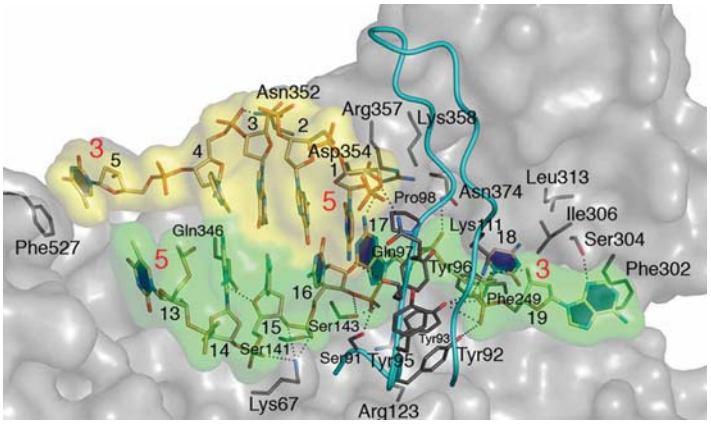


Figure 1.2. UvrB-DNA interactions. (a) Superposition of UvrB-DNA complex and apo-UvrB (pdb: 1D9X). (b) UvrB bound to DNA with β-hairpin shown in cyan (pdb: 2FDC).[5]

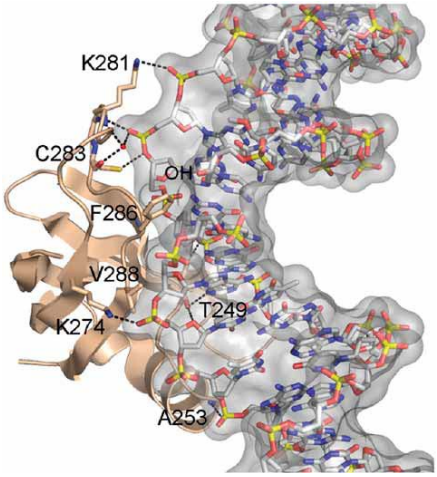


Figure 1.3. DNA interface with the OB-fold domain in Chlorella virus DNA ligase.

In a similar manner, Rad4 positions a β-hairpin through the duplex DNA which causes the two damaged bases to flip out of the duplex and be recognized by the protein.[7] Chlorella virus DNA ligase, like all ligases, contains an OB-fold domain, and this peptide recognizes nicked DNA through a hairpin which stems from the OB-fold domain and inserts into the major groove of the DNA near the nick (Figure 1.3).[6] BRCA2, the breast cancer susceptibility gene 2 protein, is involved in double-stranded DNA break repair and it has been shown to function in a similar fashion to replication protein A.[1b] In another similar recognition motif, a β-hairpin within the RNA polymerase of bacteriophage 7 intercalates DNA at the fork of the single-stranded/double-stranded junction of the promoter DNA. A valine residue at the tip of the hairpin seems to stack with the first base pair in the duplex region.[8] All of these proteins exemplify the variety of roles that these domains play in DNA recognition. While aromatic residues are key to stacking with single bases, these peptides also recognize less stable regions within duplex DNA as well as in separated DNA strands or single bases that exist temporarily due to a variety of cellular metabolic processes.

B. Targeting ssDNA – Applications

i. Molecular Recognition of Nucleotides

Many groups have employed molecular recognition processes to design peptides, proteins, or small molecules which bind biologically relevant targets. Peter Dervan's lab has designed a host of DNA binding small molecules. One example of his work is the design of hairpin polyamides which bind duplex DNA. The compounds are functionalized with a fluorophore, and a fluorescence increase is observed upon binding. This type of study has led to the development of sensors that could be used to detect certain DNA sequences such as

[8] Stano, N. M.; Patel, S. S. *J. Mol. Biol.* **2002**, *315*, 1009-1025.

DNA mismatches.[9] Schepartz and colleagues have designed mini-proteins that recognize DNA with high affinity and specificity through protein grafting. This method has afforded α-helices that are highly selective for their DNA targets.[10] Using a different approach, Robinson and colleagues have designed cyclic β-hairpin peptides which bind tightly to BIV TAR RNA and inhibit the Tat-TAR interaction of the bovine immunodeficiency virus.[11a] The group also inhibited the REV/RRE interaction.[11b] This type of binding achieved is significant because it is applicable to ongoing studies to stop HIV and other viruses. Molecules such as Robinson's peptidomimetic disrupt the Tat protein-TAR RNA interaction which inhibits the generation of full-length transcripts and decreases HIV replication.[11]

ii. Disrupting Protein-ssDNA Interactions

Robinson's peptidomimetic is one of many molecules that disrupt key protein-nucleotide interactions to block debilitating processes from occurring. Inhibiting the Simian virus 40 function can be likened to disruption of the well-known Tat-TAR interaction. Simian virus 40 large tumor antigen (SV40 LTag) is crucial for DNA unwinding during the replication stage of the virus. This protein is a hexamer which interacts with DNA through a positively charged hexameric channel. Six β-hairpins are connected to each of these subunits, and two Lys residues along with a histidine and a phenylalanine are positioned on

---

[9] Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 1195-1202.

[10] Yang, L.; Schepartz, A. *Biochemistry* **2005**, *44*, 7469-7478.

[11] (a) Athanassiou, Z.; Patora, K.; Dias, R. L. A.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007**, *46*, 741-751. (b) Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew Chem. Int. Ed.* **2007**, *46*, 9101-9104.

the tip of the hairpin.[12] Their positions indicate that those residues are important for DNA recognition. A close examination of the interactions involved in DNA binding could lead to a similar peptide designed to inhibit viral replication.[12] Using the knowledge of ssDNA binding protein characteristics, one could imagine several applications to the design of ssDNA binding ligands such as telomerase inhibition, inhibition of replication, and DNA damage detection.

iii. Mimicking the OB-fold Domain

Recognizing the importance of ssDNA and other nucleotides as target molecules, this laboratory has delved into the area of molecular recognition of ATP, ssDNA, and RNA.[13] This research began with the synthesis of peptides designed to target ATP. A β-hairpin peptide was designed with a binding pocket to recognize the adenine or other nucleotide base.[5] The binding pocket contains two aromatic Trp residues intended to form favorable stacking interactions with the base accompanied by two flanking Lys residues to form electrostatic interactions with the phosphates of the ATP. The peptide was named for its binding cleft, WKWK (Figure 1.4). Fluorescence binding studies were conducted to determine the binding affinity of the peptide for ATP. The dissociation constant for the interaction is about 170 μM. This is a good starting point for studying the interactions involved in DNA binding.

---

[12] Shen, J.; Gai, D.; Patrick, A.; Greenleaf, W. B.; Chen, X. S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 11248-11253.

[13] (a) Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544. (b) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25. (c) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70,* 1105-1114. (d) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.

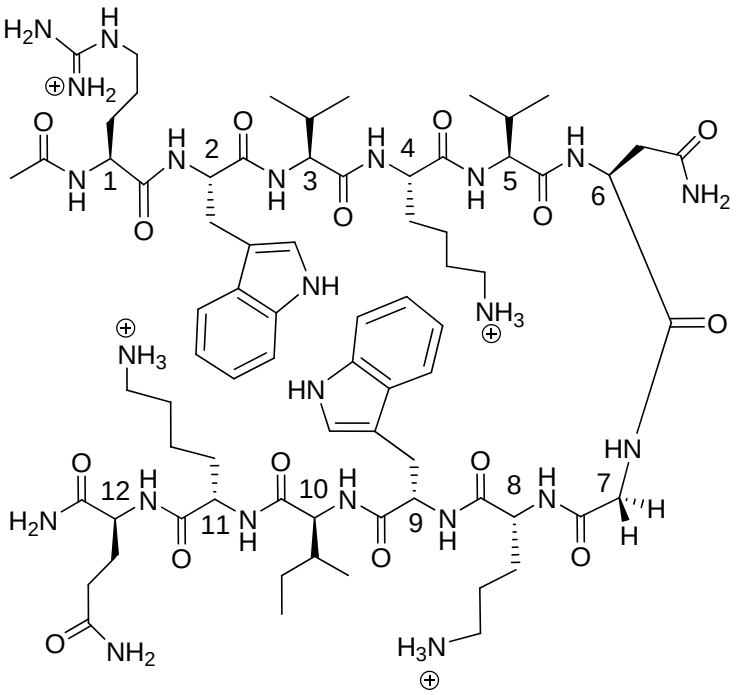


Figure 1.4. WKWK peptide designed for molecular recognition of ATP and ssDNA.

A structural model of the interaction between WKWK and ATP was generated based on NMR structural studies. The model indicates that the peptide binds ATP as intended, with the tryptophans forming stacking interactions with the adenine and the lysines making electrostatic interactions with the phosphate groups. This peptide mimics the recognition mode of OB-fold domains, and is a minimalist model system for understanding the driving forces for nucleotide binding by peptides and proteins.

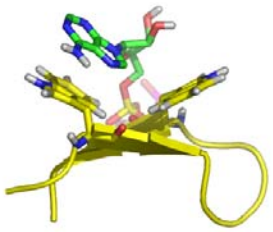


Figure 1.5. Computational model of peptide WKWK interacting with ATP. The model is based on NMR structural data with WKWK shown in yellow and Trp residues drawn. The adenine base is shown in green with the triphosphate group in red and white.

C. Model Systems for Protein Structure-Function Studies

Model systems have been used for many years to study protein structure and function. The first studies of beta-sheet peptides emerged long after that of alpha-helices due to problems with solubility and aggregation in water.[14] Many of the first model systems were designed to understand protein folding and stability. In recent years, these model systems have been applied to the study of a particular function such as DNA binding. Original model systems used short peptide sequences to probe folding. More recently, β-hairpin peptides have been used to understand specific interactions such as sidechain-sidechain interactions, π-π stacking interactions, and cation-π interactions.[15] Using small peptides to characterize such driving forces of stability can provide great insight into how these interactions work

[14] Blanco, F. J.; Jimenez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887.

[15] (a) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126*, 2028-2034. (b) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 6518-6519. (c) Hughes, R. M.; Benshoff, M. L.; Waters, M. L. *Chem. Eur. J.* **2007**, *13*, 5753-5764. (d) Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-1531. (e) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 388-395. (f) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303-1324.

together to form and stabilize larger protein complexes. When these model systems are applied to the function of particular proteins, one can gain knowledge of specific binding interactions and enzymatic activity for the full proteins using a simpler system.

WW domains have been used for many years in the study of protein folding[16] and in understanding the interactions between WW domain peptides and their natural ligand, the polyproline helix.[17] WW domains are thought to be one of the smallest classes of fully functional miniature proteins, exhibiting cooperativity in folding as well as binding their natural ligand with high affinity. These peptides are certainly some of the most well characterized peptides known. Because of this and because β-sheet peptides and proteins are

---

[16] (a) Jager, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10648-10653. (b) Fernandez-Escamilla, A. M.; Ventura, S.; Serrano, L.; Jimenez, M. A. *Protein Sci.* **2006**, *15*, 2278-2289. (c) Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3954-3959. (d) Kraemer-Pecore, C. M.; Lecomte, J. T. J.; Desjarlais, J. R. *Protein Sci.* **2003**, *12*, 2194-2205. (e) Jiang, X.; Kowalski, J.; Kelly, J. W. *Protein Sci.* **2001**, *10*, 1454-1465. (f) Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373-393. (g) Macias, M. J.; Gervais, V.; Civera, C.; Oschkinat, H. *Nat. Struct. Biol.* **2000**, *7*, 375-379. (h) Koepf, E. K.; Petrassi, M.; Ratnaswamy, G.; Huff, M. E.; Sudol, M.; Kelly, J. W. *Biochemistry* **1999**, *38*, 14338-14351. (i) Nguyen, H.; Jager, M.; Moretto, A.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3948-3953. (j) Jager, M.; Dendle, M.; Fuller, A. A.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 2306-2313.

[17] (a) Kato, Y.; Miyakawa, T.; Kurita, J.; Tanokura, M. *J. Biol. Chem.* **2006**, *281*, 40321-40329. (b) Pires, J. R.; Parthier, C.; do Aido-Machado, R.; Wiedemann, U.; Otte, L.; Bohm, G.; Rudolph, R.; Oschkinat, H. *J. Mol. Biol.* **2005**, *348*, 399-408. (c) Ball, L. J.; Kuhne, R.; Schneider-Mergener, J.; Oschkinat, H. *Angew. Chem. Int. Ed.* **2005**, *44*, 2852-2869. (d) Kato, Y.; Nagata, K.; Takahashi, M.; Lian, L.; Herrero, J. J.; Sudol, M.; Tanokura, M. *J. Biol. Chem.* **2004**, *279*, 31833-31841. (e) Bedford, M. T.; Chan, D. C.; Leder, P. *EMBO J.* **1997**, *16*, 2376-2383. (f) Otte, L.; Wiedemann, U.; Schlegel, B.; Pires, J. R.; Beyermann, M.; Schmieder, P.; Krause, G.; Volkmer-Engert, R.; Schneider-Mergener, J.; Oschkinat, H. *Protein Sci.* **2003**, *12*, 491-500. (g) Zarrinpar, A.; Lim, W. A. *Nat. Struct. Biol.* **2000**, *7*, 611-613. (h) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Peptide Sci.* **2005**, *80*, 303-311. (i) Dalby, P. A.; Hoess, R. H.; DeGrado, W. F. *Protein Sci.* **2000**, *9*, 2366-2376. (j) Chan, D. C.; Bedford, M. T.; Leder, P. *EMBO J.* **1996**, *15*, 1045-1054. (k) Macias, M. J.; Hyvonen, M.; Baraldi, E.; Schultz, J.; Sudol, M.; Saraste, M.; Oschkinat, H. *Nature* **1996**, *382*, 646-649. (l) Chen, H. I.; Sudol, M. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7819-7823.

known to bind DNA, the WW domain was employed as a model system for nucleotide recognition in this laboratory.

There are several advantages to using model systems in structure-function studies. In the case of binding, specific interactions at the binding site of large complexes are often difficult to deduce. While model systems only approximate native proteins, the task of understanding what specific forces drive the interaction of interest may be simplified by studying the interactions in a designed mini-protein or peptide. Studying the interaction between methylated DNA and the protein MBD1, for instance (Figure 1.6), can reveal many important features. An alternative to studying the full protein is to examine closely the region of the protein involved in the binding interaction. The region of the protein or the peptide that makes direct contacts with the DNA can then be synthesized quickly to study the interactions involved in the binding event. Once binding and structural studies are conducted, the knowledge gained can then be applied to the full system. Inhibitors can be made using this philosophy because learning about the interactions involved in binding can lead to a peptide or molecule designed to bind the target with high affinity and high selectivity. If this is achieved, the designed ligand may bind the target and inhibit the natural ligand from binding, as in the case of Robinson's β-hairpin peptidomimetics.[11] This method also lends itself to quickly incorporating different amino acids or functional groups for study of many peptide mutations. Another advantage to using model systems is that characterization of peptides and their interactions with other molecules is simplified because of their smaller size. This laboratory has used the WKWK peptide as a model system for

ATP, RNA, and DNA binding, and numerous peptides have been designed based on the original WKWK parent peptide sequence.[13, 18]

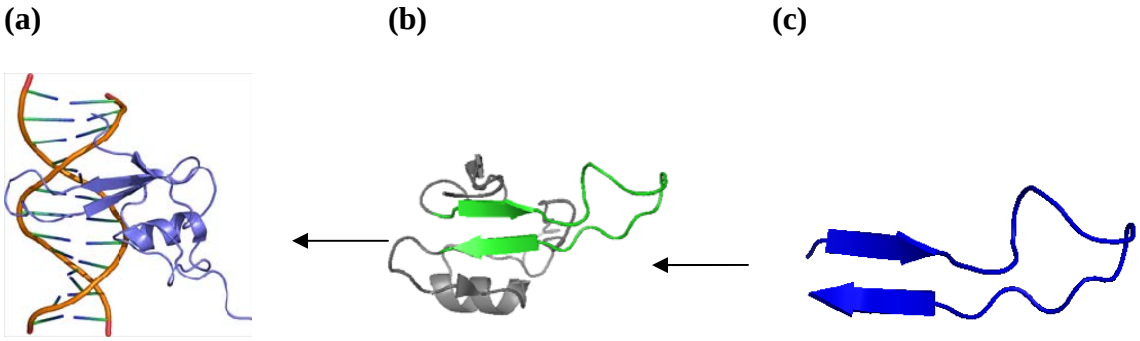


Figure 1.6. Structure of MBD1 β-hairpin, MBD1 protein, and MBD1 complexed with methylated DNA. (a) Full protein-DNA complex (pdb: lig4). (b) Full MBD1 protein. (c) β-hairpin involved in methylated DNA binding.

One advantage of designing peptides as receptors for recognition of nucleotides is that this type of design is also amenable to easily making mutations to the original design and gives the ability to add unnatural amino acids or other molecules. Peptides also allow the possibility for combinatorial chemistry which could provide attractive sequences and structures not likely considered by rational design. Cyclic peptides can be easily made by solid phase peptide synthesis, and d-amino acids can be incorporated into sequences. Both cyclic peptides[19] and peptides made with D-amino acids[20] have been shown to add stability to designed peptides.

D. Analyzing Protein-DNA Interactions

i. Binding Studies

[18] Cline, L. L.; Waters, M. L. *Org. Biomol. Chem.* **2009**, *7*, 4622-4630.

[19] Cowell, S. M.; Gu, X.; Vagner, J.; Hruby, V. J. *Methods Enzymol.* **2003**, *369*, 288-297.

[20] Samson, I.; Rozenski, J.; Samyn, B.; Van Aerschot, A.; Van Beeumen, J.; Herdewijn, P. *J. Biol. Chem.* **1997**, *272*, 11378-11383.

There are several methods in which one can determine binding interactions and recognition events between proteins and nucleotides. Here we have used β-hairpin peptides and WW domain peptides as models to study DNA recognition. The peptides used in these studies include at least one Trp residue in their sequences. This fact gives us the advantage of being able to use fluorescence quenching to quantify the binding interaction between peptides and DNA. Quenching experiments take advantage of the intrinsic fluorescence properties of a molecule. In this case, the amino acid tryptophan, which is included in the peptides and proteins studied, has fluorescent properties which allow one to observe and quantify its interactions with other proteins or nucleotides. An advantage of this technique is that nothing needs to be attached to the peptide which may affect its binding properties. A fluorophore could interact with the peptide or protein to which it is attached instead of the peptide or nucleotide which it should be binding. This type of unintended interaction leads to inaccurate binding data. The results of fluorescence quenching studies indicate a direct interaction between the two molecules without any disruption by a tag.

A disadvantage of this method when used with nucleotides is that nucleotides have absorbance spectra that overlap with the excitation wavelength of tryptophan. As seen in Figure 1.7, the DNA absorbs light in the same region as does Trp. This means that during fluorescence quenching experiments, DNA absorbs some light which decreases the excitation of the Trp residues, causing what appears to be quenching of the tryptophans. To account for this, an inner filter effect correction must be made to accurately determine binding constants. The calculations for each of these can be found in each chapter as appropriate.

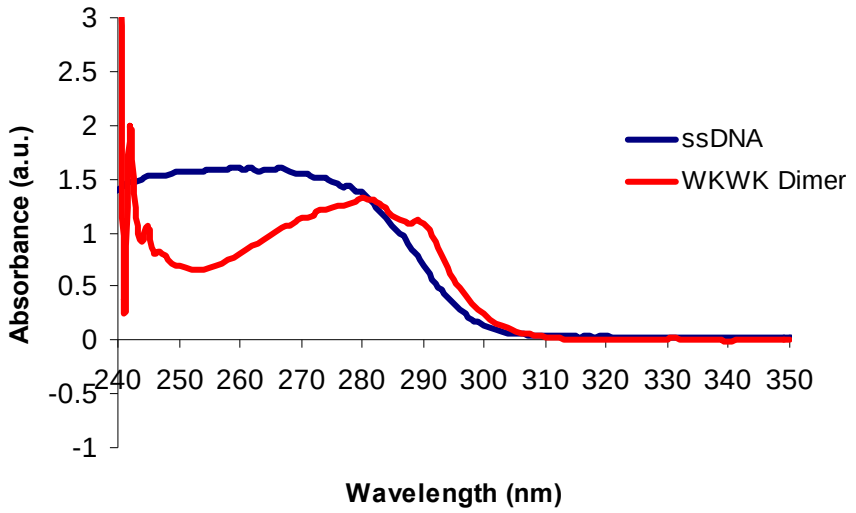


Figure 1.7. Absorbance spectra for ssDNA and WKWK peptide used in binding studies. Conditions: 10mM Na_2HPO_4, 100 mM NaCl, pH 7; ~8 μM peptide; ~ 12 μM ssDNA.

The inner filter effect calculations limit the conditions that can be used in binding studies. In order to be able to use optimal conditions, an alternative method was employed in some cases. Another method utilized to study protein-protein and protein-nucleic acid interactions is fluorescence anisotropy. Anisotropy requires a fluorophore to be added to one molecule and measures the tumbling rate of that fluorescent molecule in solution (both before and after interaction with the nucleotide). Upon binding of a ligand, the complex is larger than the individual molecule and the tumbling rate decreases, leading to an increase in anisotropy.[21]

The peptides used in these studies are made by solid phase peptide synthesis, so they are amenable to synthetic perturbation and mutation, including the addition of fluorophores. The fluorophore TAMRA was attached to a sidechain of some of the peptides, and fluorescence

[21] Bahr, M.; Valis, L.; Wagenknecht, H.-A.; Weinhold, E. *Nucleosides, Nucleotides, and Nucleic Acids* **2007**, *26*, 1581-1584.

anisotropy was conducted. Different methods were used in this report based on the conditions required for binding.

ii. Combinatorial Chemistry

There are two schools of thought regarding designing targets for nucleotides and proteins. One thought is that one should carefully study the interaction under investigation and design molecules that could inhibit the interaction by fitting into a specific pocket or by tightly binding the target molecule. The second school of thought is that there are a host of possible favorable interactions. These researchers favor the idea of creating a library of molecules to allow the best binding interactions to occur without rationally designing one specific compound. With library studies, researchers design a library rather than a particular molecule and allow for the variations in the library to provide different functional groups that will likely form favorable interactions in the target complex. In this case, the researcher rationally designs the library but allows the library to "decide" which compound is the best fit or the best binder. Subsequently, the "hits" from the library are further tested to determine the specific favorable properties involved in the binding interaction. Here we utilize methods of rational design and combinatorial chemistry in designing peptides for the recognition of ssDNA.

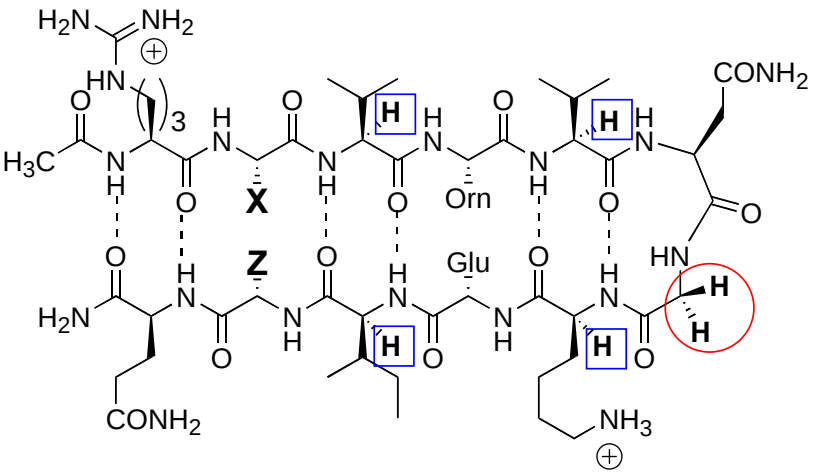
iii. Structural Studies

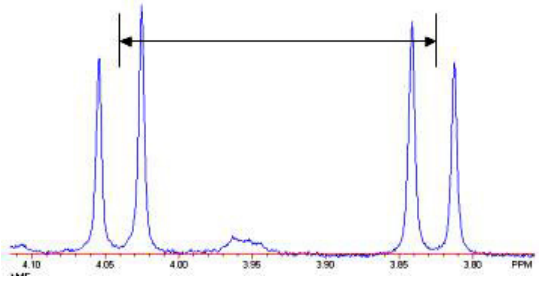
Structural studies were employed to better characterize the designed peptides. Two methods were used to gain the most insight into the stability of each peptide. Circular dichroism (CD) experiments were utilized to determine the global structure of peptides. NMR experiments were also used to attain specific information about how each amino acid contributes to the structure of the peptides. Two different NMR methods have been used,

and the first takes advantage of the glycine residue in the turn sequence of each of the peptides synthesized. The glycine hydrogens are diastereotopic. If a peptide is completely unstructured and the peptide backbone is fully extended, one signal or peak will result from a one-dimensional or two-dimensional experiment. In the case of structured peptides, however, the glycine hydrogens are in different environments and have different signals. Figure 1.8a depicts a peptide with the diastereotopic hydrogens circled in red and the resulting glycine NMR signal for that peptide. For structured peptides, the more well-folded the peptide, the larger the difference in glycine signals. This difference is termed glycine splitting and gives a measure of the folding or stability of the peptide. A comparison of the Gly splitting in the β-hairpin with that of a fully folded cyclic control peptide gives the extent of folding and overall stability for the peptide. Another method used to measure peptide folding is to determine the chemical shift differences of the α-hydrogens in the peptide. The extent of downfield shifting of the protons' chemical shifts in the hairpins relative to the corresponding random coil peptides indicates the degree of β-sheet structure for each residue along the peptide backbone. Downfield shifting of Hα protons is evidence of increased β-hairpin population, with a chemical shift difference of greater than 0.1 ppm taken to indicate β-sheet structure.[22] An example of this type of data is shown in Figure 1.8b.

[22] (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333. (b) Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996-2007.

(a)





(b)

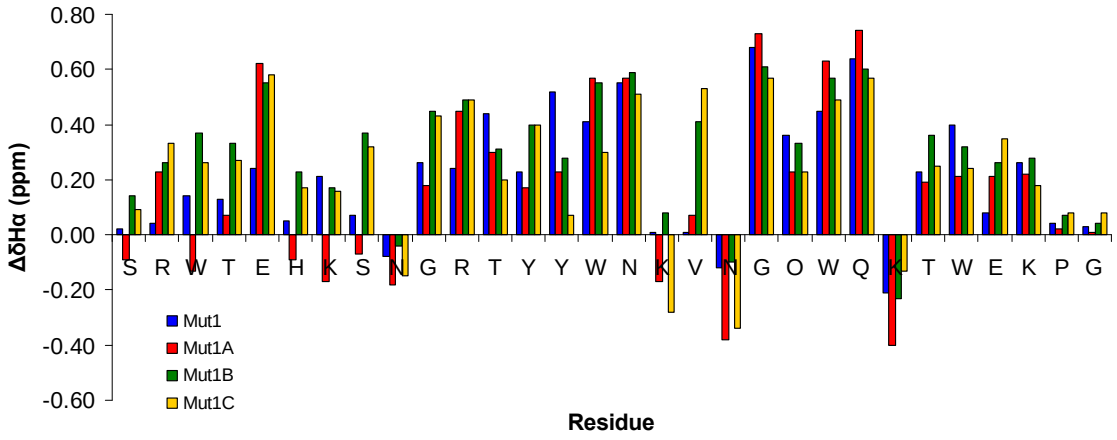


Figure 1.8. NMR characterization of peptide folding. (a) Structure of folded peptide and corresponding Gly splitting data. (b) Example of chemical shift deviations for alpha hydrogens.

E. Conclusions

The work presented here is a culmination of a variety of different experiments intended to understand the interactions that are crucial for β-sheet recognition of ssDNA. Two different model systems were utilized to study the driving forces for DNA binding by β-sheets. As stated above, there are many important proteins that have OB-fold domains or other DNA binding domains in which β-sheets or β-hairpins play key roles in recognition. A better understanding of the specific interactions involved will shed light on the mode of DNA recognition and how these domains function within the context of the full peptide or protein. Mimicking these domains may also lead to designs of inhibitors that could be used in various applications. Although we cover only a few of these examples in the following chapters, this introduction gives a broad overview of the types of interactions involved in recognition by OB-fold domains as well as the various applications which may be pursued as a result of studying these peptides and proteins.

Chapter II

Structural Effects on ss- and dsDNA Recognition by a β-Hairpin Dimer

(Reproduced, in part, with permission from Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544.)

A. Background and Significance

Protein-DNA interactions are involved in a variety of biological processes such as transcription, translation, and DNA repair. Protein interactions with single-stranded DNA (ssDNA) are particularly critical in DNA recombination, replication and repair,[1] telomere regulation,[2] and cold shock response.[3] Proteins recognize ssDNA through a combination of electrostatic, aromatic, and hydrogen-bonding interactions. Almost all proteins recognize

---

[1] (a) Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. *Nature* **1997**, *385*, 176-181. (b) Wold, M. S. *Annu. Rev. Biochem.* **1997**, *66*, 61-92.

[2] (a) Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. *Biochemistry* **2003**, *42*, 3751-3758. (b) Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. *Science* **2002**, *296*, 145-147.

[3] (a) Max, K. E. A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *FEBS J*. **2007**, *274*, 1265-1279. (b) Max, K. E. A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *J. Mol. Biol.* **2006**, *360*, 702-714. (c) Hillier, B. J.; Rodriguez, H. M.; Gregoret, L. M. *Folding Des.* **1998**, *3*, 87-93. (d) Newkirk, K.; Feng, W.; Jiang, W.; Tejero, R.; Emerson, S. D.; Inouye, M.; Montelione, G. T. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 5114-5118. (e) Schindelin, H.; Marahiel, M. A.; Heinemann, U. *Nature* **1993**, *364*, 164-168. (f) Schnuchel, A.; Wiltscheck, R.; Czisch, M.; Herrier, M.; Willimsky, G.; Graumann, P.; Marahiel, M. A.; Holak, T. A. *Nature* **1993**, *364*, 169-171.

ssDNA through an OB-fold domain (oligonucleotide/oligosaccharide-binding fold)[4] which is a solvent-exposed β-sheet surface composed of two three-stranded anti-parallel β-sheets that form a five-stranded beta-barrel in which the first strand is in both sheets (Figure 1).[3a, 4] Replication protein A (RPA)[1] and cold shock protein[3] are well-studied examples of proteins containing single-stranded DNA binding domains. This laboratory has designed a β-hairpin peptide dimer that functions as a minimalist mimic of the OB-fold for recognition of ssDNA.[5] This peptide was designed to recognize nucleic acids and to determine the favorable interactions that provide affinity and selectivity for recognition of ssDNA.

[4] (a) Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115-133. (b) Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004**, *14*, 36-42. (c) Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

[5] Butterfield, S. M.; Cooper, W. J., Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25.

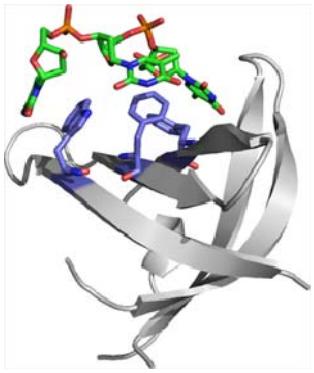


Figure 2.1. Structure of the *Bacillus caldolyticus* cold shock protein bound to oligothymidine. Cold shock protein is shown in gray; oligothymidine is highlighted in green; the aromatic sidechains involved in binding are shown in blue (pdb: 2hax).[3a]

Originally, the peptide WKWK (Figure 2.3a) was designed to bind nucleotides through a binding cleft of two tryptophan residues and two flanking cationic lysine residues. The name of the peptide reflects the residues on the binding face of the beta-hairpin.[6] The peptide is based on Gellman's 12-residue beta-hairpin peptide used to study turn nucleators.[7] The design employs a diagonal Trp-Trp pair into the sequence which is similar to that in Cochran's Trpzip peptide,[8] with the Trp pair intended to form a cleft for nucleotide base recognition. The Lys residues were placed in the opposite diagonal positions to provide electrostatic interactions with the phosphate groups. Both aromatic and cationic residues are

[6] (a) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581. (b) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70*, 1105-1114.

[7] Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236-4237.

[8] Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 5578-5583.

known to be crucial for DNA binding by OB-fold proteins,[9] so these were key factors for the design of an OB-fold mimic. NMR experiments revealed a well-folded β-hairpin conformation for the peptide.[6] Tryptophan's natural fluorescence properties were employed for this study, and the peptide was shown by fluorescence quenching of the tryptophan residues to bind ATP with a K_d of 170 μM.[6a] Through a combination of salt studies, mutation studies, and NMR and circular dichroism (CD) investigations of the peptide-ATP complex, the binding mode of WKWK was characterized. Binding was found to be driven by both aromatic and electrostatic interactions. Aromatic interactions between the nucleobase and the cleft made up of the two Trp sidechains were found to contribute about 2 kcal/mol to binding, whereas electrostatic interactions between basic residues on the peptide and the triphosphate contribute about 3 kcal/mol.[6a] Dimerization of WKWK via a disulfide bond to give $(WKWK)_2$ (Figure 2.3b) provided two aromatic pockets for binding to oligonucleotides, and the experiments were extended to short single-stranded DNA oligomers (three deoxy-pentanucleotides, A, T, and C).[5] Binding to penta-dG was not attempted because sequences of four or more guanosine bases are known to form a G-quadruplex structure.[10] The binding constant for penta-dA was determined to be 12 μM, which is greater than the binding affinities for penta-dT and penta-dC. The interaction between $(WKWK)_2$ and the dA_5 sequence was found to be between -0.3 and -0.6 kcal/mol

---

[9] (a) Kloks, C. P.; Spronk, C. A.; Lasonder, E.; Hoffmann, A.; Vuister, G. W.; Grzesiek, S.; Hilbers, C. W. *J. Mol. Biol.* **2002**, *316*, 317-326. (b) Schroder, K.; Graumann, P.; Schnuchel, A.; Holak, T. A.; Marahiel, M. A. *Mol. Microbiol.* **1995**, *16*, 699-708.

[10] (a) Neidle, S.; Parkinson, G. N. *Curr. Opin. Struct. Biol.* **2003**, *13*, 275-283. (b) Simonsson, T., *Biol. Chem.* **2001**, *382*, 621-628. (c) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718-720. (d) Sundquist, W. I.; Klug, A. *Nature* **1989**, *342*, 825-829. (e) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871-880.

stronger than the interaction with the other pentanucleotides.[5] Upfield shifting of the Trp aromatic protons and of the adenine aromatic protons in dA_5 indicates that stacking interactions contribute to binding. Comparison of the studies with the monomer and the dimer indicate that stacking of adenine between the Trp residues occurs in WKWK and that adenine bases bind within both binding pockets in the presence of the dimer, as depicted in Figure 2.2.[5-6] Mutation studies have also shown that the two Lys residues on the binding face each contribute ~ -1.5 kcal/mol to binding[6a] and that Ornithine 8, which was incorporated for water solubility, also contributes significantly to binding.[11]

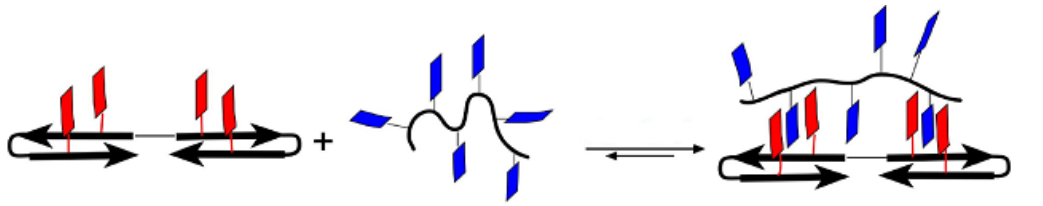


Figure 2.2. Model for beta-hairpin dimer binding to an oligonucleotide. Red indicates Trp residues; blue indicates nucleobases. Adapted from Butterfield, et al.[6b]

The binding of a randomized single-stranded 11mer (Figure 2.3g) was also analyzed, and the peptide dimer was shown to bind the ssDNA sequence with a K_d of 3 μM in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, at 298 K.[5] The higher affinity for this oligonucleotide is due to the increase in favorable contacts present in the longer sequence. This binding affinity is comparable to that found in the cold shock protein A (K_d ~ 6 μM), which consists of a single OB-fold.[3c] Binding of $(WKWK)_2$ to a duplex DNA strand of the same length (Figure 2.3h) was found to display similar affinity (K_d ~ 5-9 μM), despite the doubling of the negative charge. Studies have shown that aromatic interactions between

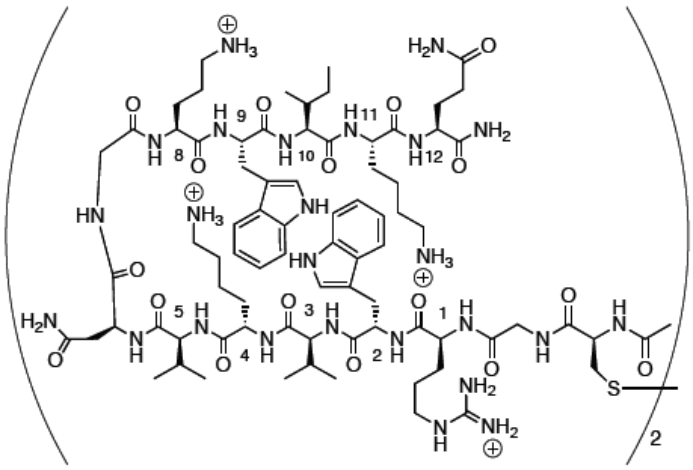
[11] Hughes R. M.; Waters, M. L. unpublished results.

$(WKWK)_2$ and the unpaired bases in ssDNA provide selectivity over dsDNA under high salt conditions in which the electrostatic interactions are screened.[5] In duplex DNA, the bases are packed inside the duplex and are not accessible to stacking with residues in the peptide. Thus, the preference for ssDNA over dsDNA at physiological salt is modest, but it is a starting point for a system to improve selectivity for ssDNA over the duplex structure. These studies aimed to explore the mode of binding to ss- and dsDNA by the hairpin dimer to provide insight into approaches for optimizing selectivity. More specifically, in this study the contributions of beta-hairpin structure, chirality, and electrostatic interactions to binding affinity and selectivity for ssDNA and double-stranded DNA (dsDNA) were investigated.

(a) WKWK: Ac-Arg-**Trp**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-NH_2

(b) $(WKWK)_2$:

(Ac-Cys-Gly-Arg-**Trp**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-$NH_2)_2$



(c) $(WKWK\text{-}N6P)_2$:

(Ac-Cys-Gly-Arg-**Trp**-Val-**Lys**-Val-<u>Pro</u>-Gly-Orn-**Trp**-Ile-**Lys**-Gln-$NH_2)_2$

(d) $(WKWK\text{-}Scrambled)_2$:

(Ac-Cys-Gly-Lys-Trp-Val-Arg-Trp-Ile-Lys-Gln-Asn-Orn-Val-Gly-$NH_2)_2$

(e) $(D\text{-}WKWK)_2$:

(Ac-DCys-Gly-DArg-D**Trp**-DVal-D**Lys**-DVal-DAsn-Gly-DOrn-D**Trp**-DIle-D**Lys**-DGln-$NH_2)_2$

(f) $(WKWK\text{-}R1Q)_2$:

(Ac-Cys-Gly-<u>Gln</u>-**Trp**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-$NH_2)_2$

(g) ssDNA: 5’-CCATCGCTACC-3’

(h) dsDNA: 5’-CCATCGCTACC-3’
3’-GGTAGCGATGG-5’

Figure 2.3. (a) β-hairpin sequence WKWK. (b) Sequence and structure of peptide $(WKWK)_2$. The bold residues are those thought to be involved in the nucleic acid binding pocket. Those residues are all on the same face of the peptide. (c-f) Mutations to $(WKWK)_2$ with specific mutations underlined. (g) ssDNA sequence. (h) dsDNA sequence.

B. Results and Discussion

i. Sequence Design

Starting with the previously reported β-hairpin dimer, $(WKWK)_2$, as a template, two peptides were designed as controls for structure. Peptide $(WKWK\text{-}N6P)_2$ (Figure 2.3c) exchanges the Asn residue with L-Pro, which has been shown to disrupt the turn,[7] and was designed to determine if structure is important for binding of the peptide to the ssDNA. The L-Pro-Gly sequence does not promote a typical turn like that of the Asn-Gly and the D-Pro-Gly turn sequences (Type I' and Type II', respectively), but rather distorts the turn such that the peptide has no secondary structure. A scrambled peptide, $(WKWK\text{-}Scrambled)_2$ (Figure 2.3d), was also made to investigate the importance of sequence and structure on binding. This peptide has no turn sequence and therefore should have no structure. In addition, the aromatic residues have been scrambled so that no aromatic pocket would be formed if it did take on a β-hairpin structure via induced fit binding, for example. This second structure control was designed to provide additional information regarding how structure affects binding to DNA. Taken together, these peptides allow us to probe directly whether the aromatic cleft in the beta-hairpin contributes to binding.

A third peptide, $(D\text{-}WKWK)_2$ (Figure 2.3e), was synthesized with all D-amino acids to determine the importance of chirality in binding. While $(WKWK)_2$ has a right-handed twist, as do all β-hairpins with all L-amino acids, $(D\text{-}WKWK)_2$ takes on a left-handed twist. Whether the macromolecular chirality of the peptide influences the binding affinity will depend on the mode of binding to ss- and dsDNA.

Finally, $(WKWK\text{-}R1Q)_2$ (Figure 2.3f), was designed with the N-terminal arginine residue mutated to a glutamine to give a neutral side chain at that position, leaving the dimer

with two fewer positive charges. This residue is on the opposite face of the hairpin as the aromatic cleft, and it is less structured due to its position at the N-terminus, which is frayed, as determined by the NMR spectrum. Initial experiments determined that Arg residues exhibited chemical shift deviations upon binding, suggesting that the Arg residues are involved in the binding event.[5] In addition, we were curious whether reducing the net charge of the peptide by +2 (since there is an Arg in each of the peptides in the dimer) would have a larger effect on dsDNA, with its greater charge, than ssDNA, as was predicted from salt studies reported previously.[5] Thus, we aimed to determine whether the arginine residues at that position contributed significantly to binding.

ii. Characterization of Structure

Circular dichroism spectra were obtained to verify the secondary structure of each of the mutants studied. The parent peptide, $(WKWK)_2$, has a characteristic β-sheet spectrum with a minimum at 210 nm (Figure 2.4). Typically, β-sheet peptides and proteins have a minimum at 215 nm. The shift for these peptides is likely due to the contributions of the Trp residues in the sequence. This is consistent with previous NMR data which indicates that both the monomer and dimer have high β-hairpin populations under physiological conditions.[5-6] The CD spectra of both $(WKWK\text{-}N6P)_2$ and $(WKWK\text{-}Scrambled)_2$, verified that they were unstructured, each having a peak with a minimum at about 200 nm (Figure 2.4). The CD spectrum of $(D\text{-}WKWK)_2$, containing all D-amino acids, displayed the expected signal for the enantiomer of the parent peptide $(WKWK)_2$, with a positive peak at 210 nm which indicates a D-peptide with β-sheet structure (Figure 2.4). The R1Q mutation in peptide $(WKWK\text{-}R1Q)_2$ did not change the overall structure of the peptide and gave a signal consistent with β-sheet structure with a minimum at 210 nm (Figure 2.4).

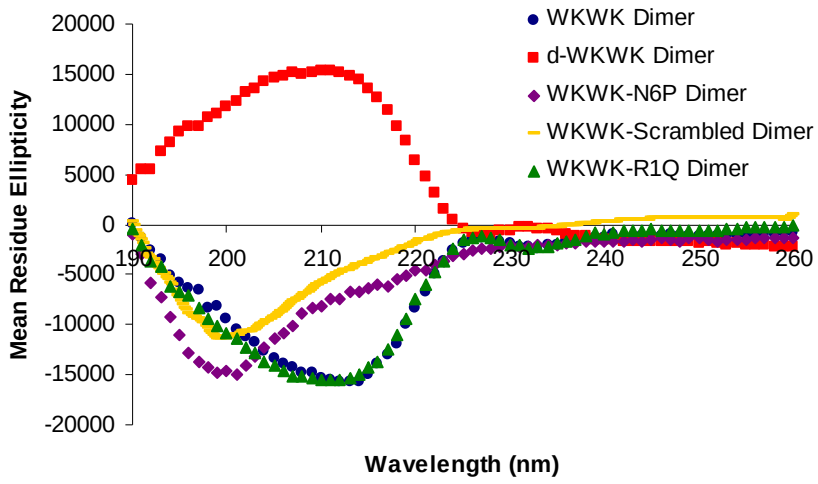


Figure 2.4. CD data for WKWK dimer and mutants. Data was obtained at 30 μM peptide concentrations in 10 mM Na_2HPO_4. Scans were performed from 260-190 nm at 298 K.

iii. Binding Measurements

Binding of ssDNA and duplex DNA to each of the peptides was determined by quenching of the Trp fluorescence as described in the Experimental Procedures. A correction for the inner filter effect arising from absorbance of the nucleobases at the excitation wavelength of Trp was performed for all binding data except for the binding of $(WKWK\text{-}N6P)_2$ to duplex DNA, where binding was weak (see Experimental Procedures). In this case, the inner filter effect was too large to correct for, so the reported dissociation constant is uncorrected and represents a lower limit. In the case of $(WKWK\text{-}N6P)_2$, CD mixing experiments were also performed to confirm that folding of the peptide was not induced upon binding to duplex DNA. This was also repeated for $(WKWK)_2$. Neither spectrum showed a change in folding due to binding (see Experimental Section).

The stoichiometry of binding was determined for $(WKWK)_2$ by fluorescence quenching using the molar variation method (see Experimental Section). A 1:1 binding stoichiometry for the interaction between $(WKWK)_2$ and the 11-mer ssDNA sequence (Figure 2.3g) was determined (Figure 2.5), in agreement with the stoichiometry reported for

the $(WKWK)_2$ interaction with dA_5.[5] The binding stoichiometry for the interaction between $(WKWK)_2$ and the duplex DNA sequence (Figure 2.3h) was also determined to be 1:1 (Figure 2.6).

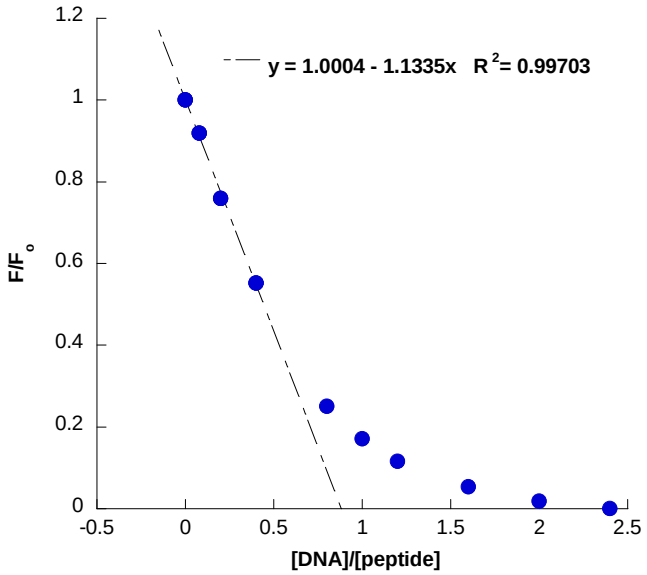


Figure 2.5. Molar variation plot for $(WKWK)_2$ (25 μM) binding to ssDNA (Figure 2.3g; 0-60 μM). A 1:1 binding interaction was demonstrated with the average of two runs in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, 298 K.

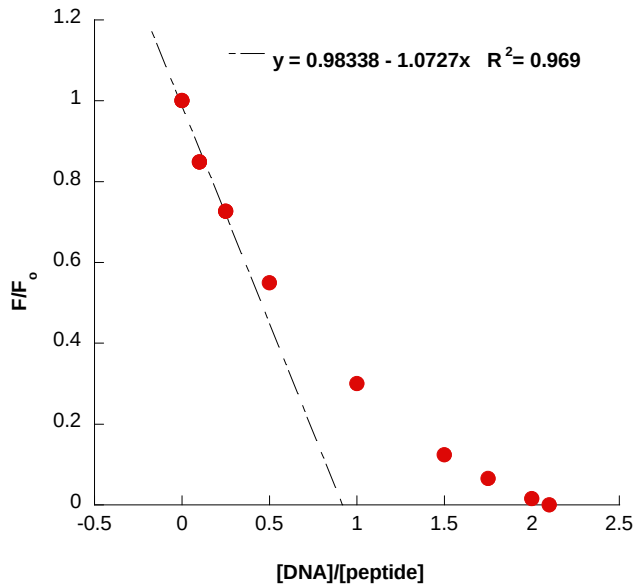


Figure 2.6. Molar variation plot for $(WKWK)_2$ (20 μM) binding to duplex DNA (Figure 2.3h; 0-42 μM). A 1:1 binding interaction was determined in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, 298 K.

iv. Influence of Folding on Binding

To probe the importance of the aromatic pocket in the β-hairpin dimer, we measured binding of the unstructured peptide, $(WKWK\text{-}N6P)_2$, to ss- and dsDNA. In both cases, binding was significantly weaker than with $(WKWK)_2$ (Figures 2.7 and 2.8b; Table 2.1). Binding is nearly 10-fold worse for both ssDNA and dsDNA. (Note that due to the small change in tryptophan fluorescence resulting from weak binding, no accurate dissociation constant could be obtained for the WKWK-N6P dimer with dsDNA). However, evidence suggests that the reason for the loss of binding upon disruption of beta-hairpin structure differs for ss- and dsDNA. In the case of ssDNA, the decrease in binding affinity associated with loss of structure is consistent with previous data indicating that interaction of ssDNA with the beta-hairpin is mediated by stacking interactions with the aromatic pocket.[5-6]

As for dsDNA, the role of structure on binding is less clear. The influence of folding on binding is surprising given that previous studies indicated that binding of $(WKWK)_2$ is primarily electrostatic in nature. Differences in charge density of the folded and unfolded peptides or differences in the entropic cost of binding may play a role. Another possibility is that binding might occur in a structurally defined manner, such as groove binding.

The peptide $(WKWK\text{-}Scrambled)_2$ also has a weaker binding affinity for ssDNA than the structured β-hairpin, $(WKWK)_2$. The scrambled peptide (K_d = 17.4 μM; Figure 2.7; Table 2.1) binds more strongly to ssDNA than does $(WKWK\text{-}N6P)_2$ but still does not bind as tightly as the structured peptide. This further suggests that the lack of structure of the peptide causes the decrease in binding affinity. The difference in binding of $(WKWK\text{-}N6P)_2$ and $(WKWK\text{-}Scrambled)_2$ may have to do with differences in their structures and with differences in spacing of charged residues. Since charged residues are known to bind tightly to phosphates in the DNA backbone, the lack of structure could allow the positively charged residues easier access to the phosphate groups which could improve the binding affinity. This result, however, likely decreases the selectivity the peptide has for ssDNA, thereby reducing its effectiveness as an OB-fold mimic. The kink in the WKWK-N6P dimer possibly prohibits easy access of cationic residues to bind the phosphates as in the scrambled sequence. Also of importance is that the conditions for binding of $(WKWK\text{-}Scrambled)_2$ to ssDNA are not ideal, as the peptide concentration of 8 μM is much higher than one-tenth the dissociation constant. This could generate a K_d value that suggests tighter binding than is actually occurring. Binding studies for the interaction between $(WKWK\text{-}Scrambled)_2$ and duplex DNA were not attempted.

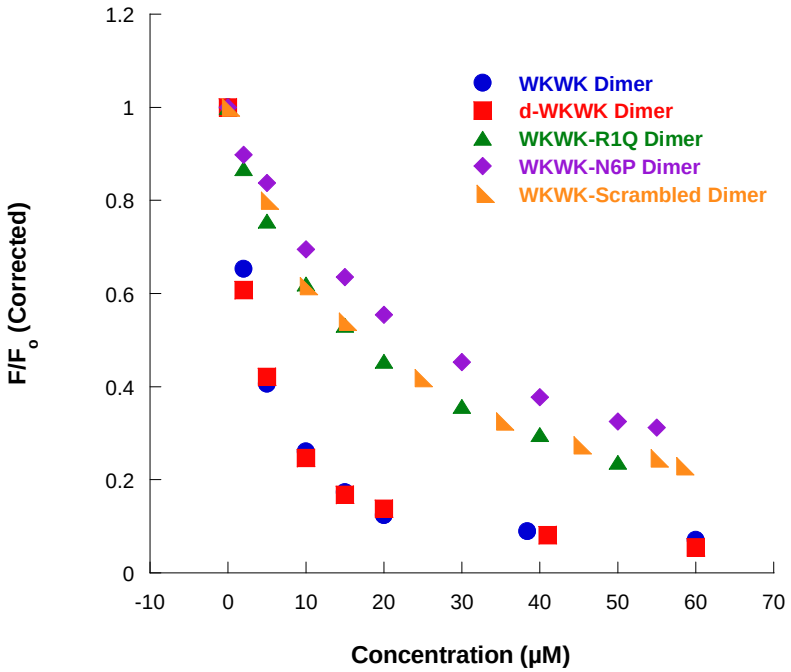


Figure 2.7. Fluorescence titrations of $(WKWK)_2$ and mutant peptide dimers with the single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 2.3g) corrected for the inner filter effect (see Experimental Section); 2-8 μM peptide concentrations; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

v. Influence of Chirality on Binding

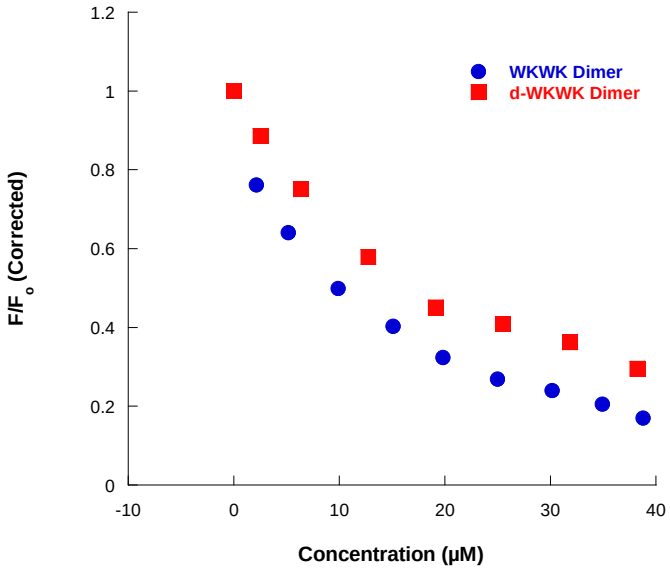
As the peptide structure is important to binding, the chirality of the peptide might be important as well. In particular, it may be more important in binding to duplex DNA due to its macromolecular chirality arising from the double helix, than to single-stranded DNA, which is unstructured.

Fluorescence binding studies of the all D-amino acid peptide $(D\text{-}WKWK)_2$ and ssDNA gave binding affinities identical to those of its enantiomer (Figure 2.7; Table 2.1). In the case of ssDNA, with no macromolecular structure, this suggests that the ribose ring is not directly involved in binding. This is reminiscent of the binding mode for cold shock protein

and ssDNA, in which the ribose rings are solvent-exposed and are not in direct contact with the protein (Figure 2.1).[3a] The fact that the all D-peptide binds with equal affinity to the ssDNA sequence is also significant because all D-peptides typically display enhanced resistance to cleavage by proteases.[12]

[12] (a) Hamamoto, K.; Kida, Y.; Zhang, Y.; Shimizu, T.; Kuwano, K. *Microbiol. Immunol.* **2002,** *46*, 741-749. (b) Tugyi, R.; Uray, K.; Ivan, D.; Fellinger, E.; Perkins, A.; Hudecz, F. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 413-418.

(a)



(b)

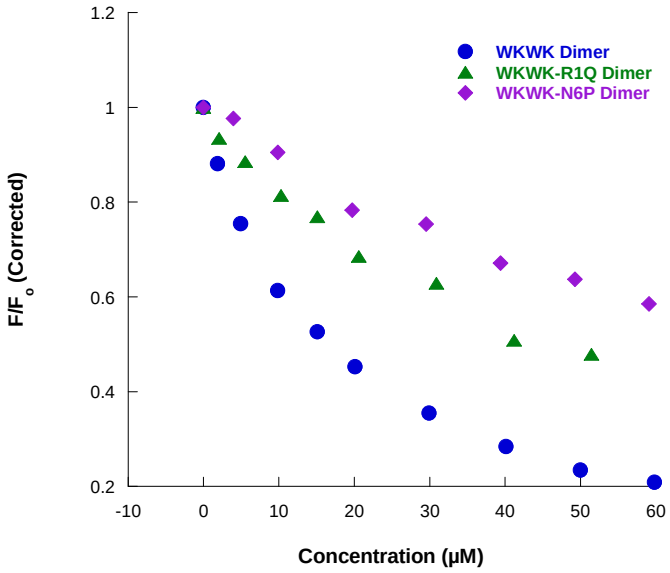


Figure 2.8. Fluorescence titrations of $(WKWK)_2$ and mutant peptide dimers with duplex DNA (Figure 2.3h); 2-7 μM peptide, 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. (a) $(WKWK)_2$ and $(D\text{-}WKWK)_2$ corrected for the inner filter effect (see Experimental Section). (b) $(WKWK)_2$, $(WKWK\text{-}R1Q)_2$, and $(WKWK\text{-}N6P)_2$ uncorrected for the inner filter effect.

In contrast to its binding to ssDNA, the enantiomeric peptide dimer (D-WKWK)$_2$ binds to duplex DNA with a somewhat weaker affinity than does (WKWK)$_2$ (Figure 2.8a; Table 2.1). Thus, chirality has some influence on binding to duplex DNA, unlike to ssDNA. For duplex DNA, the small but measurable difference in the binding affinities of the L- and D-peptides indicates that the diastereomeric complexes are of different stabilities. This is consistent with the finding that folding influences binding of the parent peptide to dsDNA and provides further support for a structurally defined binding event such as groove binding, in which the twist of the hairpin would be expected to influence its binding affinity.

Table 2.1. Affinity constants for binding of peptide dimers to ssDNA and dsDNA sequences.[a]

Peptide	ssDNA K_d, μM (error)	dsDNA K_d, μM (error)
(WKWK)$_2$	3.5 (0.2)	9.2 (0.9)[b]
(WKWK-N6P)$_2$	24.7 (1.6)	$\geq$ 80[c]
(WKWK- Scrambled)$_2$	17.4 (1.0)	n. d.
(WKWK-R1Q)$_2$	21.9 (2.4)	$\geq$ 46[c]
(D-WKWK)$_2$	3.4 (0.2)	17.5 (1.3)

(a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting. (b) This K_d was previously reported as 5 μM. See reference 5. (c) This value is not corrected for the inner filter effect and thus is only approximate. n. d. denotes a measurement that was not determined.

vi. Effect of Electrostatic Interactions on Binding

The K_d determined for the interaction between (WKWK-R1Q)$_2$ and ssDNA was 18.5 μM, as compared to 3.5 μM for (WKWK)$_2$ (Figure 2.7; Table 2.1). This amounts to a loss of 1 kcal/mol in binding affinity or 0.5 kcal/mol per Arg residue. This is less than the contribution of the Lys residues on the binding face of the hairpins, which have been

determined to contribute 1.5 kcal/mol each.[6] While Lys, Orn, and Arg are all thought to be capable of participating in similar electrostatic interactions with phosphate groups, their positioning within the peptide seems to be a primary difference and distinguishing factor in how much these cationic residues contribute to binding. The frayed N-terminus in which the Arg is placed in WKWK dimer positions the amino acid such that it does not interact with the phosphate groups as well as the Lys residues which are on the binding face and which closely contact the DNA. The Orn in $(WKWK)_2$ is also on the opposite face of the peptide from the binding pocket, but its close proximity to the binding cleft seems to contribute to its importance in binding. These results are consistent with previous research asserting that electrostatic interactions are important for DNA recognition within the OB-fold domain.[9] The results also indicate that position 1 may be a good position to reduce the net charge of the peptide in order to reduce nonspecific electrostatic interactions to other biomolecules such as duplex DNA or RNA.

Binding of $(WKWK\text{-}R1Q)_2$ to duplex DNA was found to be weak such that the inner filter effect was too large to allow for accurate correction. A lower limit of $\geq$ 46 μM was estimated from the uncorrected binding curve (Figure 2.8b; Table 1; see Experimental Section). Binding is certainly weaker than this, since the inner filter effect results in an enhanced apparent binding affinity. Thus, the reduction of net charge from +8 in $(WKWK)_2$ to +6 in $(WKWK\text{-}R1Q)_2$ results in at least a ten-fold reduction in binding affinity for duplex DNA. This compares to a reduction of only about 5-fold for ssDNA. Thus, the Arg in the first position contributes to the binding affinity for both ss- and dsDNA, but the Gln mutation increases selectivity for ssDNA over dsDNA. This is consistent with the greater salt-

dependence observed for binding of $(WKWK)_2$ to duplex DNA relative to ssDNA reported previously.[5]

vii. Structural Models for Binding ss- and dsDNA

The studies performed here, in conjunction with our previous studies, are consistent with a binding model in which ssDNA interacts with the hairpin dimer via a combination of aromatic and electrostatic interactions defined by the secondary structure of the peptide. In the parent peptide, $(WKWK)_2$, recognition of ssDNA is approximately 1.5 to 3-fold more favorable than recognition of duplex DNA, despite the fact that the net charge on the duplex is twice that of the single-stranded oligonucleotide. This selectivity is ascribed to aromatic interactions between the unpaired bases and the aromatic cleft in the beta-hairpin.[5] Disruption of the hairpin structure results in a reduction in binding affinity of nearly ten-fold and is attributed to loss of the aromatic pocket. Because the ssDNA is unstructured, the chirality of the peptide does not influence binding. Thus, $(D\text{-}WKWK)_2$ provides selectivity for ssDNA over dsDNA by a factor of five. Lastly, we found that the net charge of the peptide influences affinity of the peptide to both ss- and dsDNA, but to different degrees. With a reduction of the net positive charge on the peptide from +8 to +6, the binding affinity was decreased by 5-fold for ssDNA, but by at least 10-fold for dsDNA. Thus, decreasing the net charge on the peptide reduces binding affinity for both ss- and dsDNA, but it increases selectivity for ss- over dsDNA.

The binding mode for duplex DNA is quite different from that of ssDNA. Previous studies indicated that binding is primarily electrostatic in nature, suggesting nonspecific

binding,[5] as has been observed with unstructured cationic peptides.[13] However, the importance of both folding and chirality to binding provide a more complex picture for binding. Also of significance is that the sequence-dependent difference in the influence of charged residues on binding again reinforces the importance of structure to binding. The decrease of negative charge also decreased the peptide's affinity for the peptide, but the binding is weakened more in the case of duplex DNA, likely because of the greater negative charge due to the number of phosphates in the duplex. The structure of the peptide also seems to affect the contributions of the electrostatic interactions. Thus, it is not only the number of positive charges that a peptide or protein possesses, but the position of the charge is also important for binding. This presents a unique, additional factor in the argument that structure contributes to binding both ssDNA and duplex DNA.

Taken together, these studies suggest that binding occurs through a structurally defined complex, likely involving groove binding of the hairpin to the duplex DNA. In this case, the Trp sidechains may simply provide a well-folded structure or may position hydrophobic contacts in the major groove (alternatively, the aliphatic sidechains on the opposite face of the hairpin may provide such contacts). Groove binding by a β-hairpin is not unprecedented in protein-DNA interactions. Binding of beta-hairpins within the major groove of duplex DNA has been observed in several protein-DNA complexes such as those shown in Figure 2.9.[14]

---

[13] (a) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1992**, *31*, 8932–8946. (b) Mascotti, D. P.; Lohman, T. M. *Biochemistry* **1993**, *32*, 10568–10579.

[14] (a) Allen, M. D.; Yamasaki, K.; Ohme-Takagi, M.; Tateno, M.; Suzuki, M. *EMBO J*. **1998**, *17*, 5484-5496. (b) Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J. G.; Ikegami, T.; Nakao, M.; Shirakawa, M. *Cell* **2001**, *105*, 487-497.

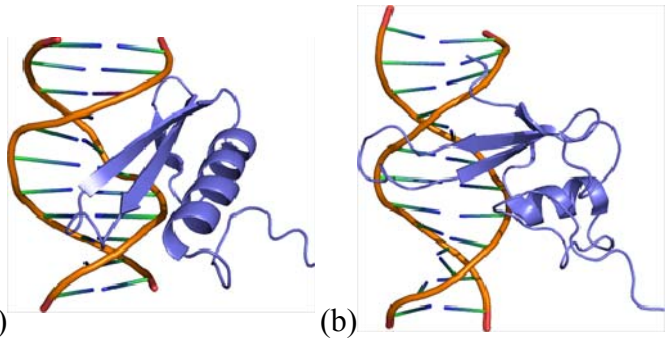


Figure 2.9. Protein-DNA complexes of (a) the GCC-box binding domain of AtERF1 (pdb: 1gcc)[14a] and (b) the methyl-CpG-binding domain of Human MBD1 (pdb: 1ig4).[14b]

A β-hairpin, for example, is key in recognizing the target sequence GCC-box in the GCC-box binding domain of AtERF1 (Figure 2.9a).[14a] The GCC-box binding domain (GBD) binds to the major groove of the GCC-box through a three-stranded β-sheet and an α-helix. The beta-sheet forms extensive interactions with the DNA, with Arg and Trp residues recognizing the DNA and forming contacts with eight base pairs in the major groove. Arginine and tryptophan also bind the sugar phosphate backbone. The α-helix is known to make only one contact with the DNA, binding a phosphate group. Interestingly, the GBD binds DNA similar to binding by zinc finger proteins except that the α-helix which binds to the major groove in the zinc finger is replaced by the β-hairpin in GBD.[14a]

In the human methylation-dependent transcriptional regulator methyl-CpG binding domain 1 (MBD1), a hairpin loop forms several contacts deep within the major groove near the methylated DNA site (Figure 2.9b).[14b] Specific recognition of the methylated bases is mediated by two arginine residues, a tyrosine residue, and an aspartic acid residue. Sidechain-sidechain interactions seem to play an important role in binding, as mutations of the Asp residue weaken DNA binding despite its lack of direct contact with the DNA.[14b]

The Chlorella virus DNA ligase recognizes nicked DNA through a β-hairpin loop that stems from an OB-fold domain to form a latch which binds the major groove flanking the nick. The OB-fold domain binds the minor groove behind the nick.[15] The homing endonuclease I-*Sce*I in yeast also recognizes DNA primarily by interactions with a long β-hairpin loop which extends into the major groove of the duplex DNA.[16]

There are recent examples of designed beta-hairpin peptides that bind to RNA,[17] but to our knowledge, this is the first example of a designed beta-hairpin that binds to duplex DNA. Moreover, there are examples of synthetic β-hairpins which mimic helical protein domains in protein-protein[18] and protein-RNA interactions,[17] but not in DNA. Our findings suggest that β-hairpins may be promising mimics for α-helical DNA binding domains. This is appealing in that β-hairpins provide a more robust structure than α-helices and can easily be cyclized to reduce entropic costs and provide increased protease resistance.

C. Conclusions

---

[15] Nair, P. A.; Nandakumar, J.; Smith, P.; Odell, M.; Lima, C. D.; Shuman, S. *Nat. Struct. Mol. Biol.* **2007,** *14*, 770-778.

[16] Moure, C. M.; Gimble, F. S.; Quiocho, F. A. *J. Mol. Biol.* **2003,** *334*, 685-695.

[17] (a) Athanassiou, Z.; Patora, K.; Dias, R. L. A.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007,** *46*, 741-751. (b) Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2007,** *46*, 9101-9104. (c) Leeper, T. C.; Athanassiou, Z.; Dias, R. L. A.; Robinson, J. A.; Varani, G. *Biochemistry* **2005,** *44*, 12362-12372. (d) Athanassiou, Z.; Dias, R. L. A.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J. Am. Chem. Soc.* **2004,** *126*, 6906-6913.

[18] (a) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2004,** *43*, 2109-2112. (b) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R.; Grutter, M. G.; Robinson, J. A. *ChemBioChem* **2006,** *7*, 515-526.

In conclusion, we investigated the contribution of folding, chirality, and net charge to binding of a folded peptide to ss- and dsDNA, as a model for the OB-fold. These studies provide insight into the factors that contribute to beta-hairpin recognition of both ss- and dsDNA, which has relevance to protein receptors for each of these types of DNA. These studies indicate that folding is required for binding of both ss- and dsDNA even though the driving forces for binding them are different. In the case of ssDNA, folding provides an aromatic cleft to form favorable stacking interactions with the nucleobases. In the case of duplex DNA, folding allows for binding in a structure-dependent manner, such as groove binding, via favorable electrostatic interactions. We found that folding contributed to binding in the case of both ss- and duplex DNA while chirality only influenced binding to duplex DNA. Reduction of charge had a larger impact on binding to duplex DNA than on binding to ssDNA. Thus, both chirality and net charge influence the selectivity for ss- over dsDNA. These results underscore the subtle features that contribute to molecular recognition of oligonucleotides, and the observed differences in binding suggest approaches to improve selectivity for ssDNA versus double-stranded DNA with biomimetic receptors.

D. Experimental Section

i. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using an Applied Biosystems Pioneer Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-

diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 10 minutes. Each amino acid was coupled on an extended cycle of 75 minutes to improve coupling. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin was performed in 94% trifluoroacetic acid (TFA), 2.5% H_2O, 2.5% ethanedithiol (EDT), and 1% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen, and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 5-35% solvent B over 25 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and the peptide sequence was confirmed by MALDI mass spectrometry. Peptide dimers were formed by oxidation of the cysteine residues with stirring in 1% DMSO in 10 mM phosphate buffer, pH 7.5 for 7-12 hrs. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

ii. DNA Sample Preparation

Two strands of DNA were purchased from IDT (Integrated DNA Technologies). The strands obtained were 5'-CCATCGCTACC-3' and its complement. All DNA samples were prepared in triethylammonium acetate (TEAA) buffer. DNA samples were dissolved in 1 mL TEAA buffer (5.6% acetic acid, 13.86% triethylamine (TEA) in water, adjusted to pH 7.0). Buffer was added individually to 1 μmole DNA. The DNA solution was pooled and an additional 1 mL of TEAA buffer was added to each DNA sample. This was then pooled and filtered before purification by reversed-phase HPLC. Purification was performed using the

same column as that used for peptides with a gradient of 25-75% B over 20 minutes with a 6 mL/min flow rate. Solvent A was 5% acetonitrile, 95% TEAA buffer and solvent B was 15% acetonitrile, 85% TEAA buffer. Samples were lyophilized after purification and the resulting product was then dissolved in water and re-lyophilized to remove any remaining salts. Concentrations of both DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficients of the two DNA strands ($\varepsilon_{260,\ ssDNA} = 95500\ M^{-1}\cdot cm^{-1}$ and $\varepsilon_{260,\ ssDNA\ complement} = 112600\ M^{-1}\cdot cm^{-1}$). Equal concentrations of the two strands (in sodium phosphate buffer, pH = 7.0) were pooled in a final concentration of 100 mM NaCl. The solution was heated at 95 ºC for 5 minutes to anneal the strands and was then allowed to cool to room temperature before storing at -20 ºC.

iii. Circular Dichroism

All peptide dimers were analyzed by CD to verify their structure. CD measurements were performed on an Applied Photophysics Pistar-180 Circular Dichroism/Fluorescence Spectrophotometer. CD data was obtained for the peptide dimers at 30 μM and 60 μM concentrations. The peptides were dissolved in 10 mM Na_2HPO_4, pH 7.0. Wavelength scans were performed at 298 K from 260-185 nm. CD mixing experiments were performed using the same method. Concentrations for the peptide and DNA were 50 μM, and the mixtures were 50 μM of each. Mixing experiments were performed in phosphate buffer containing 100 mM NaCl because NaCl is required for DNA annealing.

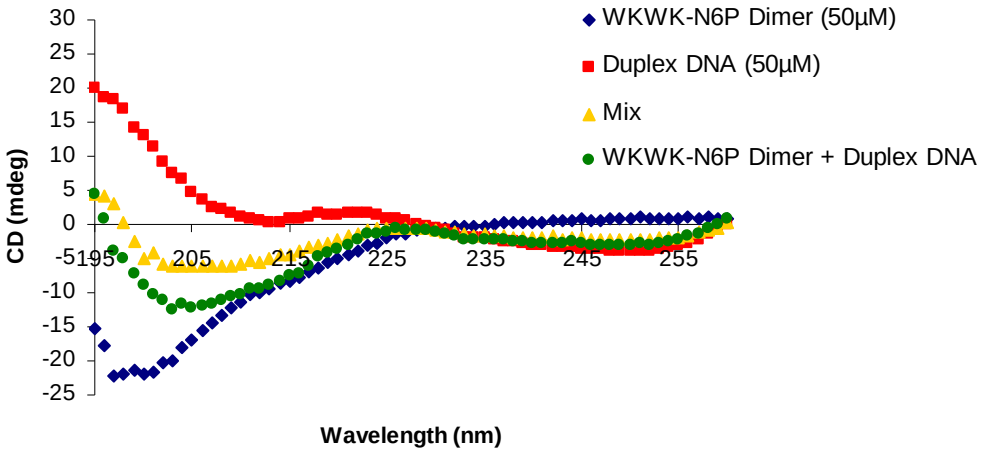


Figure 2.10. CD spectrum of $(WKWK\text{-}N6P)_2$ (blue), duplex DNA (red), the additive spectrum (green), and a 1:1 mixture (yellow). Differences in the green and yellow spectra are likely due to errors in concentration measurement.

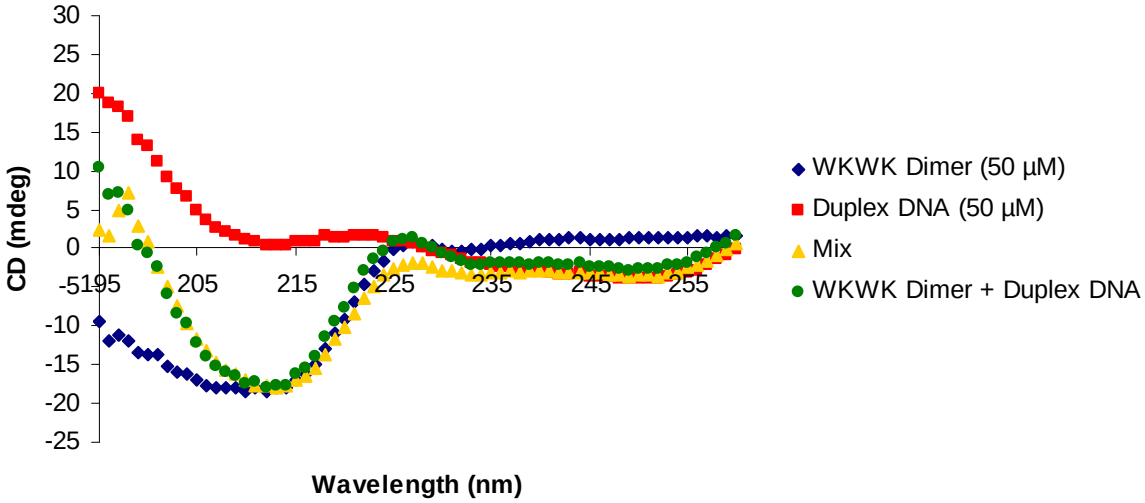


Figure 2.11. CD spectrum of $(WKWK)_2$ (blue), duplex DNA (red), the additive spectrum (green), and a 1:1 mixture (yellow).

iv. Fluorescence Titrations

To determine the recognition of single-stranded and double-stranded oligonucleotides by the peptide dimers, fluorescence titrations were performed which followed the Trp quenching with increasing oligonucleotide concentration. Peptide and nucleotide samples were prepared in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0. Peptide

concentrations were determined in 5 M guanidine hydrochloride by recording the absorbance of the Trp residues at 280 nm ($\varepsilon = 5690\ M^{-1}cm^{-1}$) by UV/vis spectroscopy. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. Experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 348 nm were fit as a function of nucleotide concentration to the binding equation (Equation 2.1) on Kaleidagraph using non-linear least squares fitting.[19]

Equation 2.1. $I = [I_o + I_\infty([L]/K_d)]/[1 + ([L]/K_d)]$

where I is the observed fluorescence intensity, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, and K_d is the dissociation constant.

Oligonucleotides have an observable absorbance at the excitation wavelength of Trp (297 nm), and therefore there is an inner filter effect for which one must take account. The absorbance of the oligonucleotides at 297 nm was monitored at known concentrations and the extinction coefficient was determined. Absorbance values were determined for each oligonucleotide concentration. Corrected fluorescence values were determined from the following equations (Equation 2.2 and Equation 2.3).[20]

Equation 2.2. $F_c = F_o/C_i$

Equation 2.3. $C_i = (1\text{-}10^{-Ai})/(2.303)(A_i)$

---

[19] Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994,** *3*, 1261-1266.

[20] Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424-458.

where F_c is the corrected fluorescence, F_o is the fluorescence observed, and C_i is the correction factor for each absorbance value (i). A_i is the absorbance value for each concentration determined by the extinction coefficient.

An example of the fluorescence raw data observed for these peptides is given in Figure 2.12. The plots in the DNA Binding Measurements section show overlays of both ssDNA and duplex DNA data, but individual plots with the binding curve included are provided in the Experimental Section (Figures 2.13-2.21).

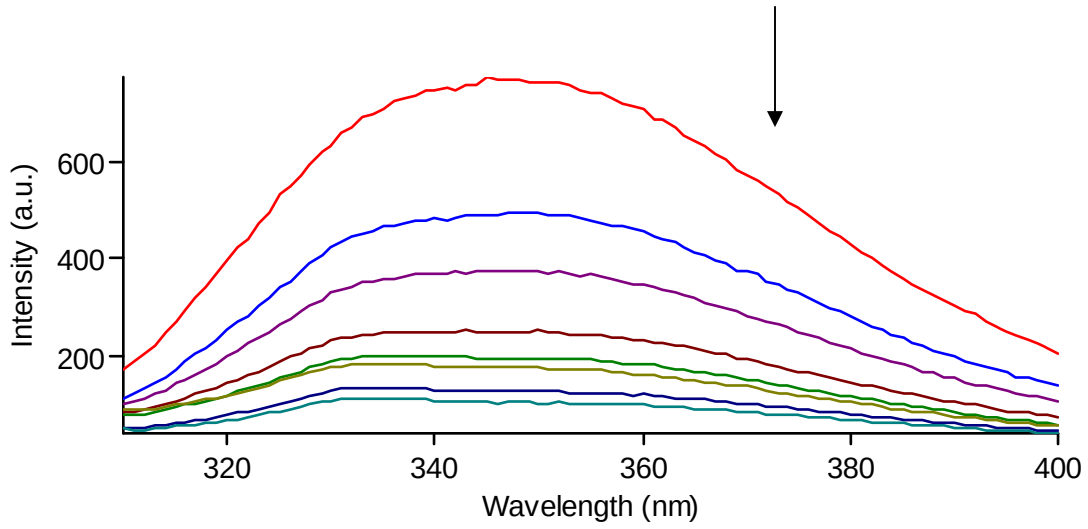


Figure 2.12. Fluorescence quenching of (D-WKWK)$_2$ (2 μM) by ssDNA (Figure 2.3g, 0-60 μM), in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298K. The arrow indicates the direction of change in fluorescence intensity as ssDNA concentration increases.

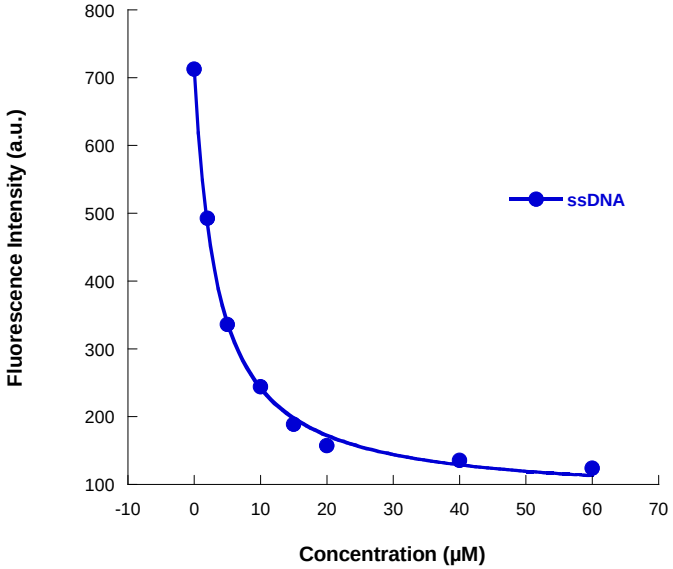


Figure 2.13. Fluorescence titration of $(WKWK)_2$ with single-stranded DNA sequence 5'-CCATCGCTACC-3' corrected for the inner filter effect (see Experimental Section); 2 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

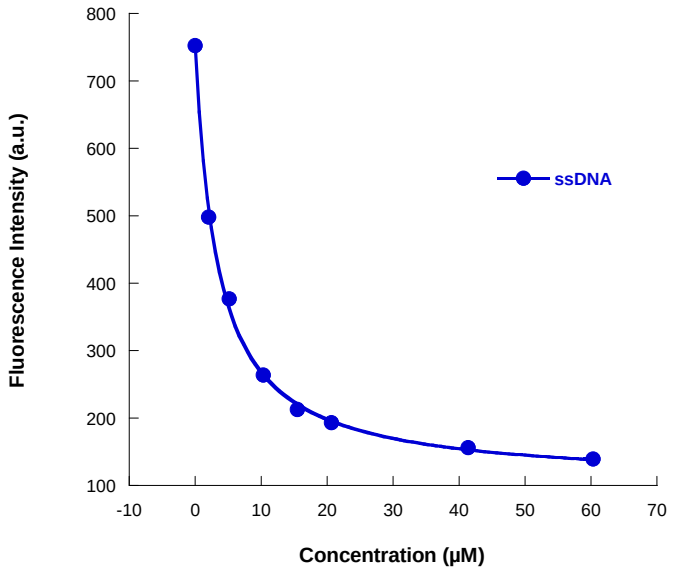


Figure 2.14. Fluorescence titration of $(D\text{-}WKWK)_2$ with single-stranded DNA sequence 5'-CCATCGCTACC-3' corrected for the inner filter effect (see Experimental Section); 2 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

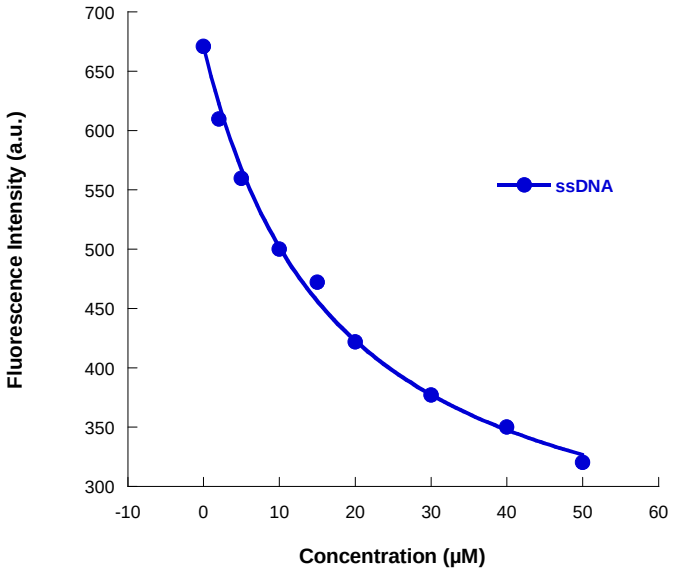


Figure 2.15. Fluorescence titration of $(WKWK\text{-}R1Q)_2$ with single-stranded DNA sequence 5'-CCATCGCTACC-3' corrected for the inner filter effect (see Experimental Section); 5 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

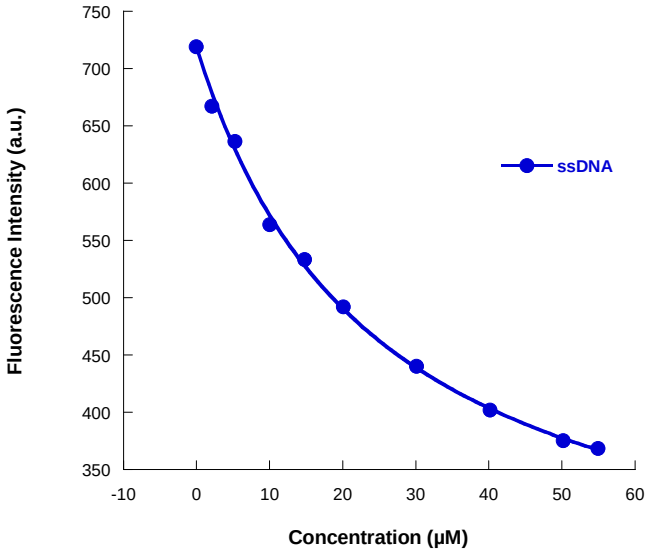


Figure 2.16. Fluorescence titration of $(WKWK\text{-}N6P)_2$ with single-stranded DNA sequence 5'-CCATCGCTACC-3' corrected for the inner filter effect (see Experimental Section); 5 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

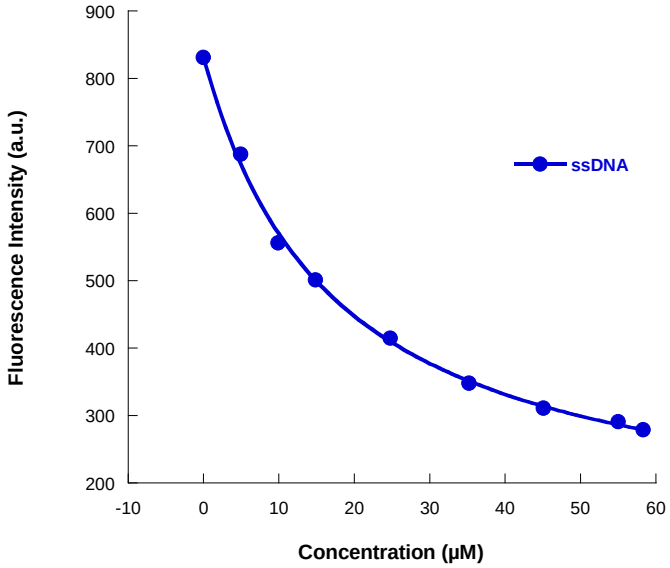


Figure 2.17. Fluorescence titration of $(WKWK\text{-}Scrambled)_2$ with single-stranded DNA sequence 5'-CCATCGCTACC-3' corrected for the inner filter effect (see Experimental Section); 8 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

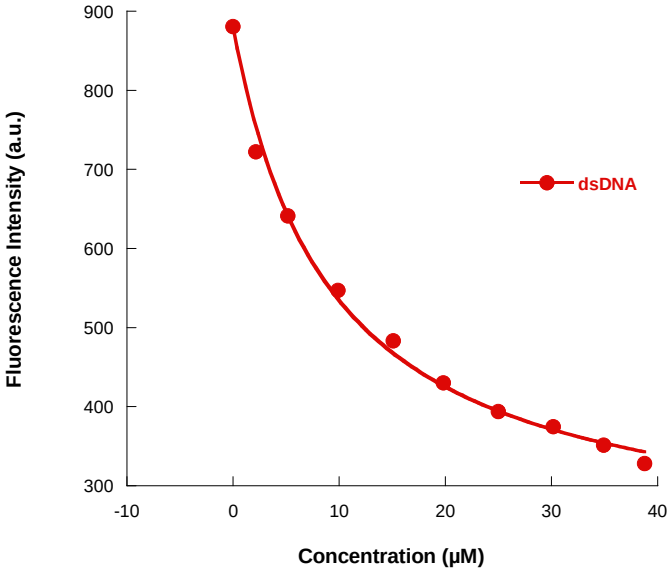


Figure 2.18. Fluorescence titration of $(WKWK)_2$ with duplex DNA (Figure 2.3h) corrected for the inner filter effect (see Experimental Section); 2 μM peptide, 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

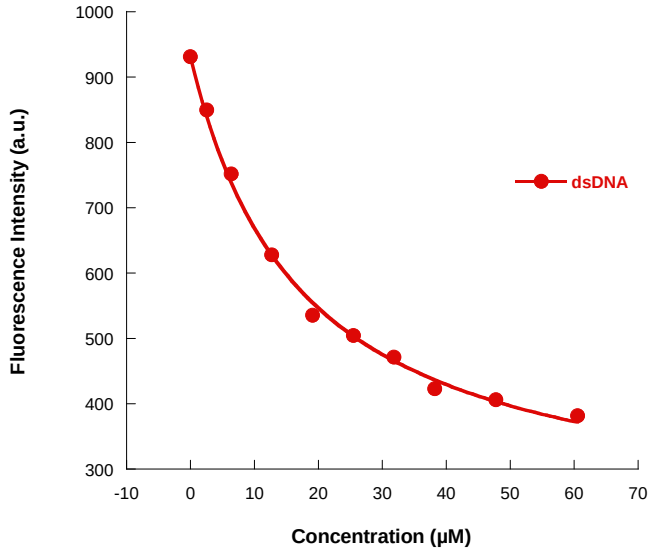


Figure 2.19. Fluorescence titration of (D-WKWK)$_2$ with duplex DNA (Figure 2.3h) corrected for the inner filter effect (see Experimental Section); 5 μM peptide, 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

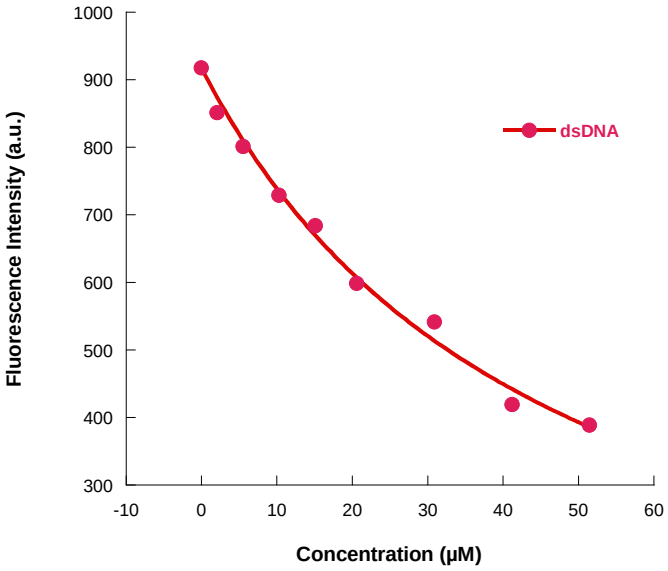


Figure 2.20. Fluorescence titration of (WKWK-R1Q)$_2$ with duplex DNA (Figure 2.3h) uncorrected for the inner filter effect (see Experimental Section); 7 μM peptide, 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

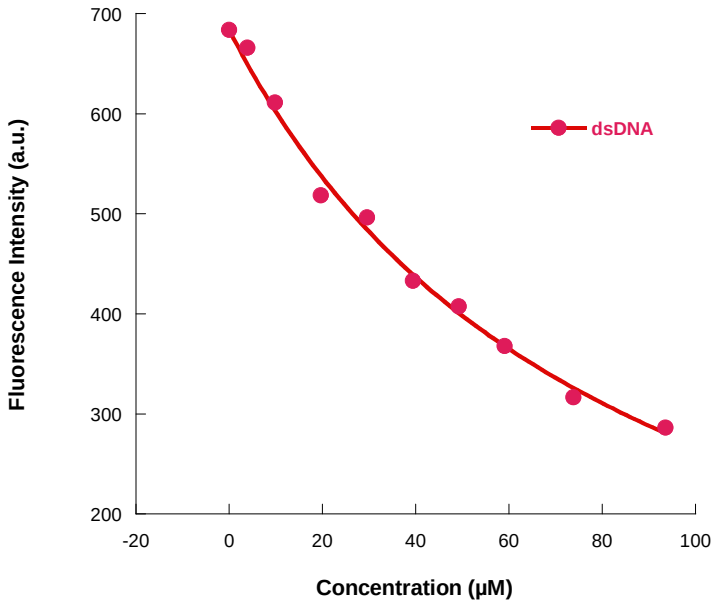


Figure 2.21. Fluorescence titration of $(WKWK\text{-}N6P)_2$ with duplex DNA (Figure 2.3h) uncorrected for the inner filter effect (see Experimental Section); 5 μM peptide, 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

v. Stoichiometry of Binding

The stoichiometry of binding was determined by the molar variation method following the quenching of tryptophan fluorescence. Conditions were such that the peptide and DNA should be fully bound. Peptide concentrations were in the range of 25-45 μM, depending on the maximum DNA concentrations used. The conditions were limited to low DNA concentrations (~60 μM for ssDNA and ~40 μM for duplex DNA) because of the inner filter effect. After correction for the inner filter effect, the fluorescence intensity was plotted against the ratio of DNA/peptide concentrations to give the stoichiometry of binding. The stoichiometry of binding is shown in plots (Figures 2.5 and 2.6) with the X-intercept of the dashed lines indicating the stoichiometry for each.

Chapter III

Optimization of the Binding Face of the Designed β-Hairpin Peptide Receptor for Selective Recognition of ssDNA through Library Screening

A. Background and Significance

The binding capabilities of $(WKWK)_2$ have been explored previously, and the results show that it is a strong binder of oligonucleotides. Much has been learned about the system and about the residues involved in binding.[1] While strong binding has been determined, the extent to which one can gain significant selectivity with this peptide is unknown. Analysis of various protein structures has shown, however, that selectivity can be achieved by incorporating amino acids that can hydrogen-bond selectively to a particular base.[2] For example, hydrogen-bonding is favored between guanine and Arg, Lys, His, and Ser, while Asn and Gln sidechains form hydrogen bonds more favorably with adenine bases.[2a] Figure 3.1 shows favorable hydrogen-bonding patterns between nucleotide bases and amino acid

---

[1] (a) Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544. (b) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25. (c) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70,* 1105-1114. (d) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.

[2] (a) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860-2874. (b) Moodie, S. L.; Mitchell, J. B. O.; Thornton, J. M. *J. Mol. Biol.* **1996**, *263*, 486-500. (c) Cheng, A. C.; Chen, W. W.; Fuhrmann, C. N.; Frankel, A. D. *J. Mol. Biol.* **2003**, *327*, 781-796.

sidechains which have favored interactions with those bases. With this in mind, we employed combinatorial chemistry to produce peptides that mimic the OB-fold domain.[3]

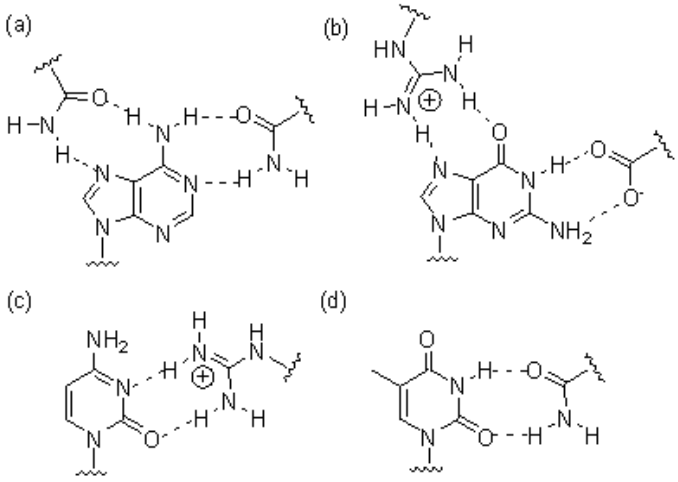


Figure 3.1. H-bonding patterns between amino acid sidechains and (a) adenine, (b) guanine, (c) cytosine, and (d) thymine. Adapted from Cheng, et al.[2c]

Combinatorial chemistry has been used for many years to determine high affinity, high selectivity binders for a number of different targets[4] and has been a crucial tool for many pharmaceutical companies in their drug discovery efforts.[5] Combinatorial libraries of RNA,[6, 7] DNA,[7, 8] proteins and antibodies,[9] peptides,[10] and small molecules[11] have all been

[3] Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

[4] Liu, D. R.; Schultz, P. G. *Angew. Chem. Int. Ed.* **1999**, *38*, 36-54.

[5] Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P.; Gordon, E. M. *J. Med. Chem.* **1994,** *37*, 1233-1251.

[6] (a) Famulok, M.; Jenne, A. *Curr. Opin. Chem. Biol.* **1998**, *2*, 320-327. (b) Joyce, G. F. *Gene* **1989**, *82*, 83-87. (c) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510. (d) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818-822.

[7] (a) Hahn, W. C.; Dunn, I. F.; Kim, S. Y.; Schinzel, A. C.; Firestein, R.; Guney, I.; Boehm, J. S. *Biochim. Biophys. Acta.* **2009,** *1790*, 478-484. (b) Perrin, D. M. *Comb. Chem. High Throughput Screen.* **2000,** *3*, 243-269. (c) Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*,

used for high throughput screens of hundreds to millions of compounds. Combinatorial chemistry has been useful in determining not only ligand-binding activity[5, 9a] but also in

---

3548-3553. (d) Ahn, D. R.; Yu, J. *Bioorg. Med. Chem.* **2005**, *13*, 1177-1183. (e) Tisne, C.; Dardel, F. *Comb. Chem. High Throughput Screen.* **2002**, *5*, 523-529.

[8] (a) Gunawardane, R. N.; Sgroi, D. C.; Wrobel, C. N.; Koh, E.; Daley, G. Q.; Brugge, J. S. *Cancer Res.* **2005**, *65*, 11572-11580. (b) Koh, E. Y.; Chen, T.; Daley, G. Q. *Nucleic Acids Res.* **2002**, *30*, e142. (c) Wang, B.; Dickinson, L. A.; Koivunen, E.; Ruoslahti, E.; Kohwi-Shigematsu, T., *J. Biol. Chem.* **1995**, *270*, 23239-23242. (d) Krook, M.; Mosbach, K.; Lindbladh, C. *Biochem. Biophys. Res. Commun.* **1994**, *204*, 849-854. (e) Pollock, R.; Treisman, R. *Nucleic Acids Res.* **1990**, *18*, 6197-6204. (f) Blackwell, T. K..; Weintraub, H. *Science* **1990**, *250*, 1104-1110. (g) Horwitz, M. S. Z.; Loeb, L. A. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 7405-7409. (h) Oliphant, A. R.; Nussbaum, A. L.; Struhl, K. *Gene* **1986**, *44*, 177-183.

[9] (a) Kijanka, G.; Murphy, D. *J. Proteomics* **2009**, *72*, 936-944. (b) Hughes, M. D.; Zhang, Z. R.; Sutherland, A. J.; Santos, A. F.; Hine, A. V. *Nucleic Acids Res.* **2005**, *33*, e32. (c) Bussow, K.; Cahill, D.; Nietfeld, W.; Bancroft, D.; Scherzinger, E.; Lehrach, H.; Walter, G. *Nucleic Acids Res.* **1998**, *26*, 5007-5008. (d) Barbas, C. F., 3rd; Bain, J. D.; Hoekstra, D. M.; Lerner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 4457-4461.

[10] (a) Murray, J. K.; Farooqi, B.; Sadowsky, J. D.; Scalf, M.; Freund, W. A.; Smith, L. M.; Chen, J.; Gellman, S. H. *J. Am. Chem. Soc.* **2005**, *127*, 13271-13280. (b) Camperi, S. A.; Marani, M. M.; Iannucci, N. B.; Cote, S.; Albericio, F.; Cascone, O. *Tettrahedron Lett.* **2005**, *46*, 1561-1564. (c) Halkes, K. M.; Gotfredsen, C. H.; Grotli, M.; Miranda, L. P.; Duus, J. O.; Meldal, M. *Chem. Eur. J.* **2001**, *7*, 3584-3591. (d) Vetter, D.; Thamm, A.; Schlingloff, G.; Schober, A. *Mol. Divers.* **2000**, *5*, 111-116. (e) Gotfredsen, C. H.; Grotli, M.; Willert, M.; Meldal, M.; Duus, J. O. *J. Chem. Soc., Perkin Trans. 1* **2000**, 1167-1171. (f) Huang, P. Y.; Carbonell, R. G. *Biotechnol. Bioeng.* **1999**, *63*, 633-641. (g) Hiemstra, H. S.; Duinkerken, G.; Benckhuijsen, W. E.; Amons, R.; de Vries, R. R.; Roep, B. O.; Drijfhout, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 10313-10318. (h) Kramer, A.; Volkmer-Engert, R.; Malin, R.; Reineke, U.; Schneider-Mergener, J. *Pept. Res.* **1993**, *6*, 314-319. (i) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86. (j) Scott, J. K.; Smith, G. P. *Science* **1990**, *249*, 386-390.

[11] (a) Hu, L. A.; Tang, P. M.; Eslahi, N. K.; Zhou, T.; Barbosa, J.; Liu, Q., *J. Biomol. Screen.* **2009**, *14*, 789-797. (b) Yu, P.; Liu, B.; Kodadek, T. *Nat. Biotechnol.* **2005**, *23*, 746-751. (c) Liu, R.; Marik, J.; Lam, K. S. *Methods Enzymol.* **2003**, *369*, 271-287. (d) Schouten, J. A.; Ladame, S.; Mason, S. J.; Cooper, M. A.; Balasubramanian, S. *J. Am. Chem. Soc.* **2003**, *125*, 5594-5595. (e) Lam, K. S.; Lehman, A. L.; Song, A.; Doan, N.; Enstrom, A. M.; Maxwell, J.; Liu, R. *Methods Enzymol.* **2003**, *369*, 298-322.

identifying catalysts[12] and in studying folding of peptides and proteins.[13] Of the various combinatorial methods utilized, many involve targeting nucleic acids. These include the methods of split-and-pool on-bead libraries,[7c-7e] phage display,[14] dynamic combinatorial chemistry,[15] and various selection and randomization methods.[7]

To apply the concept of combinatorial chemistry to this project, a peptide library was designed. One advantage of designing a β-hairpin peptide library is that it can be synthesized by solid phase peptide synthesis, rendering it amenable to a variety of combinatorial methods. By using solid phase peptide synthesis, any amino acid can be incorporated into a library, allowing natural amino acids, D-amino acids,[16] or non-natural amino acids[16] to be included. This approach can produce a large variety of functional groups that may provide different types of interactions not observed in libraries comprised of only natural amino acids. Peptides can be synthesized on solid support, and cyclic[13] and dimeric[17] peptide libraries, which could improve affinity, are also possible using on-bead libraries. Because of these advantages, a peptide library was synthesized through split-and-pool synthesis to give a

[12] (a) Taylor, S. J.; Morken, J. P. *Science* **1998**, *280*, 267-270. (b) Danek, S. C.; Queffelec, J.; Morken, J. P. *Chem. Commun.* **2002**, 528-529.

[13] Cowell, S. M.; Gu, X.; Vagner, J.; Hruby, V. J. *Methods Enzymol.* **2003**, *369*, 288-297.

[14] Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391-410.

[15] Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652-3711.

[16] Samson, I.; Rozenski, J.; Samyn, B.; Van Aerschot, A.; Van Beeumen, J.; Herdewijn, P. *J. Biol. Chem.* **1997**, *272*, 11378-11383.

[17] Aggarwal, S.; Harden, J. L.; Denmeade, S. R. *Bioconjugate Chem.* **2006**, *17*, 335-340.

"one-bead-one-compound" library.[18] The one-bead-one-compound method was first described in 1991 and generates an on-bead library in which each bead contains a unique peptide sequence, allowing for screening of individual hits.[18] This split-and-pool synthesis method is depicted in Figure 3.2.

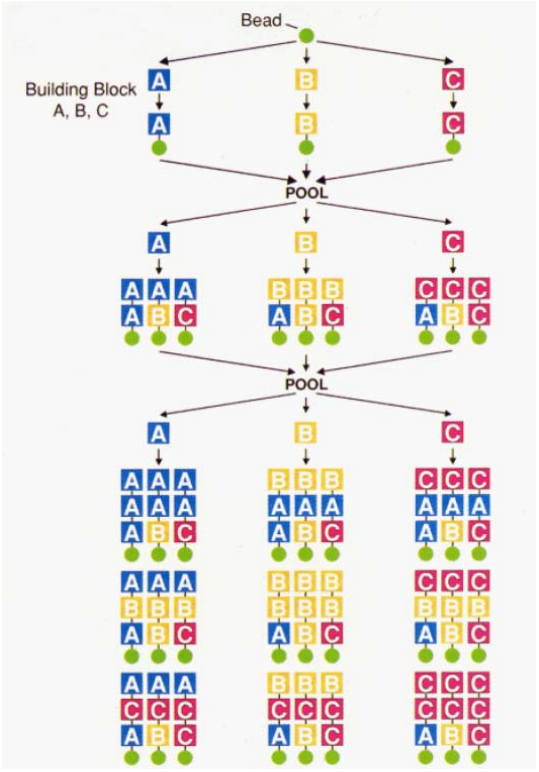


Figure 3.2. Split-and-Pool synthesis method for producing a one-bead-one-compound combinatorial library using solid phase peptide synthesis. Adapted from Gallop, et al.[5]

B. Results

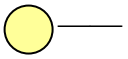
i. Library 1 Design

[18] (a) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84. (b) Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411-448. (c) Lam, K. S.; Lehman, A. L.; Song, A.; Doan, N.; Enstrom, A. M.; Maxwell, J.; Liu, R. *Methods Enzymol.* **2003**, *369*, 298-322. (d) Morken, J. P.; Kapoor, T. M.; Feng, S.; Shirai, F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1998**, *120,* 30-36. (e) Combs, A. P.; Kapoor, T. A.; Feng, S.; Chen, J. K.; Daude-Snow, L. F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1996**, *118,* 287-288.

To obtain high affinity sequence selectivity for single-stranded DNA or RNA sequences, a peptide library has been designed using WKWK as a template and varying amino acids within the binding pocket of the peptide. The twelve-residue peptide library was synthesized through split-and-pool synthesis on TentaGel Macrobeads. The peptide was originally designed to mimic the OB-fold domain, with aromatic residues to stack with nucleotide bases and cationic residues to form electrostatic interactions with the phosphate groups.[1] We wanted to maintain these characteristics in the library since these types of interactions are known to be important in OB-fold domains.[19] For this library, we kept one side of the binding pocket constant and varied the other side to provide additional diversity near the aromatic and basic residues. The peptide was varied at N-terminal positions 2 and 4 for Library 1 (Figure 3.3) to determine if we could improve the binding pocket. Polar, cationic, and anionic residues were included at position 4, and various hydrophobic residues were included at position 2 (Table 3.1). Unnatural amino acids incorporated into Library 1 are shown in Figure 3.4. The variations were intended to maintain similar types of functionalities in those positions, with hydrophobic and aromatic residues in the original Trp position 2 and residues that will form hydrogen bonds or electrostatic interactions at Lys 4.

[19] (a) Kloks, C. P.; Spronk, C. A.; Lasonder, E.; Hoffmann, A.; Vuister, G. W.; Grzesiek, S.; Hilbers, C. W. *J. Mol. Biol.* **2002**, *316*, 317-326. (b) Schroder, K.; Graumann, P.; Schnuchel, A.; Holak, T. A.; Marahiel, M. A. *Mol. Microbiol.* **1995**, *16*, 699-708.

(a) WKWK Monomer: Ac-R-**W**-V-**K**-V-N-G-O-**W**-I-**K**-Q-NH_2

(b)  βAla-βAla-Gln-Lys-Ile-Trp-Orn-Gly-Asn-Val-**Lys**^{4}-Val-**Trp**^{2}-Arg-NH_2

(c)

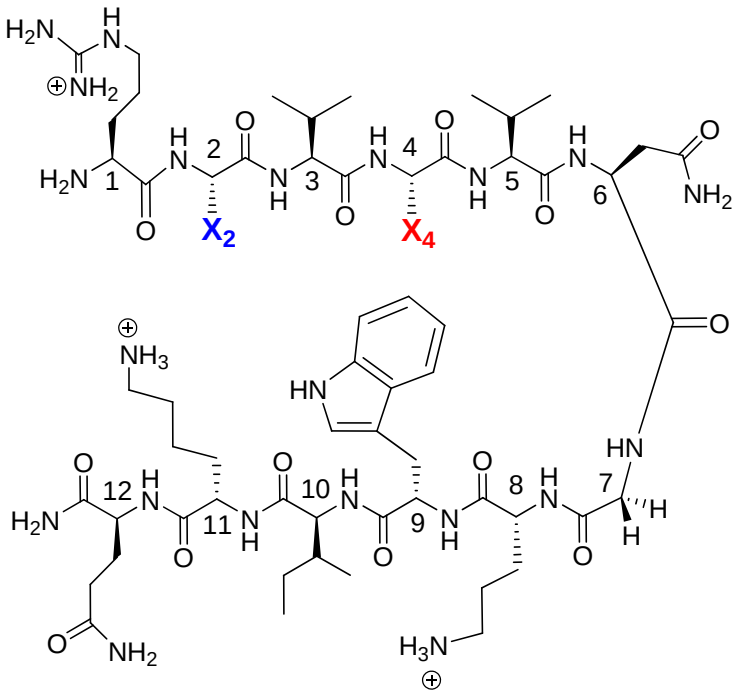


Figure 3.3. (a) Sequence of WKWK with amino acids thought to be involved in the binding pocket shown in bold. (b) On-bead representation of Library 1. (c) Structure of Library 1 peptide with position 2 shown as X_2 in blue and position 4 shown as X_4 in red. Positions 2 and 4 were varied with five different residues each, as shown in Table 3.1. The βAla-βAla linker was added as a spacer between the bead and the peptide.

Table 3.1. Varied positions and residues used for Library 1.

Varied Residue	Amino Acids
Tryptophan (2)	Trp, Phe, Leu, Val, PhG
Lysine (4)	Lys, Gln, Orn, Glu, Arg

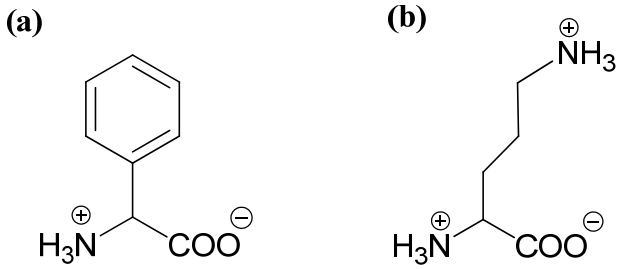


Figure 3.4. Structures of unnatural amino acids incorporated into Library 1. (a) phenylglycine (PhG). (b) ornithine (Orn).

ii. Library 1 Screen and Identification of Hits

The library was screened against the single-stranded nucleotide sequence 5'-d(AAAAA)-3' (penta-dA). A colorimetric assay reported by the Schreiber group which takes advantage of the strong binding interaction between biotin and streptavidin was used to screen the libraries for hits.[18d-18e] In the screen, biotinylated penta-dA was incubated with a streptavidin/alkaline phosphatase complex. After incubation, the Biotin-penta-dA/SAAP complex was allowed to react with the on-bead peptide library. Once the reaction was complete, the resin was washed with high salt screening buffer to remove any nonspecific binders, and then the resin was washed with a dye. The dye reacts with alkaline phosphatase, which is complexed with the DNA, and this reaction produces an insoluble dye that colors the beads which are bound to DNA. Additional details of the screen are included in the Experimental Section. Hits were identified from the assay and selected for sequence analysis by Edman degradation.

Thirty hits were identified in screen 1, and eight of those were sequenced. The results of the screen and subsequent Edman degradation analysis are shown in Table 3.2. The first screen identified charged residues in position 4 for each of the hits and gave aromatic residues (either Trp or Phe) in position 2 for each hit. This suggests that the initial design

(with Trp 2 and Lys 4) was already optimized. Basic residues seem to increase binding affinity to a higher degree than the other hydrogen-bonding residues for position 4. Aromatic residues apparently provide necessary stacking interactions with the DNA bases that are missing in the other hydrophobic residues.

Table 3.2. Sequences of hits from Library 1 screen (Varied positions shown in bold).

Bead Number	Sequence of Hits
Bead 1	Arg-**Trp**-Val-**Orn**-Val
Bead 2	Arg-**Phe**-Val-**Lys**-Val
Bead 3	Arg-**Phe**-Val-**Lys**-Val
Bead 4	Arg-**Trp**-Val-**Arg**-Val
Bead 5	Arg-**Phe**-Val-**Arg**-Val
Bead 6	Arg-**Trp**-Val-**Orn**-Val
Bead 7	Arg-**Trp**-Val-**Orn**-Val
Bead 8	Arg-**Phe**-Val-**Arg**-Val

iii. Fluorescence Binding of Library 1 Hit

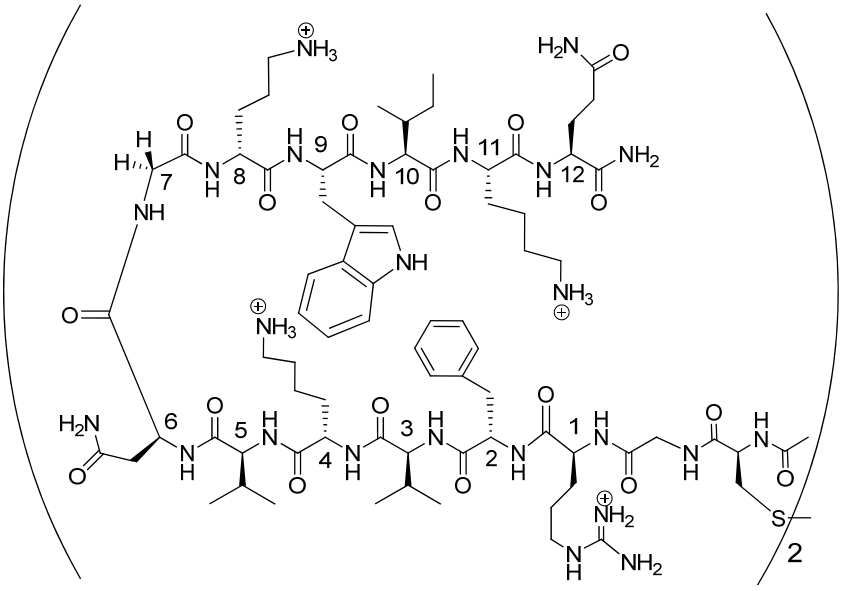
Based on the results of screen 1, the FKWK hit (Beads 2 and 3 from Table 3.2; Figure 3.5a) was studied further. The FKWK dimer (Figure 3.5b) was synthesized, and fluorescence titrations were performed with ssDNA. We were interested in understanding the difference in binding by a Phe-Trp pocket versus the Trp-Trp binding pocket. Although the library was made of peptide monomers, the FKWK dimer was made because the original WKWK dimer binds ssDNA with higher affinity than does the monomer.[1] The dissociation constant for the interaction between $(FKWK)_2$ and the randomized 11 base pair single-

stranded DNA sequence (Figure 3.5c) was determined to be around 32 μM. This is weaker than that of the WKWK dimer, which has a K_d of 3 μM for ssDNA (Table 3.3; Figure 3.6).[1]

(a) FKWK: Ac-Arg-**<u>Phe</u>**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-NH_2

(b) $(FKWK)_2$:

(Ac-Cys-Gly-Arg-**<u>Phe</u>**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-NH_2)$_2$



(c) ssDNA: 5'-CCATCGCTACC-3'

Figure 3.5. Sequence and structure of FKWK monomer and dimer. (a) β-hairpin sequence FKWK with the Phe mutation underlined. (b) Sequence and structure of peptide $(FKWK)_2$ with the Phe mutation underlined. Note that the numbering excludes the Cys-Gly residues added to form the disulfide. The bold residues are those thought to be involved in the nucleic acid binding pocket. Those residues are all on the same face of the peptide. (c) ssDNA sequence.

Table 3.3. Dissociation constants for the binding interaction between WKWK and FKWK dimers and ssDNA[a]

Peptide	ssDNA K_d, µM (error)
$(WKWK)_2$	3.5 (0.2)[b]
$(FKWK)_2$	32 (6)

a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting. (b) Reported previously.[1]

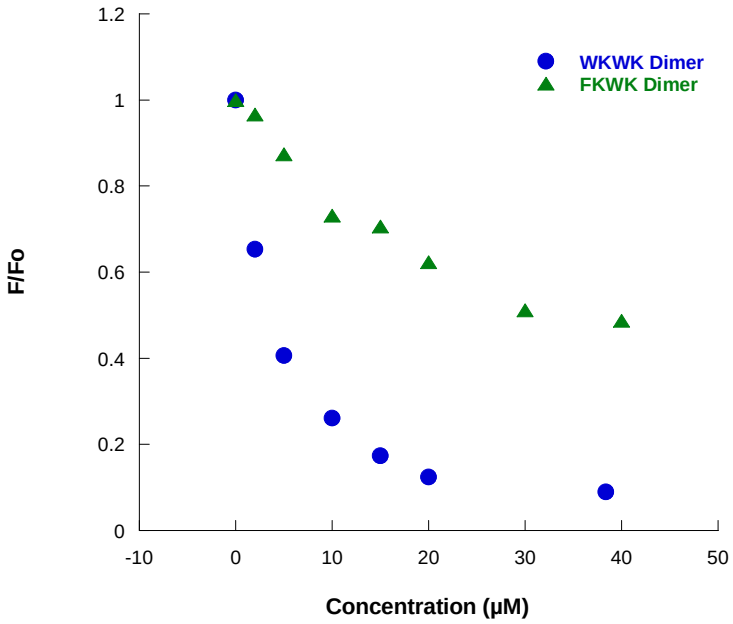
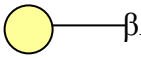


Figure 3.6. Fluorescence titrations of $(FKWK)_2$ and $(WKWK)_2$ with the single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 3.5c) corrected for the inner filter effect (see Experimental Section); 2-4 µM peptide concentrations; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

iv. Library 2 Design

A second library was designed with variations in the Arg 1, Val 3, and Lys 4 residues (Table 3.4; Figure 3.7). The library should reveal possible additional binding interactions as well as other amino acids that could increase binding affinity. This library varied the first

and third residues positioned on the opposite face of the binding pocket while keeping the Trp residues constant. The Lys 4 position was varied as well. Non-natural amino acids in the library are shown in Figure 3.8.

(a)  βAla-βAla-Gln-Lys-Ile-Trp-Orn-Gly-Asn-Val-**Lys**^{4}-**Val**^{3}-Trp-**Arg**^{1}-NH_2

(b)

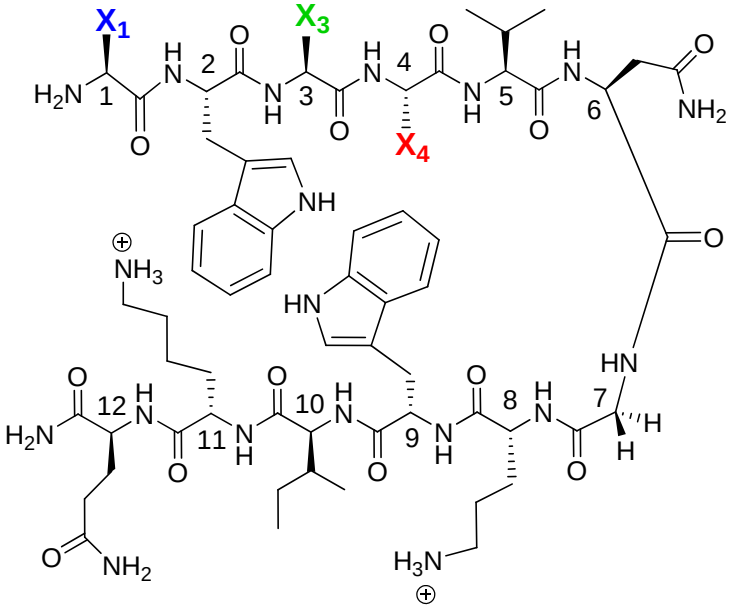


Figure 3.7. Library 2. (a) Depiction of on-bead Library 2. Lys, Val, and Arg residues shown were varied with several different residues each, shown in Table 3.4. The βAla-βAla linker was added as a spacer between the bead and the peptide. (b) Structure of the WKWK peptide template for Library 2 with positions X_1, X_3, and X_4 as varied positions in the library.

Table 3.4. Varied positions and residues used for Library 2.

Varied Residue	Amino Acids
Arginine (1)	Arg, Gln, Lys, Orn, Asn, Cit, Dab, Ser, Thr, Glu, His
Valine (3)	Arg, Gln, Lys, Orn, Asn, Cit, Dab, Ser, Thr, Glu, His, Val, Tyr
Lysine (4)	Arg, Gln, Lys, Orn, Asn, Cit, Dab, Ser, Thr, Glu, His, Val, Tyr, Trp, Phe

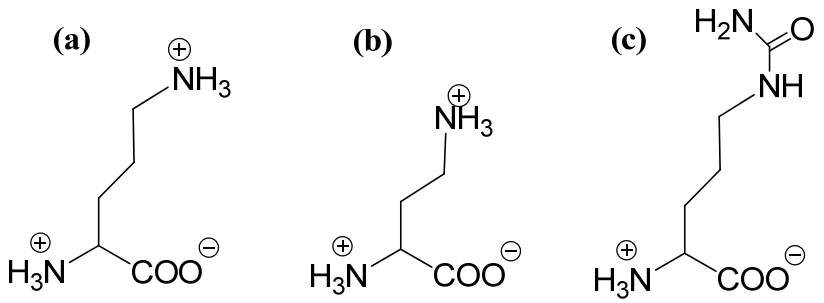


Figure 3.8. Structures of unnatural amino acids used in Library 2. (a) ornithine (Orn). (b) diaminobutyric acid (Dab). (c) citrulline (Cit).

v. Library 2 Screen and Identification of Hits

For the second library, 30 hits were identified and again eight were sequenced. The results of the sequencing showed that binding is primarily due to electrostatics, with mostly positively charged residues selected. As shown in Table 3.5, non-charged residues were found only on the opposite face of the hairpin as Trp. This suggests that the face of the peptide containing the Trp residues is the nucleotide binding face. Bead 7 (Table 3.5) was selected as a hit in the second library screen and is the parent sequence, WKWK. This once again indicates that this peptide is optimized for binding ssDNA. The positively charged residues are very important in binding of ssDNA, consistent with data from the (WKWK-R1Q)$_2$ peptide, which gave a K_d of about 22 μM versus 3 μM for (WKWK)$_2$.[1a] The Lys residues on the binding face[1d] as well as the Orn in the sequence[20] have also been shown to be critical for binding ssDNA, with the Lys residues each contributing -1.5 kcal/mol to the binding interaction.[1d] The presence of different basic residues chosen suggests that the difference in the length and type of cationic sidechains makes little difference, with a small preference for Arg and Lys. Arg may be favored in some cases for its guanidinium group which provides additional hydrogen-bonding capabilities.

[20] Hughes R. M.; Waters, M. L. unpublished results.

Table 3.5. Sequences of hits from Library 2 screen with varied positions shown in bold.

Bead Number	Sequence of Hits
Bead 1	**Arg**-Trp-**Arg**-**Arg**-Val
Bead 2	**Dab**-Trp-**Orn**-**Arg**-Val
Bead 3	**Dab**-Trp-**Arg**-**Arg**-Val
Bead 4	**Lys**-Trp-**Tyr**-**Arg**-Val
Bead 5	**Cit**-Trp-**Val**-**Arg**-Val
Bead 6	**Orn**-Trp-**Orn**-**Arg**-Val
Bead 7	**Arg**-Trp-**Val**-**Lys**-Val
Bead 8	**Orn**-Trp-**Arg**-**Dab**-Val
Bead 9	**Dab**-Trp-**Arg**-**Orn**-Val

The screens were performed at physiological pH and salt conditions, but magnesium, which is in the cell, was not included. To determine whether the presence of magnesium affects the binding of single-stranded DNA sequences to the peptides, a 1 mM solution of $MgCl_2$ was added to another screen of Library 2. The screen was performed with the same concentrations, including the NaCl concentration of 500 mM. As expected, the beads chosen from this screen (Table 3.6) were very similar to those in the screen without magnesium (Table 3.5).

Another assay was performed in 1M NaCl (Table 3.7) to further screen the electrostatic interactions and to determine other favorable interactions between the hairpin and the DNA. This screen did show some change in residues at each position, indicating that the high salt concentration screened the electrostatic interactions to some extent. The general conclusion from these screens is that the original design of the peptide is optimized with

binding occurring through an aromatic pocket with electrostatic interactions providing increased affinity. Little or no selectivity was observed with the screens.

Table 3.6. Sequences of hits from the Library 2 screen with 1 mM $MgCl_2$ (varied positions shown in bold).

Bead Number	Sequence of Hits
Bead 1	**Arg**-Trp-**Orn**-**Tyr**-Val
Bead 2	**Arg**-Trp-**Dab**-**Arg**-Val
Bead 3	**Lys**-Trp-**Tyr**-**Arg**-Val
Bead 4	**Lys**-Trp-**Dab**-**Arg**-Val
Bead 5	**Arg**-Trp-**Orn**-**Arg**-Val
Bead 6	**Arg**-Trp-**Dab**-**Lys**-Val
Bead 7	**Lys**-Trp-**Orn**-**Dab**-Val
Bead 8	**Arg**-Trp-**Arg**-**Dab**-Val
Bead 9	**Arg**-Trp-**Dab**-**Tyr**-Val

Table 3.7. Sequences of hits from the Library 2 screen at 1 M NaCl with varied positions shown in bold.

Bead Number	Sequence of Hits
Bead 1	**Ser**-Trp-**Arg-Dab**-Val
Bead 2	**Cit**-Trp-**Asn-Thr**-Val
Bead 3	**Orn**-Trp-**His**-**Tyr**-Val
Bead 4	**Orn**-Trp-**Tyr**-**Arg**-Val
Bead 5	**Ser**-Trp-**Thr-His**-Val
Bead 6	**Orn**-Trp-**Gln-Dab**-Val
Bead 7	**Arg**-Trp-**His-Orn**-Val
Bead 8	**Arg**-Trp-**Asn-Arg**-Val

vi. KFFK Library Design

An alternative method of analysis for the library screen was pursued in an attempt to build a library which included varied positions throughout the peptide. Because Edman degradation sequences peptides from the N-terminus, it becomes more expensive with each additional position sequenced. As a result, making variations at many positions would become costly. Using tandem mass spectrometry for analyzing hits would allow for variation in any position without added expense. Due to their advantages in peptide libraries, the TentaGel Macrobeads were used as the resin for the library. A photocleavable linker (See Experimental Section; Figure 3.10) was purchased for use in the library so that the non-cleavable linker of the TentaGel resin could be functionalized with a linker which could be

cleaved after screening.[21] This linker allows for a library screen to be performed and for analysis to be made on each bead "hit." In this method, hits can be extracted and individually cleaved by light for mass spec analysis and sequencing. A similar method was used by Zuckermann and colleagues in which they used a cleavable linker for their one-bead-one-compound library, and they analyzed their hits by tandem mass spectrometry.[22] Our method differs from this only in that it employs a photocleavable linker for cleavage of the hits from the beads. Single bead cleavage of a template peptide was achieved and the mass spectrum showed the correct molecular weight of the peptide.

The peptide made for analysis was KFFK (Figure 3.9), a peptide based on the WKWK template but which has cross-strand phenylalanine residues to mimic some OB-fold proteins. The peptide design retains the Asn-Gly turn sequence that has provided well folded β-hairpins in previous peptides.[1] It also maintains its +4 charge by switching the Lys near the turn (Lys 4) to the original Trp position (position 2) and placing the Phe residues in the Lys 4 position and in the Trp 9 position. Overall the change in sequence is a displacement of the Trp binding pocket with a cross-strand Phe-Phe pair near the turn (Figure 3.9). This could mimic certain OB-fold domains because many are known to have a pocket of Phe residues at the DNA binding site.[23]

---

[21] Burgess, K.; Martinez, C. I.; Russell, D. H.; Shin, H.; Zhang, A. J. *J. Org. Chem.* **1997**, *62*, 5662-5663.

[22] Paulick, M. G.; Hart, K. M.; Brinner, K. M.; Tjandra, M.; Charych, D. H.; Zuckermann, R. N. *J. Comb. Chem.* **2006**, *8,* 417-426.

[23] (a) Chan, K.-W.; Lee, Y.-J.; Wang, C.-H.; Huang, H.; Sun, Y.-J. *J. Mol. Biol.* **2009**, *388*, 508–519. (b) Warren, E. M.; Huang, H.; Fanning, E.; Chazin, W. J.; Eichman, B. F. *J. Biol. Chem.* **2009**, *284*, 24662-24672.

(a) KFFK: Ac-Arg-**Lys**-Val-**Phe**-Val-Asn-Gly-Orn-**Phe**-Ile-**Lys**-Gln-NH_2

(b)

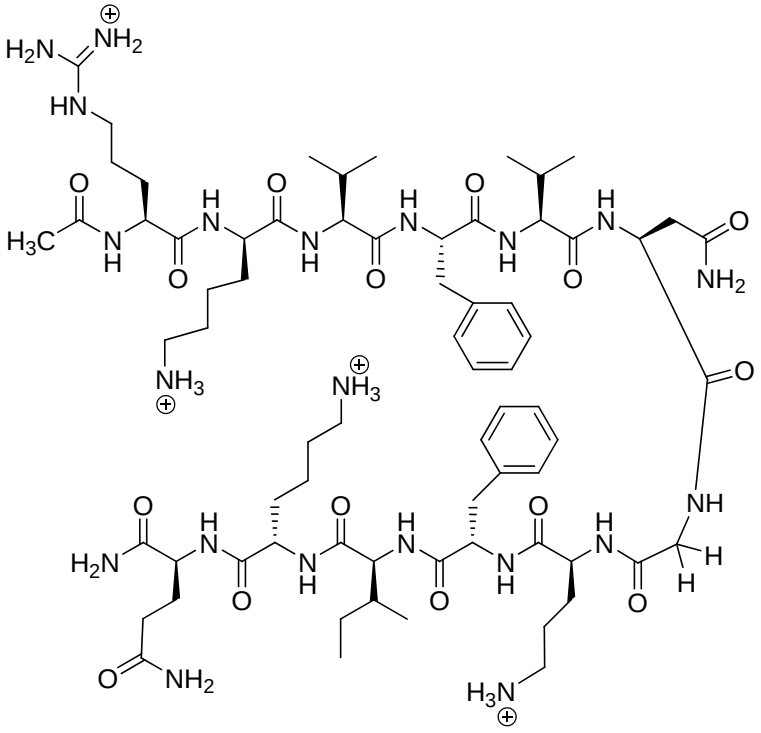


Figure 3.9. (a) Sequence of peptide KFFK with the new binding pocket residues shown in bold. (b) Structure of KFFK.

If ATP binds KFFK with comparable affinity to that of WKWK, this peptide could be used as a template for creating a new library. Additional mutations could be made to the template as in the previous library to determine residues that could provide more favorable binding interactions with ssDNA. To determine KFFK's binding potential, we conducted NMR titrations with ATP. The KFFK peptide does not have any Trp residues, so fluorescence quenching could not be used.

An NMR 1D experiment was performed to determine if any chemical shift change was evident after addition of ATP to the KFFK peptide sample. Very little chemical shift difference was observed upon titration with ATP (Table 3.8). If binding was occurring, it

was too weak to be useful in a library, so this peptide was not pursued as a template for DNA binding in library screens.

Table 3.8. Chemical shift difference for ATP titration of KFFK peptide.[a]

Peptide	Chemical Shift Difference (ppm)
KFFK	6.801, 7.185, 7.224, 7.335
ATP + KFFK	6.761, 7.149, 7.204, 7.294

(a) Conditions: Values calculated from data obtained at 298 K, 10 mM potassium dideuterium phosphate, pD 7.0 (uncorrected), referenced to DSS. Error is $\pm$ 0.005 ppm, determined by chemical shift accuracy on the NMR spectrometer. The ATP concentration was 6.7 mM; the KFFK concentration was 0.67 mM.

Tandem mass spectrometry was attempted for the KFFK peptide to determine if the photocleavage of single hits along with the MS/MS method would provide unambiguous identification of the library hits' sequences. Unfortunately, the tandem mass spectrometry did not provide data that could give sequence information. Some of the amino acids could be identified while others were ambiguous. One difference between this method and Zuckermann's is that KFFK is a 12 residue peptide while the Zuckermann group used a pentamer. The difference in peptide length could affect the mass spectrometry analysis due to overlapping signals for amino acids with similar masses. Thus, this method was not pursued further.

C. Discussion and Conclusions

The peptide libraries resulted in several different hits for each, but the overwhelming result was the selection of basic residues that prevented the desired result of favorable hydrogen-bonding interactions to improve selectivity. In this case, one step that could be taken is to exclude charged residues from the library. That could produce favorable

hydrogen-bonding residues as hits and prevent nonspecific binding. The overall conclusion from both library studies is that the original WKWK peptide is optimized for binding with its Trp-Trp binding pocket and the two cationic Lys residues to form electrostatic interactions with the phosphates. Other aromatic residues were selected in place of Trp, but fluorescence studies have shown that Phe incorporated into the binding pocket does not recognize ssDNA as well as Trp within this particular peptide. There seems to be no selectivity for a specific basic residue. This is likely because of nonspecific electrostatics. The KFFK peptide was determined to be a weaker ATP binder than WKWK, so this peptide was not pursued for further library studies. We also found that tandem mass spectrometry does not give unambiguous results for library members of this size, so that method will no longer be utilized for library studies of this kind.

D. Experimental Section

i. Library Design and Synthesis

The libraries were designed such that the WKWK template was used. Both libraries were synthesized manually by solid phase peptide synthesis using the split-and-pool synthesis method.[18] For the first library, varied residues were positions two and four. For the second library, the varied positions were residues one, three, and four.

Peptides studied in the library screens were synthesized manually using Fmoc protected amino acids and TentaGel S NH_2 Macrobeads resin (0.21 mmol/g loading with 0.21 mmol scale). Amino acid residues (5 equivalents each) were activated with 5 equivalents of HOBT/HBTU solution in DMF along with 5 equivalents of DIPEA for 90 minutes. The peptide was washed with DMF and methanol between couplings and deprotections. Amino acids were Fmoc deprotected with 20% piperidine in DMF (two 15

minute washes). The Kaiser test was performed after each deprotection and after each coupling to confirm that each residue was deprotected and coupled appropriately. Deprotection of the sidechains was completed in 95% TFA, 2.5% water, and 2.5% TIPS for 4 hrs.

ii. Library Screening

Both libraries were screened against the single-stranded nucleotide sequence 5'-d(AAAAA)-3' (penta-dA). A colorimetric assay reported by the Schreiber group which employs the use of biotin and streptavidin was used to screen the libraries for hits.[18d-18e] The pentanucleotide was purchased from Integrated DNA Technologies, Inc., with the 5' end biotinylated, and streptavidin was purchased from Pierce Biotechnology, Inc., in conjugate with alkaline phosphatase (SAAP). The alkaline phosphatase is essential for the dye that is utilized in the screen. All screens were performed by incubating the biotinylated penta-dA with the streptavidin/alkaline phosphatase complex in high salt (500 mM NaCl) phosphate buffer, pH 7.4. (Screening the library with lower salt concentrations resulted in too many hits due to nonspecific electrostatic interactions.) The resin containing the peptide library (~75 mg) was also incubated in high salt PBS. Once the biotin was complexed with the streptavidin, the Biotin-penta-dA/SAAP was allowed to bind to the peptides on the resin. After reaction, the resin was washed with high salt screening buffer four times and then washed with a dye for 10 minutes. The dye is a mixture of nitro blue tetrazolium chloride/5-Bromo-4-Chloro-3'-Indolyl phosphate p-Toluidine Salt (NBT/BCIP). When in the presence of alkaline phosphatase, the phosphate group of the BCIP is hydrolyzed and then a tautomerization occurs, followed by the dimerization of the compound by NBT. The resulting compound produces an insoluble indigo dye that colors the beads which are bound

to the biotinylated DNA. After quenching the reaction with EDTA, the beads were washed with phosphate buffer and then plated in screening buffer. The beads were various colors of purple, and the darkest beads were picked as hits and sequenced by Edman degradation. Control experiments were performed to show that the screen was choosing true hits due to binding.

iii. Photocleavable Linker

A photocleavable linker, Fmoc-aminoethyl photolinker (4-{4-[1-(Fmoc-amino)ethyl]-2-methoxy-5-nitrophenoxy}butanoic acid) (Figure 3.10), was coupled to the TentaGel Macrobeads and manual synthesis of the peptide KFFK (Figure 3.9) was completed as described for a regular manual synthesis. The N-terminus was capped for this peptide because Edman degradation was not performed. The capping of the N-terminus was performed by acetylating with 50% acetic anhydride in DMF. Single bead cleavages were performed by subjecting one bead to 365 nm light for 3 hours in 5 μL 1:2 acetonitrile:water.[21] Longer cleavages (~12 hours) were performed to get better mass spectral data.

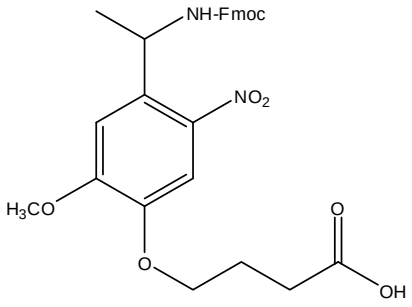


Figure 3.10. Structure of the Fmoc aminoethyl photolinker used for KFFK synthesis.

iv. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using an Applied Biosystems Pioneer Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 10 minutes. Each amino acid was coupled on an extended cycle of 75 minutes to improve coupling. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin was performed in 94% trifluoroacetic acid (TFA), 2.5% H_2O, 2.5% ethanedithiol (EDT), and 1% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen, and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 5-35% solvent B over 25 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and the peptide sequence was confirmed by MALDI mass spectrometry. Peptide dimers were formed by oxidation of the cysteine residues with stirring in 1% DMSO in 10 mM phosphate buffer, pH 7.5 for 7-12 hrs. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

v. DNA Sample Preparation

DNA sequences were purchased from IDT (Integrated DNA Technologies). All DNA samples were dissolved in 10mM Na_2HPO_4, 100 mM NaCl, adjusted to pH 7.0. Concentrations of both DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficients of the two DNA strands ($\varepsilon_{260,\ ssDNA} = 95500\ M^{-1}\cdot cm^{-1}$ and $\varepsilon_{260,\ dsDNA} = 112600\ M^{-1}\cdot cm^{-1}$). Equal concentrations of the two strands (in sodium phosphate buffer, pH 7.0) were pooled in a final concentration of 100 mM NaCl. The solution was heated at 95 ºC for 5 minutes to anneal the strands and was then allowed to cool to room temperature before storing at -20 ºC.

vi. Fluorescence Titrations

To determine the recognition of single-stranded and double-stranded oligonucleotides by the peptide dimers, fluorescence titrations were performed which followed the Trp quenching with increasing oligonucleotide concentration. Peptide and nucleotide samples were prepared in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0. Peptide concentrations were determined in 5 M guanidine hydrochloride by recording the absorbance of the Trp residues at 280 nm ($\varepsilon = 5690\ M^{-1}cm^{-1}$) by UV/vis spectroscopy. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. Experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 348 nm were fit as a function of

nucleotide concentration to the binding equation (Equation 3.1) on Kaleidagraph using non-linear least squares fitting.[24]

Equation 3.1. $I = [I_o + I_\infty([L]/K_d)]/[1 + ([L]/K_d)]$

where I is the observed fluorescence intensity, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, and K_d is the dissociation constant.

Oligonucleotides have an observable absorbance at the excitation wavelength of Trp (297 nm), and therefore there is an inner filter effect for which one must take account. The absorbance of the oligonucleotides at 297 nm was monitored at known concentrations and the extinction coefficient was determined. Absorbance values were determined for each oligonucleotide concentration. Corrected fluorescence values were determined from the following equations (Equation 3.2 and Equation 3.3).[25]

Equation 3.2. $F_c = F_o/C_i$

Equation 3.3. $C_i = (1\text{-}10^{-Ai})/(2.303)(A_i)$

where F_c is the corrected fluorescence, F_o is the fluorescence observed, and C_i is the correction factor for each absorbance value (i). A_i is the absorbance value for each concentration determined by the extinction coefficient.

[24] Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994**, *3*, 1261-1266.

[25] Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424-458.

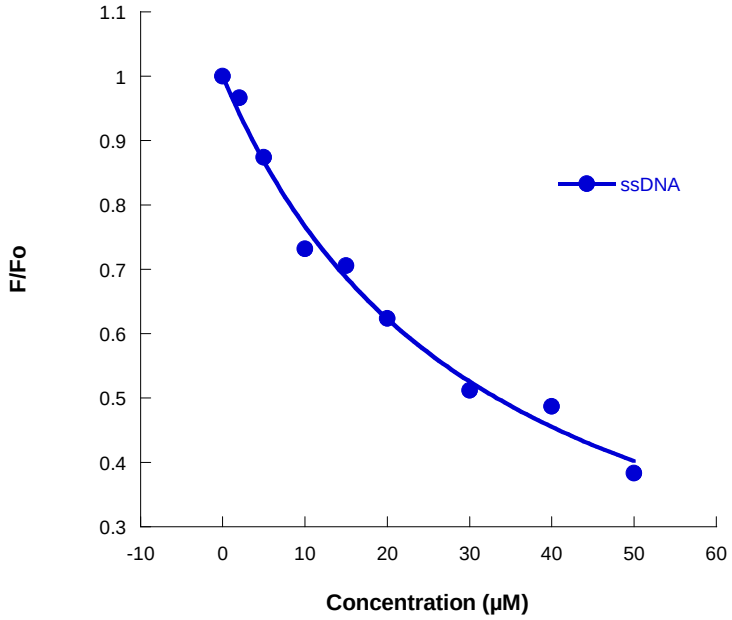


Figure 3.11. Fluorescence titration of FKWK dimer (Figure 3.5b) with ssDNA (Figure 3.5c) corrected for the inner filter effect (see Experimental Section); 4 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

vii. NMR Spectroscopy

Both KFFK and ATP were dissolved in 10 mM K_2DPO_4, 0.5 mM DSS, and buffered to pD 7.0 (uncorrected) with sodium deuteroxide and deuterated phosphoric acid. Samples were analyzed on a Varian Inova 600 MHz spectrometer at room temperature. All one-dimensional NMR spectra were collected using 32K data points and between 16 and 64 scans with 1-3 s presaturation or solvent suppression. The ATP concentration (6.7 mM) was ten times the concentration of peptide (0.67 mM) to ensure that complete binding could occur. A 1D NMR experiment was conducted for solutions of ATP, KFFK and an ATP/KFFK mixture. These were analyzed for chemical shift changes that could indicate binding (Table 3.8).

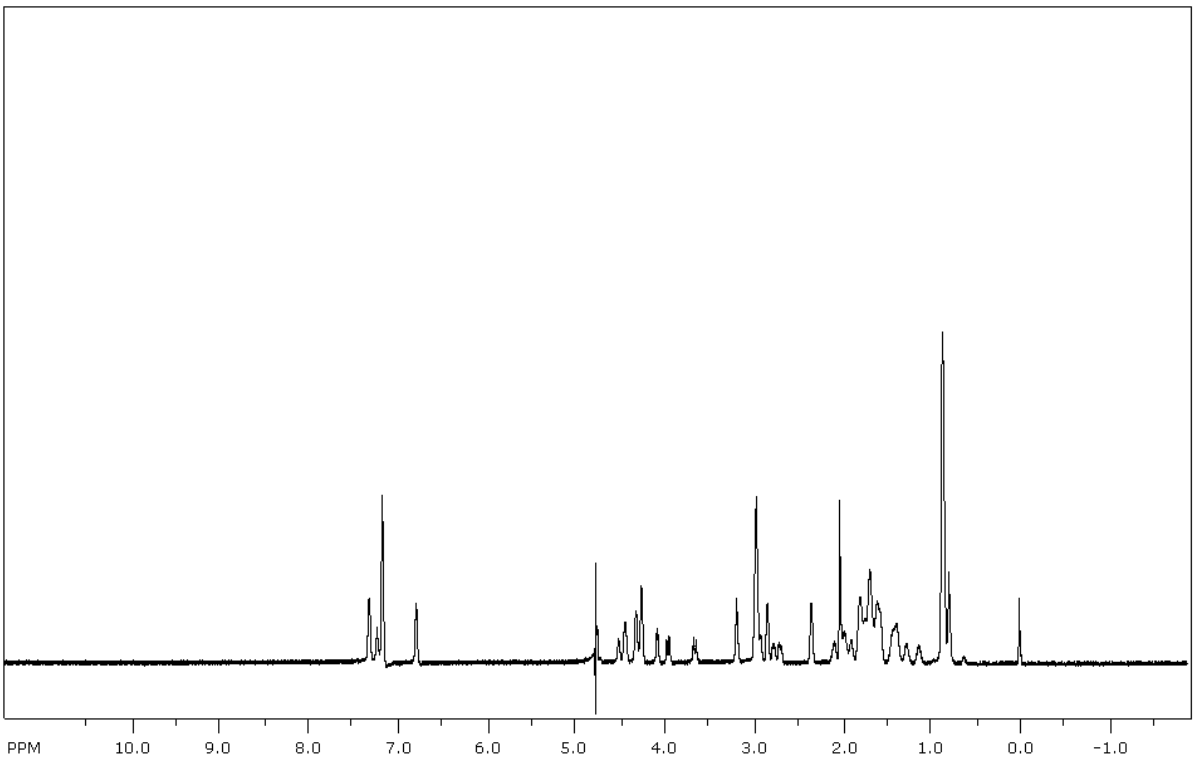


Figure 3.12. [1]HNMR of peptide KFFK.

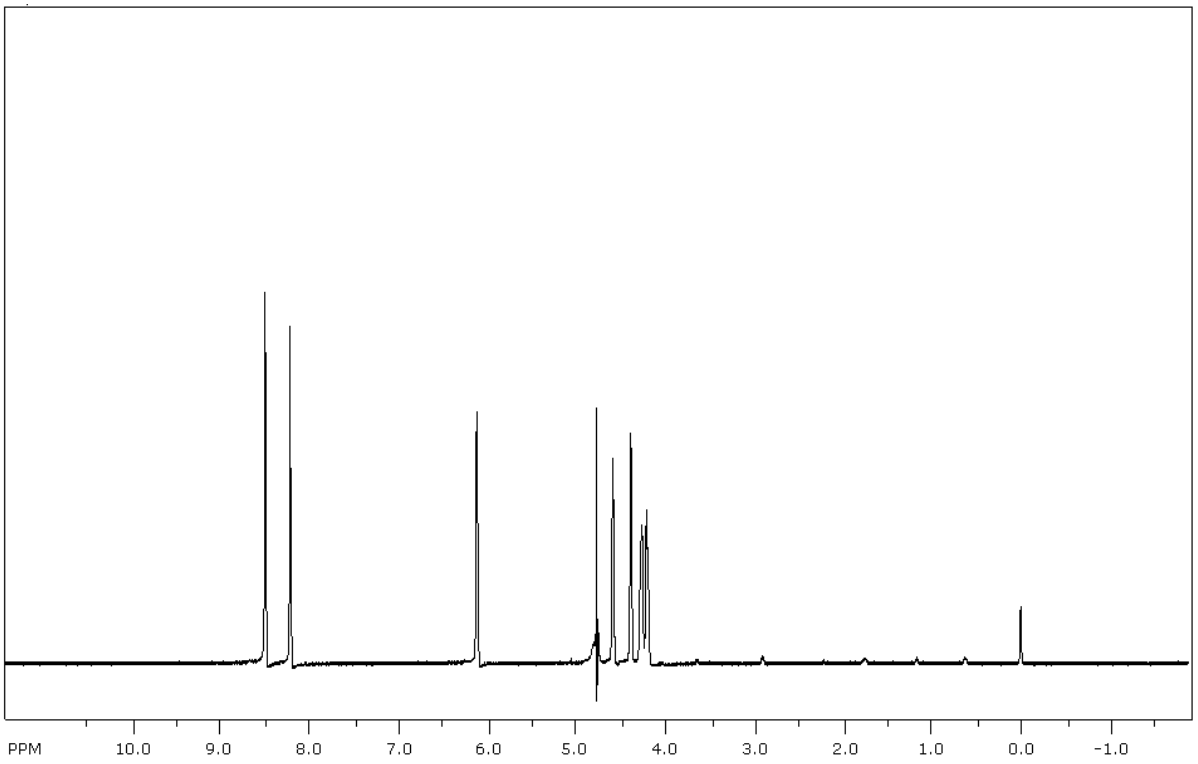


Figure 3.13. [1]HNMR of ATP.

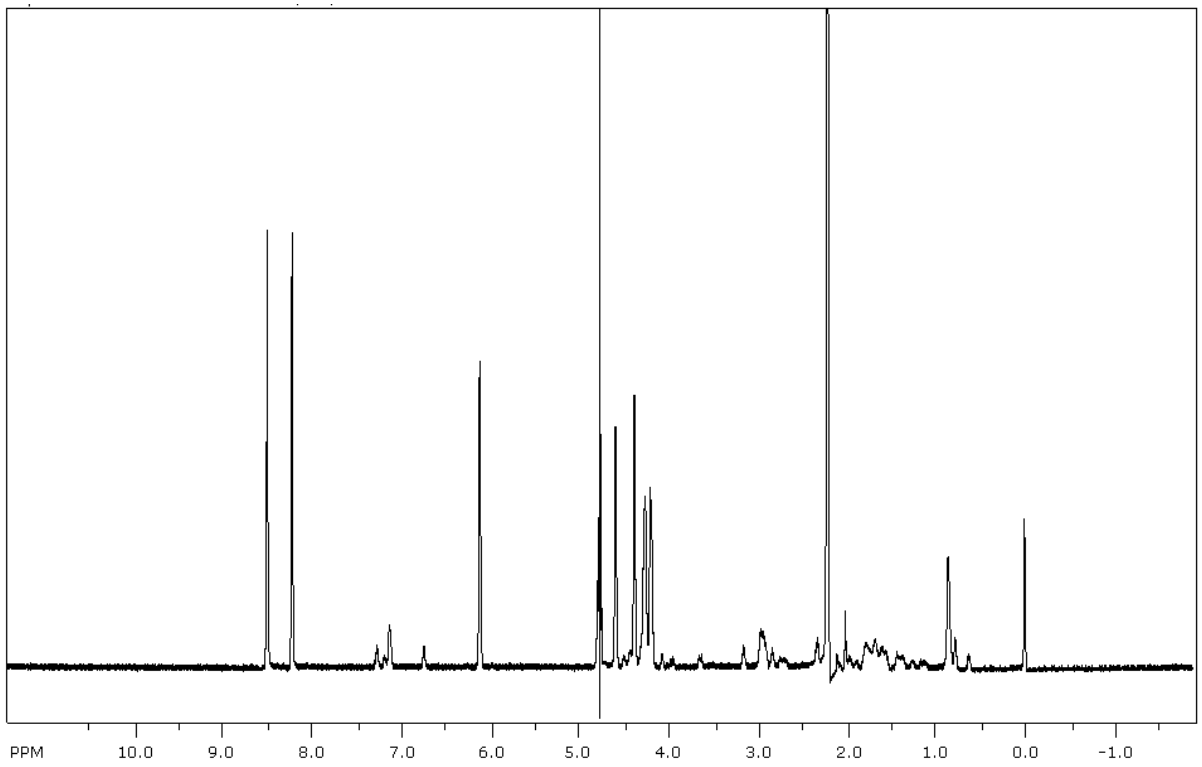


Figure 3.14. ^{1}HNMR Titration of peptide KFFK with ATP.

Chapter IV

Redesign of a WW Domain Peptide for Selective Recognition of ssDNA

A. Background and Significance

While much research has been performed to determine the mechanisms by which proteins interact with double-stranded DNA,[1] the interactions between β-sheet proteins and single-stranded DNA are less well studied. This laboratory previously reported a β-hairpin peptide dimer designed as a minimalist oligonucleotide/oligosaccharide binding motif (OB-fold, Figure 4.1[2]) mimic. The peptide $(WKWK)_2$ is composed of two identical well folded beta-hairpins, each with a binding cleft made up of two diagonal Trp sidechains and two Lys residues, which binds ssDNA with a K_d of 3 μM.[3] This is comparable to the 6 μM K_d of cold shock protein A with ssDNA; the cold shock protein A consists of a single OB-fold.[4] $(WKWK)_2$ mimics an OB-fold in that it uses a combination of aromatic and electrostatic

---

[1] (a) Krajewska, W. M. *Int. J. Biochem.* **1992**, *24*, 1885-1898. (b) Kaptein, R. *Curr. Opin. Struct. Biol.* **1993**, *3*, 50-56. (c) Latchman, D. S. *Int. J. Biochem. Cell. Biol.* **1997**, *29*, 1305-1312. (d) Harrison, S. C. *Nature* **1991**, *353*, 715-719. (e) Travers, A. *DNA—Protein Interactions.* Chapman and Hall: London, 1993.

[2] Max, K. E. A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *J. Mol. Biol.* **2006**, *360*, 702-714.

[3] Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25.

[4] Hillier, B. J.; Rodriguez, H. M.; Gregoret, L. M. *Folding Des.* **1998**, *3*, 87-93.

interactions on the surface of a β-sheet to bind to the unpaired nucleotides in ssDNA.[3, 5] However, $(WKWK)_2$ binds dsDNA with a similar affinity, albeit primarily through electrostatic interactions.[3] Structure-function studies demonstrated additional differences in the mechanism of binding to ss- and dsDNA.[6] In particular, these studies indicated that binding to duplex DNA may be occurring via groove binding. Thus, we hypothesized that addition of a third strand may inhibit binding to duplex DNA while maintaining or increasing affinity for ssDNA. To this end, a three-stranded β-sheet peptide based on the FBP11 (formin binding protein 11) WW1 domain peptide has been redesigned for binding to ssDNA. While WW domains have been studied extensively in the areas of protein folding and protein design,[7] the potential for DNA binding has never been explored. The WW domain proteins are three-stranded β-sheet mini proteins known for their conserved tryptophan residues. The natural ligands for FBP WW domains are proline-rich sequences

[5] (a) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581. (b) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70,* 1105-1114.

[6] Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544.

[7] (a) Jager, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10648-10653. (b) Fernandez-Escamilla, A. M.; Ventura, S.; Serrano, L.; Jimenez, M. A. *Protein Sci.* **2006**, *15*, 2278-2289. (c) Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3954-3959. (d) Kraemer-Pecore, C. M.; Lecomte, J. T. J.; Desjarlais, J. R. *Protein Sci.* **2003**, *12*, 2194-2205. (e) Jiang, X.; Kowalski, J.; Kelly, J. W. *Protein Sci.* **2001**, *10*, 1454-1465. (f) Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373-393. (g) Macias, M. J.; Gervais, V.; Civera, C.; Oschkinat, H. *Nat. Struct. Biol.* **2000**, *7*, 375-379. (h) Koepf, E. K.; Petrassi, M.; Ratnaswamy, G.; Huff, M. E.; Sudol, M.; Kelly, J. W. *Biochemistry* **1999**, *38*, 14338-14351. (i) Nguyen, H.; Jager, M.; Moretto, A.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3948-3953. (j) Jager, M.; Dendle, M.; Fuller, A. A.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 2306-2313.

which often form polyproline helices,[8] with the FBP11 WW1 domain ligand being PPLP.[8a-f] Knowledge of the structural features of this class of proteins combined with data from previous ssDNA binding peptides from this laboratory led to this redesign of a WW domain mutant (Figure 4.2) which binds ssDNA in the low micromolar range with 20-fold selectivity over duplex DNA. This chapter describes the redesign of a WW domain peptide as a molecular receptor selective for ssDNA, mimicking the natural OB-fold domain. These model systems provide a method to reveal factors contributing to protein-nucleic acid recognition.

[8] (a) Kato, Y.; Miyakawa, T.; Kurita, J.; Tanokura, M. *J. Biol. Chem.* **2006**, *281*, 40321-40329. (b) Pires, J. R.; Parthier, C.; do Aido-Machado, R.; Wiedemann, U.; Otte, L.; Bohm, G.; Rudolph, R.; Oschkinat, H. *J. Mol. Biol.* **2005**, *348*, 399-408. (c) Ball, L. J.; Kuhne, R.; Schneider-Mergener, J.; Oschkinat, H. *Angew. Chem. Int. Ed.* **2005**, *44*, 2852-2869. (d) Kato, Y.; Nagata, K.; Takahashi, M.; Lian, L.; Herrero, J. J.; Sudol, M.; Tanokura, M. *J. Biol. Chem.* **2004**, *279*, 31833-31841. (e) Bedford, M. T.; Chan, D. C.; Leder, P. *EMBO J.* **1997**, *16*, 2376-2383. (f) Otte, L.; Wiedemann, U.; Schlegel, B.; Pires, J. R.; Beyermann, M.; Schmieder, P.; Krause, G.; Volkmer-Engert, R.; Schneider-Mergener, J.; Oschkinat, H. *Protein Sci.* **2003**, *12*, 491-500. (g) Zarrinpar, A.; Lim, W. A. *Nat. Struct. Biol.* **2000**, *7*, 611-613. (h) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Peptide Sci.* **2005**, *80*, 303-311. (i) Dalby, P. A.; Hoess, R. H.; DeGrado, W. F. *Protein Sci.* **2000**, *9*, 2366-2376. (j) Chan, D. C.; Bedford, M. T.; Leder, P. *EMBO J.* **1996**, *15*, 1045-1054. (k) Macias, M. J.; Hyvonen, M.; Baraldi, E.; Schultz, J.; Sudol, M.; Saraste, M.; Oschkinat, H. *Nature* **1996**, *382*, 646-649. (l) Chen, H. I.; Sudol, M. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7819-7823.

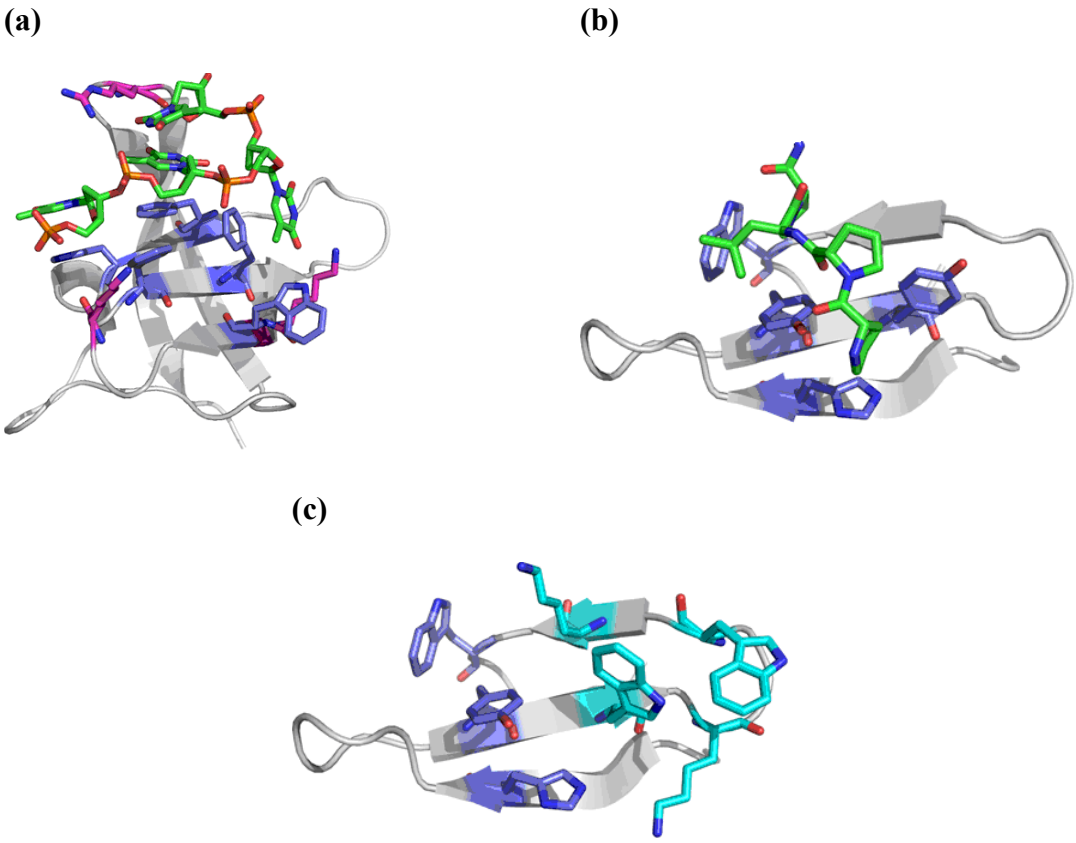


Figure 4.1. (a) Structure of a single OB-fold in cold shock protein B from *Bacillus subtilis* bound to dT_6 (pdb: 2es2).[2] (b) Structure of the FBP11 WW1 domain bound to a polyproline helix (pdb: 2dyf). (c) FBP11 WW1 domain showing the WKWK binding pocket mutations (cyan) in Mut1. Aromatic residues are shown in blue, basic residues are shown in magenta, and ligands are shown in green. Mutations were generated with Pymol.

B. Results

i. Sequence Design

In further attempts to design a well folded beta-sheet peptide to selectively bind ssDNA, we explored the WW domain class of three-stranded β-sheets for use as a scaffold. Results from previous studies showed that structure is important for DNA binding.[6] Because we wanted to design a peptide that binds single-stranded DNA selectively over duplex DNA, the additional strand in the WW domain peptides was an attractive method for gaining selectivity by possibly inhibiting duplex DNA binding. Since proteins recognize ssDNA primarily through an OB-fold motif (Figure 4.1) via a combination of electrostatic, aromatic,

and hydrogen-bonding interactions,[9] we wanted to maintain a peptide template that included structural features like that of an OB-fold domain. We also endeavored to maintain a small scaffold amenable to both solid phase peptide synthesis and to the structure-function studies necessary for a DNA-binding β-sheet model system.

The FBP11 WW1 domain peptide (Figure 4.3a) used in these studies corresponds to residues 15-42 of the native peptide (residues 144-171 of the full mammalian FBP11 WW1 protein),[8f-8g] with an additional glycine residue coupled to the C-terminus for ease in synthesis. Using the same principles as with WKWK and $(WKWK)_2$,[3, 5-6] the FBP11 WW1 domain was mutated to form a putative binding cleft for ssDNA to give Mutant 1 (Mut1; Figure 4.2; 4.3b). The C-terminal β-hairpin was mutated such that a Trp binding pocket was placed on the binding face with two flanking Lys residues cross-strand to the two tryptophans. The turn sequences were changed to Asn-Gly turns since these gave well folded turns in the hairpins used previously for nucleotide binding studies,[3, 5-6] and the mutations taken together gave the peptide a +5 net charge. The C-terminal β-hairpin sequence mimics the binding face of the β-hairpin peptide, WKWK. The amino acid numbering for the mutant peptides is 1-30. Position 16 (Position 2 in Mut1) is varied in several known FBP WW domains.[8f-g] Arginine was placed in that position for Mut1 since Arg is known to provide favorable contacts in DNA binding.[10] Residues 2 and 4 are cross-strand from the binding

---

[9] (a) Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115-133. (b) Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004**, *14*, 36-42. (c) Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

[10] (a) Allen, M. D.; Yamasaki, K.; Ohme-Takagi, M.; Tateno, M.; Suzuki, M. *EMBO J*. **1998**, *17*, 5484-5496. (b) Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J.-G.; Ikegami, T.; Nakao, M.; Shirakawa, M. *Cell* **2001**, *105*, 487-497. (c) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860-2874. (d) Moodie, S. L.; Mitchell, J. B. O.; Thornton, J. M. *J. Mol. Biol.* **1996**, *263*, 486-500.

pocket in strands 2 and 3 and may be positioned to make favorable contacts with the DNA to improve binding affinity and selectivity. These additional interactions were not present in the WKWK monomer and dimer.

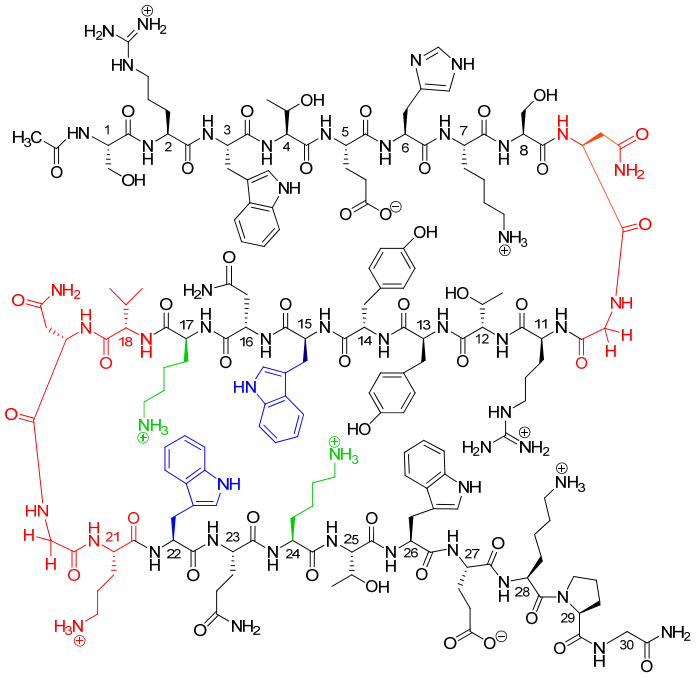


Figure 4.2. Structure of WW domain Mut1. Turn mutations are shown in red and the WKWK binding cleft is shown in blue and green.

Two additional mutants were designed to further understand the characteristics that drive the interactions of interest. Mut2 (Figure 4.3c) was designed with no proline on the C-terminus. Proline is conserved in many different WW domains and has been shown to be crucial in the folding of native WW domain peptides.[8a] In the original mutant, the Pro was retained because of its proposed importance in β-sheet stability. As noted in previous work by this laboratory, folding is crucial for binding both ssDNA and duplex DNA.[6] It is therefore thought that the C-terminal Pro may be important for binding ssDNA using the WW Domain template as well. To determine if this residue is still important for folding and ssDNA recognition in Mut1, structure and binding studies of Mut2 were performed.

Research by Kelly, et al., has shown that the native turn sequence between strands 1 and 2 is important for structure and stability,[7a] leading to another mutant made with the native turn one sequence replacing the Asn-Gly turn that was introduced in the original mutant. The other mutations remained the same to give Mut3 (Figure 4.3d).

(a) Native FBP11 WW1 domain: Ac-S-E-W-T-E-H-K-S-P-D-G-R-T-Y-Y-Y-N-T-E-T-K-Q-S-T-W-E-K-P-**G**-NH_2

(b) Mut1: Ac-S-**R**-W-T-E-H-K-S-**N**-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-**O**-**W**-Q-**K**-T-W-E-K-P-**G**-NH_2

(c) Mut2: Ac-S-**R**-W-T-E-H-K-S-**N**-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-**O**-**W**-Q-**K**-T-W-E-K-**G**-NH_2

(d) Mut3: Ac-S-**R**-W-T-E-H-K-S-**P**-**D**-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-**O**-**W**-Q-**K**-T-W-E-K-P-**G**-NH_2

(e) Mut1-S1: Ac-S-**R**-W-T-E-H-K-S-**N**-**G**-NH_2

(f) Mut1-S2: Ac-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-NH_2

(g) Mut1-S3: Ac-**N**-**G**-**O**-**W**-Q-**K**-T-W-E-K-P-**G**-NH_2

(h) Mut1-S12: Ac-S-**R**-W-T-E-H-K-S-**N**-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-NH_2

(i) Mut1-S23: Ac-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-**O**-**W**-Q-**K**-T-W-E-K-P-**G**-NH_2

(j) Mut1-S23-E27Q: Ac-**G**-R-T-Y-Y-**W**-N-**K**-**V**-**N**-**G**-**O**-**W**-Q-**K**-T-W-**Q**-K-P-**G**-NH_2

(k) WKWK: Ac-R-**W**-V-**K**-**V**-**N**-**G**-**O**-**W**-I-**K**-Q-NH_2

(l) Y-G-G-G-P-P-P-P-P-P-P-L-P-P

(m) ssDNA: 5'-CCATCGCTACC-3'

(n) dsDNA: 5'-CCATCGCTACC-3'
3'-GGTAGCGATGG-5'

Figure 4.3. Sequences of native FBP11 WW1 domain, WW domain mutants and controls, polyproline helix, and DNA sequences used in binding studies. Residues in the turn sequences and in the binding pocket are highlighted as in Figure 4.2, with turn sequences in red, Trp residues in the binding pocket in blue, and Lys residues in the binding pocket in green. Other mutations are shown in bold.

The isolated hairpins made up of strands 1 and 2 (Mut1-S12) or strands 2 and 3 (Mut1-S23) as well as single strands (Mut1-S1, Mut1-S2, and Mut1-S3; Figure 4.3e-i) were also characterized as additional means to understand the role of each strand in folding and binding to ssDNA. An additional hairpin composed of strands 2 and 3 with an E27Q

mutation, Mut1-S23-E27Q (Figure 4.3j), was also characterized to better understand the role of stability and electrostatic interactions in ssDNA recognition by Mut1.

A polyproline sequence based on a peptide used in studies as an FBP11 WW1 domain ligand (Figure 4.3l)[8b, 8m] was synthesized to determine if binding to this sequence is affected by the mutations made. This polyproline helix contains the PPLP motif known as the ligand specific to the FBP11 WW1 domain.[8b, 8h, 8i, 8m]

ii. CD Characterization of Folding

Structural studies on Mutants 1-3 were performed by CD and compared to the native FBP11 WW1 domain to determine the effect of the mutations on β-sheet structure (Figure 4.4). The CD spectra of the mutants differ from that of the native WW domain. In particular, the native protein displays a larger exciton coupling peak at 227 nm than the mutants. This is likely due to differences in orientations of the aromatic sidechains. Nonetheless, the mutants clearly exhibit β-sheet structure, as indicated by the minima at 210 nm. Although β-sheets typically have minima at about 215 nm, the shift to 210 nm is likely due to the contributions of the Trp residues, and the differences in these spectra could be due to the number and relative conformations of tryptophans in each peptide. The minima near 195 nm may represent random coil due to some degree of fraying or some extent of polyproline helix character due to the KPG sequence at the C-terminus of each peptide, as a minimum is also observed in that region for the native protein. Mut3 displays a more significant minimum near 195 nm, which may be attributed to the more flexible turn sequence between strands 1 and 2.

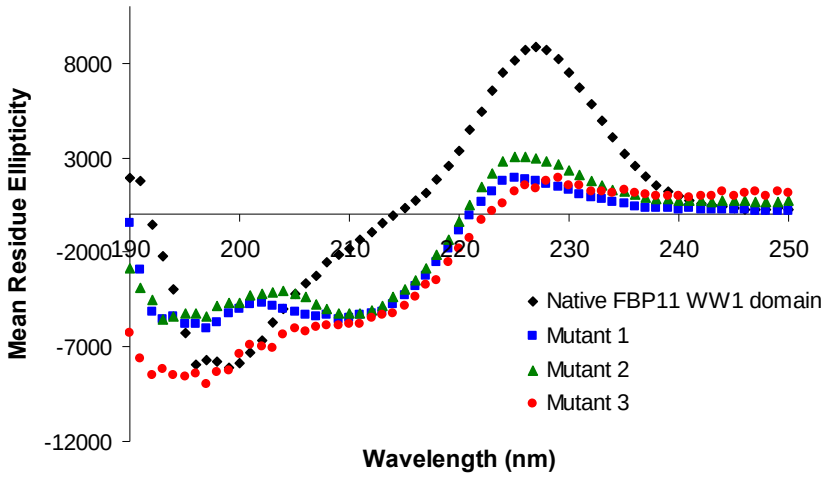
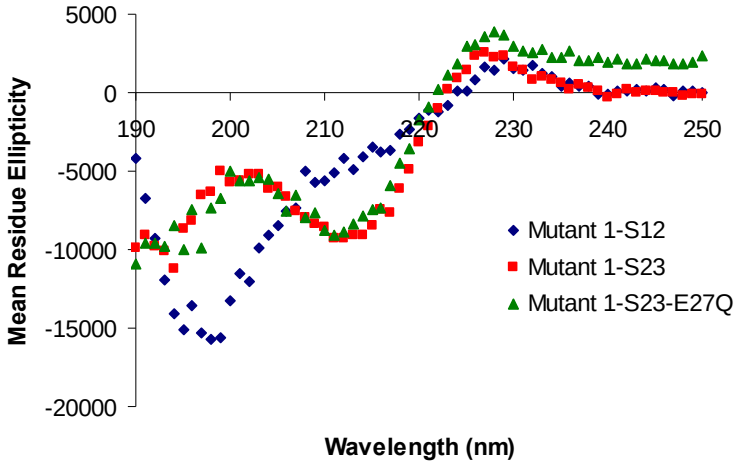


Figure 4.4. CD spectra for the native WW domain and WW domain mutants. Data was recorded at 30 μM peptide, 10 mM Na_2HPO_4, 298 K, pH 7.0, with a wavelength range from 250-190 nm.

The Mut1 control peptides Mut1-S12 and Mut1-S23 were also analyzed by CD (Figure 4.5a). Mut1-S12 is primarily unstructured with a minimum at 195 nm corresponding to random coil but with a shoulder at 215 nm which is indicative of β-sheet structure. Mut1-S23 gives a minimum CD signal at about 215 nm which is consistent with beta-sheet peptides and proteins. This peptide also has a minimum at about 190 nm, which may be due to fraying or the KPG tail at the C-terminus. The CD spectrum of Mut1-S23-E27Q is almost identical to that of Mut1-S23, as expected, with primarily β-sheet character not disrupted by the E27Q mutation. Thus, it appears that the random coil nature of the three-stranded sheets primarily comes from strand one. Each of the three control peptides, Mut1-S1, Mut1-S2, and Mut1-S3, produced CD wavelength data consistent with random coil peptides, as expected (Figure 4.5b). CD wavelength data was also obtained for the polyproline helix (Figure 4.3l), used for binding studies with the native WW domain and the mutants, to confirm that it adopts a polyproline helix structure (Figure 4.6).

(a)



(b)

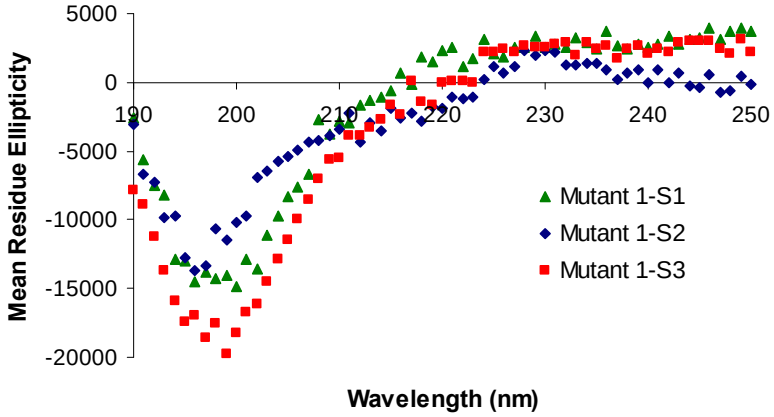


Figure 4.5. CD spectra for WW domain mutants (a) control hairpins and (b) individual strand control peptides. Wavelength data was recorded at 30 μM peptide concentrations, 10 mM Na_2HPO_4, 298 K, pH 7.0 with a wavelength range from 250-190 nm.

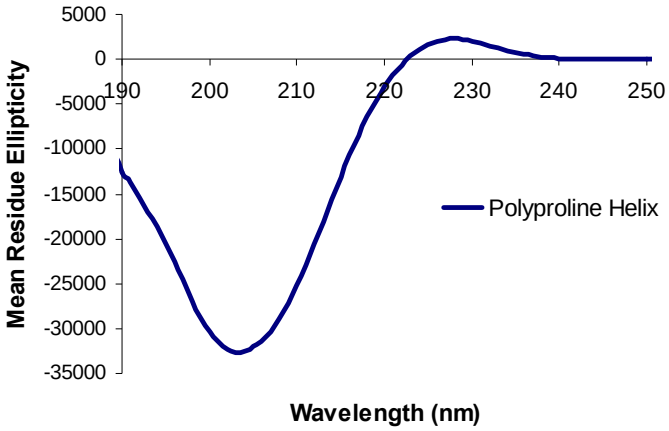


Figure 4.6. CD spectra for the polyproline helix (Figure 4.3l). CD wavelength data was recorded at 75 μM with a wavelength range from 250-190 nm. The scan was performed at 298 K.

Thermal denaturation studies were conducted for the native WW domain and for Mutants 1-3 as well as for Mut1-S23 to determine the stability of the mutants compared to the native peptide (Figure 4.7). Thermal denaturations were performed at 227 nm, following the Trp exciton coupling peak over a range of temperatures. The native FBP11 WW1 domain has a melting temperature similar to that of the hPin1 and FBP28[7a] WW domains. Mutants 1 and 3 show similar stability as the native WW domain, but Mut2 appears to be less folded (Figure 4.7). The reduced stability of Mut2 verifies the role of Pro 27 in stabilizing the mutants, as has been observed in the native protein.[8a] Mut1-S23 seems to have similar thermal stability as Mut2, so Strand 1 may increase the stability of S23 to some extent.

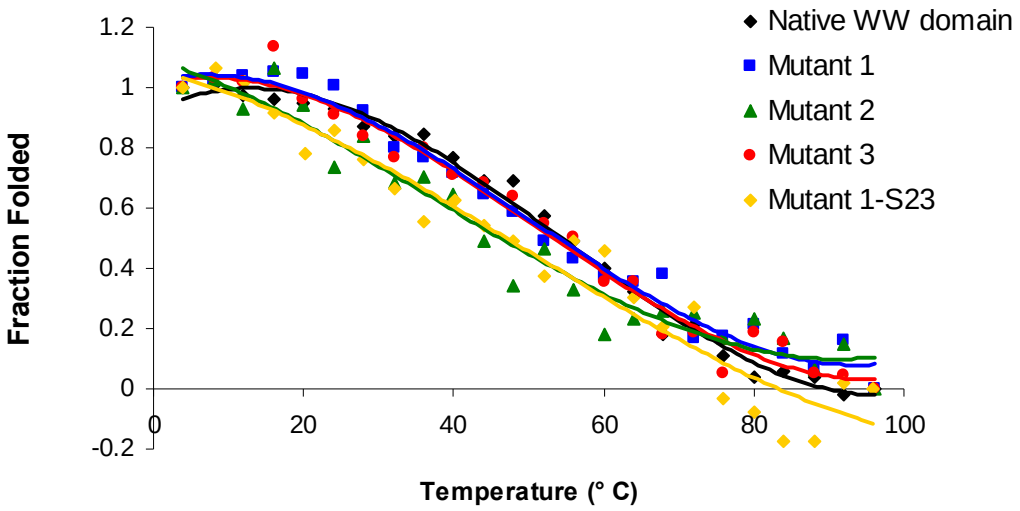


Figure 4.7. Thermal denaturation plots for the native WW domain and the mutant peptides. Thermal denaturations were performed at 227 nm from 4-96 °C with 4°/step. Data was recorded at 30 μM peptide, 10 mM Na_2HPO_4, pH 7.0. Lines are meant to guide the eye.

iii. Characterization of the Peptide-ssDNA Interactions

The stoichiometry of binding was determined for Mut1 as well as Mut1-S23 and ssDNA by fluorescence quenching using the molar variation method (see Experimental Section). A 1:1 binding stoichiometry for the interaction between Mut1 and the 11-mer ssDNA sequence (Figure 4.3m) was determined, in agreement with the stoichiometry reported for the $(WKWK)_2$ interaction with dA_5 and with the same 11-mer ssDNA (see Experimental Section; Figure 4.8).[3, 6] The binding stoichiometry for the interaction between Mut1-S23 (Figure 4.3i) and ssDNA was also determined to be 1:1 (see Experimental Section; Figure 4.9). Molar variation plots were performed using different peptide concentrations to attain optimal conditions of DNA/peptide ratios while using the highest DNA concentrations possible under the inner filter effect limitations.

Binding of ssDNA to each of the peptides was determined by quenching of the tryptophan fluorescence as described in the Experimental Procedures. A correction for the

inner filter effect arising from absorbance of the nucleobases at the excitation wavelength of Trp was performed for all binding data (see Experimental Procedures).

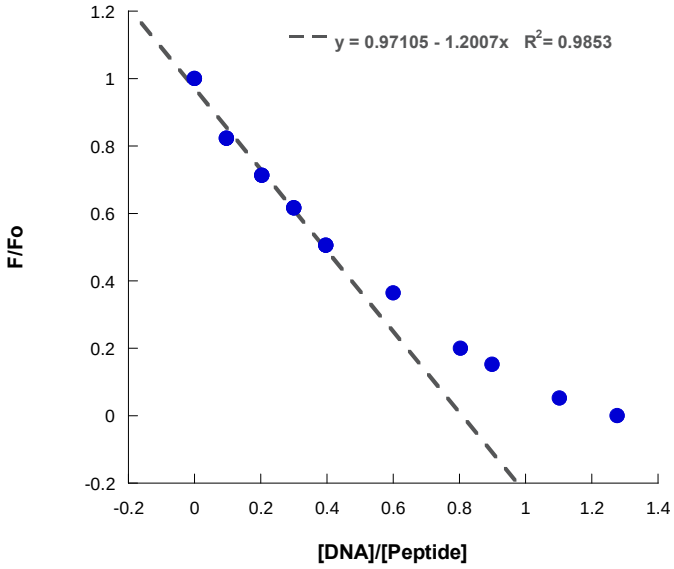


Figure 4.8. Molar variation plot for Mutant 1 (50 µM) binding to ssDNA (0-64 µM). A 1:1 binding interaction was demonstrated with the average of two runs in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, 298 K.

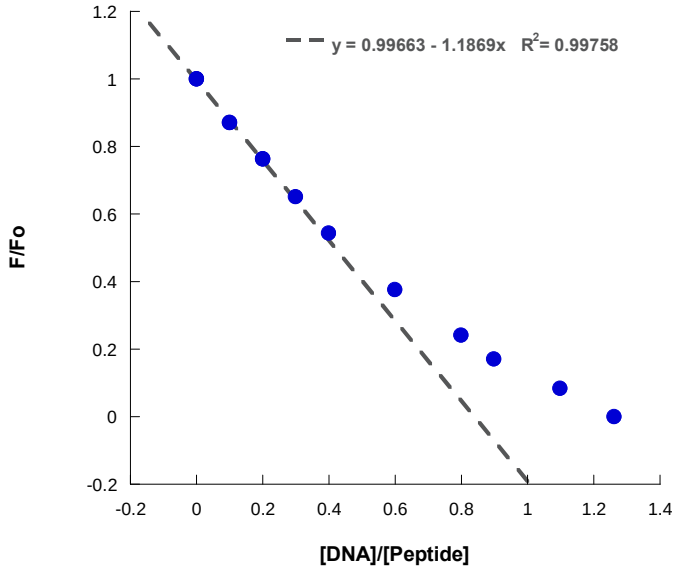


Figure 4.9. Molar variation plot for Mut1-S23 (50 µM) binding to ssDNA (0-63 µM). A 1:1 binding interaction was demonstrated with the average of two runs in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, 298 K.

As a reference point, the dissociation constant of the WKWK monomer binding to the 11-mer ssDNA sequence was determined to be 39.3 µM (Table 4.1; Figure 4.10). This affinity is much weaker than that of $(WKWK)_2$, due to the lack of the second DNA binding pocket provided by the dimer as well as a lower net charge. WKWK can be compared to Mut1 and Mut1-S23 to gain information regarding structural and sequence related aspects of ssDNA binding for each.

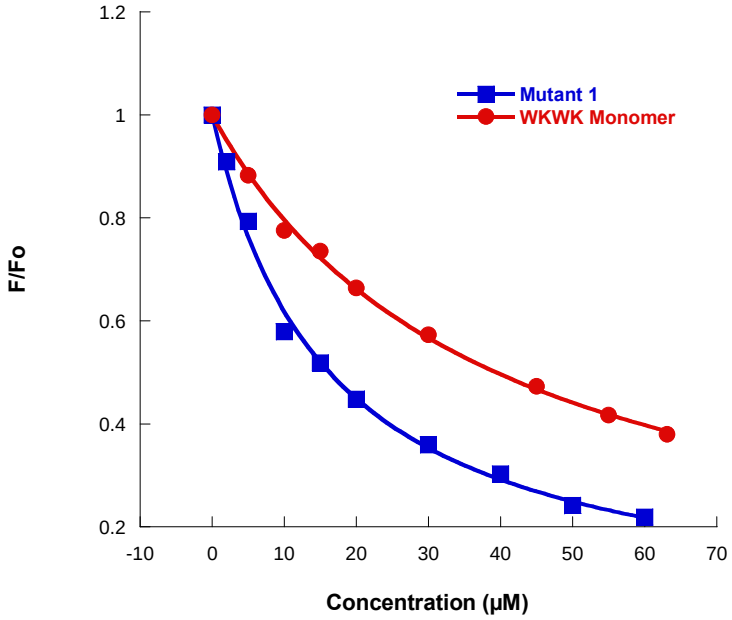


Figure 4.10. Fluorescence titrations of WW domain Mutant 1 and WKWK with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 4.3m) corrected for the inner filter effect (see Experimental Section); 3-5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

Mut1 was designed with two strands mimicking the binding pocket of the β-hairpin WKWK and has a net charge of +5. The additional N-terminal strand was intended to allow additional contacts to improve the binding to ssDNA. Fluorescence quenching experiments determined that Mut1 binds ssDNA with an affinity of 16.6 μM (Table 4.1; Figures 4.10-4.12). This peptide does not bind ssDNA as well as the WKWK dimer, which has a net charge of +8 and two aromatic binding pockets.[3] However, Mut1 binds ssDNA more than 2-fold more tightly than the WKWK hairpin monomer, which has a similar aromatic binding cleft as Mut1 and a +4 charge (Figure 4.10). This suggests that strand 1 contributes to the binding affinity and provides some additional contacts favorable for binding. Whether these

additional contacts are electrostatic or not was explored further, as described in the following section.

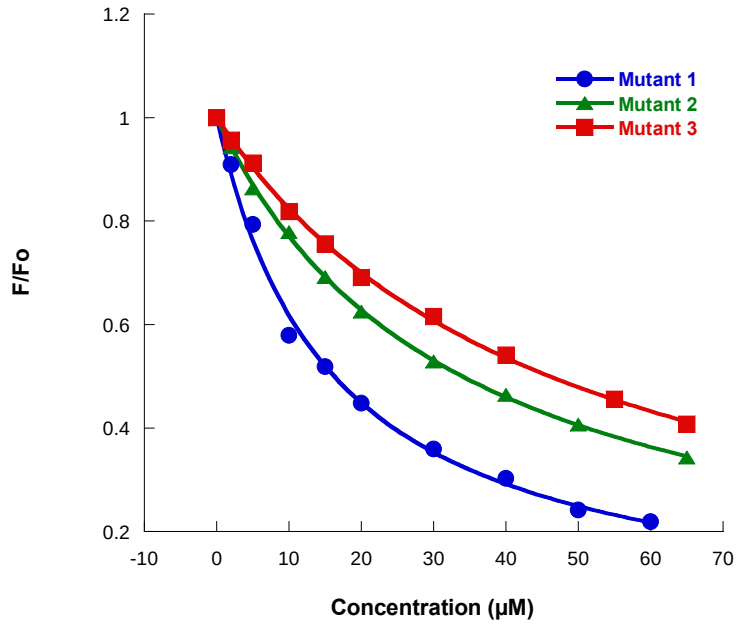


Figure 4.11. Fluorescence titrations of WW domain mutants 1-3 with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 4.3m) corrected for the inner filter effect (see Experimental Section); 5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

Table 4.1. Dissociation constants for the binding interaction between WW domain peptides and ssDNA[a]

Peptide	ssDNA K_d, μM (error)
Native WW domain	12500[b]
WKWK	39.3 (2.4)
WKWK Dimer	3.5 (0.2)[c]
Mut1	16.6 (1.5)
Mut1-S23	32.4 (1.8)
Mut1-S23-E27Q	22.9 (0.8)
Mut2	34.2 (0.9)
Mut3	45.8 (2.7)

(a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting. (b) This data was collected using fluorescence anisotropy. The conditions are the same as with the fluorescence quenching binding measurements. (c) Reported previously.[3, 6]

The C-terminal proline was removed in Mut2 to determine if the residue is important in binding since it has been shown to be critical for folding of the native peptide.[8a] Pro 37 in the native WW domain (Pro 29 in Mutant1) interacts with the N-terminal Trp 12 as well as Tyr 24 to form a small hydrophobic pocket which stabilizes the structure.[8a] Thermal denaturations suggest that it contributes to the stability of the mutants as well. Binding of Mut2 to ssDNA is weaker than Mut1, with a K_d of 34.2 μM (Table 4.1; Figure 4.11). This indicates that stability of the three-stranded sheet influences binding, confirming the importance of the folded structure.

Replacing the Asn-Gly turn sequence in turn 1 of Mut1 with the native turn sequence in Mut3 also weakened the binding as compared to the original mutant by about three-fold

(Table 4.1; Figure 4.11). It may be that strand 1 is less folded in Mut3, which would explain the more negative peak at 195 nm in the CD spectrum for Mut3 relative to Mut1 (Figure 4.4).

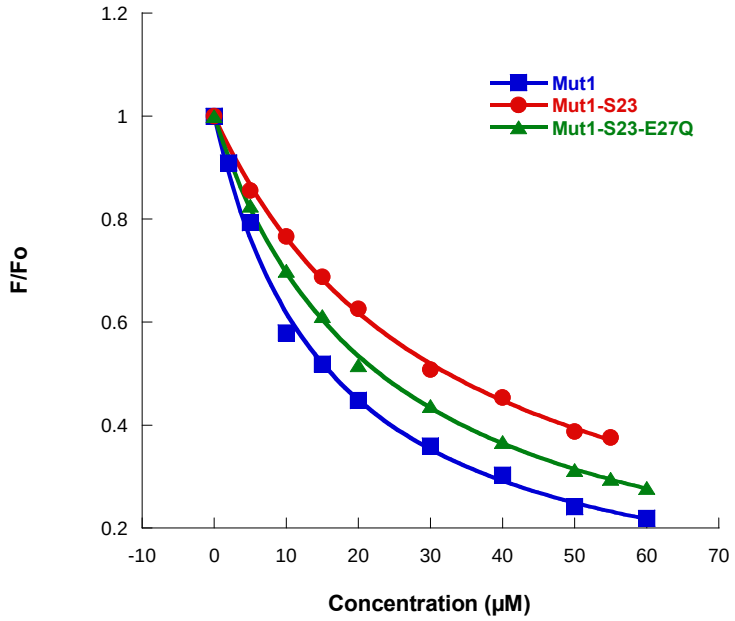


Figure 4.12. Fluorescence titrations of WW domain Mutant 1, Mut1-S23, and Mut1-S23-E27Q with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 4.3m) corrected for the inner filter effect (see Experimental Section); 5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

To further explore the role of strand 1 in Mut1, the binding of Mut1-S23 to ssDNA was analyzed by fluorescence quenching. The affinity was reduced by about two-fold (K_d = 32.4 μM; Table 4.1; Figure 4.12). The overall charge of the peptide is +4 versus the +5 charge for the full peptide, which could explain the difference in binding. However, Mut1-S23 binds ssDNA more strongly than does WKWK even though the charge is the same. Mut1-S23 has a longer sequence and the number of available contacts may enhance binding.

To determine whether charge is the only factor affecting the difference in binding of ssDNA to Mut1 versus Mut1-S23, a glutamic acid in strand 3 of Mut1-S23 was mutated to a

glutamine (E27Q), resulting in a peptide with the same net charge as Mut1. The binding affinity of Mut1-S23-E27Q was intermediate to the full peptide and strands 2&3, with a K_d of 22.9 μM (Table 4.1; Figure 4.12). This indicates that the addition of the single positive charge improves the binding affinity, but the data suggests that strand 1 in Mut1 also contributes to binding in some way other than simply net charge.

As a control, we measured the binding of the native FBP11 WW1 domain to ssDNA. While all of the data for the mutants binding to ssDNA and duplex DNA were obtained by fluorescence quenching, the binding affinities of the native peptide for these DNA sequences were too weak to be determined using that method due to inner filter effects. Fluorescence anisotropy was therefore employed to determine the binding of the native peptide to the DNA sequences. A fluorophore, 5'-Bodipy 630/650-X NHS Ester, was attached to the ssDNA sequence. Fluorescence anisotropy data gave a K_d in the millimolar range, providing evidence that the native peptide has little affinity for ssDNA (Table 4.1; Figure 4.13). Mut1 binds Bodipy-labeled ssDNA with a K_d of about 26 μM, similar to that determined by fluorescence quenching. This provides further evidence that the native WW domain has been redesigned to bind ssDNA (Figure 4.13). Control experiments were performed to confirm that the peptide does not bind to the fluorophore directly.

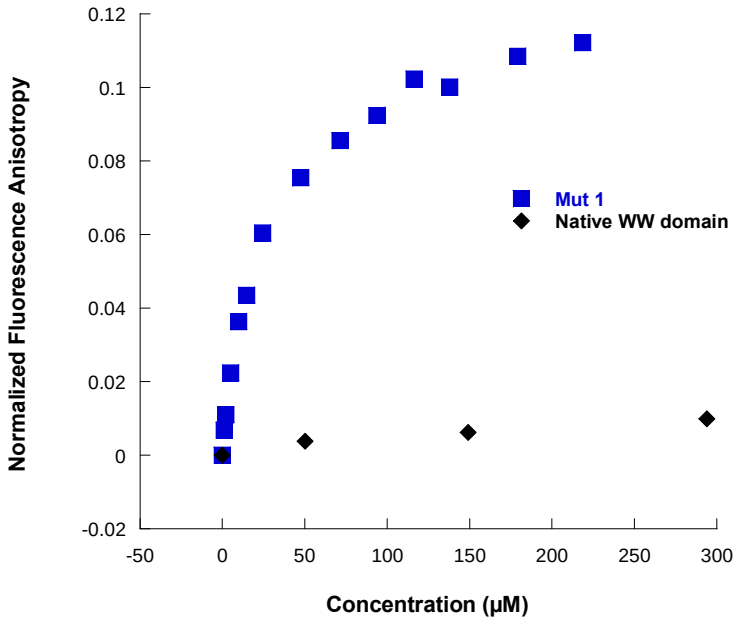


Figure 4.13. Bodipy-ssDNA titrated with the native WW domain and Mut 1. Fluorescence monitored by fluorescence anisotropy; 1.5 μM Bodipy-ssDNA for Mut 1; 50 μM Bodipy-ssDNA for native WW domain; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

iv. Characterization of the Peptide-dsDNA Interactions

Because of inner filter effects due to DNA absorption at the excitation wavelength of Trp, fluorescence quenching of Trp could not be used for accurate duplex DNA binding studies with Mut1. Instead, fluorescence anisotropy was utilized to determine a more accurate dissociation constant for the binding interaction between Mut1 and duplex DNA. (5-(and -6)-Carboxytetramethylrhodamine (TAMRA) was coupled onto the peptide Mut1 at position 21 for these binding experiments via attachment to a Lys sidechain (see Experimental Procedures). This amino acid substitution has been shown to have minimal

effects on stability and folding in model β-hairpins.[11] As a control, fluorescence anisotropy was also used to measure the binding of Mut1 to ssDNA. The fluorescence anisotropy data for binding to ssDNA is consistent with the fluorescence quenching data, giving a dissociation constant of 20.4 μM (Table 4.2). The anisotropy data shows that Mut1 binds ssDNA with about 20-fold selectivity over duplex DNA (K_d = 390 μM; Table 4.2; Figure 4.14). This selectivity of Mut1 for ssDNA versus dsDNA is much greater than that observed for $(WKWK)_2$ and ssDNA over duplex DNA.[3,6] This is consistent with earlier studies that suggested WKWK dimer is a groove binder.[6] The three-stranded sheet in Mut1 is likely too big to function as a groove binder and results in significantly greater selectivity for ssDNA.

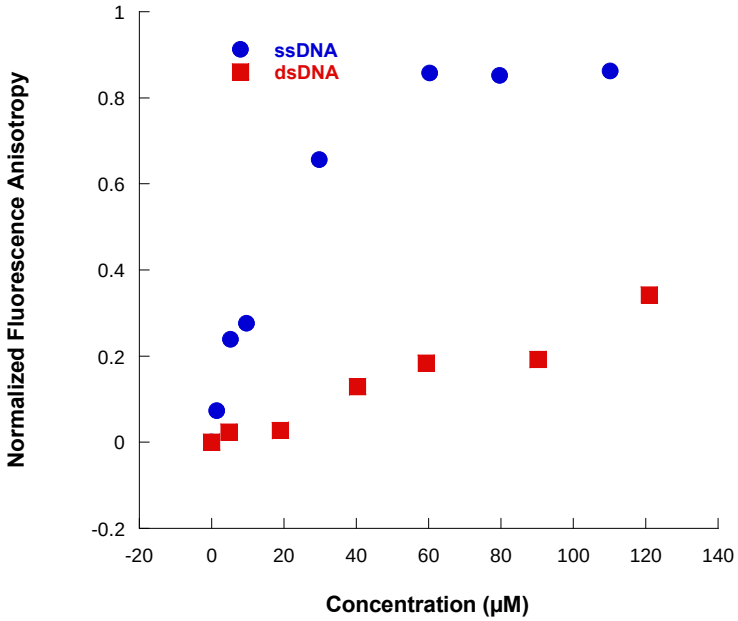


Figure 4.14. Fluorescence anisotropy titrations of TAMRA-WW domain Mut1 and ssDNA sequence 5'-CCATCGCTACC-3' (Figure 4.3m) and the complementary duplex DNA (Figure 4.3n); 2 and 20 μM peptide, respectively; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

[11] Cooper, W. J.; Waters, M. L. *Org. Lett.* **2005**, *7*, 3825-3828.

Table 4.2. Comparison of dissociation constants for the binding interactions between WW domain peptides and ssDNA and dsDNA sequences and polyproline helix[a]

Peptide	ssDNA K_d, μM (error)	dsDNA K_d, μM (error)	Polyproline Helix K_d, μM (error)
Native FBP11 WW1 domain	>6000[b]	>1100[b]	59.6 (5.9)
WKWK dimer	3.5 (0.2)[c]	9.2 (0.9)[c]	n. d.
TAMRA-Mut1	20.4 (3.8)	390 (45)	No binding observed

(a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting. n.d. denotes a measurement that was not determined. (b) This data was collected using fluorescence anisotropy with Bodipy-labeled DNA. (c) These values were reported previously.[3, 6]

Control experiments showed that binding of the native WW domain to duplex DNA was in the millimolar range, as with ssDNA. This provides further evidence that Mut1 has been successfully redesigned to bind DNA and that the native peptide has little affinity for any DNA sequence.

v. Binding to the Native Polyproline Helix

The ability of Mut1 to bind the natural ligand of the FBP11 WW1 domain was investigated by conducting fluorescence experiments with a polyproline helix containing a PPLP consensus sequence (Figure 4.3l) and comparing that to binding of the consensus sequence for the native WW domain peptide. The binding affinity for the polyproline helix used in this laboratory was 59.6 μM (Table 4.2; Figure 4.15). Pires, et al., found that the interaction between FBP11 WW1 domain and a similar polyproline helix had a K_d of 145 μM.[8b] The difference between the two polyproline sequences is that the one used in this study contains a glycine spacer of three residues and a tyrosine tag at the N-terminus (Figure 4.3l). The difference in K_ds could be due to small structural differences. The tyrosine could

contribute to the increased binding affinity to give the lower K_d. Tyrosine residues are known to form favorable interactions with other aromatic residues, and this may contribute to the binding affinity.[12] In contrast to binding by the native WW domain, no measurable binding was observed between Mut1 and the polyproline helix (Table 4.2; Figure 4.15). This demonstrates that the mutant does not bind its natural ligand and has been redesigned to bind ssDNA. No binding interaction was observed between Mutant 3 and the polyproline helix, providing additional evidence that this class of mutants has been redesigned for the purpose of binding ssDNA.

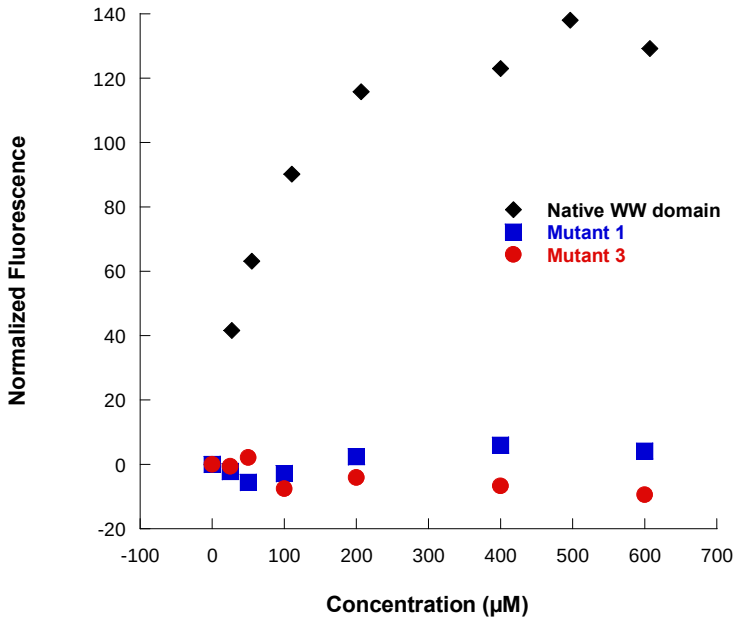


Figure 4.15. WW domain peptide titrations with polyproline helix (Figure 4.3l); 15 μM WW domain peptides, 0-600 μM polyproline helix; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. This data was normalized by subtracting the initial peptide observed fluorescence from each fluorescence value.

[12] Sillerud, L. O.; Larson, R. S. *Current Protein and Peptide Science* **2005**, *6*, 151-169.

vi. NMR Characterization of Mut1

NMR experiments were conducted to more fully characterize the folding of Mut1 and the role of strand 1 in stability and binding. Mut1-S1, Mut1-S12, and Mut1-S23 were studied by one- and two-dimensional NMR. Glycine splitting and α-hydrogen (Hα) chemical shift values were determined to characterize the full peptide, both hairpins, the three individual strands, and their respective stabilities. The Hα chemical shifts of the hairpins relative to the random coil peptides indicate the degree of β-sheet structure for each residue along the peptide backbone. Downfield shifting of Hα protons is evidence of increased hairpin population, with a chemical shift difference of greater than 0.1 ppm taken to indicate β-sheet structure.[13] A second method to determine the extent of folding of each peptide is to determine the Gly Hα splitting values. The glycine in the turn possesses two diastereotopic hydrogens. As folding of the hairpin increases, the splitting of the hydrogens increases as well, giving a measure of stability.[13b] A comparison of the Gly splitting in the β-hairpin with that of a fully folded cyclic control peptide gives the extent of folding and overall stability for the peptide (see Experimental Section). NMR peak assignments from individual strands S1, S2, and S3 are consistent with random coil values, as expected.

NMR data revealed that the N-terminal β-hairpin, Mut1-S12, has a glycine splitting value of 0.11, giving a fraction folded of 27% (Table 4.3). The C-terminal β-hairpin, Mut1-S23, is significantly more folded, with a glycine splitting value of 0.67 and a fraction folded of 89% (Table 4.3). This is similar to the WKWK peptide, which has a reported percent folding value of 96%.[5]

---

[13] (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333. (b) Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996-2007.

Comparison of Gly splitting for the full Mut1, Mut1-S12, and Mut1-S23 indicate that while the hairpin formed by strands 1&2 alone is poorly folded, addition of strand 3 in the full-length peptide adds stability to this hairpin, exhibiting an increase in structure from 27% to 63% folded (Table 4.3). This data suggests some degree of cooperative folding of the three-stranded sheet. This type of cooperativity is not unprecedented.[14] For example, Searle et al., found that the N-terminal strand of their three-stranded beta-sheet (peptide 1-24) cooperatively stabilized the C-terminal hairpin of the peptide.[14a-c] Likewise, Kelly and coworkers have shown that the hPin1 WW domain and various mutants exhibit cooperative unfolding.[14e-f] In contrast to Mut1-S12, strands 2&3 alone are as well folded in Mut1-S23 as in the full peptide, with a percent folded of 89% and 91%, respectively (Table 4.3).

Table 4.3. Fraction folded for Mut1 and control peptides.[a]

Peptide	Gly chemical shifts, ppm[b]	ΔδGly, ppm[b]	Fraction Folded (Gly Splitting)[c]
Mut1-S12	3.91, 4.02	0.11	0.27
Mut1-S23	3.45, 4.12	0.67	0.89
Mut1	3.82, 4.08;	0.26;	0.63;
	3.38, 4.06	0.68	0.91

(a) Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 7.0 (uncorrected), referenced to DSS. (b) Error is $\pm$ 0.005 ppm, determined by chemical shift accuracy on the NMR spectrometer. (c) Error is $\pm$ 0.01, determined by chemical shift accuracy on the NMR spectrometer.

Hα chemical shift differences for the two hairpins and the full peptide are consistent with both the CD data and the NMR glycine splitting data (Figure 4.16). This data indicates that Mut1-S12 is only marginally folded at best, but when incorporated into the full length

[14] (a) Sharman, G. J.; Searle, M. S. *Chem. Commun*. **1997**, 1955-1956. (b) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc*. **1998**, *120*, 5291-5300. (c) Griffiths-Jones, S. R.; Searle, M. S. *J. Am. Chem. Soc*. **2000**, *122*, 8350-8356. (d) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253-256. (e) Kaul, R.; Angeles, A. R.; Jager, M.; Powers, E. T.; Kelly, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 5206-5212. (f) Nguyen, H.; Jager, M.; Kelly, J. W.; Gruebele, M. *J. Phys. Chem. B* **2005**, *109*, 15182-15186.

peptide, strand 1 exhibits modestly increased downfield shifting, corresponding to an increase in folding. In contrast, strands 2 and 3 are well folded in both Mut1 and Mut1-S23.

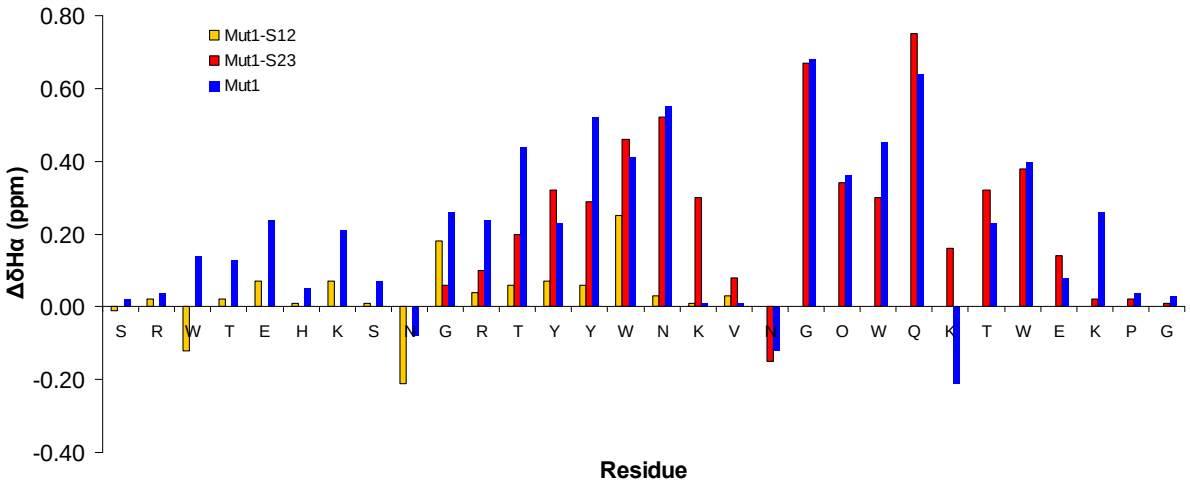


Figure 4.16. NMR chemical shift differences for WW domain Mut1 peptides. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 7.0 (uncorrected), referenced to DSS. Error is $\pm$ 0.005 ppm, determined by chemical shift accuracy on the NMR spectrometer.

C. Discussion & Conclusions

We have found that introduction of a WKWK binding pocket into a three-stranded β-sheet via mutation of a native WW domain leads to a change in function of the native protein from a polyproline helix binder to a ssDNA receptor. In contrast, the native WW domain exhibits no binding to ssDNA. Both sets of data clearly indicate that the binding pocket designed with two Trp residues to recognize and stack with a DNA base and two flanking Lys residues for electrostatic contacts with the phosphate backbone result in a redesigned peptide which binds ssDNA with high affinity and knocks out binding to the polyproline helix. The mutations which form the binding pocket lead to DNA recognition as the native peptide does not bind DNA. This result is consistent with research by Kelly and colleagues

who found that adding a Trp–Trp pair in a beta-sheet structure adds stability but at the cost of its natural function. They showed that the addition of a cross-strand Trp-Trp pair in a non hydrogen-bonding position in the hPin1 WW domain increased the peptide's thermodynamic stability but in turn decreased its phosphorylated peptide (YSPTpSPS) binding capacity.[7j]

The designed binding pocket mimics that of the WKWK peptide, but its position in this three-stranded β-sheet peptide gives it additional contacts to enhance binding that are not present in WKWK. The first strand provides residues that should increase the binding affinity and selectivity for the peptide to ssDNA through hydrogen bonding, aromatic, and electrostatic interactions. A comparison of the dissociation constants of Mutant 1 and WKWK reveals that Mutant 1 binds ssDNA with a higher affinity than does WKWK, with K_ds of 16.6 μM and 39.3 μM, respectively (Table 4.1; Figure 4.10). The new design is an improvement over the WKWK peptide as well as the WKWK dimer, which has a binding affinity of 3 μM for ssDNA but little selectivity over dsDNA.[3, 6] While the binding constant for the interaction between $(WKWK)_2$ and ssDNA is better than that of the WW domain Mutant 1, the goal of this research is to design a ssDNA binding peptide that displays selectivity for ssDNA over duplex DNA. The binding constants for $(WKWK)_2$ recognition of ssDNA and duplex DNA are both in the low micromolar range, with the dimer binding duplex DNA with a K_d of 5-9 μM.[3, 6] The difference in binding between the two is minimal, giving only a starting point for obtaining high affinity, high selectivity ssDNA binding peptides. The Mutant 1 peptide is significant because the difference in selectivity is nearly twenty-fold, much greater than the previously designed peptides.[3, 5-6] Of particular importance is that the binding affinity of Mutant 1 for ssDNA is similar to that of cold shock

protein A (CspA), a natural OB-fold protein whose full structure is a single OB-fold.[3, 15] This peptide is the most similar to CspA because the peptide forms a three-stranded β-sheet structure which contains a single binding pocket for DNA base recognition. CspA binds ssDNA with a dissociation constant of ~6 μM[4] while Mutant 1 binds ssDNA with a K_d of about 20 μM. Although this value is somewhat higher than the natural protein, it is within that range of affinities and binds with selectivity for ssDNA which is requisite for a mimic of the OB-fold domain.

Structure-function studies indicate that Pro 29 helps to stabilize the folded state of the mutant, as has been observed for the native protein, and this also impacts ssDNA binding affinity. In addition, Mut1, which has an Asn-Gly turn between strands 1 and 2, was found to have a stronger binding affinity to ssDNA than did Mut3, which has the native turn sequence at that position. The Pro deletion in Mutant 2 supports the previous conclusion that structure is crucial for binding both ssDNA and duplex DNA.[6] Proline 37 (Pro 29 in Mutant1) is known to maintain the structure of the full FBP11 WW1 domain peptide because it interacts with the N-terminal Trp 12 as well as Tyr 24 to form a small hydrophobic pocket which stabilizes the structure.[8a] While Mutant 1 was designed to have a vastly different function than the native peptide, presumably Pro maintains its interaction with Trp 12 and Tyr 24 to stabilize the peptide, as has been shown in native WW domains.[8a] It should be noted that Mutant 2 does give some selectivity for ssDNA. Mutant 3 demonstrates that the

[15] (a) Max, K. E. A.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *FEBS J.* **2007**, *274*, 1265-1279. (b) Newkirk, K.; Feng, W.; Jiang, W.; Tejero, R.; Emerson, S. D.; Inouye, M.; Montelione, G. T. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 5114-5118. (c) Schindelin, H.; Marahiel, M. A.; Heinemann, U. *Nature* **1993**, *364*, 164-168. (d) Schnuchel, A.; Wiltscheck, R.; Czisch, M.; Herrier, M.; Willimsky, G.; Graumann, P.; Marahiel, M. A.; Holak, T. A. *Nature* **1993**, *364*, 169-171.

original Asn-Gly turn sequence for Mutant 1 is important for nucleating a turn that best orients strand 1 to interact with strands 2&3 and which gives the higher binding affinity for ssDNA. While the native turn one sequence was shown by Kelly and co-workers to give a more stable peptide,[7a] this did not result in a higher affinity DNA binding peptide. Interestingly, CD studies suggest that Mut3 has a higher random coil population than does Mut1, presumably due to a less structured turn region and less well folded strand 1, which may explain the weaker binding of Mut3. Thus, the native turn 1 sequence may maintain higher stability in the native peptide but not when combined with the other mutations which produced Mut1. The original Mut1 design gave the most stable peptide as well as the one with the highest binding affinity and selectivity for ssDNA. Comparison of Mut1 to the truncated peptides, Mut1-S23 and Mut1-S23-E27Q, confirms that strand 1 contributes to ssDNA binding, albeit weakly. NMR studies indicate that Mut1 is well folded at strands 2 and 3, but that strand 1 is not as well folded. Nonetheless, strand 1 influences binding to ssDNA.

This three-stranded motif provides substantial (20-fold) selectivity for ssDNA over dsDNA, unlike the $(WKWK)_2$, which has only approximately 2-fold selectivity for ssDNA.[3] Previous studies suggest that $(WKWK)_2$ binds to duplex DNA via groove binding.[6] Thus, the selectivity observed for Mut1 may be due to inhibition of groove binding due to the third strand (strand 1).

These studies provide insight into features that can provide structure-selectivity protein-DNA interactions. This model system represents a conceptual mimic of the OB-fold, which consists of a beta-sheet surface that binds ssDNA. As such, it may have applications as an inhibitor of OB-fold-containing proteins such as replication protein A, which is

involved in DNA replication and repair,[16] or the telomere-binding protein *Cdc13*, which is involved in telomere regulation.[17] Further mutation studies were continued to optimize ssDNA binding affinity and selectivity for this peptide.

D. Experimental Section

i. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using an Applied Biosystems Pioneer Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 10 minutes. Each amino acid was coupled on an extended cycle of 75 minutes to improve coupling. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin as well as sidechain deprotection was performed in 95% trifluoroacetic acid (TFA), 2.5% H_2O, and 2.5% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen,

---

[16] (a) Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. *Nature* **1997**, *385*, 176-181. (b) Wold, M. S. *Annu. Rev. Biochem.* **1997**, *66*, 61-92. (c) Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115-133. (d) Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004,** *14*, 36-42.

[17] (a) Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. *Biochemistry* **2003**, *42*, 3751-3758. (b) Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. *Science* **2002**, *296*, 145-147. (c) Chandra, A.; Hughes, T. R.; Nugent, C. I.; Lundblad, V. *Genes Dev.* **2001**, *15*, 404-414.

and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 5-35% solvent B over 25 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and peptide sequence was confirmed by MALDI mass spectrometry. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

ii. DNA Sample Preparation

DNA sequences were purchased from IDT (Integrated DNA Technologies, Inc.). All DNA samples were dissolved in 10mM Na_2HPO_4, 100 mM NaCl, adjusted to pH 7.0. Concentrations of both DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficients of the two DNA strands ($\varepsilon_{260,\ ssDNA} = 95500$ $M^{-1}\cdot cm^{-1}$ and $\varepsilon_{260,\ dsDNA} = 112600$ $M^{-1}\cdot cm^{-1}$). Equal concentrations of the two strands (in sodium phosphate buffer, pH 7.0) were pooled in a final concentration of 100 mM NaCl. The solution was heated at 95 ºC for 5 minutes to anneal the strands and was then allowed to cool to room temperature before storing at -20 ºC.

iii. Fluorescence Titrations

To determine the recognition of single-stranded and double-stranded oligonucleotides by the peptides, fluorescence titrations were performed which followed the Trp quenching with increasing oligonucleotide concentration. Peptide and nucleotide samples were prepared in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0. Peptide concentrations

were determined in 5 M guanidine hydrochloride by recording the absorbance of the Trp residues at 280 nm ($\varepsilon = 5690\ M^{-1}cm^{-1}$) by UV/vis spectroscopy. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. The experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 348 nm were fit as a function of nucleotide concentration to the binding equation (Equation 4.1) with Kaleidagraph using non-linear least squares fitting.[18]

Equation 4.1. $I = [I_o + I_\infty([L]/K_d)]/[1 + ([L]/K_d)]$

where I is the observed fluorescence intensity, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, and K_d is the dissociation constant. Oligonucleotides have an observable absorbance at the excitation wavelength of Trp (297 nm), and therefore there is an inner filter effect for which one must take account. The absorbance of the oligonucleotides at 297 nm was monitored at known concentrations and the extinction coefficient was determined. New absorbance values were determined for each oligonucleotide concentration. Corrected fluorescence values were determined from the following equations (Equation 4.2 and Equation 4.3).[19]

Equation 4.2. $F_c = F_o/C_i$

Equation 4.3. $C_i = (1\text{-}10^{-Ai})/(2.303)(A_i)$

[18] Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994**, *3*, 1261-1266.

[19] Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424-458.

where F_c is the corrected fluorescence, F_o is the fluorescence observed, and C_i is the correction factor for each absorbance value (i). A_i is the new absorbance value for each concentration determined by the extinction coefficient.

iv. Stoichiometry of Binding

The stoichiometry of binding was determined by the molar variation method following the quenching of tryptophan fluorescence. Conditions were such that the peptide and DNA should be fully bound. Peptide concentrations were in the range of 25-50 μM, depending on the maximum DNA concentrations used. The conditions were limited to low DNA concentrations (~60 μM ssDNA) because of the inner filter effect. After correction for the inner filter effect, the fluorescence intensity was plotted against the ratio of DNA/peptide concentrations to give the stoichiometry of binding. The stoichiometry of binding is shown in Figure 4.8 for Mut1 and Figure 4.9 for Mut1-S23 with the X-intercept of the dashed lines indicating the stoichiometry for each.

v. Polyproline Helix Binding

Fluorescence titrations using the polyproline helix were performed using the same procedures as with DNA. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. The experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 340 nm were fit as a function of polyproline concentration to the following equation (Equation 4.4) with Kaleidagraph using non-linear least squares fitting[8b]

Equation 4.4. $F_{Obs} = [F_{free} + (F_{sat} - F_{free})[W_L]/[W_0]$

where F_{free} is the fluorescence intensity without ligand added and F_{sat} is the fluorescence intensity of a saturating concentration of ligand titrated. $[W_L]$ is the fraction of WW domain

bound to ligand and is obtained by Equation 4.5. $[W_0]$ is the total WW domain concentration used.

Equation 4.5. $$[W_L] = \frac{[W_0]+[L_0]+K_d}{2} - \sqrt{\left(\frac{[W_0]+[L_0]+K_d}{2}\right)^2 - [W_0][L_0]}$$

vi. Fluorescence Anisotropy

A fluorophore, 5'-Bodipy 630/650-X NHS Ester, was attached to the ssDNA sequence for all binding studies. For duplex DNA binding, the labeled DNA was annealed to the complementary sequence as above stated. The fluorophore has an absorbance maximum at 638 nm and an emission maximum at 653 nm. Its molar extinction coefficient is 101000 $M^{-1} \cdot cm^{-1}$. Fluorescence anisotropy was measured using a Varian Cary Eclipse Fluorescence Spectrophotometer with a temperature controller. Bodipy-labeled DNA samples (low micromolar concentrations) were titrated with peptide samples (0-low millimolar concentrations) in 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0. Fluorescence samples were analyzed at 298 K and were excited at 638 nm with excitation and emission slit widths of 2.5 nm. Fluorescence was observed at 653 nm, and the anisotropy was determined by the software that came with the instrument. The anisotropy was fit to the following equation (Equation 4.6) using Kaleidagraph to determine the binding constant

Equation 4.6. $$F = \frac{(((-(-K_d - [L] - [P]) - (\sqrt{((-K_d - [L] - [P])^2 - (4 \times [L] \times [P]))})))}{(2 \times [P])) \times (I_\infty - I_o)) + I_o}$$

where F is the fluorescence anisotropy, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added

nucleotide, [P] is the peptide concentration for each fraction, and K_d is the dissociation constant. Equation 4.6 was derived from equations given by Wang and coworkers.[20]

Another fluorophore employed was 5-(and -6)-Carboxytetramethylrhodamine, mixed isomers (TAMRA), which was purchased from Biotium, Inc. TAMRA was coupled onto the peptide Mut1 at Orn 21. The synthesis was completed by coupling Lys(ivDde) in the original ornithine position (Orn21Lys). The ivDde protecting group was orthogonally deprotected by treatment with 2% hydrazine in DMF. Manual coupling of TAMRA was performed with two equivalents of TAMRA (100 mg bottle used for 0.1 mmol peptide) and four equivalents of HOBT, HBTU, and DIPEA in DMF. Cleavage from the resin and sidechain deprotection was completed as with all other peptides. The resulting peptide was purified by HPLC and its sequence and purity determined by mass spectrometry. Peptide concentrations were determined by UV-Vis using TAMRA's extinction coefficient of 91000 $M^{-1} \cdot cm^{-1}$ at wavelength 559 nm. This extinction coefficient was supplied by Integrated DNA Technologies, www.idtdna.com. The excitation wavelength used in the experiments was 559 nm, and the observed emission wavelength was 583 nm. Anisotropy experiments were performed using the same methods as with the Bodipy-labeled DNA experiments.

[20] Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768-779.

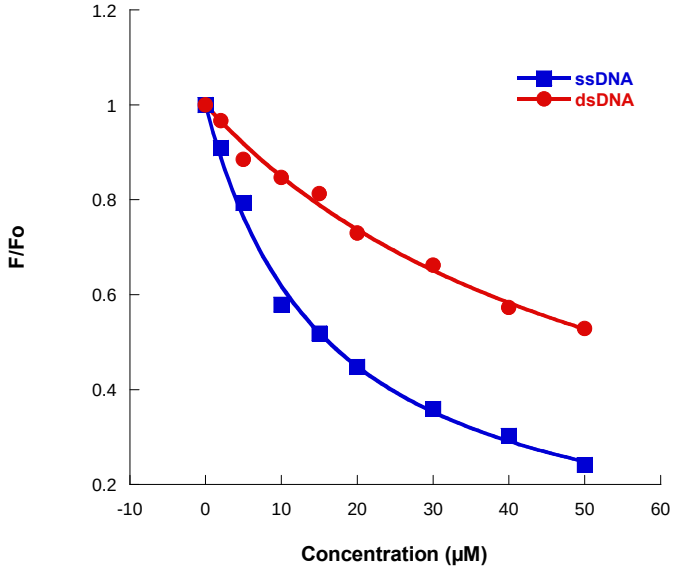


Figure 4.17. Fluorescence titrations of WW domain Mutant 1 with ssDNA (Figure 4.3m) and duplex DNA (Figure 4.3n) by the Trp quenching method, corrected for the inner filter effect (see Experimental Section); 5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

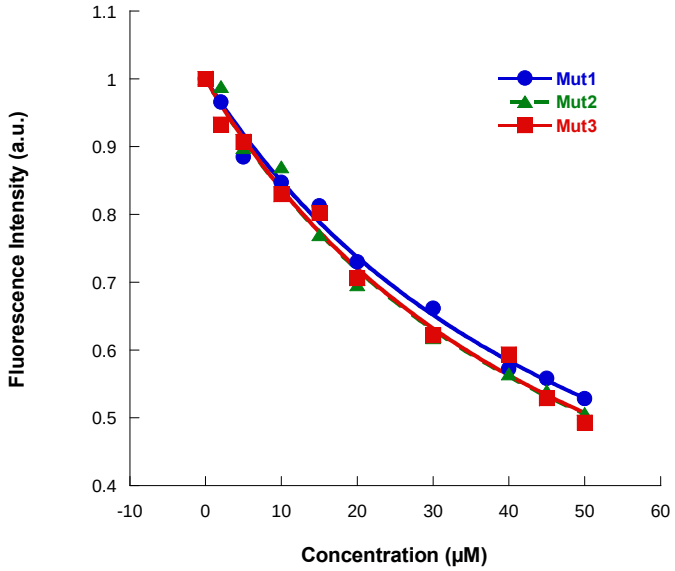


Figure 4.18. Fluorescence titrations of WW domain mutants with duplex DNA sequence (Figure 4.3n) corrected for the inner filter effect (see Experimental Section); 5-8 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

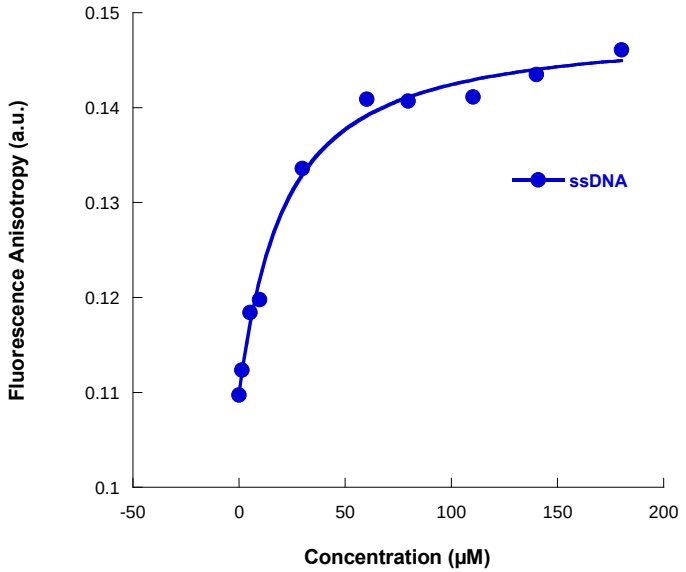


Figure 4.19. Fluorescence anisotropy titrations of TAMRA-WW domain Mut1 and ssDNA sequence 5'-CCATCGCTACC-3' (Figure 4.3m); 2 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

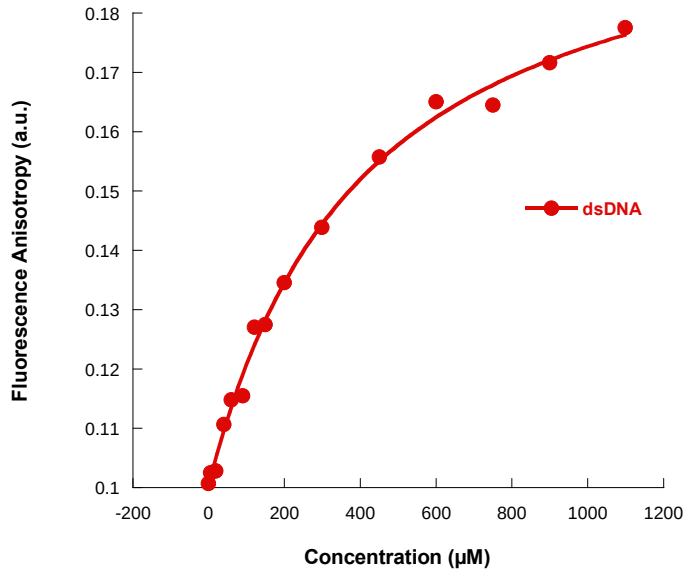


Figure 4.20. Fluorescence anisotropy titrations of TAMRA-WW domain Mut1 and duplex DNA sequence (Figure 4.3n); 20 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

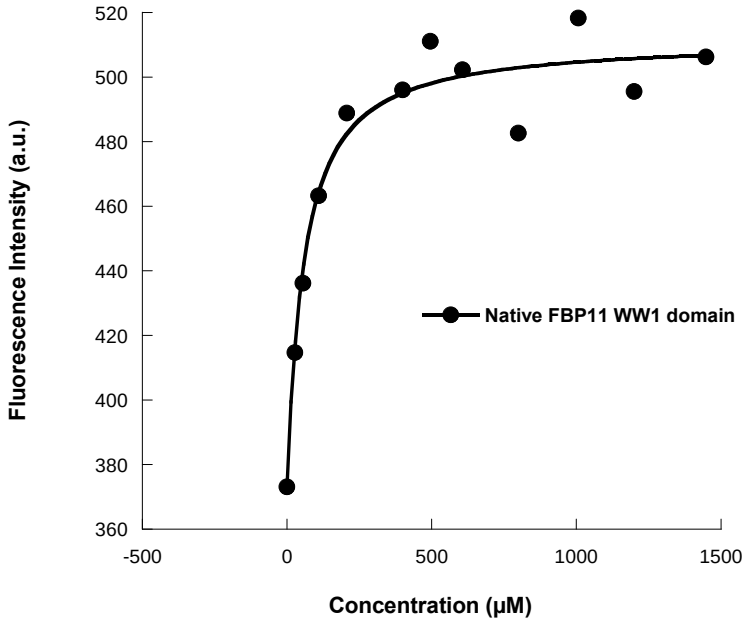


Figure 4.21. Native FBP11 WW1 domain peptide (Figure 4.3a) titrated with polyproline helix (Figure 4.3l); 15 μM WW domain peptide, 0-1450 μM polyproline helix; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

vii. Circular Dichroism

All CD measurements were obtained using an AVIV 62 DS Circular Dichroism Spectrometer. CD data was obtained for the WW domain peptides at 30 μM concentrations from 260-185 nm. The peptides were dissolved in 10 mM Na_2HPO_4, pH 7.0. Wavelength scans were performed at 25 ºC. Thermal denaturations were performed at 227 nm from 4 to 96 ºC, with 4 degree temperature steps. Equilibration times were ten minutes at each step.

viii. NMR Spectroscopy

NMR samples were made in about 1 mM concentrations (650 μL) and were dissolved in 50 mM KD_2PO_4, 0.5 mM DSS, and buffered to pD 7.0 (uncorrected) with sodium deuteroxide. Samples were analyzed on a Varian Inova 600 MHz spectrometer. All one-dimensional NMR spectra were collected using 32K data points and between 16 and 64 scans

and 1-3 s presaturation or solvent suppression. The two-dimensional TOCSY NMR spectra used pulse sequences from the Chempack software. The 2D NMR scans were obtained using 16-64 first dimension scans and 128-512 second dimension scans. All spectra were analyzed by using standard window functions (Sinebell and Gaussian). Peptide proton assignments were made using standard methods as described by Wuthrich.[21] Deviations in alpha hydrogen chemical shifts from random coil values, $\Delta\delta H\alpha$, were calculated according to Equation 4.7,

Equation 4.7. $\Delta\delta H_\alpha = \delta H_{\alpha,obs} - \delta H_{\alpha,RC}$

where $\delta H_{\alpha,obs}$ is the observed chemical shift of a given alpha hydrogen in the peptide, and $\delta H_{\alpha,RC}$ is the random coil chemical shift of the corresponding proton determined from unstructured control peptides, strand 1, strand 2, and strand 3. The extent of folding for each peptide can be determined by calculating the fraction folded, as in equation 4.8,

Equation 4.8. Fraction folded = $\Delta\delta Gly_{obs}/\Delta\delta Gly_{100}$

where $\Delta\delta Gly_{obs}$ is the observed glycine diastereotopic protons splitting for the peptide, and $\Delta\delta Gly_{100}$ is the glycine splitting for the fully folded control peptide that is presumed to take on a 100% fold.

[21] Wuthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley-Interscience: New York, 1986.

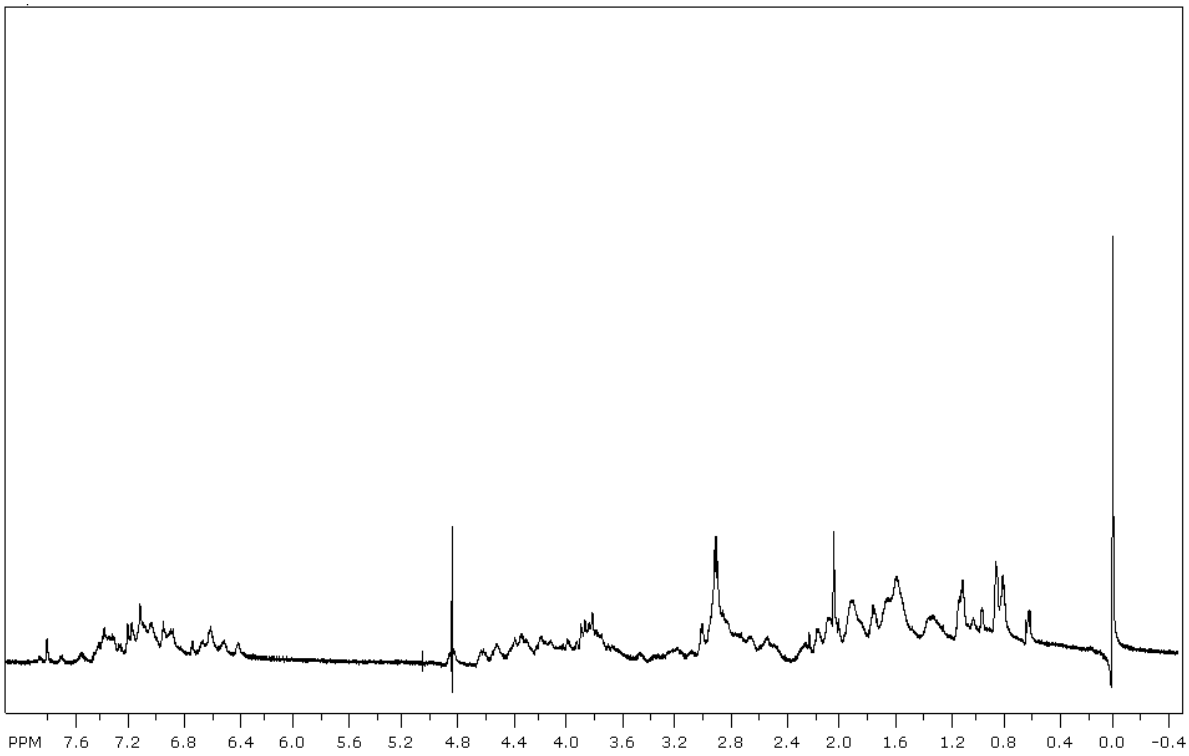


Figure 4.22. ^{1}HNMR spectrum of Mutant 1.

Table 4.4. Alpha proton chemical shifts (in ppm) of WW domain Mutant 1 in 50 mM KD_2PO_4 buffer, pD 7.0.

Residue	α	β	γ	δ	ε
Ser	4.40	3.75			
Arg	4.32	1.61	1.35	2.90	
Trp	4.86	3.30, 3.10			
Thr	4.32	4.07	1.04		
Glu	4.36	1.83	2.16		
His	4.67	3.22			
Lys	4.53	2.13, 1.85	1.12	1.70	2.98
Ser	4.49	3.85			
Asn	4.64	2.83, 2.64			
Gly	0.26				
Arg	4.57	1.87	1.69	3.00	
Thr	4.64	3.97	0.95		
Tyr	4.63	2.47			
Tyr	5.03	2.70			
Trp	4.92	3.42, 3.07			
Asn	5.21	2.68, 2.51			
Lys	4.16	0.61, 0.4	1.33	2.07, 1.78	2.53
Val	4.03	1.86	0.85		
Asn	4.53	2.98, 2.75			
Gly	0.68				
Orn	4.73	1.67, 1.63	1.35	2.90	
Trp	5.10	3.36, 3.18			
Gln	4.84	2.12, 1.91	2.27		
Lys	4.06	1.83, 0.59	1.27	1.57	2.85
Thr	4.43	4.10	1.09		
Trp	5.00	3.16, 2.95			
Glu	4.22	1.88, 1.76	2.09		
Lys	4.34	0.40, 0.16			2.50
Pro	4.41	2.26, 2.01	1.96, 1.83	3.82, 3.66	
Gly	3.89				

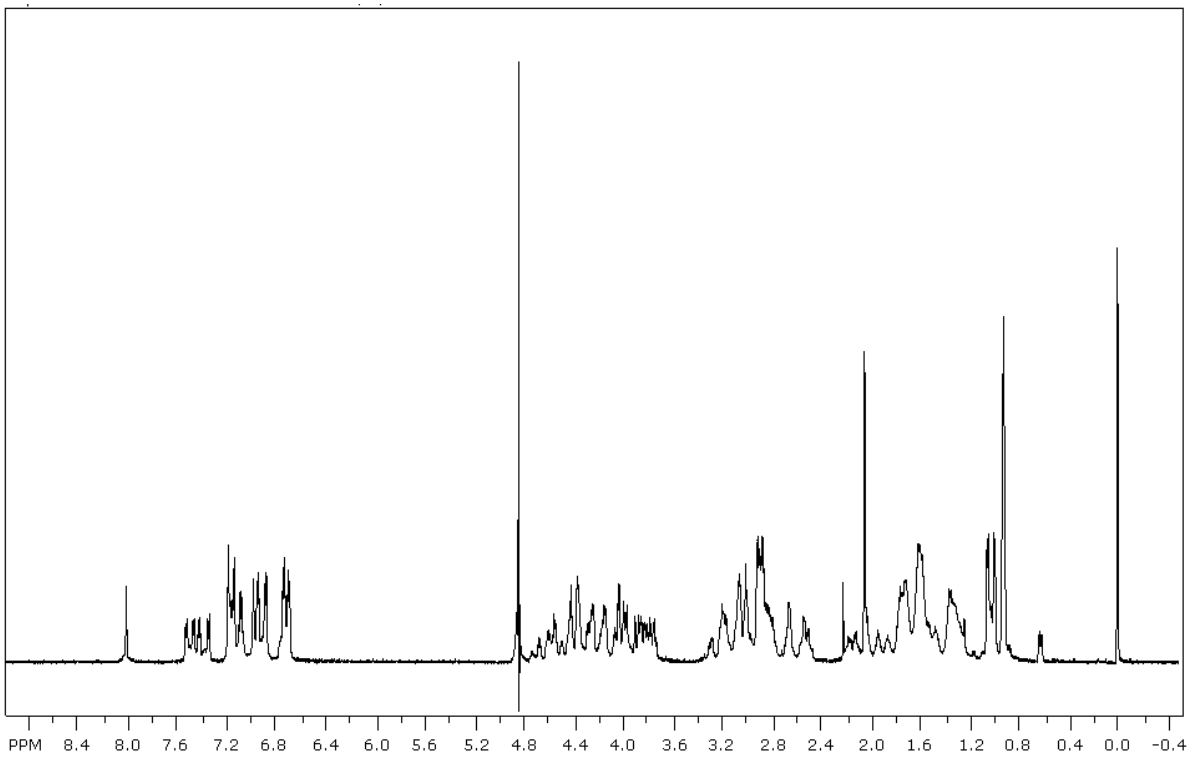


Figure 4.23. ^{1}HNMR spectrum of Mutant 1 Strands 1&2.

Table 4.5. Alpha proton chemical shift assignments (in ppm) for Mutant 1 Strands 1&2, Ac-Ser-Arg-Trp-Thr-Glu-His-Lys-Ser-Asn-Gly-Arg-Thr-Tyr-Tyr-Trp-Asn-Lys-Val-NH_2

Residue	α	β	γ	δ	ε
Ser	4.37	3.76			
Arg	4.30	1.57	1.37	3.00	
Trp	4.60	3.18			
Thr	4.21	4.00	1.03		
Glu	4.19	1.89	2.14		
His	4.63	3.02			
Lys	4.39	1.72	1.33	1.63	2.90
Ser	4.43	3.84			
Asn	4.51	2.84			
Gly	3.91				
Arg	4.37	1.70, 1.55	1.33	3.06	
Thr	4.26	4.01	1.00		
Tyr	4.47	2.66			
Tyr	4.57	2.51			
Trp	4.76	3.21			
Asn	4.69	2.84			
Lys	4.16	2.13	1.29	1.59	2.89
Val	4.05	2.02	0.92		

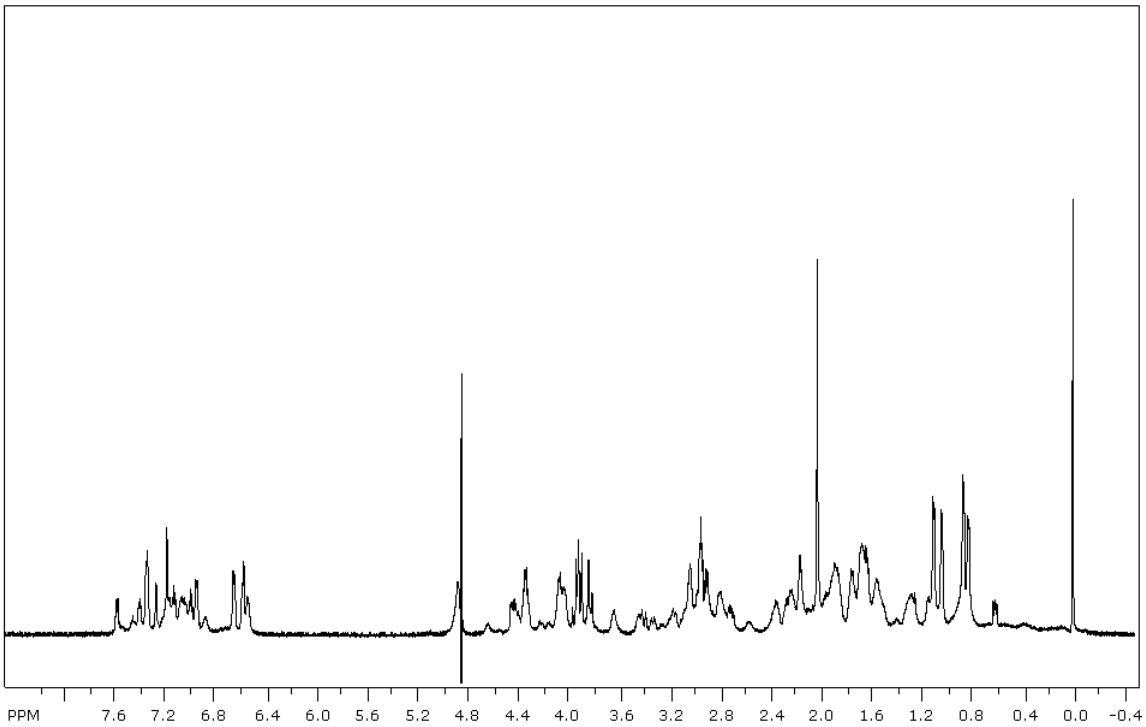


Figure 4.24. ^{1}HNMR spectrum of Mutant 1 Strands 2&3.

Table 4.6. Alpha proton chemical shift assignments (in ppm) for Mutant 1 Strands 2&3, Ac-Gly-Arg-Thr-Tyr-Tyr-Trp-Asn-Lys-Val-Asn-Gly-Orn-Trp-Gln-Lys-Thr-Trp-Glu-Lys-Pro-Gly-NH_2

Residue	α	β	γ	δ	ε
Gly	3.91				
Arg	4.43	2.21	1.92	3.01	
Thr	4.40	4.12	1.08		
Tyr	4.72	3.08			
Tyr	4.80	2.85			
Trp	4.97	3.38			
Asn	5.18	2.82, 2.61			
Lys	4.45	1.97	1.43	1.68	2.99
Val	4.10	1.91	0.90		
Asn	4.50	2.96, 2.76			
Gly	4.12				
Orn	4.71	1.93	1.75	3.08	
Trp	4.95	3.14			
Gln	4.95	2.19, 2.00	2.35		
Lys	4.43	1.98	1.70	1.89	2.76
Thr	4.52	3.90	1.16		
Trp	4.98	3.14			
Glu	4.28	1.89, 1.84	2.15		
Lys	4.10	1.68	1.33	1.70	3.10
Pro	4.39	2.27, 2.07	1.93	3.69, 3.50	
Gly	3.87				

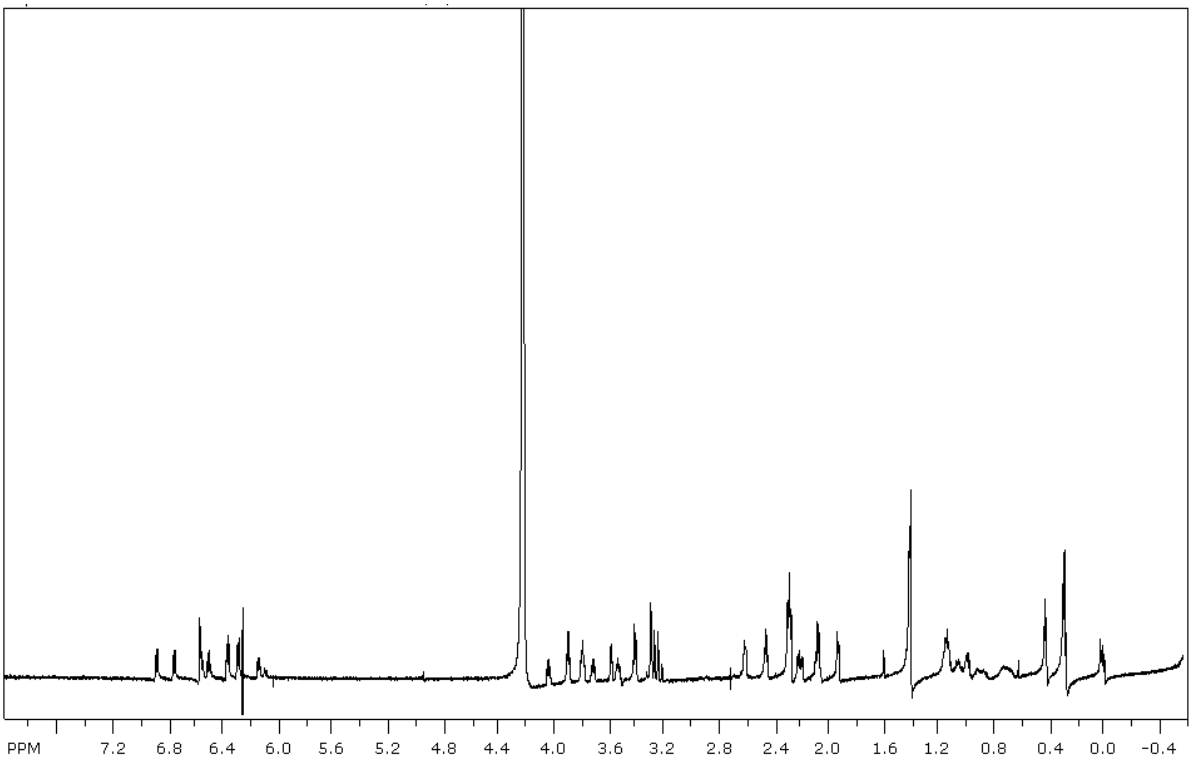


Figure 4.25. ^{1}HNMR spectrum of Mutant 1 Strand 1.

Table 4.7. Alpha proton chemical shift assignments (in ppm) for random coil peptide Mutant 1 Strand 1, Ac-Ser-Arg-Trp-Thr-Glu-His-Lys-Ser-Asn-Gly-NH_2

Residue	**α**
Ser	4.38
Arg	4.28
Trp	4.72
Thr	4.19
Glu	4.12
His	4.62
Lys	4.32
Ser	4.42
Asn	4.72
Gly	3.89

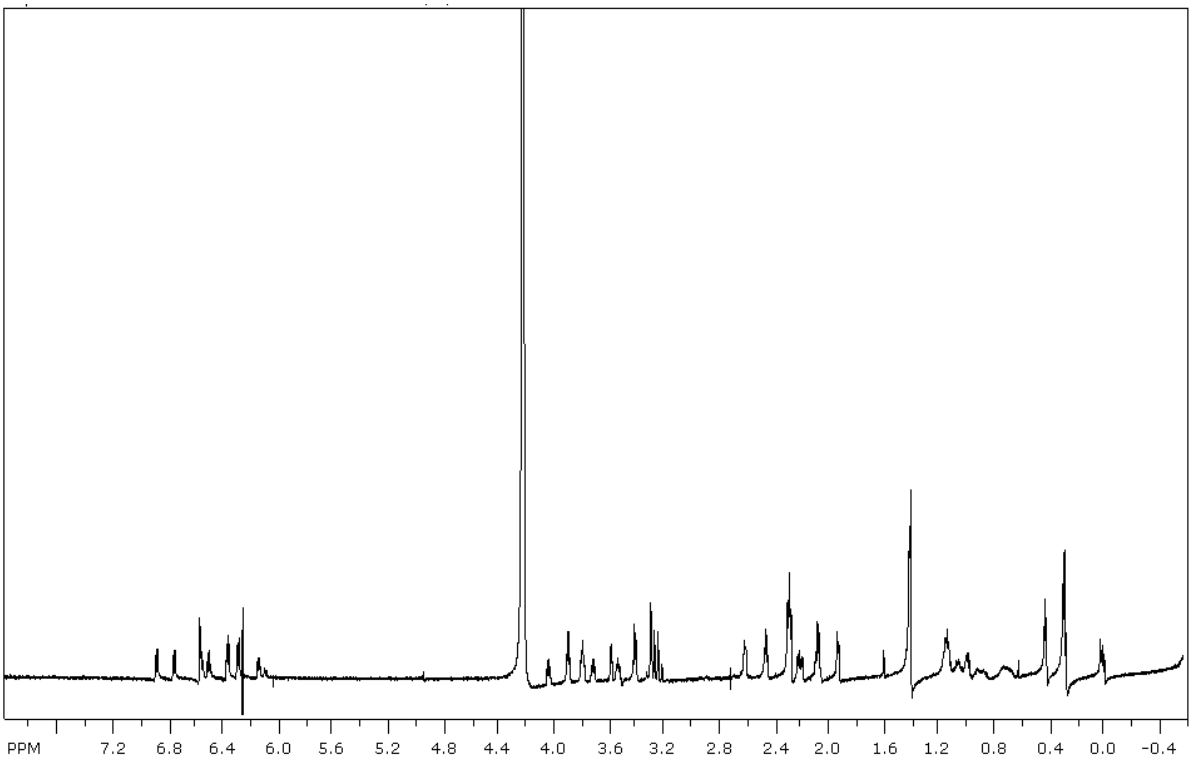


Figure 4.26. ^{1}HNMR spectrum of Mutant 1 Strand 2.

Table 4.8. Alpha proton chemical shift assignments (in ppm) for random coil peptide Mutant 1 Strand 2, Ac-Gly-Arg-Thr-Tyr-Tyr-Trp-Asn-Lys-Val-Asn-Gly-NH_2[a]

Residue	**α**
Gly	3.85
Arg	4.33
Thr	4.20
Tyr	4.40
Tyr	4.51
Trp	4.51
Asn	4.66
Lys	4.15
Val	4.02
Asn	4.40
Gly	3.91

[a]Tyrosine residues could not be definitively distinguished.

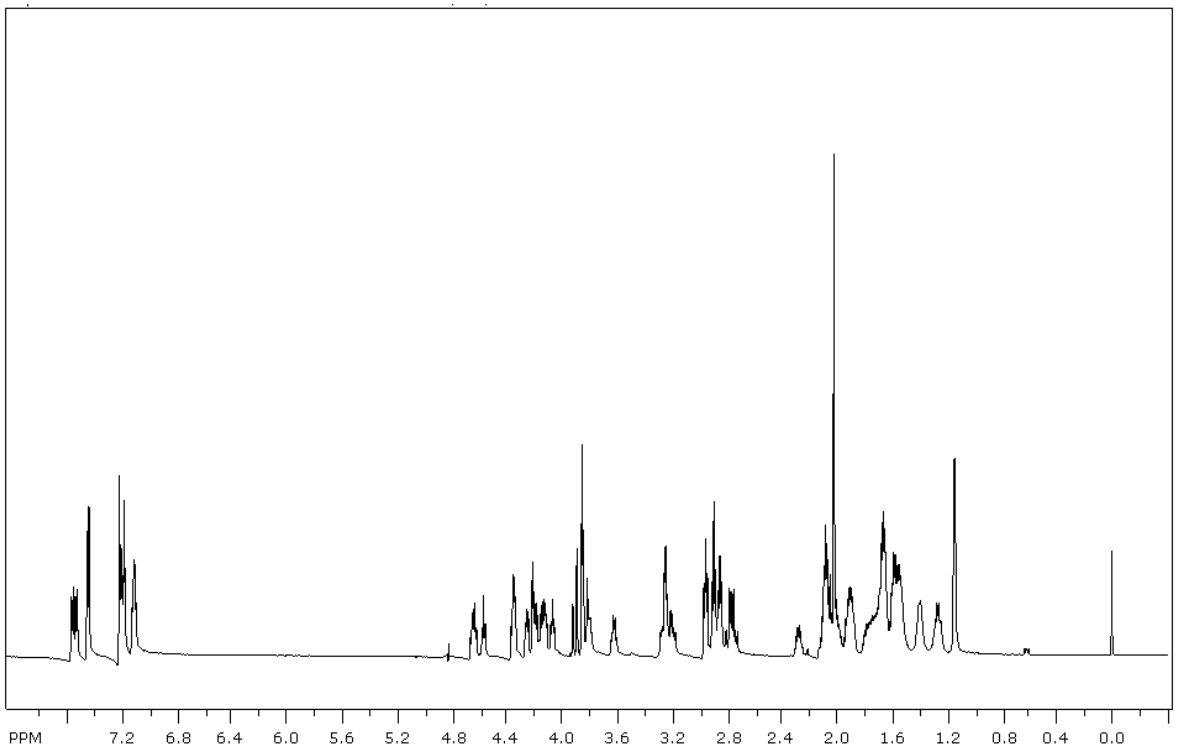


Figure 4.27. ^{1}HNMR spectrum of Mutant 1 Strand 3.

Table 4.9. Alpha proton chemical shift assignments (in ppm) for random coil peptide Mutant 1 Strand 3, Ac-Asn-Gly-Orn-Trp-Gln-Lys-Thr-Trp-Glu-Lys-Pro-Gly-NH_2

Residue	**α**
Asn	4.65
Gly	3.91
Orn	4.37
Trp	4.65
Gln	4.20
Lys	4.27
Thr	4.20
Trp	4.60
Glu	4.14
Lys	4.08
Pro	4.37
Gly	3.86

Chapter V

Designing a WW Domain Peptide Library to Optimize the Binding Face of the OB-Fold Mimic Mut1

A. Background and Significance

In previous studies, the FBP11 WW1 domain peptide (Figure 5.1a) was mutated to form a putative ssDNA binding pocket based on the WKWK peptide (Mut1, Figure 5.1b). The peptide was designed such that the WKWK binding pocket was positioned within strands two and three with the first strand intended to add favorable contacts for binding. This peptide was the latest design to mimic the OB-fold domain. The peptide displayed tight binding and good selectivity for ssDNA. The binding, however, was weaker than with the $(WKWK)_2$, with binding constants of about 20 μM for Mut1 and 3 μM for $(WKWK)_2$.[1a, b] The difference in binding affinities for the two peptides may be due to the fact that the peptide dimer has two DNA binding pockets while Mut1 has only one. Thus, we endeavored to gain affinity and selectivity by mutating residues on the third strand that might add favorable contacts to the DNA. Based on the fluorescence and CD data for Mutant 1, a library was designed in an attempt to find a WW domain mutant which is a high affinity ssDNA binder and which demonstrates high selectivity for ssDNA over duplex DNA and over other peptides in the library.

---

[1] (a) Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544. (b) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25. (c) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70*, 1105-1114. (d) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.

B. Results

i. Library Design

The library was based on Mut1 and contained two varied positions. Because of electrostatic interactions overwhelming previous library screens (see Chapter 3), all charged amino acids were left out of the library to eliminate binding simply due to non-specific electrostatic interactions. The library was designed to determine specific hydrogen-bonding interactions that give enhanced selectivity. Unlike previous libraries designed with WKWK, this library keeps the binding pocket constant and utilizes the third strand to incorporate additional favorable contacts to increase binding affinity and selectivity. These interactions were not possible to this extent in the WKWK peptides because there was no third strand to provide additional interactions to the binding pocket. We wanted to maintain the putative binding pocket because much evidence suggests that a DNA base stacks with the Trp residues and the Lys residues form electrostatic interactions with the phosphate groups.[1] Previous library studies also indicate that the WKWK peptide binding pocket is optimized for ssDNA binding. Because the goal is to mimic the OB-fold domain, it is also important to maintain the basic and aromatic residues known to be crucial for ssDNA binding,[2] as was intended in the original peptide design and in the design of Mut1.

[2] (a) Kloks, C. P.; Spronk, C. A.; Lasonder, E.; Hoffmann, A.; Vuister, G. W.; Grzesiek, S.; Hilbers, C. W. *J. Mol. Biol.* **2002**, *316*, 317-326. (b) Schroder, K.; Graumann, P.; Schnuchel, A.; Holak, T. A.; Marahiel, M. A. *Mol. Microbiol.* **1995**, *16*, 699-708. (c) Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

(a) Native FBP11 WW1 domain:
Ac-S-E-W-T-E-H-K-S-P-D-G-R-T-Y-Y-Y-N-T-E-T-K-Q-S-T-W-E-K-P-**G**-NH_2

(b) Mut1:
Ac-S-**R**-W-**T**-E-H-K-S-N-G-R-T-Y-Y-**W**-N-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(c) S-X_2-W-X_4-E-H-K-S-N
N-V-**K**-N-**W**-Y-Y-T-R-G
G-O-**W**-Q-**K**-T-W-E-K-P-G-βAla-βAla

(d)

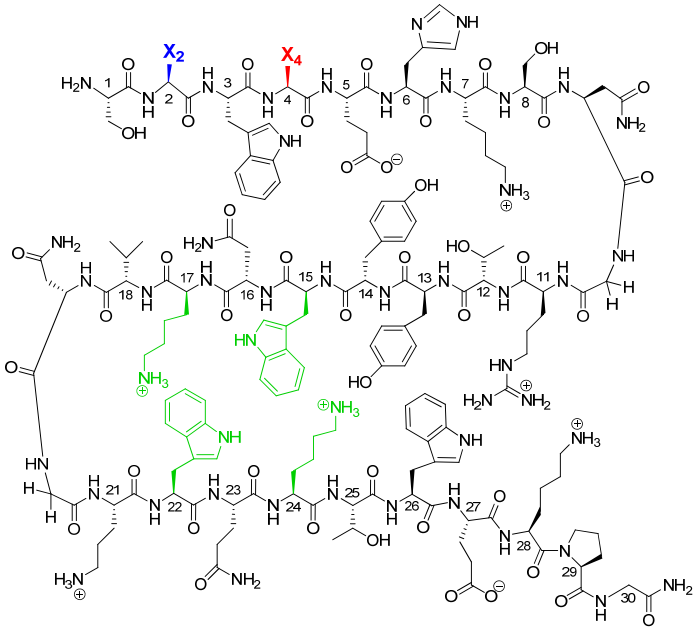


Figure 5.1. (a) Sequence of native FBP11 WW1 domain. (b) Sequence of Mutant 1 with position 2 (Arg 2) in blue and position 4 (Thr 4) in red. (c) Illustration of WW domain library. The βAla-βAla linker was added as a spacer between the bead and the peptide. (d) Structure of Mutant 1 with varied position X_2 shown in blue and position X_4 in red. The DNA binding pocket cross-strand from these positions is shown in green.

Mut1 has the potential to make more contacts near the binding pocket which may increase binding affinity and improve selectivity, and we can take advantage of this by varying residues across from the Trp pocket on the binding face to attain the peptide with the highest ssDNA affinity and selectivity. As noted previously, the analysis of protein

structures has shown that hydrogen bonding is important for selective protein recognition of nucleotides.[3] Figure 5.2 shows common examples of favorable interactions between specific bases and different amino acid sidechains. Amino acids known to form favorable hydrogen bonds with DNA bases were included in the library. The positions varied in the peptide were residues 2 and 4, which are Arg 2 and Thr 4 in Mut1 (Table 5.1; Figure 5.1). As in previous studies using Mutant 1, the numbering for the mutant peptide is from 1-30, corresponding to the native FBP11 WW1 domain peptide residues 15-42. The library was synthesized using solid phase peptide synthesis as before on TentaGel Macrobeads. Split-and-pool synthesis was completed to give a one-bead-one-compound library,[4] as shown in Figure 5.3.[5]

[3] (a) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860-2874. (b) Moodie, S. L.; Mitchell, J. B. O.; Thornton, J. M. *J. Mol. Biol.* **1996**, *263*, 486-500. (c) Cheng, A. C.; Chen, W. W.; Fuhrmann, C. N.; Frankel, A. D. *J. Mol. Biol.* **2003**, *327*, 781-796.

[4] (a) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84. (b) Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411-448. (c) Lam, K. S.; Lehman, A. L.; Song, A.; Doan, N.; Enstrom, A. M.; Maxwell, J.; Liu, R. *Methods Enzymol.* **2003**, *369*, 298-322. (d) Morken, J. P.; Kapoor, T. M.; Feng, S.; Shirai, F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1998**, *120*, 30-36. (e) Combs, A. P.; Kapoor, T. A.; Feng, S.; Chen, J. K.; Daude-Snow, L. F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1996**, *118*, 287-288.

[5] Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233-1251.

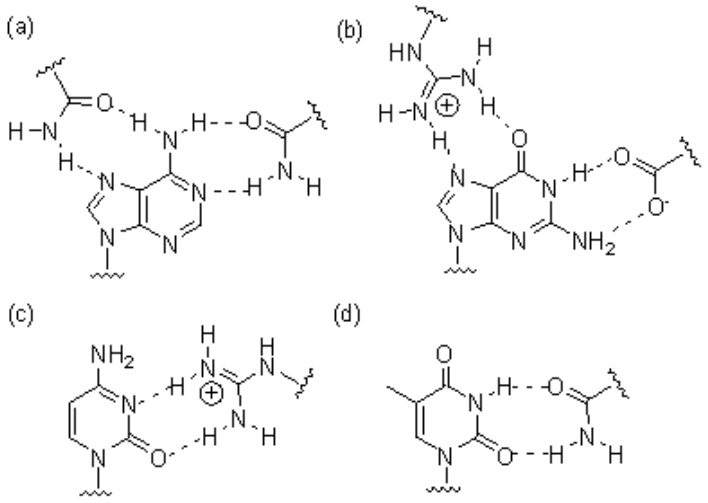


Figure 5.2. Hydrogen-bonding patterns between amino acid sidechains and (a) adenine, (b) guanine, (c) cytosine, and (d) thymine. Adapted from Cheng, et al.[3c]

Varied residues included hydrogen-bonding amino acids as well as aromatic amino acids. The hydrogen-bonding residues were thought to add selectivity and improve binding affinity by increasing direct favorable contacts to the DNA bases. Aromatic residues were included because they could form additional stacking interactions with the bases to improve the binding affinity. Both positions were varied with the same eight residues (Table 5.1) to generate a 64-member on-bead peptide library as depicted in Figure 5.1c.

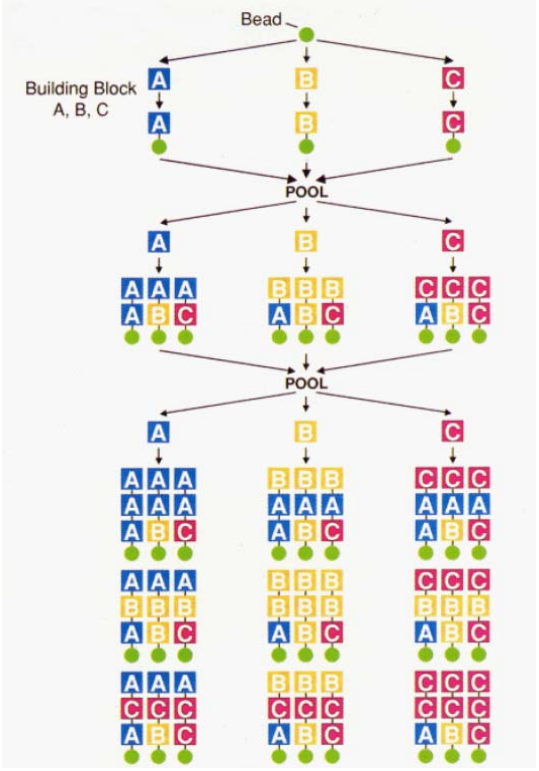


Figure 5.3. Split-and-pool synthesis method for producing a one-bead-one-compound combinatorial library using solid phase peptide synthesis. Figure adapted from Gallop, et al.[5]

Table 5.1. Amino acids incorporated into the WW domain peptide library.

Varied Residue	Amino Acids
Arginine (2)	Q, N, S, H, Y, W, F, T
Threonine (4)	Q, N, S, H, Y, W, F, T

ii. Library Screen and Identification of Hits

This peptide library was screened against the single-stranded nucleotide sequence 5'-d(AAAAA)-3' (penta-dA) as in previous library screens. A colorimetric assay reported by the Schreiber group involving biotin and streptavidin was used to screen the libraries for hits as in Chapter 3 (See Experimental Section).[4d-4e] Hits were identified from the assay and selected for sequence analysis by Edman degradation.

Several library screens were performed before optimal conditions were reached. Once those conditions were determined (2 nM Penta-dA; 0.5 nM SAAP; 500 mM NaCl in 10 mM Na_2HPO_4, pH 7.5), a library screen was performed and about 45 hits were obtained. Of these, ten hits were analyzed by Edman degradation and amino acid sequencing. The results of the screen are reported in Table 5.2 and gave a consensus sequence of 'S-F-W-Q' for the first four residues. Position two resulted in primarily aromatic residues with five phenylalanines and a Tyr as well as three His residues and an asparagine. The fourth position gave a variety of different amino acids but included mostly glutamines and asparagines and a few aromatic amino acids. This is consistent with the report by Luscombe and coworkers which showed that Asn and Gln residues are favored to hydrogen bond with adenine bases as seen in Figure 5.2a.[3] With charged amino acids excluded from the library, the screens resulted in presumably the most favorable amino acids hydrogen-bonding with the adenine bases of the penta-adenosine. Phenylalanine was selected most for position 2 while Gln was most favored at position 4, so these two amino acids were included in the peptide to give the final consensus sequence, Mutant 4. The numbering of the mutants continues from the previous WW domain mutant research (Chapter 4). Mutant 4 (Mut4; Figure 5.4) was synthesized, and binding studies were performed as with previous fluorescence quenching experiments (see Experimental Section).

Table 5.2. Sequences of hits from WW domain library screen with varied positions shown in blue and red for positions 2 and 4, respectively.

Number of Hits	Sequence of Hits
2 Beads	Ser-**Phe**-Trp-**Gln**-Glu
2 Beads	Ser-**Phe**-Trp-**Asn**-Glu
1 Bead	Ser-**Phe**-Trp-**Thr**-Glu
1 Bead	Ser-**His**-Trp-**His**-Glu
1 Bead	Ser-**His**-Trp-**Trp**-Glu
1 Bead	Ser-**Tyr**-Trp-**Phe**-Glu
1 Bead	Ser-**Asn**-Trp-**Phe**-Glu
1 Bead	Ser-**His**-Trp-**Gln**-Glu

iii. Fluorescence Binding Studies of Consensus Sequence and Mutants

The interaction between Mutant 4 and ssDNA was tested by fluorescence quenching, and the binding affinity was determined to be 34 μM (Table 5.3; Figure 5.5). This is about a two-fold lower affinity than that of Mut1, which has a K_d of 16.6 μM as determined by fluorescence quenching. The mutations R2F and T4Q weakened the binding moderately, indicating that the loss of a positive charge is likely decreasing the affinity by a lack of some electrostatic interactions between the arginine's guanidinium group and the phosphate group of the DNA backbone. Another possibility for a decrease in binding affinity is that the Arg could bind favorably to a DNA base and that specific interaction may be lost with phenylalanine in that position. The difference in Thr and Gln may have some effect, but that is not thought to be significant.

(a) WW domain Mut4:
Ac-S-**F**-W-**Q**-E-H-K-S-N-G-R-T-Y-Y-W-N-K-V-N-G-O-W-Q-K-T-W-E-K-P-G-NH_2

(b) WW domain Mut5:
Ac-S-**G**-W-**G**-E-H-K-S-N-G-R-T-Y-Y-W-N-K-V-N-G-O-W-Q-K-T-W-E-K-P-G-NH_2

(c) ssDNA: 5'-CCATCGCTACC-3'

Figure 5.4. (a) Sequence of WW domain Mutant 4. (b) WW domain Mutant 5 sequence. (c) ssDNA sequence used in binding studies. Peptide mutations are shown in bold.

Because the varied positions 2 and 4 are directly across from Lys 17 and Trp 15 of the binding pocket, these two positions are most likely to contribute to binding if the peptide is structured. Results of the binding studies with Mut4 led to more structural and binding studies to further investigate the stability and the binding of Mut1 and subsequent mutants. As a negative control, Mutant 5 (Mut5) was synthesized, replacing the amino acids at positions 2 and 4 with glycines (Figure 5.4b). If the peptide is unstructured or is in some conformation in which the first strand is not in close proximity to the second strand, the two peptides should give similar binding constants. The R2F mutation also allows an easier comparison between Mutants 4 and 5 rather than comparing Mutants 1 and 5 because the former mutants have the same net charge. Mutant 1 could have greater affinity for ssDNA based on its charge alone.

Table 5.3. Dissociation constants for the binding interaction between WW domain mutants and ssDNA.[a]

Peptide	K_d, μM (error)
Mutant 1	16.6 (1.5)
Mutant 4	34.0 (2.2)
Mutant 5	38.8 (4.3)

(a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting.

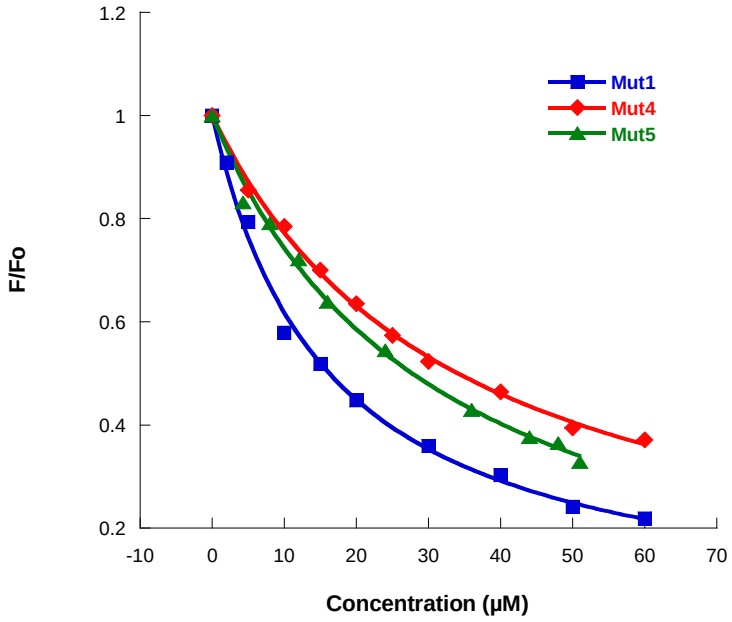


Figure 5.5. Fluorescence titrations of WW domain mutants 1, 4, and 5 with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 5.4c) corrected for the inner filter effect (see Experimental Section); 5-8 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

The negative control peptide Mut5 with R2G and T4G surprisingly binds with similar affinity to Mut4. This indicates that the particular residues in these positions are not very important. This result suggests that the overall net charge is particularly important but that the residues in strand 1 do not contribute significantly to ssDNA binding. These results suggest that strand 1 may not be interacting with the DNA. If strand 1 is structured, it is more likely to contribute to binding due to hydrogen-bonding and other possible interactions. If strand 1 is not structured, it is possibly not close enough to the binding pocket to contribute to the binding interaction.

iv. CD Characterization of Folding

WW Domain Mutant 1 was characterized by CD, thermal denaturation, as well as one- and two-dimensional NMR data (See Chapter 4). It seems that the peptide structure comes from strands 2&3, while strand 1 adds little to the stability. This could be the reason that Mutant 4 exhibits a lower affinity for ssDNA than Mutant 1. Strand 1 of Mut1 may improve binding largely due to electrostatic interactions. If stand 1 is not close in space to the second strand, strand 1 may not be able to contribute to binding. This would explain why Mutants 4 and 5 have similar binding affinities for ssDNA and why glycines can be added in seemingly important or proximal positions and have no effect on binding affinity. Structural studies of Mut4 and Mut5 were conducted by CD and compared to Mut1 to gain further information about the stability of each mutant. All three mutants have similar CD spectra, each exhibiting a minimum around 210 nm which is characteristic of β-sheet peptides and proteins (Figure 5.6). Typically β-sheets are displayed at 215 nm, and this difference is attributed to the contributions of the tryptophan residues. This has been seen in previous WW domain mutants designed by this laboratory. Mut1 is shifted slightly more than Mut4 and Mut5, and this could be due to some change in orientation of the Trp residues based on the mutations made. Each peptide also has a similar exciton coupling peak near 227 nm due to the aromatic amino acids present in each sequence. Consistent with previous CD and NMR structural studies with this class of WW domain mutants, these peptides also have some unstructured region as each displays a minimum around 195 nm, consistent with random coil peptides. Along with previous structural evidence for Mut1, this data further indicates that a portion of the peptide is very stable and well folded while part of the peptide is unstructured. All evidence indicates that the unstructured region is Strand 1.

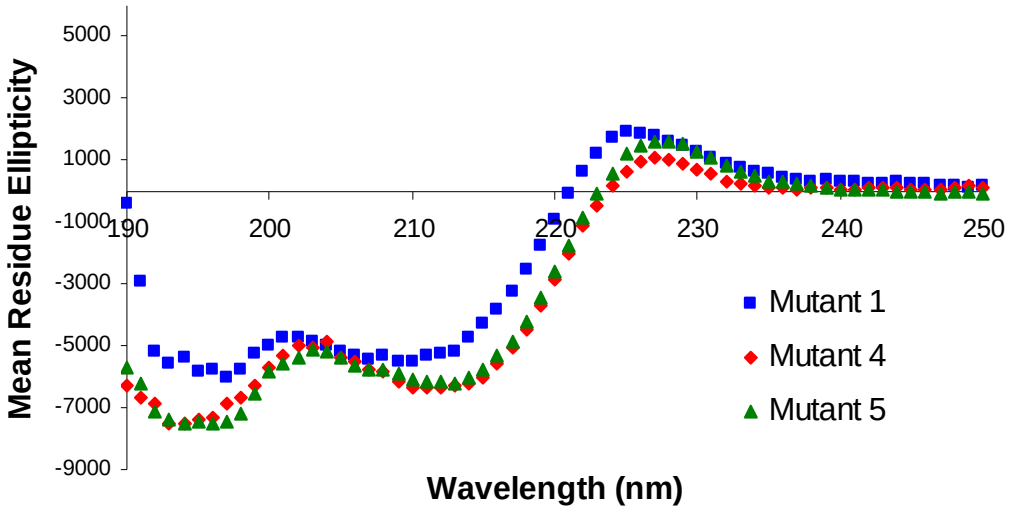


Figure 5.6. CD spectra for WW domain mutants 1, 4, and 5. Data was recorded at 30 μM peptide, 10 mM Na_2HPO_4, 298 K, pH 7.0, with a wavelength range from 250-190 nm.

C. Discussion and Conclusions

The designed peptide was based on the idea that the first strand would provide additional contacts to improve affinity and selectivity for ssDNA. The two residues cross-strand from the binding pocket in Mutant 1, Arg 2 and Thr 4, were thought to add affinity and selectivity by potential hydrogen-bonding interactions not available in the WKWK monomer and dimer peptides. Initial results supported that hypothesis in part because the peptide provided selectivity for ssDNA versus duplex DNA. Further studies found, however, that residues in strand 1 do not contribute to binding as much as expected. When glycine residues replaced residues 2 and 4 of strand one, which are directly across from the binding pocket in strands two and three, there was little change in binding affinity as compared to Mut4, which has a Phe at position 2 and a Gln at position 4. Of importance is that Mutants 4 and 5 have the same net charge, so charge is not a factor here. If strands 1&2 form a structured β-hairpin within Mutant 1, then residues at those two positions should provide

favorable contacts that influence binding. The data, however, suggests that this is not the case. To further study the structural features of Mut1 and to increase the stability of strand 1 within the full peptide, additional structural studies and mutations will be conducted.

D. Experimental Section

i. Library design and screening

The WW domain library was synthesized as in the library synthesis in Chapter 3 and was screened against the single-stranded nucleotide sequence 5'-d(AAAAA)-3' (penta-dA). A screen reported by the Schreiber group which takes advantage of the strong binding interaction between biotin and streptavidin was used to screen the libraries for hits.[4d-4e] The pentanucleotide was purchased from Integrated DNA Technologies, Inc., with the 5' end biotinylated, and streptavidin was purchased from Pierce Biotechnology, Inc., in conjugate with alkaline phosphatase (SAAP). The alkaline phosphatase is essential for the dye that is utilized in the screen. All screens were performed by incubating the biotinylated penta-dA with the streptavidin/alkaline phosphatase complex in high salt (500 mM NaCl) phosphate buffer, pH 7.4. The resin containing the peptide library (~75 mg) was also incubated in high salt PBS. Once the biotin was complexed with the streptavidin, the Biotin-penta-dA/SAAP was allowed to bind to the peptides on the resin. After reaction, the resin was washed with high salt screening buffer four times and then washed with a dye for 10 minutes. The dye is a mixture of Nitro Blue Tetrazolium chloride/5-Bromo-4-Chloro-3'-Indolyl Phosphate p-toluidine salt (NBT/BCIP). When in the presence of alkaline phosphatase, the phosphate group of the BCIP is hydrolyzed and then a tautomerization occurs, followed by the dimerization of the compound by NBT. The resulting compound produces an insoluble indigo dye that colors the beads which are bound to the biotinylated DNA. After quenching

the reaction with EDTA, the beads were washed with phosphate buffer and then plated in screening buffer. The beads were various colors of purple, and the darkest beads were picked as hits and sequenced by Edman degradation. Control experiments were performed to show that the screen was selecting true hits due to binding.

ii. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using an Applied Biosystems Pioneer Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 10 minutes. Each amino acid was coupled on an extended cycle of 75 minutes to improve coupling. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin was performed in 95% trifluoroacetic acid (TFA), 2.5% H_2O, and 2.5% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen, and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 5-35% solvent B over 25 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and the peptide sequence was confirmed by MALDI mass spectrometry. Peptide dimers were formed by oxidation of the cysteine residues with

stirring in 1% DMSO in 10 mM phosphate buffer, pH 7.5 for 7-12 hrs. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

iii. DNA Sample Preparation

DNA sequences were purchased from IDT (Integrated DNA Technologies). All DNA samples were dissolved in 10mM Na_2HPO_4, 100 mM NaCl, adjusted to pH 7.0. Concentrations of DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficient of the ssDNA ($\varepsilon_{260,\ ssDNA} = 95500\ M^{-1}\cdot cm^{-1}$).

iv. Fluorescence Titrations

To determine the recognition of single-stranded oligonucleotides by the peptide dimers, fluorescence titrations were performed which followed the Trp quenching with increasing oligonucleotide concentration. Peptide and nucleotide samples were prepared in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0. Peptide concentrations were determined in 5 M guanidine hydrochloride by recording the absorbance of the Trp residues at 280 nm ($\varepsilon = 5690\ M^{-1}cm^{-1}$) by UV/vis spectroscopy. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. Experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 348 nm were fit as a function of nucleotide concentration to the binding equation (Equation 5.1) with Kaleidagraph using non-linear least squares fitting.[6]

Equation 5.1. $I = [I_o + I_\infty([L]/K_d)]/[1 + ([L]/K_d)]$

[6] Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994**, *3*, 1261-1266.

where I is the observed fluorescence intensity, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, and K_d is the dissociation constant.

Oligonucleotides have an observable absorbance at the excitation wavelength of Trp (297 nm), and therefore there is an inner filter effect for which one must take account. The absorbance of the oligonucleotides at 297 nm was monitored at known concentrations and the extinction coefficient was determined. Absorbance values were determined for each oligonucleotide concentration. Corrected fluorescence values were determined from the following equations (Equation 5.2 and Equation 5.3).[7]

Equation 5.2. $F_c = F_o/C_i$

Equation 5.3. $C_i = (1\text{-}10^{-Ai})/(2.303)(A_i)$

where F_c is the corrected fluorescence, F_o is the fluorescence observed, and C_i is the correction factor for each absorbance value (i). A_i is the absorbance value for each concentration determined by the extinction coefficient.

Binding curves were shown comparing Mut1, Mut4, and Mut5. The fits of the corrected, un-normalized data are included in this section.

[7] Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424-458.

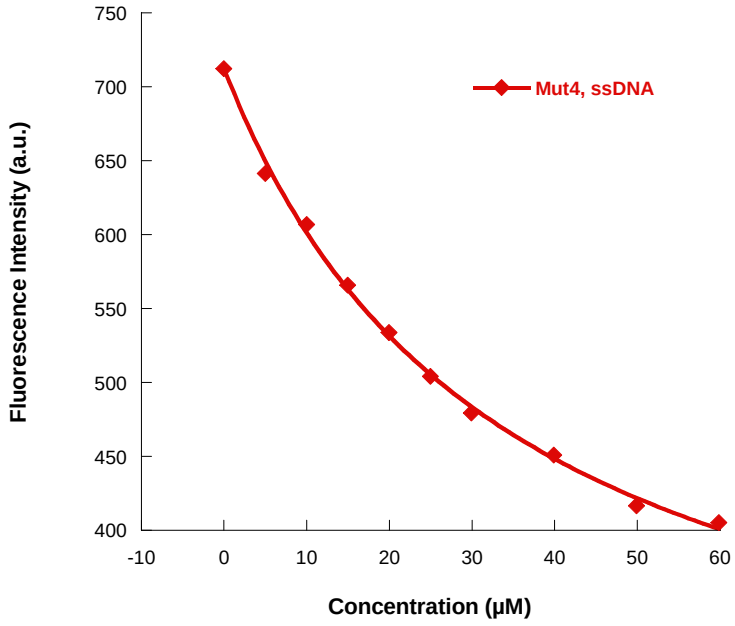


Figure 5.7. Fluorescence titration of WW domain Mut4 with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 5.4c) corrected for the inner filter effect (see Experimental Section); 8 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

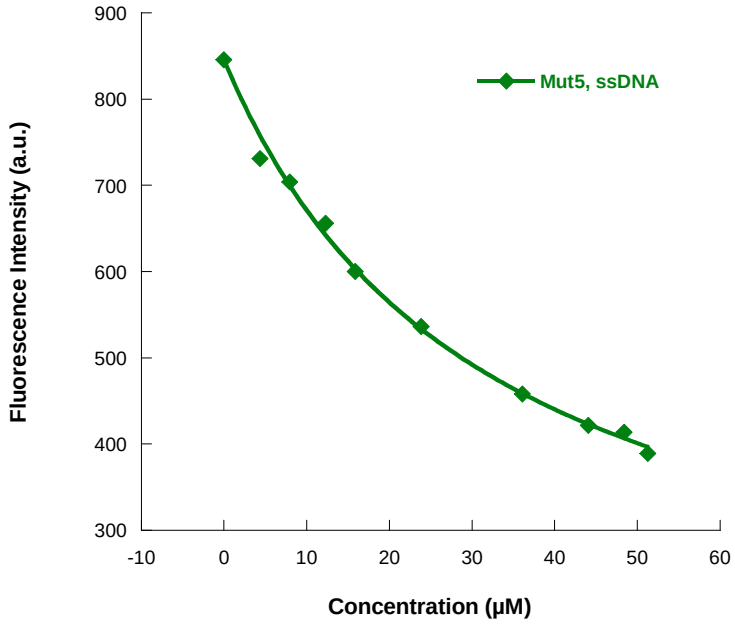


Figure 5.8. Fluorescence titration of WW domain Mut5 with single-stranded DNA sequence 5'-CCATCGCTACC-3' (Figure 5.4c) corrected for the inner filter effect (see Experimental Section); 5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

v. Circular Dichroism

The WW domain peptides were analyzed by CD to verify their structure. CD measurements were performed on an AVIV 62 DS Circular Dichroism Spectrometer. CD data was obtained for the peptides at 30 μM concentrations. The peptides were dissolved in 10 mM Na_2HPO_4, pH 7.0. Wavelength scans were performed at 298 K from 260-185 nm.

Chapter VI

Stabilization of the WW Domain Mutant 1 for Single-Stranded DNA Binding

A. Background and Significance

In an effort to gain enhanced affinity and improve upon the selectivity of Mut1 for ssDNA, further structural studies of Mut1 were conducted. This chapter focuses on the structural features of the peptide as we endeavor to optimize Mut1 as an OB-fold mimic. Although the study of β-sheet structure has lagged behind that of α-helices, much research has been done to advance our understanding of β-sheet structure and to improve *de novo* design of small structured beta-sheet peptides and proteins in recent years.[1] The WW domain

[1] (a) Jager, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10648-10653. (b) Fernandez-Escamilla, A. M.; Ventura, S.; Serrano, L.; Jimenez, M. A. *Protein Sci.* **2006**, *15*, 2278-2289. (c) Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3954-3959. (d) Kraemer-Pecore, C. M.; Lecomte, J. T. J.; Desjarlais, J. R. *Protein Sci.* **2003**, *12*, 2194-2205. (e) Jiang, X.; Kowalski, J.; Kelly, J. W. *Protein Sci.* **2001**, *10*, 1454-1465. (f) Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373-393. (g) Macias, M. J.; Gervais, V.; Civera, C.; Oschkinat, H. *Nat. Struct. Biol.* **2000**, *7*, 375-379. (h) Koepf, E. K.; Petrassi, M.; Ratnaswamy, G.; Huff, M. E.; Sudol, M.; Kelly, J. W. *Biochemistry* **1999**, *38*, 14338-14351. (i) Nguyen, H.; Jager, M.; Moretto, A.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3948-3953. (j) Jager, M.; Dendle, M.; Fuller, A. A.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 2306-2313. (k) Zarrinpar, A.; Lim, W. A. *Nat. Struct. Biol.* **2000**, *7*, 611-613. (l) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Peptide Sci.* **2005**, *80*, 303-311. (m) Dalby, P. A.; Hoess, R. H.; DeGrado, W. F. *Protein Sci.* **2000**, *9*, 2366-2376. (n) Chan, D. C.; Bedford, M. T.; Leder, P. *EMBO J.* **1996**, *15*, 1045-1054. (o) Macias, M. J.; Hyvonen, M.; Baraldi, E.; Schultz, J.; Sudol, M.; Saraste, M.; Oschkinat, H. *Nature* **1996**, *382*, 646-649. (p) Chen, H. I.; Sudol, M. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7819-7823. (q) Sharman, G. J.; Searle, M. S. *Chem. Commun*. **1997**, 1955-1956. (r) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc*. **1998**, *120*, 5291-5300. (s) Griffiths-Jones, S. R.; Searle, M. S. *J. Am. Chem. Soc*. **2000**, *122*, 8350-8356. (t) Kortemme, T.; Ramirez

is one class of peptides that is especially well studied because its three-stranded β-sheet makes it one of the smallest known functional native peptides.[1a-p] This laboratory has also contributed significantly to this promising field through studies of β-hairpins.[2] Research has been conducted in the areas of protein design and redesign, folding and stability, and structure-function studies.[1, 2] The knowledge gained from these studies and other significant breakthroughs have led to our attempt to stabilize Mut1 to optimize binding affinity and selectivity for ssDNA.

B. Results

i. Sequence Design

Beginning with the Mut1 (Figure 6.1a-b) peptide from previous studies, mutations were made to improve binding and selectivity for ssDNA. The original mutant was designed with the WKWK nucleotide binding pocket in the second hairpin of a three-stranded beta-sheet peptide based on the FBP11 WW1 domain. The insertion of these mutations into this peptide produced a folded peptide that binds ssDNA with a K_d of about 20 μM and which demonstrates selectivity for ssDNA over duplex DNA by about 20-fold.

Upon determination that strand 1 of Mut1 was not well folded and could not contribute to binding as was intended in the original design, we used our knowledge of β-

---

-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253-256. (u) Kaul, R.; Angeles, A. R.; Jager, M.; Powers, E. T.; Kelly, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 5206-5212. (v) Nguyen, H.; Jager, M.; Kelly, J. W.; Gruebele, M. *J. Phys. Chem. B* **2005**, *109*, 15182-15186. (w) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 388-395. (x) Ramirez-Alvarado, M.; Blanco, F. J.; Niemann, H.; Serrano, L. *J. Mol. Biol.* **1997**, *273*, 898-912.

[2] (a) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126*, 2028-2034. (b) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 6518-6519. (c) Hughes, R. M.; Benshoff, M. L.; Waters, M. L. *Chem. Eur. J.* **2007**, *13*, 5753-5764. (d) Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-1531.

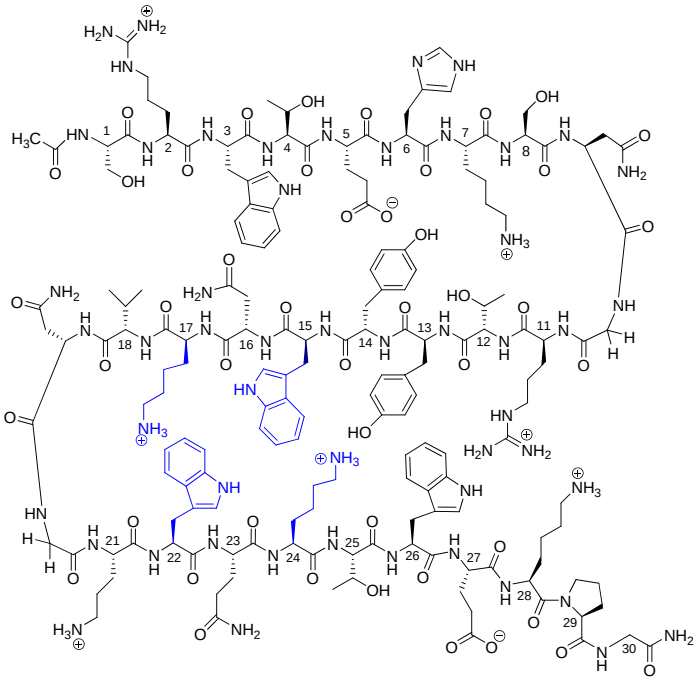
sheet structure to design more well folded peptides without disrupting the favorable interactions afforded by the Trp binding pocket. Maintaining hairpin 2 stability is crucial because disruption of the binding pocket would decrease binding affinity and likely selectivity. Three positions were mutated to give Mut1A (Figure 6.1c). Ser8 was substituted for a Val because Ser is thought to be destabilizing near a turn sequence and Val has been shown to be a stabilizing residue in that position in the WKWK peptide and other systems as well.[1x, 3] Phe was substituted for Tyr 14 to ease in NMR analysis. Phe is also known to be important in the recognition of ssDNA and damaged bases by OB-fold proteins,[4] so this may contribute to binding. The Asn at position 16 was replaced with a Leu residue as leucine is known to pack with cross-strand Trp residues to stabilize beta-sheet peptides, and thus should pack with Trp 3 to increase the stability of hairpin 1.[1w]

[3] (a) Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009**, *10*, 539-544. (b) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25. (c) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005,** *70,* 1105-1114. (d) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.

[4] (a) Chan, K.-W.; Lee, Y.-J.; Wang, C.-H.; Huang, H.; Sun, Y.-J. *J. Mol. Biol.* **2009**, *388*, 508–519. (b) Warren, E. M.; Huang, H.; Fanning, E.; Chazin, W. J.; Eichman, B. F. *J. Biol. Chem.* **2009**, *284*, 24662-24672. (c) Kunkel, T. A.; Erie, D. A. *Annu. Rev. Biochem.* 2005, *74*, 681-710. (d) Yamamoto, A.; Schofield, M. J.; Biswas, I.; Hsieh, P. *Nucleic Acids Res.* **2000**, *28*, 3564-3569. (e) Malkov, V. A.; Biswas, I.; Camerini-Otero, R. D.; Hsieh, P. *J. Biol. Chem.* **1997**, *272*, 23811-23817.

(a) Mut1:
Ac-S-R-W-T-E-H-K-S-N-G-R-T-Y-Y-**W**-N-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(b)



(c) Mut1A:
Ac-S-R-W-T-E-H-K-**V**-N-G-R-T-Y-**F**-**W**-**L**-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(d) Mut1B:
Ac-S-R-W-T-E-**V**-**F**-**V**-N-G-R-**F**-**I**-**K**-**W**-**L**-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(e) Mut1C:
Ac-S-R-W-T-E-**V**-**F**-**V**-N-G-R-**F**-**L**-**K**-**W**-**L**-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(f) ssDNA: 5'-CCATCGCTACC-3'

(g) dsDNA: 5'-CCATCGCTACC-3'
3'-GGTAGCGATGG-5'

Figure 6.1. WW domain Mut1. (a) Sequence of Mut1 with residues in the binding pocket highlighted. (b) Structure of Mut1. The WKWK binding cleft is shown with Trp and Lys residues in blue. (c-e) Mutants 1A-C. Mutations made at each generation are highlighted. (f) ssDNA sequence. (g) dsDNA sequence.

We wanted to further stabilize the β-sheet, so a second generation of mutations was made to give Mut1B (Figure 6.1d). Strands 1&2 of Mut1B contain elements from two stable hairpins designed in this laboratory. The N-terminal portion of the peptide contains the previously reported WKL sequence in non-hydrogen bonded positions 3, 14, and 16,[2c] and the residues near the turn at positions 5, 7, 12, and 14 mimic the EFFK peptide from a previous report.[5] His 6 and Tyr 13 were mutated to Val and Ile, respectively, because a cross-strand hydrophobic interaction on the H-bonded face has been shown to be important for folding in β-hairpin model systems.[1w] The cross-strand phenylalanines at positions 7 and 12 were intended to be stabilizing residues near the turn to increase folding throughout the hairpin. It is important that the charge is maintained at 5 to compare to previous mutants because electrostatic interactions are important for binding to the phosphates on the DNA backbone. For this reason, Lys that was replaced by Phe at position 7 was incorporated at position 14. These mutations made up Mut1B to determine if the mutations that were stabilizing for previous β-hairpins would stabilize the full Mut1 peptide. Strand 3 was not changed for this mutant.

A third generation mutant was designed with Ile 13 mutated to Leu (Mut1C; Figure 6.1e). It was thought that the Ile which should be stabilizing to hairpin 1 may be destabilizing to hairpin 2 due to steric clashes with Trp 26. Additionally, Ile may alter the conformation of Trp 26 and influence binding to ssDNA. Leucine has been shown to pack favorably with Trp residues.[1w]

In addition, cyclization was explored to induce stability within the peptide. Cyclization has been used in designed peptides to stabilize structures and to improve

[5] Kiehna, S. E.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2657-2667.

binding.[6] Because strand 1 was less structured than hairpin 2, we chose to cyclize hairpin 1 of Mut1 in an attempt to move the residues cross-strand from the binding pocket in close proximity to the cleft such that they might contribute to the binding event (Mut1D; Figure 6.2).

(a) Mut1D:
Ac-**C**-R-W-T-E-H-K-S-N-G-R-T-Y-Y-**W**-N-**K**-**C**-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(b)

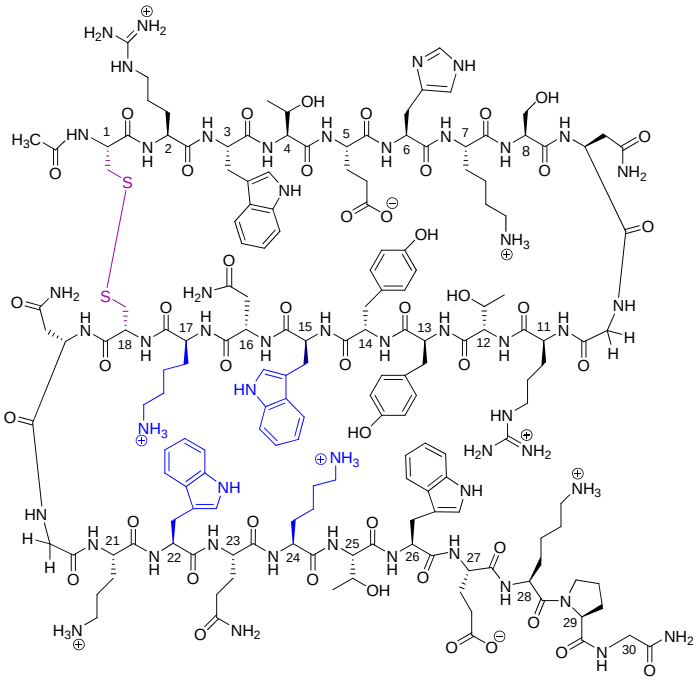


Figure 6.2. Sequence and structure of WW domain Mut1D. (a) Sequence of Mut1D. (b) Structure of Mut1. The WKWK binding cleft is shown with Trp and Lys residues in blue. The disulfide bond formed to cyclize strands 1&2 is highlighted in purple.

[6] (a) Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew. Chem. Int. Ed.* **2007**, *46*, 9101-9104. (b) Moreno, R.; Jiang, L.; Moehle, K.; Zurbriggen, R.; Gluck, R.; Robinson, J. A.; Pluschke, G. *ChemBioChem* **2001**, *2*, 838-843.

ii. CD Characterization of Folding

Structural insight was first gained from the CD spectra for all of the full peptides as well as the truncated peptides composed of Strands 1&2 (S12) and those forming Strands 2&3 (S23). The full peptides all exhibit β-sheet structure with minima around 210 nm (Figure 6.3). The CD signal for these peptides occurs closer to 210 nm than the typical 215 nm, as was seen previously (Chapter 4), and this is likely due to the contributions of the Trp residues in the peptide. Mut1 displays two distinct minima, one which is consistent with beta-sheet structure and another which indicates that the peptide has some random coil character. Mutants 1A-1C each have a broad minimum around 210 nm that overlaps with the region for random coil peptides at 195 nm. This suggests that they all have some random coil character as well, but this is manifested differently in the mutants' CD spectrum than in that of Mut1. All of the mutants 1A-1C display a stronger CD signal at 210 nm than does Mut1. The mutants no longer have the exciton coupling peak at 227 nm, which is consistent with a mutation of an aromatic residue. The orientation of the aromatic species within the peptide likely changed with the first generation mutation. The exciton coupling peak is present, however, with Mut1D, which is the same Mut1 sequence with hairpin 1 cyclized. In the case of Mut1D, the peak at 227 nm is larger than for the original Mut1. This suggests that the cyclization changes the structure such that the orientation of the aromatic residues within the peptide is different from the uncyclized mutant.

A close inspection of the CD spectra for the peptides composing hairpin 1 (Figure 6.4a) reveals a measured increase in structure with each generation of mutations. With each round of mutations, the minima around 210 nm grow more distinct and the minima at 195 nm

decrease in intensity. This data suggests that the mutations made to Strands 1&2 have increased the β-sheet structure in hairpin 1.

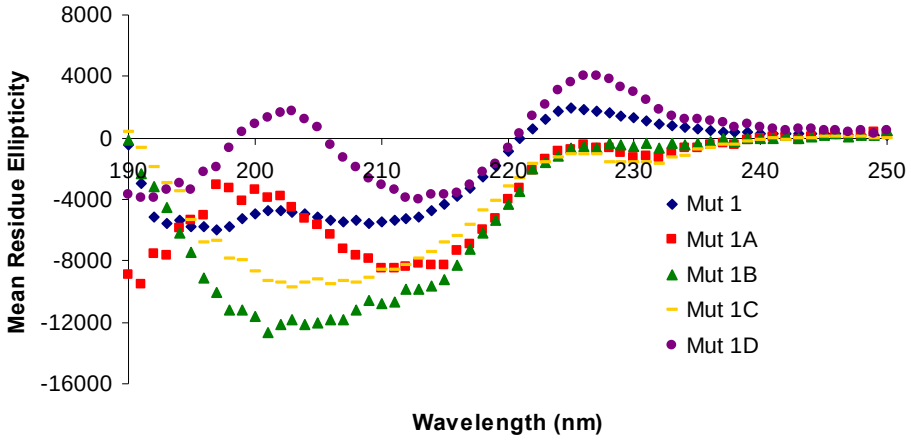
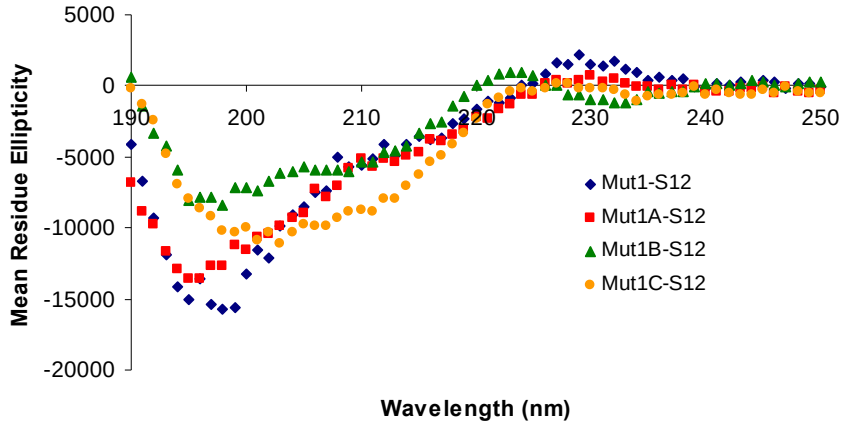


Figure 6.3. CD spectra for the WW domain Mut1 and the structural mutants 1A-D. Data was recorded at 30 μM peptide, 10 mM Na_2HPO_4, 298 K, pH 7.0, with a wavelength range from 250-190 nm.

(a)



(b)

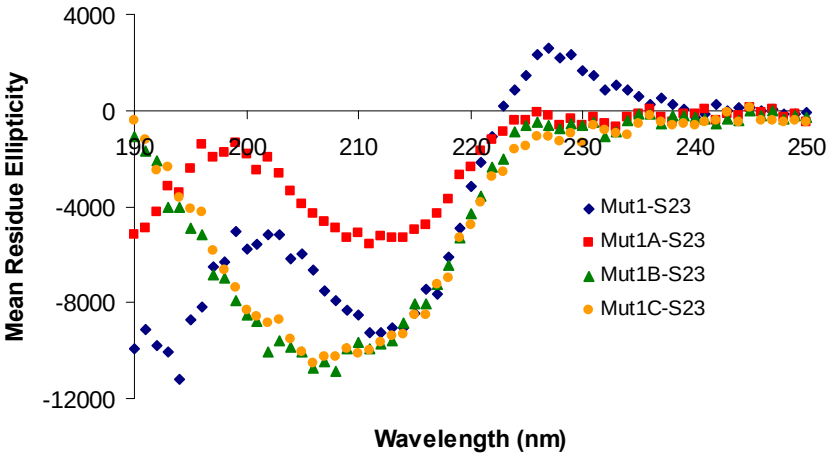


Figure 6.4. CD spectra for (a) Mut1-S12 and structural S12 mutants and (b) Mut1-S23 and S23 mutant hairpins. Data was recorded at 30 μM peptide, 10 mM Na_2HPO_4, 298 K, pH 7.0, with a wavelength range from 250-190 nm.

Since the binding pocket of Mut1 is in Strands 2&3, it is important that the mutations do not decrease the stability of hairpin 2. Thus, we characterized the folding of hairpin 2 for each of the mutants as well. The exciton coupling peak is lost with Mut1A in the case of Strands 2&3 as well. This is likely due to a change in structure which could change the relative orientations of the aromatic residues in the peptide. The mutations do not seem to decrease the stability significantly, however. Mut1A-S23 seems to be destabilized but still

maintains its primarily β-sheet character. The other mutants maintain a strong minimum around 210 nm indicating that there was no destabilization of hairpin 2 due to the mutations.

iii. NMR Characterization of Structure

To better understand the specific interactions occurring in Mut1 and in the mutants, the peptides were more fully characterized by one- and two-dimensional NMR experiments. Figure 6.5 shows the Hα chemical shift differences relative to random coil for Mut1 and its two hairpins, S12 and S23. One method to determine the extent of folding of each peptide is to determine the Gly Hα splitting values.[7] The glycine in the turn possesses two diastereotopic hydrogens. As folding of the hairpin increases, the splitting of the hydrogens increases as well, giving a measure of stability.[7b] A comparison of the Gly splitting in a β-hairpin with that of a fully folded cyclic control peptide gives the extent of folding and overall stability for the peptide (see Experimental Section). Glycine splitting and α-hydrogen (Hα) chemical shift values were determined to characterize the full peptides, both hairpins for each mutant, and the individual strands composing each. The Hα chemical shifts of the hairpins and the full mutants relative to the random coil peptides indicate the degree of β-sheet structure for each residue along the peptide backbone. Downfield shifting of Hα protons is evidence of increased hairpin population, with a chemical shift difference of greater than 0.1 ppm taken to indicate β-sheet structure.[7] As observed in previous studies (see Chapter 4), the first hairpin is not structured and cannot contribute to DNA binding. The mutations were intended to increase the peptide stability, specifically in Strands 1&2. Glycine splitting data also indicates that Strand 1 is not fully structured.

---

[7] (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333. (b) Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996-2007.

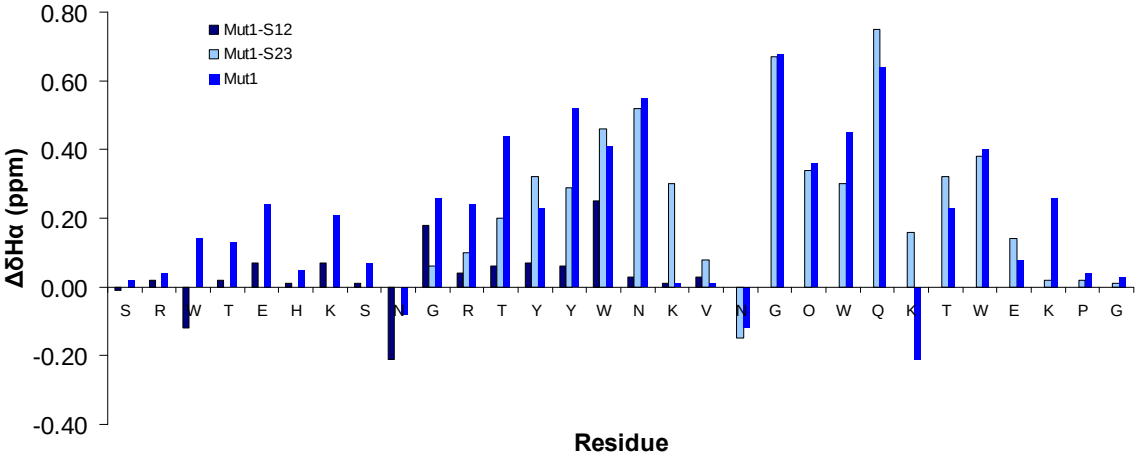


Figure 6.5. NMR chemical shift differences for WW domain Mut1 peptides. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 7.0 (uncorrected), referenced to DSS. Error is ± 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

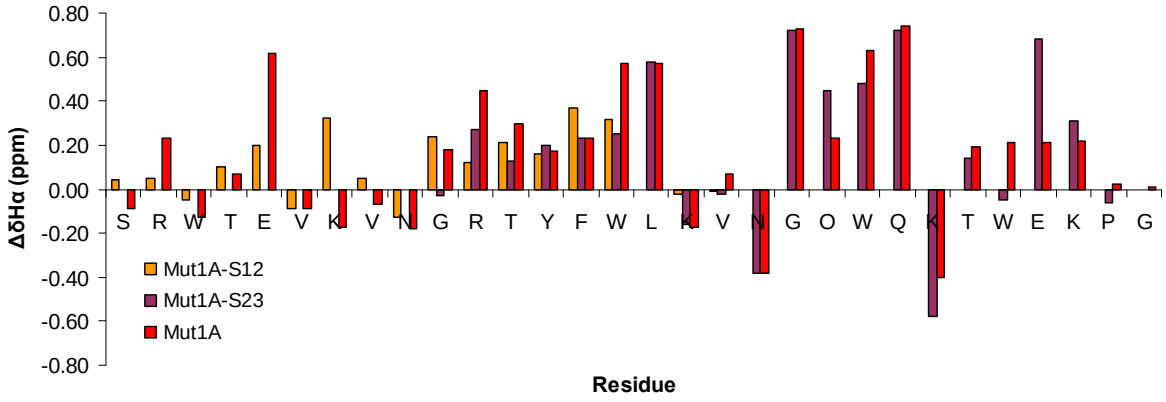


Figure 6.6. NMR Hα chemical shift differences for Mut1A peptides. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is ± 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

The NMR data for Mut1A (Figure 6.6) indicates that the structure of hairpin 1 increased to some extent, but not significantly. Strand 1 is still much less structured than Strands 2&3. Glycine splitting data is consistent with the chemical shift data, with an increase in fraction folded for hairpin 1 from 27 to 39 percent. In the full peptide, the Mut1A folding decreased compared to Mut1, with Mut1-S12 increasing in stability in the full peptide

and Mut1A-S12 stability decreasing in the full peptide. With the second generation mutant, a marked increase in structure was observed (Figure 6.7). Strands 2&3 seem to be somewhat more stable than Strands 1&2, but the overall stability for the peptide is much greater than for previous mutants. Fraction folded values indicate this as well, with a jump from 39 to 92 percent folded in hairpin 1 and a fraction folded of 81 percent in the full peptide.

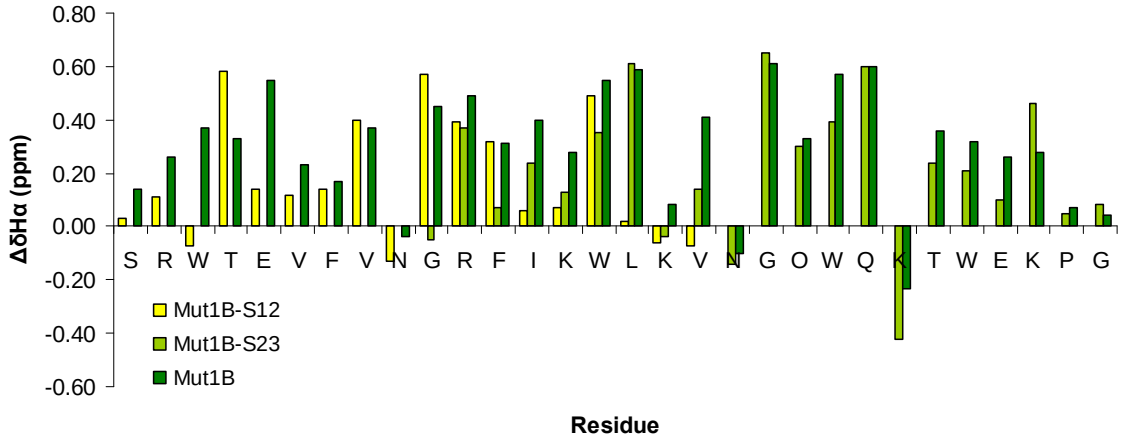


Figure 6.7. NMR Hα chemical shift differences for Mut1B peptides. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is $\pm$ 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

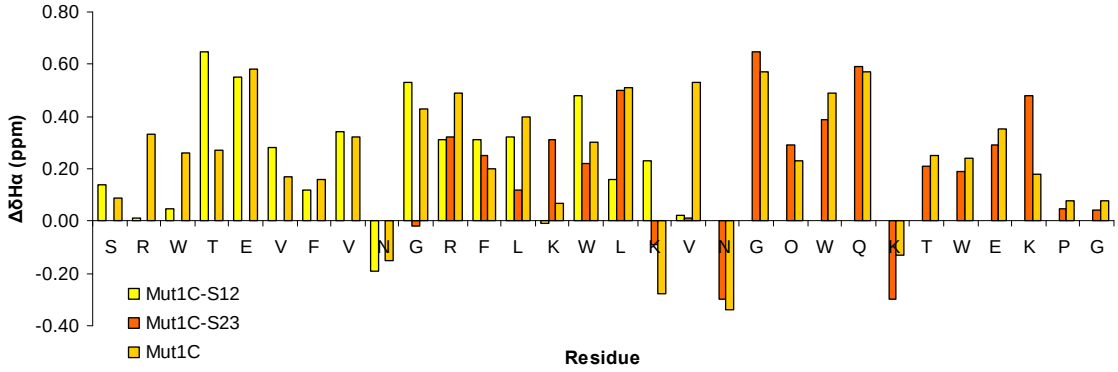


Figure 6.8. NMR Hα chemical shift differences for Mut1C peptides. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is $\pm$ 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

There was little difference between the NMR spectra for Mut1B and Mut1C (Figures 6.7 and 6.8). The chemical shift differences were about the same for both peptides, and fraction folded data indicates that the folding may have been slightly destabilized with this Ile to Leu mutation. Mut1B has a fraction folded of 81% for both hairpins within the full peptide, and the fraction folded is slightly higher for the hairpins alone (Table 6.1). Mut1C seems to be less structured for the full peptide and the individual hairpins, with fraction folded values of 84% for both of the individual hairpins and 68 and 74% folded for S12 and S23, respectively within the full peptide. No chemical shift data is available for Mut1D yet, but glycine splitting data (Table 6.1) indicates that one hairpin is fully folded (98%) and the other hairpin is moderately well folded (69% folded).

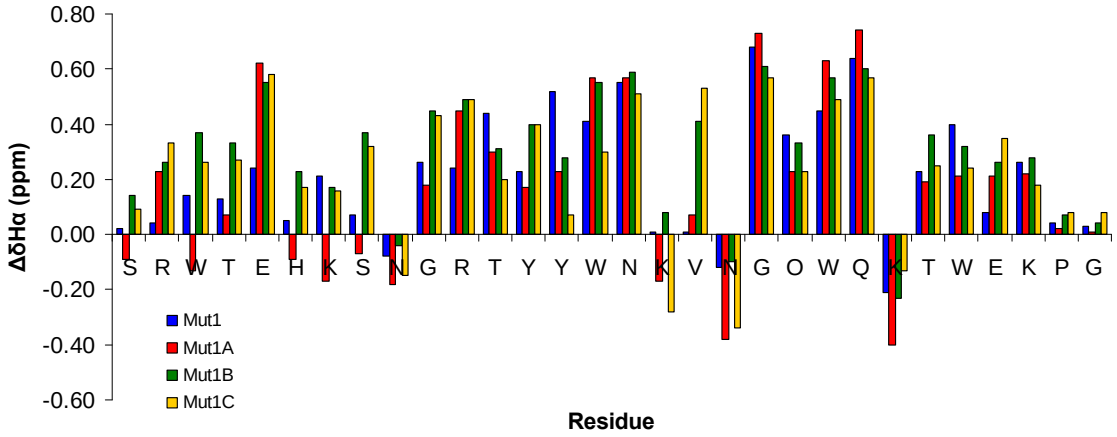


Figure 6.9. NMR chemical shift differences for Mut1 full peptide and structural mutants 1A-D. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is $\pm$ 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

A comparison of all of the NMR data is consistent with the CD data, indicating that the mutations resulted in an increase in the stability of the peptide. The chemical shift differences also indicate that structure has been gained in Strand 1. Overlays of the full peptides and the individual hairpins for each mutant are provided (Figures 6.9-6.11), and they

suggest that the chemical shift differences for many of the amino acids within strands 2&3 are very similar for all peptides. Residues in Strands 1&2 exhibit the largest differences among the mutants.

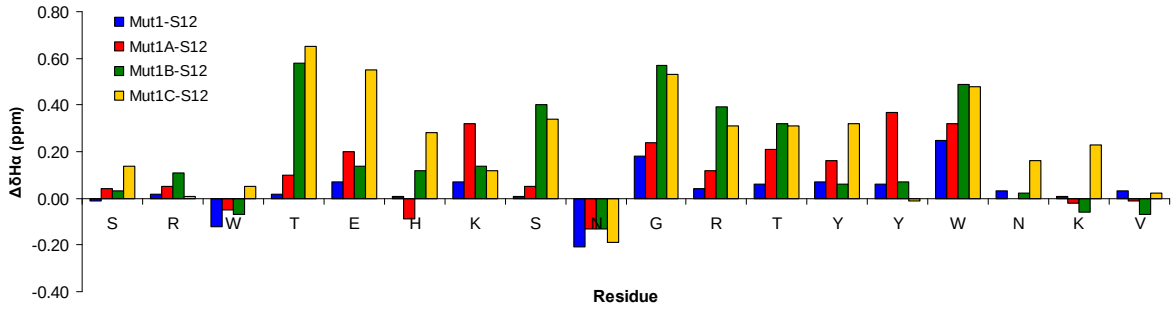


Figure 6.10. NMR Hα chemical shift differences for Mut1-S12 and S12 structural mutants. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is ± 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

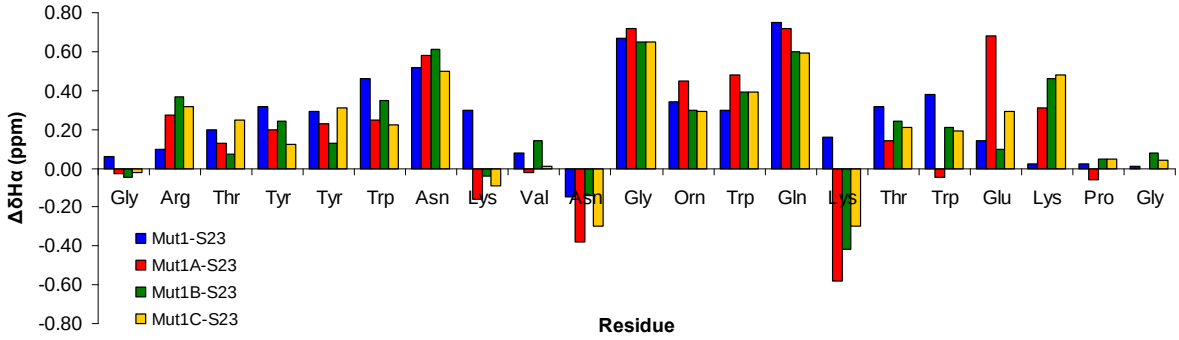


Figure 6.11. NMR Hα chemical shift differences for Mut1-S23 and S23 structural mutants. Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. Error is ± 0.005 ppm, determined by the chemical shift accuracy on the NMR spectrometer.

Table 6.1. Fraction folded values for Mut1 and structural mutants.[a]

Peptide	ΔδGly, ppm[b]	Fraction Folded (Gly Splitting)[c]
Mut1 S12	0.11	27
Mut1 S23	0.67	89
Mut1	0.26; 0.68	63; 91
Mut1A S12	0.24	39
Mut1A S23	0.73	96
Mut1A	0.18; 0.73	30; 96
Mut1B-S12	0.57	92
Mut1B-S23	0.65	87
Mut1B	0.46; 0.61	81; 81
Mut1C-S12	0.53	84
Mut1C-S23	0.65	84
Mut1C	0.43; 0.57	68; 74
Mut1D S12[d]	0.41	100
Mut1D-S23[e]	0.67	96
Mut1D	0.48; 0.40	98; 69

(a) Conditions: Values calculated from data obtained at 298 K, 50 mM potassium dideuterium phosphate, pD 4.0 (uncorrected), referenced to DSS. (b) Error is $\pm$ 0.005 ppm, determined by chemical shift accuracy on the NMR spectrometer. (c) Error is $\pm$ 0.01, determined by chemical shift accuracy on the NMR spectrometer. (d) Mut1D-S12 = Mutant 1 S1&2 Cyclic (same peptide); cyclic peptides are taken to be fully folded. (e) Mut1D-S23 = Mutant 1 S2&3; used as an estimate for Gly splitting calculations.

iv. Characterization of the Binding Interactions Between Mutants and DNA

To complete the study, we wanted to determine if the mutations increased the binding and selectivity for ssDNA. As seen in previous studies, structure is crucial for binding,[3a] so fluorescence binding studies between the mutants and ss- and dsDNA would determine if increased structure in strand 1 would improve ssDNA binding. Fluorescence anisotropy binding studies were conducted as in the past except these experiments were performed using a plate reader (see Experimental Section).

Table 6.2. Comparison of dissociation constants for the binding interaction between Mut1 peptide structural mutants and ssDNA and dsDNA sequences [a]

Peptide	ssDNA K_d, μM (error)	dsDNA K_d, μM (error)
Mut1	48.7 (3.1)[b]	235 (18)[b]
Mut1A	51.1 (2.1)	263 (7)
Mut1B	53.0 (3.5)	271 (12)
Mut1C	91.9 (3.4)	436 (28)
Mut1D	35.6 (1.8)	210 (8)

(a) Conditions: 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of three measurements. The error is from the fitting. This data was collected using fluorescence anisotropy with TAMRA as the fluorophore. (b) This data was reported previously (see Chapter 4) and the difference is due to the instrument used.

As can be seen in Table 6.2 and Figure 6.12, ssDNA binding is fairly similar for most of these mutants. Mut1, Mut1A, and Mut1B all bind ssDNA with a K_d of about 50 μM. Mut1C does not bind as tightly, with a dissociation constant of 91.9 μM. Mut1D, however, displays a somewhat higher affinity, with a K_d of 35.6 μM. The data suggests that the structural changes either did not improve the structure enough for strand 1 to contribute to binding or the mutations affected the binding pocket to the extent that the mutants did not bind in the same manner. Mut1D may keep the binding pocket the same and may bring Strand 1 closer to the binding pocket in such a way that binding is improved.

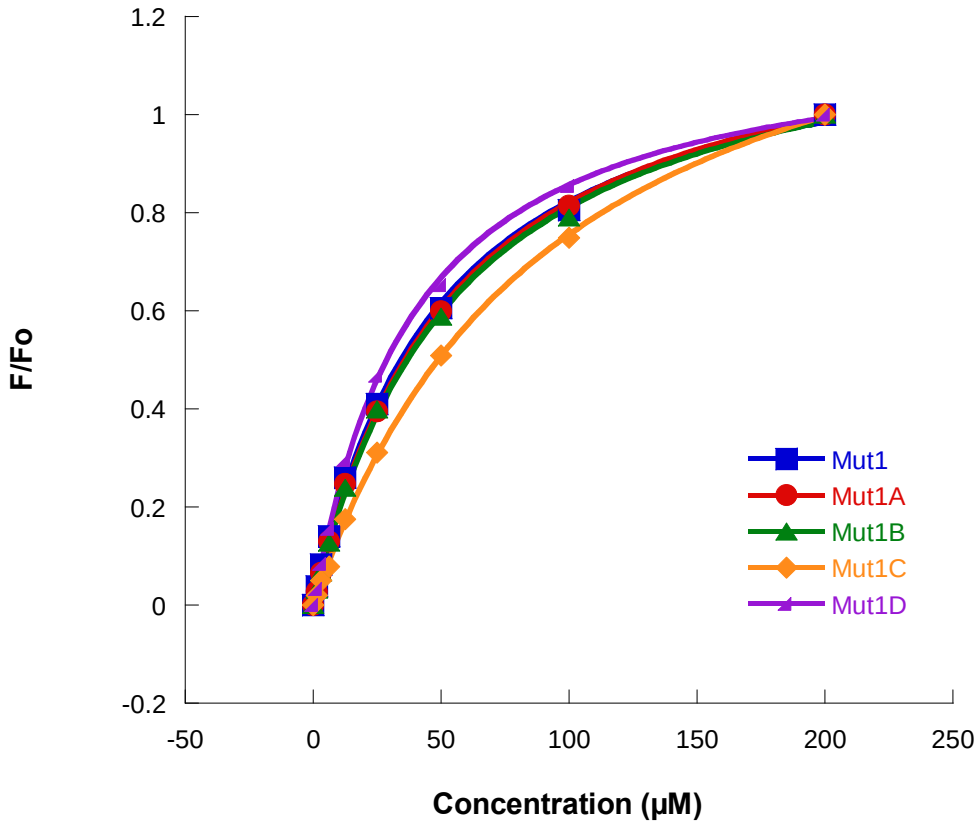


Figure 6.12. Fluorescence titrations of Mut1 and structural mutants with ssDNA; 5 µM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

For duplex DNA, binding by each peptide is weak (Figure 6.13), but the trends for binding affinities are the same as for ssDNA. Mutants 1, 1A, and 1B have very similar binding constants, while 1C is a weaker dsDNA binder, as in the case of ssDNA. Mut1D binds dsDNA with a slightly higher affinity, with a K_d of 210 µM. This makes Mut1D the peptide with the highest affinity and selectivity for ssDNA, with a preference for ssDNA of nearly 6-fold.

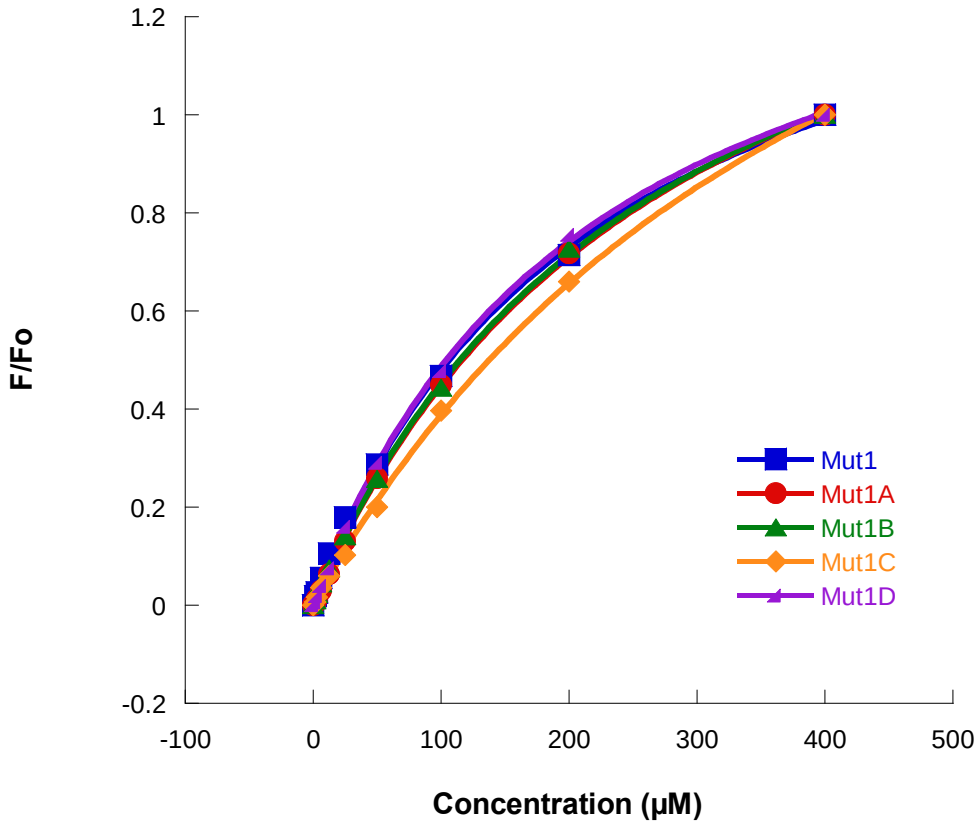


Figure 6.13. Fluorescence titrations of Mut1 and structural mutants with dsDNA; 5 μM peptide; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

C. Discussion and Conclusions

Structural mutations made to the WW domain Mut1 produced increased stability with each new mutant designed. Varying degrees of stability resulted with each mutation, but in each case greater stability was achieved in Strands 1&2. Based on Gly splitting alone, Mut1C seems to be less stable than Mut1B. A closer examination of the chemical shift data reveals, however, that greater downfield shifting of some amino acids occurs with Mut1C, indicating that the mutations increased the stability to some extent. One can see from the

chemical shift data that the ends of the peptide are frayed, and this could be the reason that binding did not improve for Mutants 1A-1C. The Arg 2 position in the full peptide incrementally increases in structure with each mutation, but the position is still not as structured as other residues in hairpin 1. The Thr at position 4 exhibits large chemical shift changes in the S12 peptides, but that does not occur to the same extent in the full length mutants. These data indicate that the N-terminus of the peptide is still not fully folded and may not be positioned to interact favorably with DNA near the binding pocket. It may be that Mut1D has the advantage of the best sequence for overall stability of Strands 2&3, and specifically the binding pocket region, while maintaining Strand 1 near the binding pocket due to the rigidity of the cyclic hairpin 1. This would explain the improved binding to ssDNA. The Gly splitting data indicates that one hairpin is nearly 100% folded while the other is only about 70% folded in Mut1D. It could be that structure is induced upon binding. Further structural and binding data could provide more insight into these interactions, but the cyclization of hairpin 1 seems to have improved the stability and subsequently the binding of the peptide as has been seen in other systems. CD and NMR structural studies are underway to determine if structure is induced upon binding and to determine the role of structure in Mut1D's recognition of ssDNA.

D. Experimental Section

i. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using a Tetras Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with

DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 15 minutes. Each amino acid was coupled on a double coupling cycle of 30 minutes each. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin as well as sidechain deprotection was performed in 95% trifluoroacetic acid (TFA), 2.5% H_2O, and 2.5% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen, and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 0-100% solvent B over 60 or 100 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and peptide sequence was confirmed by MALDI mass spectrometry. Cyclic peptides were formed by oxidation of the cysteine residues with stirring in 1% DMSO in 10 mM phosphate buffer, pH 7.5 for 7-12 hrs. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

ii. DNA Sample Preparation

DNA sequences were purchased from IDT (Integrated DNA Technologies, Inc.). All DNA samples were dissolved in 10mM Na_2HPO_4, 100 mM NaCl, and adjusted to pH 7.0. Concentrations of both DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficients of the two DNA strands ($\varepsilon_{260,\ ssDNA} = 95500\ M^{-1}\cdot cm^{-1}$ and $\varepsilon_{260,\ dsDNA} = 112600\ M^{-1}\cdot cm^{-1}$). Equal concentrations of the two strands (in

sodium phosphate buffer, pH 7.0) were pooled in a final concentration of 100 mM NaCl. The solution was heated at 95 ºC for 5 minutes to anneal the strands and was then allowed to cool to room temperature before storing at -20 ºC.

iii. Peptide Concentration Determination

To determine the recognition of single-stranded and double-stranded oligonucleotides by the peptides, fluorescence titrations were performed by fluorescence anisotropy of the WW domain mutants titrated with increasing oligonucleotide concentrations. Peptide samples were prepared in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0. Peptide concentrations were determined in 5 M guanidine hydrochloride by UV-Vis by recording the absorbance of the fluorophore attached to the peptide, 5-(and -6)-Carboxytetramethylrhodamine, mixed isomers. TAMRA's extinction coefficient is 91000 $M^{-1} \cdot cm^{-1}$ at wavelength 559 nm. This extinction coefficient was supplied by Integrated DNA Technologies, www.idtdna.com. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm.

iv. Fluorescence Anisotropy

As in previous chapters, a fluorophore, 5-(and -6)-Carboxytetramethylrhodamine, mixed isomers (TAMRA), was purchased from Biotium, Inc., and was employed to provide an alternative to fluorescence quenching experiments. Because of the inner filter effect, many of the binding interactions cannot be measured accurately using the fluorescence quenching method. TAMRA was coupled onto the peptide Mut1 at Orn 21. The synthesis was completed by coupling Lys(ivDde) in the original ornithine position (Orn21Lys). The ivDde protecting group was orthogonally deprotected by 3 treatments with 2% hydrazine in DMF for 3 minutes each. Manual coupling of TAMRA was performed with two equivalents

of TAMRA (100 mg bottle used for 0.1 mmol peptide) and four equivalents of HOBT, HBTU, and DIPEA in DMF, and the reaction was allowed to bubble with nitrogen overnight. Cleavage from the resin and sidechain deprotection was completed as with all other peptides using 95% TFA, 2.5% H_2O, and 2.5% TIPS. The resulting peptide was purified by HPLC and its sequence and purity determined by mass spectrometry as in previous chapters (see Experimental Section, Chapter 4). The excitation wavelength used in the experiments is 559 nm, and the observed emission wavelength is 583 nm. In a departure from the usual anisotropy experiments, a plate reader was used in these studies. The peptide concentration for each study was 5 μM; DNA concentrations ranged from 0-400 μM. DNA samples were made by serial dilutions, and a multi-channel pipet was used. DNA samples were placed into a prep 96 well plate and then transferred using the multi-channel pipet to a Corning 96 well plate. The experiments were performed using a POLARstar Omega plate reader. Anisotropy was determined by the software that came with the instrument. The anisotropy was fit to the following equation (Equation 6.1) using Kaleidagraph to determine the binding constant

Equation 6.1. $$F = \frac{(((-(-K_d - [L] - [P]) - (\sqrt{((-K_d - [L] - [P])^2 - (4 \times [L] \times [P]))}))}{(2 \times [P])) \times (I_\infty - I_o)) + I_o}$$

where F is the fluorescence anisotropy, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, [P] is the peptide concentration for each fraction, and K_d is the dissociation constant. Equation 6.1 was derived from equations given by Wang and coworkers.[8]

v. Circular Dichroism

All CD measurements were obtained using an AVIV 62 DS Circular Dichroism Spectrometer. CD data was obtained for the WW Domain peptides at 30 μM concentrations

[8] Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768-779.

from 250-190 nm. The peptides were dissolved in 10 mM Na_2HPO_4, pH 7.0. Wavelength scans were performed at 298 K.

vi. NMR Spectroscopy

NMR samples for Mutant 1 peptides were made in about 1 mM concentrations (650 μL) and were dissolved in 50 mM KD_2PO_4 (in D_2O), 0.5 mM DSS, and buffered to pD 7.0 (uncorrected) with sodium deuteroxide. Control peptides and other mutants were dissolved in 50 mM d3-NaOAc (in D_2O), 0.5 mM DSS, and buffered to pD 4.0. These samples were made with a lower pD because of solubility problems. A Mutant 1 sample was made and analyzed by NMR to determine if the change in pD affected the NMR signal. NMR data for Mutant 1 at pD 4.0 and at pD 7.0 was not significantly different. Samples were analyzed on a Varian Inova 600 MHz spectrometer. All one-dimensional NMR spectra were collected using 32K data points and between 16 and 64 scans and 1-3 s presaturation or solvent suppression. The two-dimensional TOCSY NMR spectra used pulse sequences from the Chempack software. The 2D NMR scans were obtained using 16-64 first dimension scans and 128-512 second dimension scans. All spectra were analyzed by using standard window functions (Sinebell and Gaussian). Peptide proton assignments were made using standard methods as described by Wuthrich.[9] Deviations in alpha hydrogen chemical shifts from random coil values, $\Delta\delta H\alpha$, were calculated according to Equation 6.2,

Equation 6.2. $\Delta\delta H_\alpha = \delta H_{\alpha,obs} - \delta H_{\alpha,RC}$

where $\delta H_{\alpha,obs}$ is the observed chemical shift of a given alpha hydrogen in the peptide, and $\delta H_{\alpha,RC}$ is the random coil chemical shift of the corresponding proton determined from

[9] Wuthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley-Interscience: New York, 1986.

unstructured control peptides, strand 1, strand 2, and strand 3. The extent of folding for each peptide can be determined by calculating the fraction folded, as in Equation 6.3,

Equation 6.3. Fraction folded = $\Delta\delta Gly_{obs}/\Delta\delta Gly_{100}$

where $\Delta\delta Gly_{obs}$ is the observed splitting of the glycine diastereotopic protons for the peptide, and $\Delta\delta Gly_{100}$ is the glycine splitting for the fully folded control peptide that is presumed to take on a 100% fold.

The following pages include tables of proton chemical shift assignments and ^{1}HNMR spectra for each mutant. Chemical shift values that are highlighted were not able to be unambiguously assigned. Amino acids are colored to indicate mutations as in Figures 6.1 and 6.2.

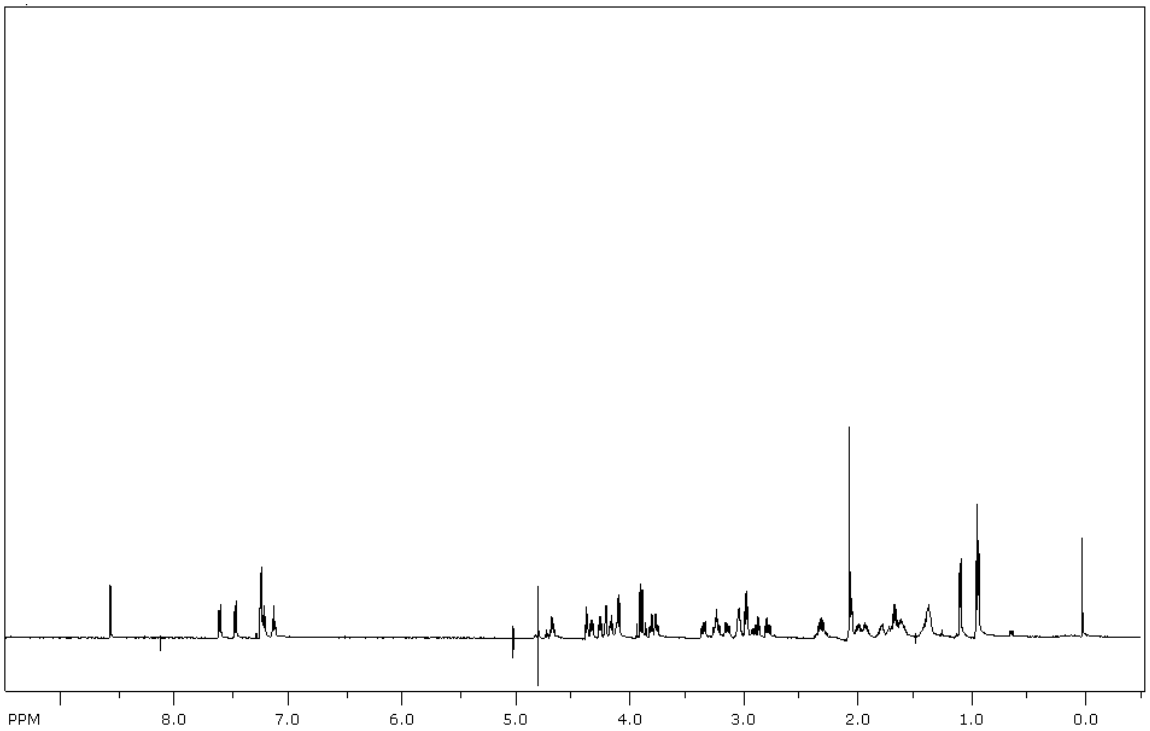


Figure 6.14. ^{1}HNMR spectrum of Mut1A-S1.

Table 6.3. Proton Chemical Shift Assignments for Peptide Mut1A-S1 (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.38	3.80			
Arg	4.26	1.61	1.36	3.04	
Trp	4.74	3.34, 3.23			
Thr	4.21	4.10	1.08		
Glu	4.17	1.94	2.32		
His	4.69	3.24			
Lys	4.34	1.73	1.39	1.68	2.97
Val	4.09	2.06	0.94		
Asn	4.71	2.86, 2.78			
Gly	3.92				

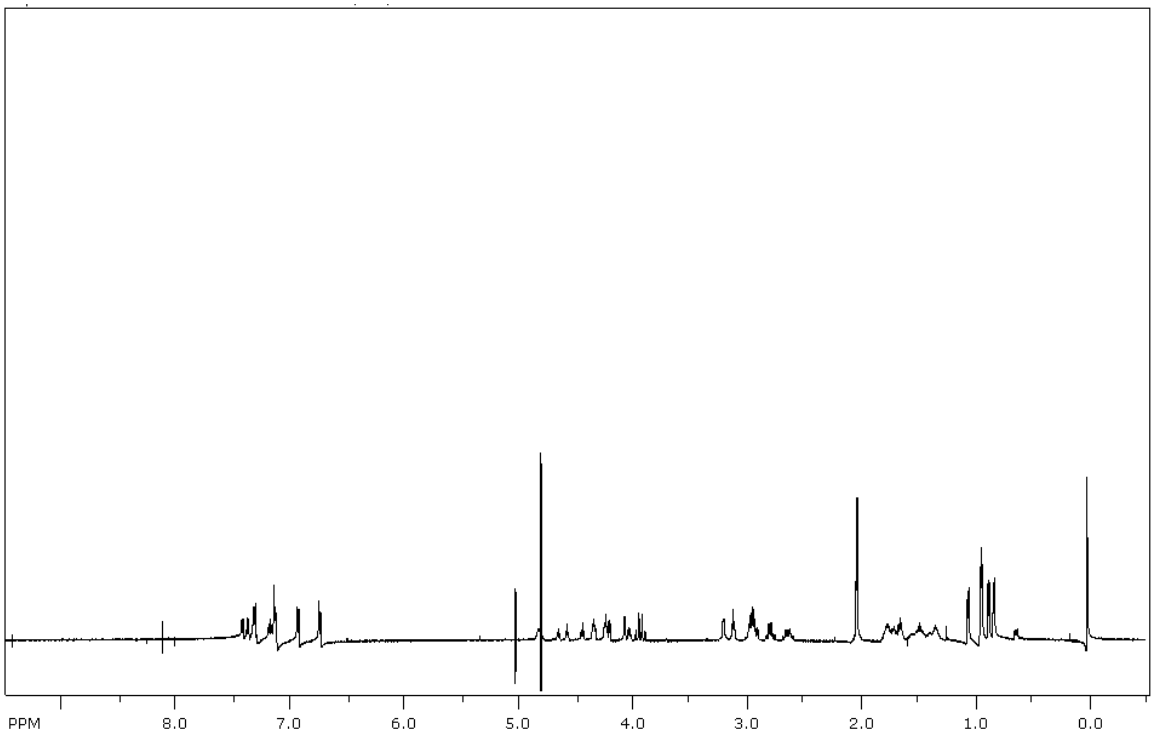


Figure 6.15. ^{1}HNMR spectrum of Mut1A-S2.

Table 6.4. Proton Chemical Shift Assignments for Peptide Mut1A-S2 (in ppm).

Residue	α	β	γ	δ	ε
Asn	4.66	2.79			
Gly	3.93				
Arg	4.34	1.74	1.54	3.12	
Thr	4.21	4.03	1.06		
Tyr	4.34	2.65			
Phe	4.44	2.97			
Trp	4.57	3.21			
Leu	4.24	1.72	1.69	0.85	
Lys	4.24	1.67	1.35	1.47	2.94
Val	4.06	2.04	0.94		

Mut1A-S3 (and S3 for all other mutants in this chapter) is the same as Mut1-S3, shown in Chapter 4.

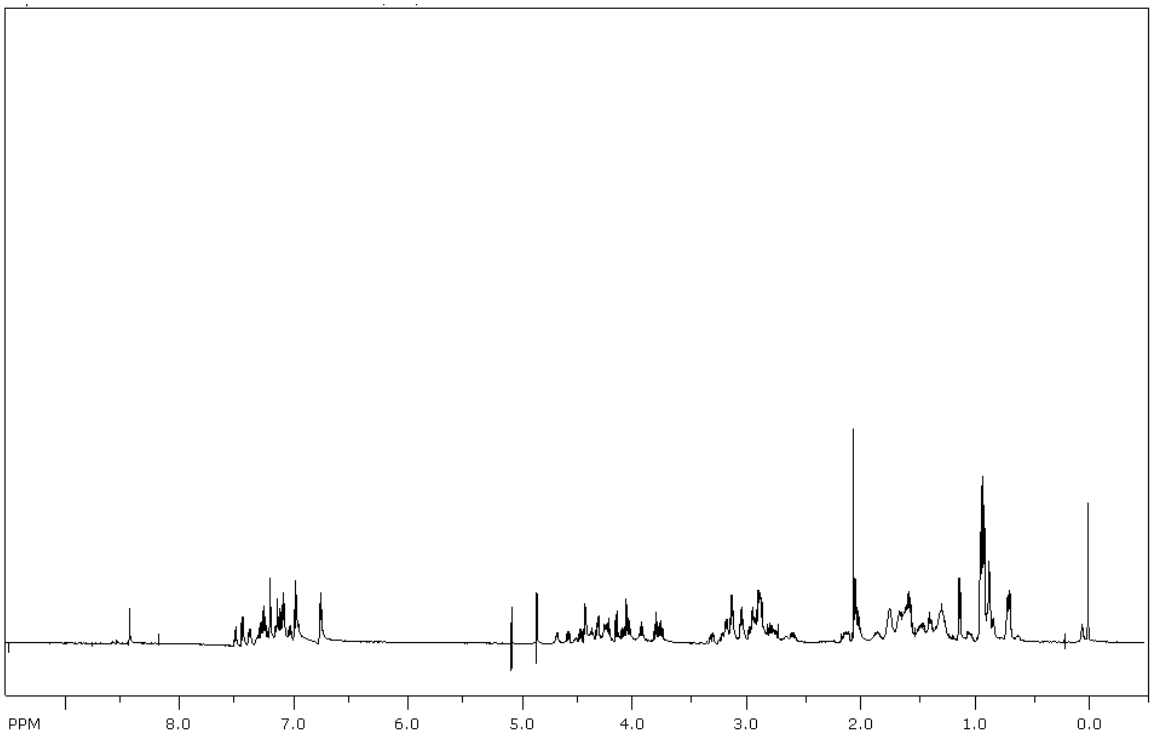


Figure 6.16. ^{1}HNMR spectrum of Mut1A-S12.

Table 6.5. Proton Chemical Shift Assignments for Peptide Mut1A-S12 (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.42	3.78			
Arg	4.31	1.61, 1.56	1.39	3.05	
Trp	4.69	3.05, 2.85			
Thr	4.31	3.92	0.94		
Glu	4.37	1.85	2.13		
His	4.60	3.19			
Lys	4.66	1.67	1.27	1.56	2.99
Val	4.14	2.03	0.81		
Asn	4.58	2.95, 2.79			
Gly	4.04, 3.80				
Arg	4.46	1.74	1.60, 1.47	3.13	
Thr	4.42	4.10	1.13		
Tyr	4.50	2.76, 2.59			
Phe	4.81	2.90			
Trp	4.89	3.31, 3.18			
Leu	4.24	1.46, 1.30	1.02	0.76, 0.69	
Lys	4.22	1.73	1.30	1.61	2.91
Val	4.05	2.01	0.93		

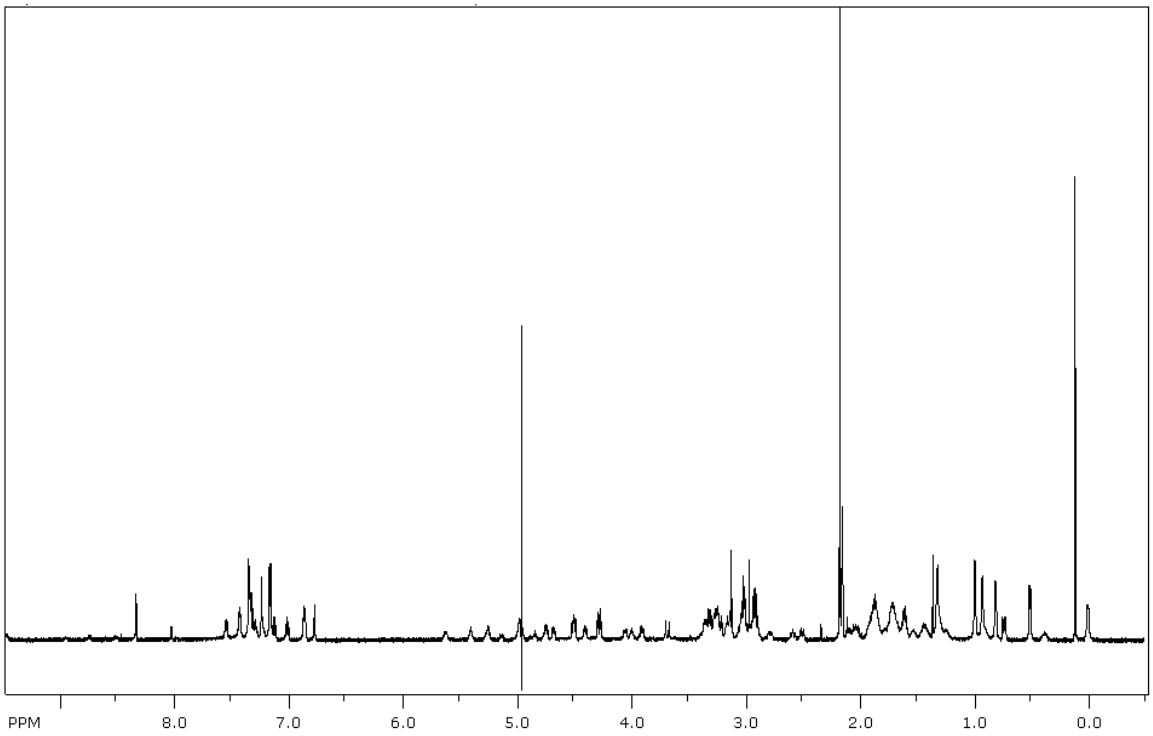


Figure 6.17. ^{1}HNMR spectrum of Mut1A-S12 Cyclic.

Table 6.6. Proton Chemical Shift Assignments for Peptide Mut1A-S12 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.73	3.02, 2.47			
Arg	4.57	1.77, 1.61	1.39	3.13	
Trp	5.29	3.22			
Thr	4.38	3.78	0.78		
Glu	4.64	1.77	1.90		
His	4.89	3.13			
Lys	5.15	1.54	1.18		2.82
Val	4.17	1.96	085, 0.69		
Asn	4.42	3.09, 2.78			
Gly	4.16, 3.55				
Arg	4.63	1.77	1.57	3.20	
Thr	4.77	3.87	1.19		
Tyr	4.73	2.87, 2.66			
Phe	5.02	2.98, 2.93			
Trp	5.52	2.86			
Leu	3.95	1.31	0.77	0.31, 0.23	
Lys	4.30	1.70	1.26, 1.15	1.56	2.79
Cys	5.17	3.01, 2.38			

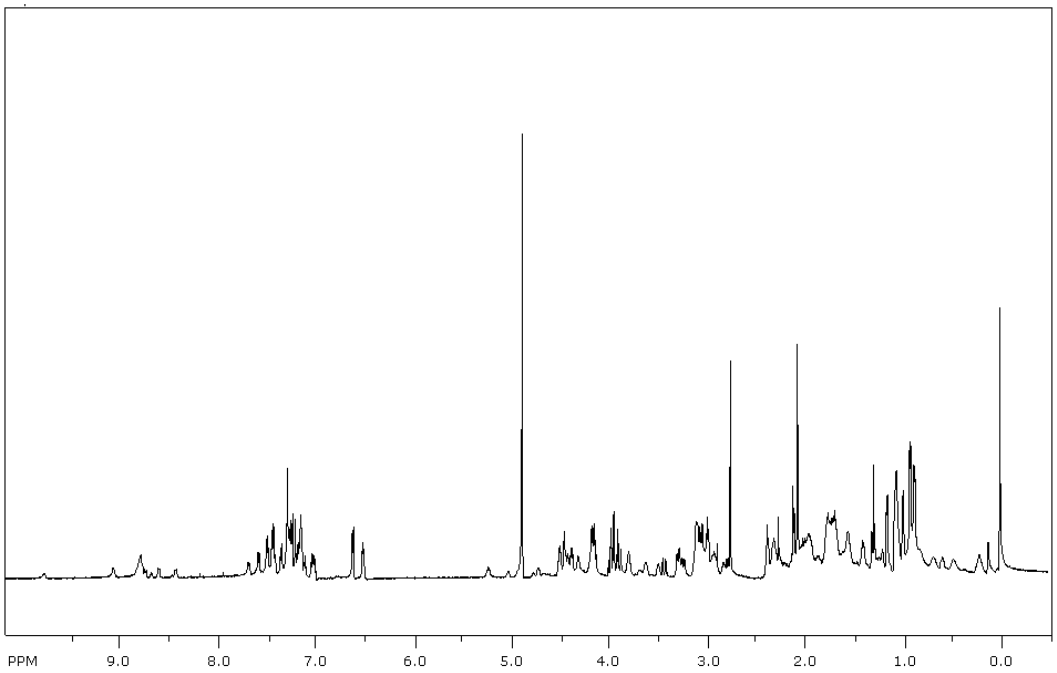


Figure 6.18. ^{1}HNMR spectrum of Mut1A-S23.

Table 6.7. Proton Chemical Shift Assignments for Peptide Mut1A-S23 (in ppm).

Residue	α	β	γ	δ	ε
Gly	3.9				
Arg	4.61	1.90	1.74	3.08	
Thr	4.34	4.13	1.06		
Tyr	4.54	3.24			
Phe	4.67	3.04			
Trp	4.82	3.21	3.04		
Leu	4.82	1.72	1.54	0.98	
Lys	4.08	1.15, -0.02	0.42	0.19	2.24
Val	4.04	1.84	0.88		
Asn	4.27	3.04, 2.73			
Gly	0.72				
Orn	4.82	2.30	1.96	3.20	
Trp	5.13	3.23, 3.04			
Gln	4.92	1.92	2.10		
Lys	3.69	0.82, -0.02		0.17	~2.1[a]
Thr	4.34	3.47	1.06		
Trp	4.55	3.23			
Glu	4.82	1.96	2.30		
Lys	4.39	2.09, 1.93	1.53	1.71	2.33
Pro	4.31	2.29, 1.96	2.02, 1.97	3.69, 3.51	
Gly	3.86				

[a] Highlighted values are those that could not be determined unambiguously.

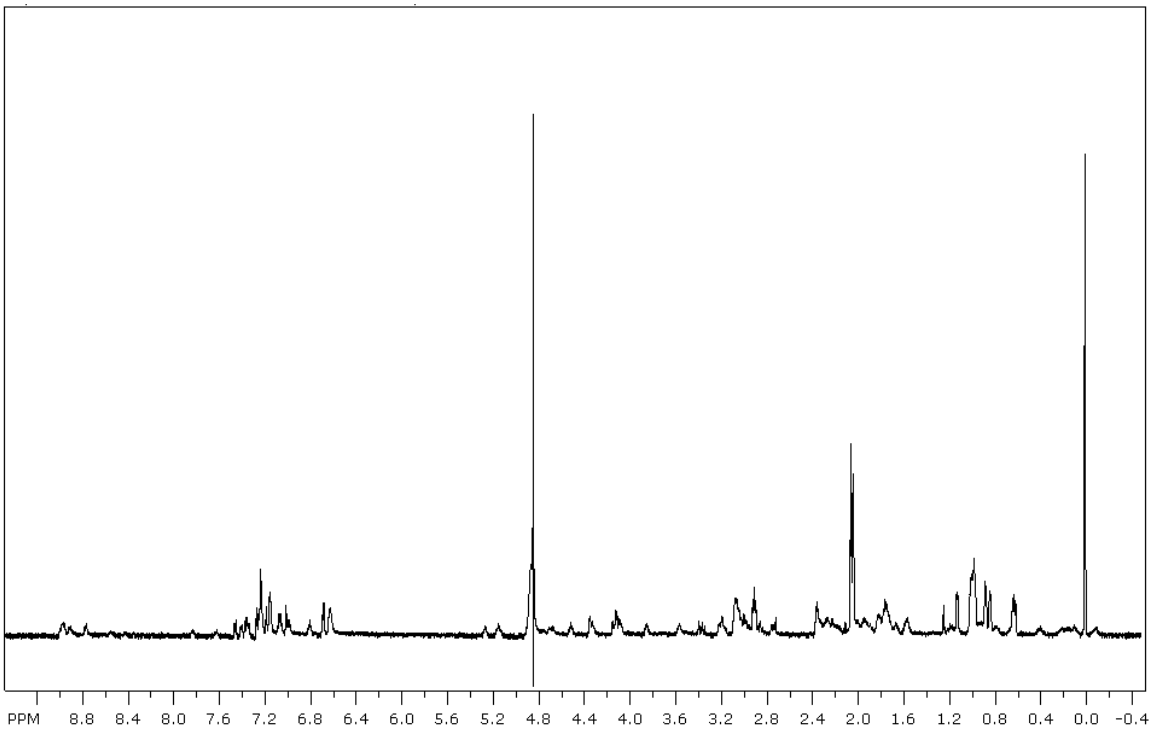


Figure 6.19. ^{1}HNMR spectrum of Mut1A-S23 Cyclic.

Table 6.8. Proton Chemical Shift Assignments for Peptide Mut1A-S23 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.82	3.17, 2.93			
Thr	4.36	4.10	0.98		
Tyr	4.76	3.00			
Phe	4.90	3.06			
Trp	4.96	3.03, 2.86			
Leu	4.92	1.58		0.98	
Lys	4.13	0.84, -0.096	0.091	0.37	2.23
Val	4.11	1.81	0.84		
Asn	4.32	3.03, 2.74			
Gly	4.13, 3.37				
Orn	4.72	1.90	1.74	3.07	
Trp	5.16	3.01			
Gln	5.28	1.81	2.24		
Lys	3.86	0.88	0.16	0.65	1.89
Thr	4.52	3.55	1.11		
Trp	4.82	3.17			
Glu	4.90	1.97	2.30, 2.23		
Cys	5.16	3.19, 2.33			

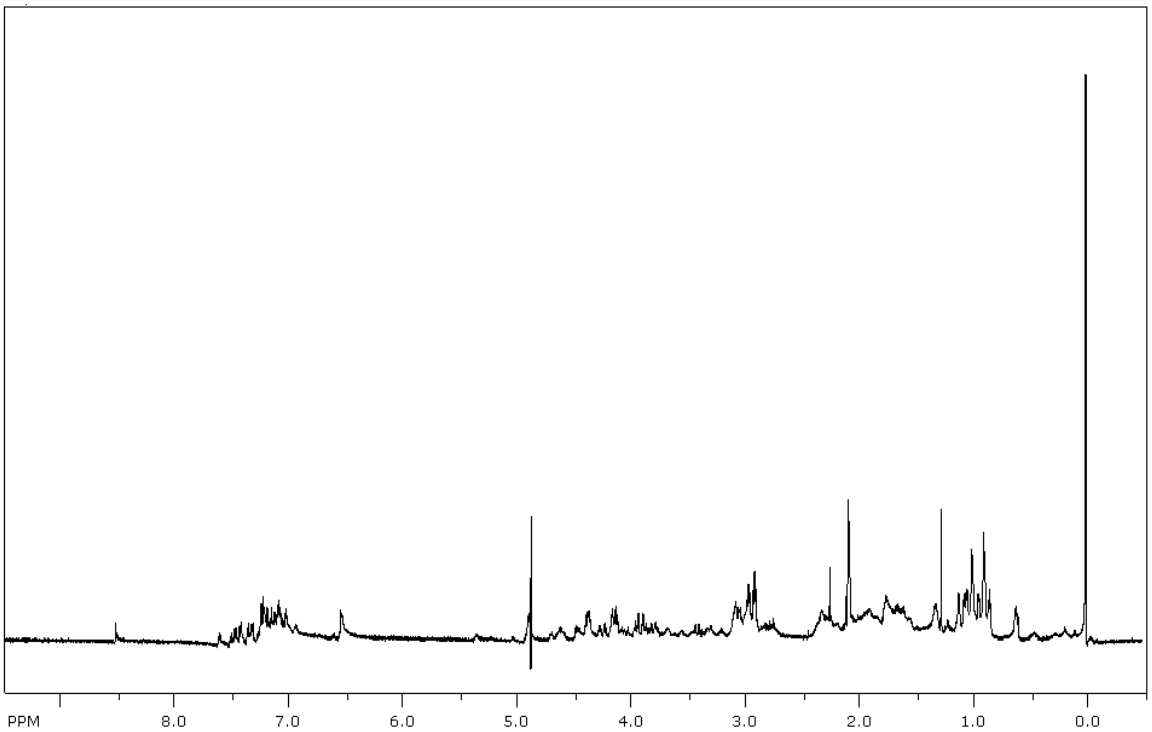


Figure 6.20. ^{1}HNMR spectrum of Mut1A.

Table 6.9. Proton Chemical Shift Assignments for Peptide Mut1A (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.29	3.79			
Arg	4.49	1.80	1.59	2.95	
Trp	4.61	2.78			
Thr	4.28	4.13	1.10		
Glu	4.79	2.18, 1.98	2.32		
His	4.60	3.07			
Lys	4.17	1.52	1.3		2.93
Val	4.02	1.83	0.86		
Asn	4.51	2.96, 2.78			
Gly	3.78, 3.51				
Arg[a]	4.79			3.04	
Thr	4.51	4.06	1.01		
Tyr	4.51	2.69			
Phe	4.67	3.02			
Trp	5.14	3.00			
Leu	4.81	1.60	1.40	0.90	
Lys	4.07	2.23, -0.048	0.18	1.63, 0.42	2.94
Val	4.13	2.10	0.97		
Asn	4.27	3.03, 2.74			
Gly	4.13, 3.40				
Orn	4.60	1.72	1.30	3.07	
Trp	5.28	3.28, 3.03			
Gln	4.94	2.15	1.94		
Lys	3.87	1.85, -0.015	0.25	0.87, 0.64	2.06
Thr	4.39	3.54	1.05		
Trp	4.81	3.28, 3.00			
Glu	4.35	2.06, 1.86	2.32		
Lys[a]	~4.3				
Pro	4.39	2.25	1.95	3.57,3.36	
Gly	3.87				

[a] Highlighted values are those that could not be determined unambiguously.

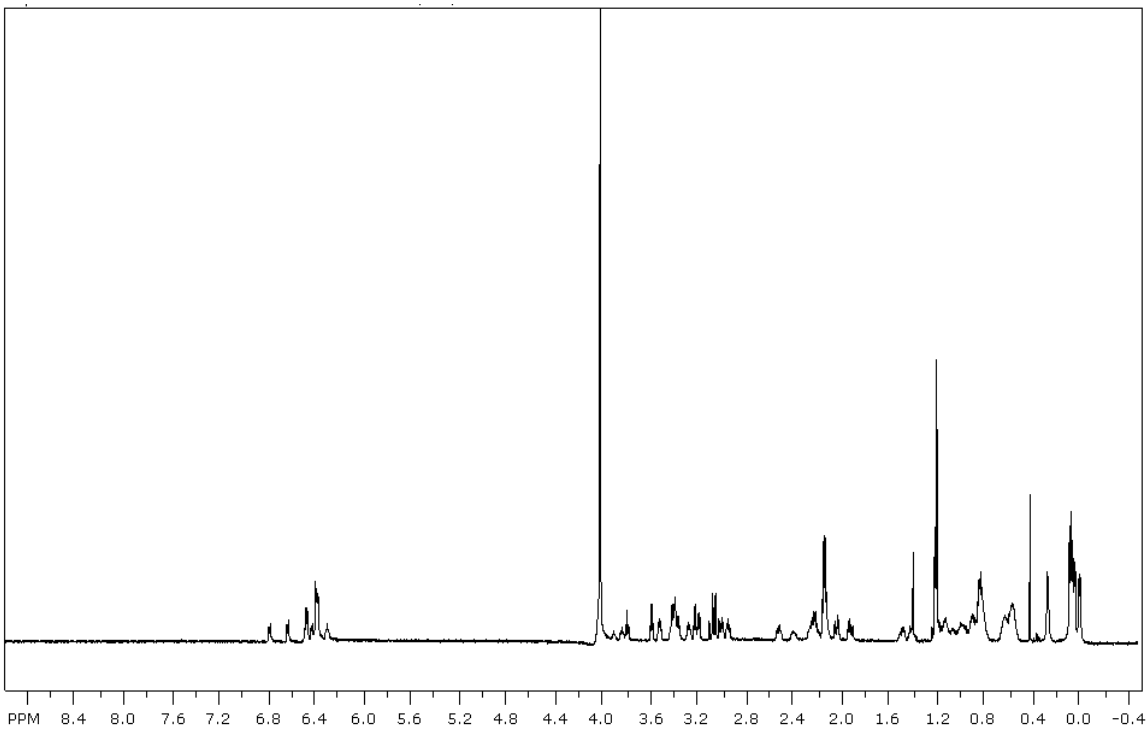


Figure 6.21. ^{1}HNMR spectrum of Mut1B-S1-KK. This peptide was used in place of Mut1B-S1 due to solubility problems.

Table 6.10. Proton Chemical Shift Assignments for Peptide Mut1B-S1-KK (in ppm).

Residue	α	β	γ	δ	ε
Lys	4.15		1.41	1.69	3.01
Lys	4.17		1.40	1.68	3.06
Ser	4.33	3.81			
Arg	4.27	1.70	1.44	2.96	
Trp	4.65	3.34, 3.23			
Thr	4.13	4.24	1.10		
Glu	4.11	1.90	2.28		
Val	3.93	1.96	0.86		
Phe	4.53	2.85, 2.76			
Val	3.97	2.01	0.91		
Asn	4.58	3.08, 2.95			
Gly	3.84				

Highlighted values are those that could not be determined unambiguously.

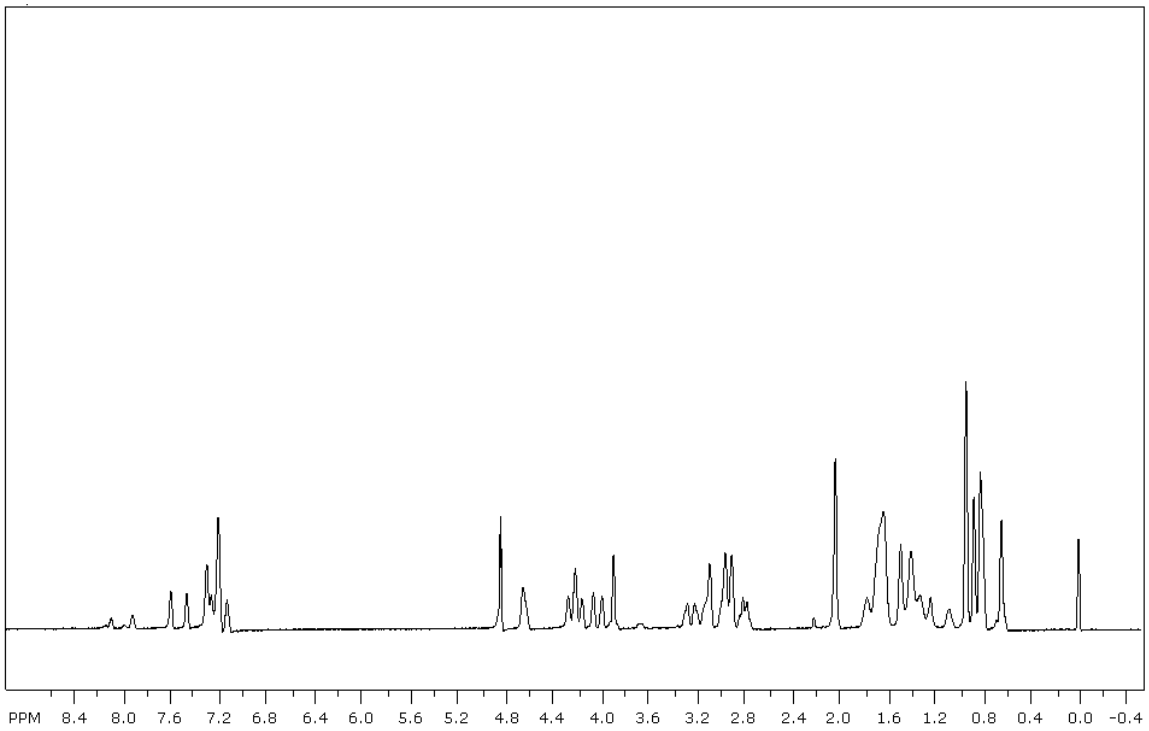


Figure 6.22. ^{1}HNMR spectrum of Mut1B-S2.

Table 6.11. Proton Chemical Shift Assignments for Peptide Mut1B-S2 (in ppm).

Residue	α	β	γ	δ	ε
Gly	3.93				
Arg	4.12	1.67	1.42	3.12	
Phe	4.56	2.81			
Ile	3.91	1.71	1.41, 1.10; (γMe=0.82)	0.67	
Lys	4.06	1.66	1.30		2.93
Trp	4.56	3.23			
Leu	4.18	1.51		0.87	
Lys	4.13	1.69	1.42		2.99
Val	3.97	2.06	0.96		
Asn	4.53	3.19, 3.01			
Gly	3.70				

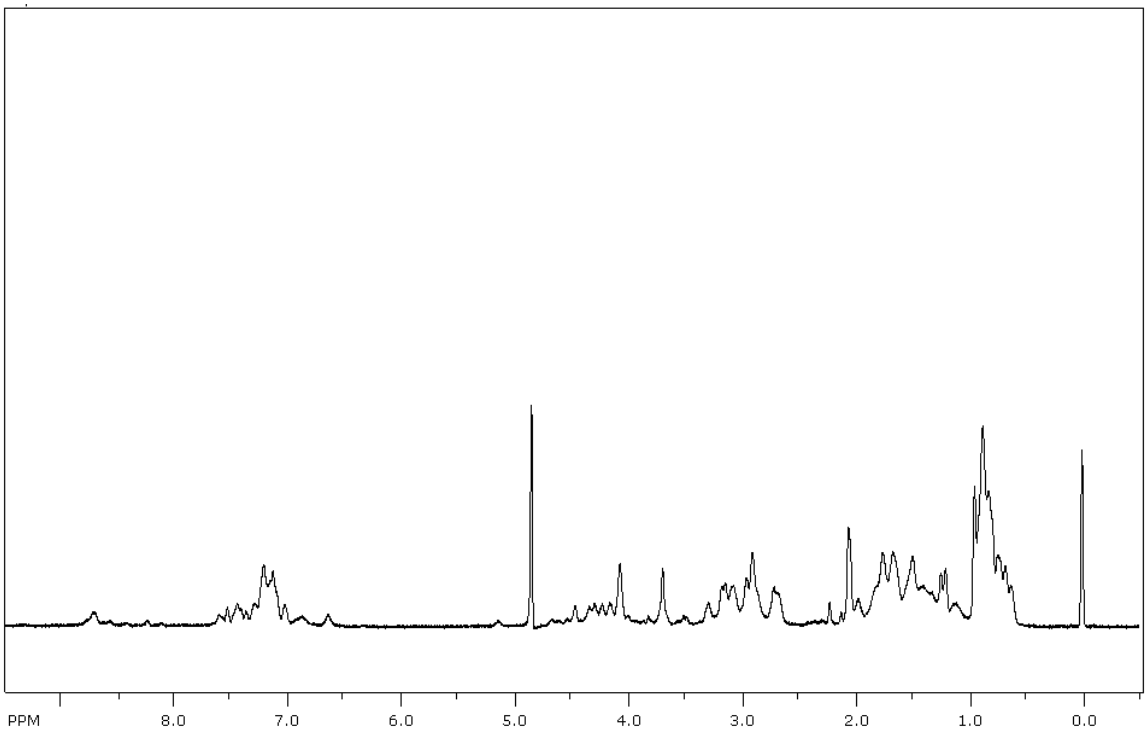


Figure 6.23. [1]HNMR spectrum of Mut1B-S12.

Table 6.12. Proton Chemical Shift Assignments for Peptide Mut1B-S12 (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.36	3.71			
Arg	4.38	1.76, 1.66	1.53	3.16	
Trp	4.58	3.22, 2.94			
Thr	4.71	4.08	1.28		
Glu	4.25	2.70[a]			
Val	4.05	1.83	0.87		
Phe	4.67	2.89			
Val	4.37	1.66	0.68		
Asn	4.45	2.91, 2.75			
Gly	4.08, 3.51				
Arg	4.51	1.76	1.56	3.18	
Phe	4.88	2.72			
Ile	3.97	2.01	1.49, (γMe=0.92)	0.68	
Lys	4.13	1.69	0.73	1.39	2.97
Trp	5.05	3.31, 3.08			
Leu	4.20		1.49	0.82	
Lys	4.07	1.65		1.25	2.92
Val	3.90	1.68	0.64		

[a] Highlighted values are those that could not be determined unambiguously.

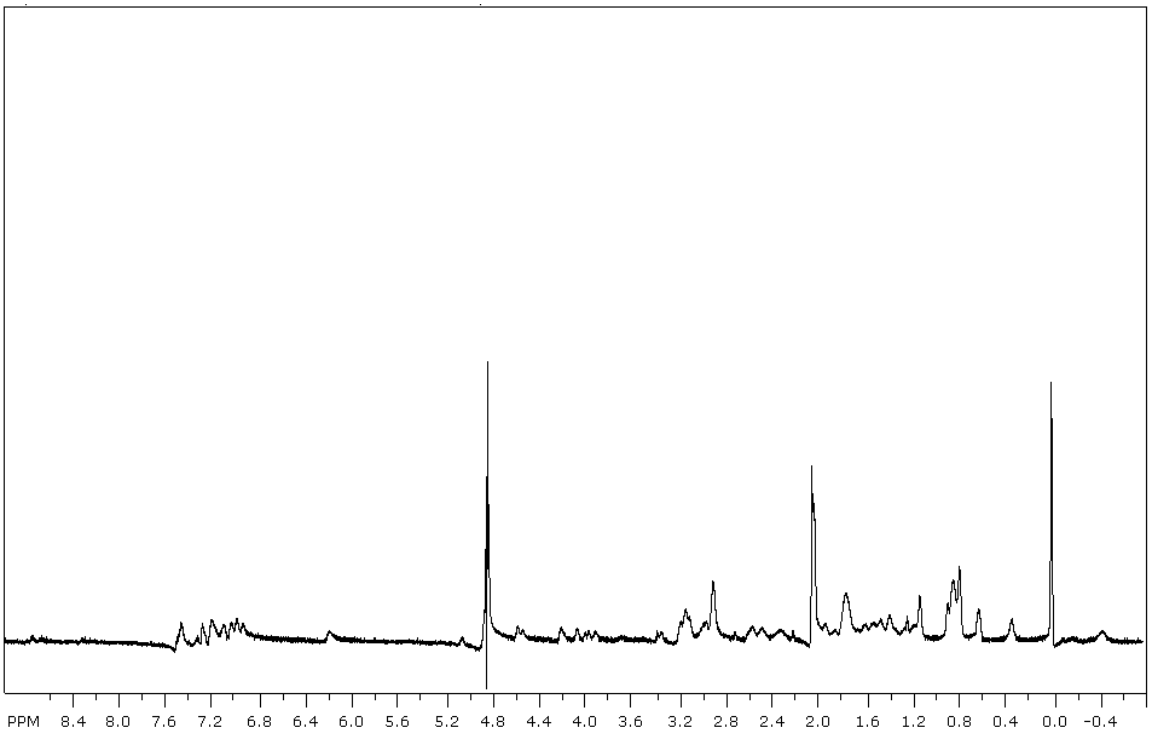


Figure 6.24. ^{1}HNMR spectrum of Mut1B-S12 Cyclic.

Table 6.13. Proton Chemical Shift Assignments for Peptide Mut1B-S12 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.73	2.87, 2.47			
Arg	4.55	1.73	1.50	3.14	
Trp	4.81	3.08			
Thr	4.77	3.92	1.12		
Glu	5.54	1.95	2.32		
Val	4.10	1.76	0.78		
Phe	4.73	2.87			
Val	4.66	1.99	0.84		
Asn	4.23	2.97, 2.59			
Gly	3.98, 3.36				
Arg	4.60	1.95, 1.74		3.18	
Phe	5.08	3.04			
Ile	4.60	1.95	1.11	0.88	
Lys	3.70	1.17	-0.16	0.32	
Trp	5.00	2.98, 2.91			
Leu	5.25	1.81		1.39	
Lys	4.21	1.76	1.25	1.61	2.40
Cys	5.17	2.95, 2.32			

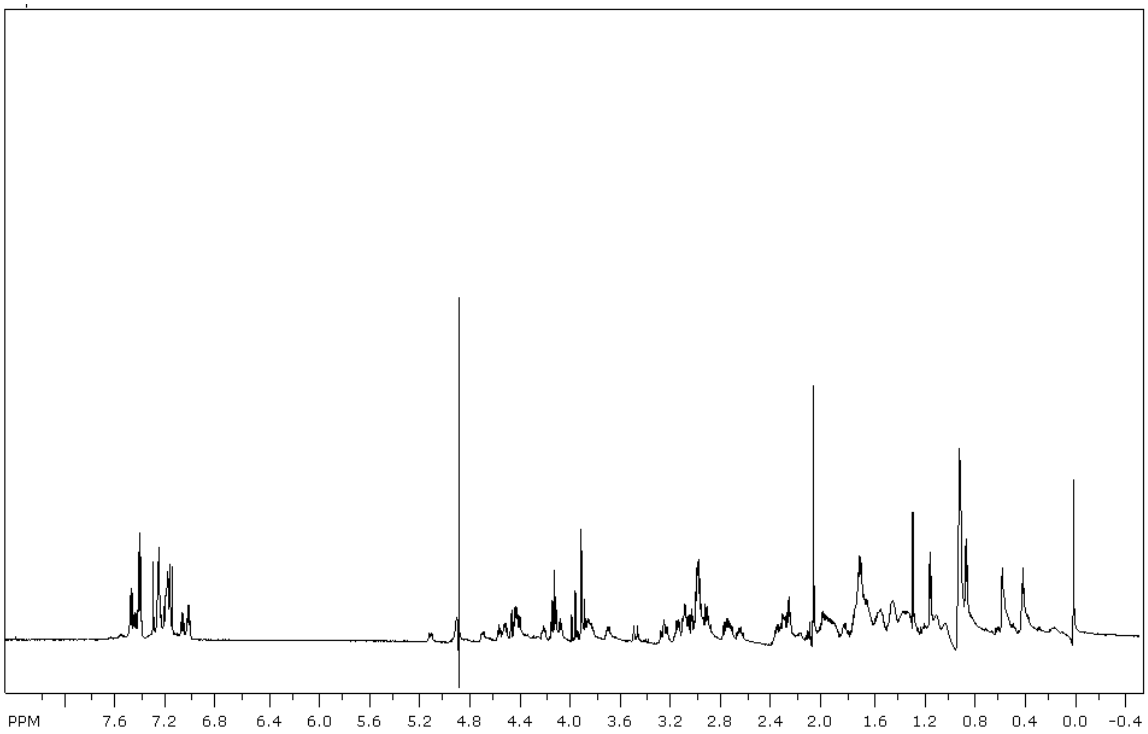


Figure 6.25. ^{1}HNMR spectrum of Mut1B-S23.

Table 6.14. Proton Chemical Shift Assignments for Peptide Mut1B-S23 (in ppm).

Residue	α	β	γ	δ	ε
Gly	3.88				
Arg	4.49	1.64	1.34	2.91	
Phe	4.63	3.21			
Ile	4.15		1.59, 1.15		
Lys	4.19	1.25, 0.13	0.49, 0.33	1.07	2.29
Trp	4.91	3.20, 3.07			
Leu	4.79	1.51		0.86	
Lys	4.09	1.49	0.35	0.53	
Val	4.11	1.84	0.84		
Asn	4.39	2.97, 2.73			
Gly	4.09, 3.44				
Orn	4.67	1.88	1.72	3.05	
Trp	5.04	3.20, 3.00			
Gln	4.80	2.30, 2.15			
Lys	3.85	1.30	0.52, 0.36	0.98	2.65
Thr	4.44	4.03	1.10		
Trp	4.81	3.10, 2.95			
Glu	4.24	1.76	2.16		
Lys	4.54	1.66	1.40		2.95
Pro	4.42	2.28, 1.95	1.85, 2.23		3.82, 3.68
Gly	3.94				

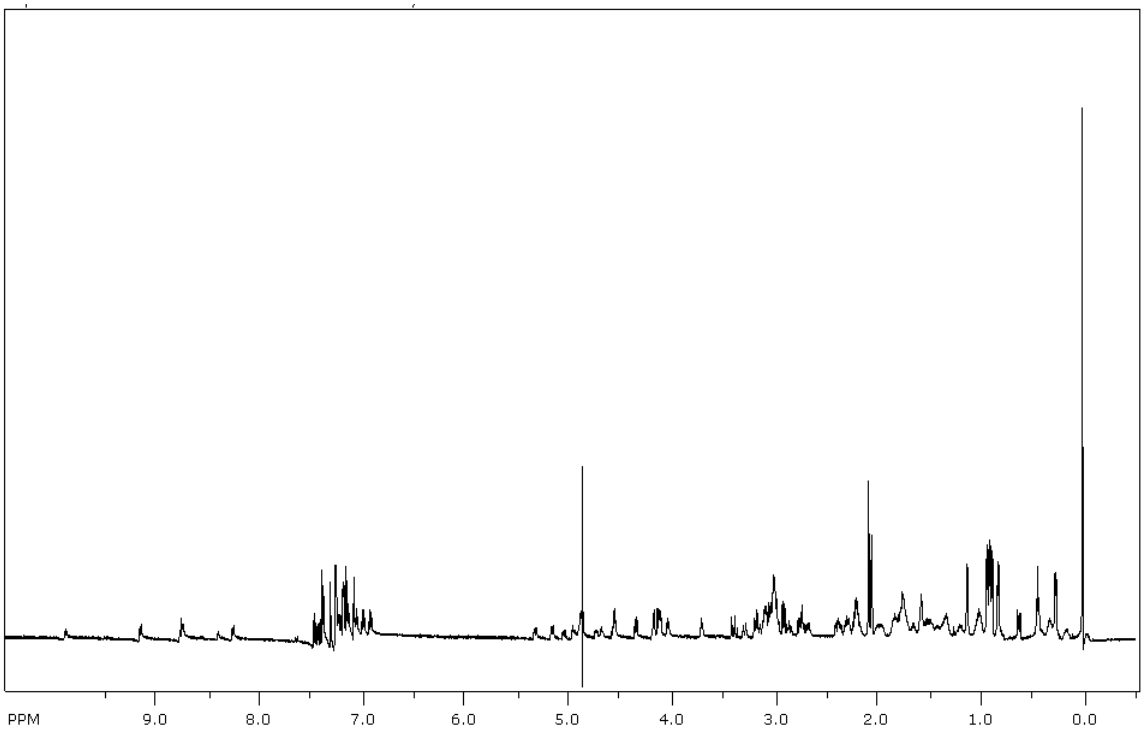


Figure 6.26. ^{1}HNMR spectrum of Mut1B-S23 Cyclic.

Table 6.15. Proton Chemical Shift Assignments for Peptide Mut1B-S23 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.97	3.16, 3.03			
Phe	4.92	3.00			
Ile	4.12	1.45	0.26	0.45, 0.85	
Lys	4.69	1.73	1.38	1.57	3.01
Trp	5.05	3.02, 2.87			
Leu	4.92	1.56		0.90	
Lys	4.19	1.19, -0.023	0.39, 0.16	1.00	
Val	4.13	1.83	0.85		
Asn	4.36	3.07, 2.76			
Gly	4.15, 3.40				
Orn	4.74	1.92	1.77	3.09	
Trp	5.17	3.19, 3.00			
Gln	4.90	2.19, 2.00	2.35		
Lys	3.72	1.31	0.32	0.85	2.70
Thr	4.56	4.04	1.12		
Trp	5.05	3.02, 2.87			
Glu	4.58	2.21	1.64, 1.50		
Cys	5.33	2.98, 2.39			

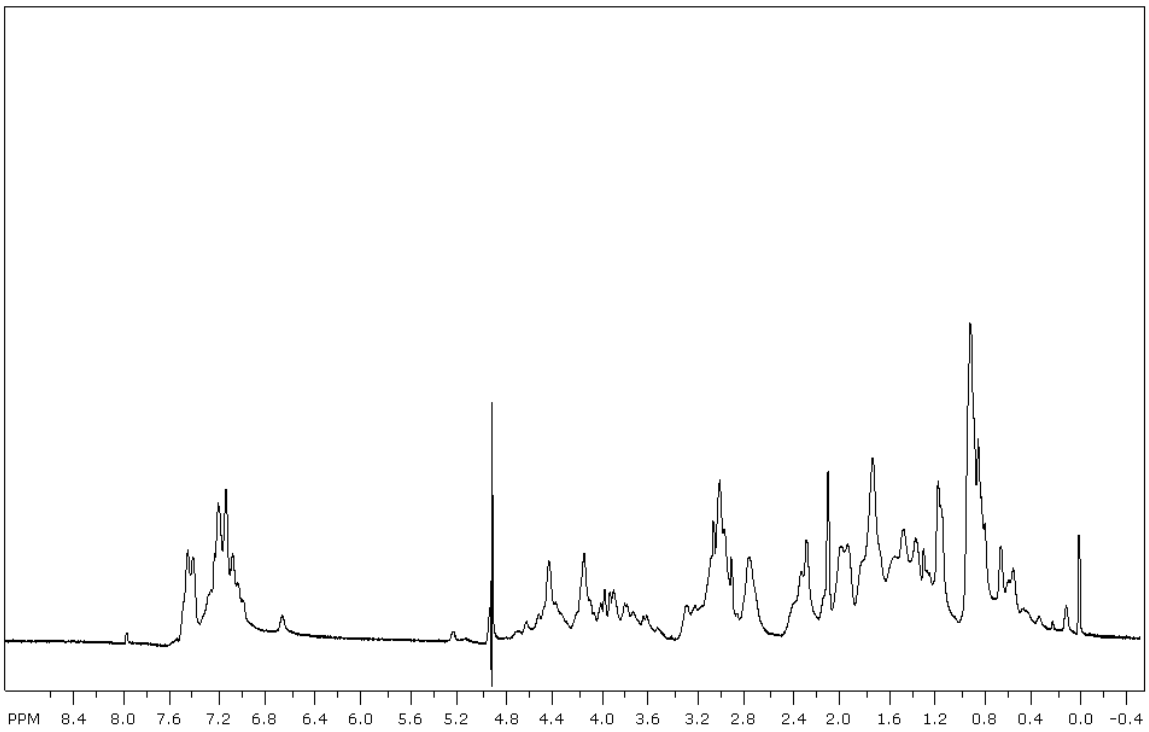


Figure 6.27. ^{1}HNMR spectrum of Mut1B.

Table 6.16. Proton Chemical Shift Assignments for Peptide Mut1B (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.47	3.79			
Arg	4.53	1.70	1.46	2.98[a]	
Trp	5.02	3.27, 3.15			
Thr	4.46	4.08	1.14		
Glu	4.66	2.20	2.29		
Val	4.16	1.93	0.88		
Phe	4.70	3.25, 3.10			
Val	4.34	1.62	0.55		
Asn	4.54	2.96, 2.78			
Gly	0.45				4.06, 3.61
Arg	4.61	1.75	1.52	3.06	
Phe	4.87	2.94, 2.72			
Ile	4.31	1.83	1.61		
Lys	4.34	1.62	0.55		
Trp	5.11	3.06			
Leu	4.77		1.38	0.78	
Lys	4.21	1.35	0.62	1.20	2.90[a]
Val	4.38	1.95	0.83		
Asn	4.43	3.01, 2.72			
Gly	0.61				4.13, 3.52
Orn[a]	4.70				3.20
Trp	5.22	3.28, 3.07			
Gln	4.80	1.96	2.31		
Lys	4.04	1.44	0.80	1.25	
Thr	4.56	4.13	1.12		
Trp	4.92	3.16			
Glu[a]	4.40	2.00	2.25		
Lys	4.36	1.60	1.37		2.97
Pro	4.44	2.28, 2.01	2.12, 1.99	3.87, 3.72	
Gly	3.9				

[a] Highlighted values are those that could not be determined unambiguously.

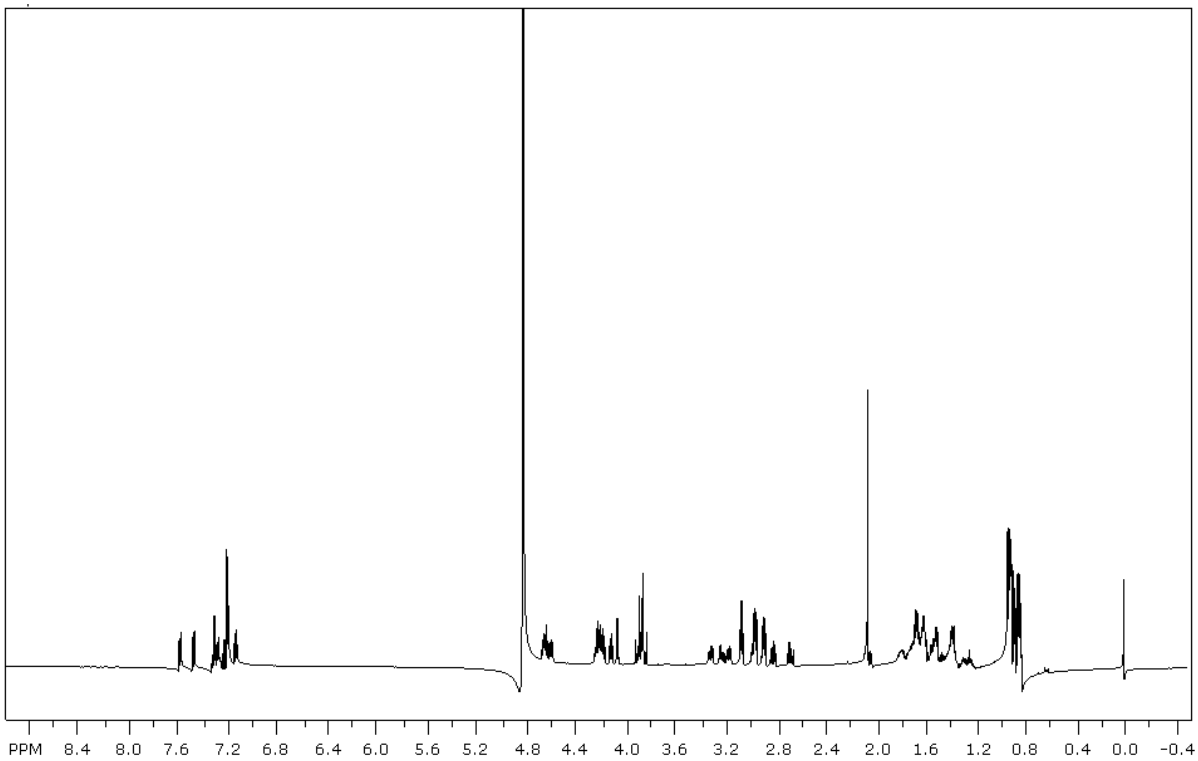


Figure 6.28. ^{1}HNMR of Mut1C-S2.

Table 6.17. Proton Chemical Shift Assignments for Peptide Mut1C-S2 (in ppm).

Residue	α	β	γ	δ	ε
Gly	3.88				
Arg	4.19		1.56	3.08	
Phe	4.60	3.17, 2.99			
Leu	4.22		1.56	0.89	
Lys	4.14		1.25	1.64	2.90
Trp	4.67	3.28			
Leu	4.26	1.51		0.86	
Lys	4.24			1.71	2.98
Val	4.09	2.05	0.92		
Asn	4.69	2.82, 2.69			
Gly	3.67				

Mut1C-S1 is the same sequence as Mut1B-S1.

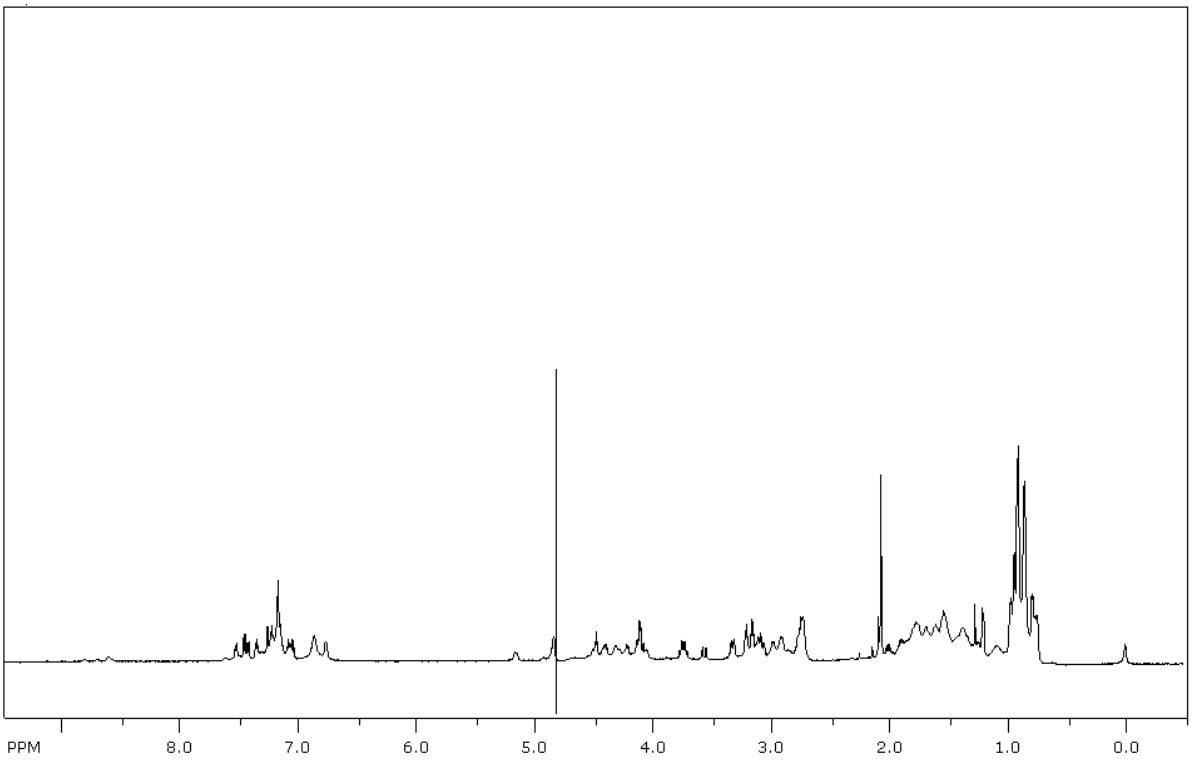


Figure 6.29. ^{1}HNMR of Mut1C-S12.

Table 6.18. Proton Chemical Shift Assignments for Peptide Mut1C-S12 (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.47	3.73			
Arg	4.28	1.70	1.48		
Trp	4.70	3.27			
Thr	4.78	4.05			
Glu	4.66	1.79, 1.61			
Val	4.21	1.89	0.89		
Phe	4.65	3.21, 2.97			
Val	4.31	1.86			
Asn	4.39	3.06, 2.75			
Gly	4.10, 3.57				
Arg	4.50	1.72	1.55	3.15	
Phe	4.91	2.76			
Leu	4.54	1.26			
Lys	4.13	1.48, 1.43	1.10		2.72
Trp	5.15	3.30, 3.10			
Leu	4.42	1.34		0.80	
Lys	4.47	1.72	1.55	1.07	2.72
Val	4.11	2.00	0.91		

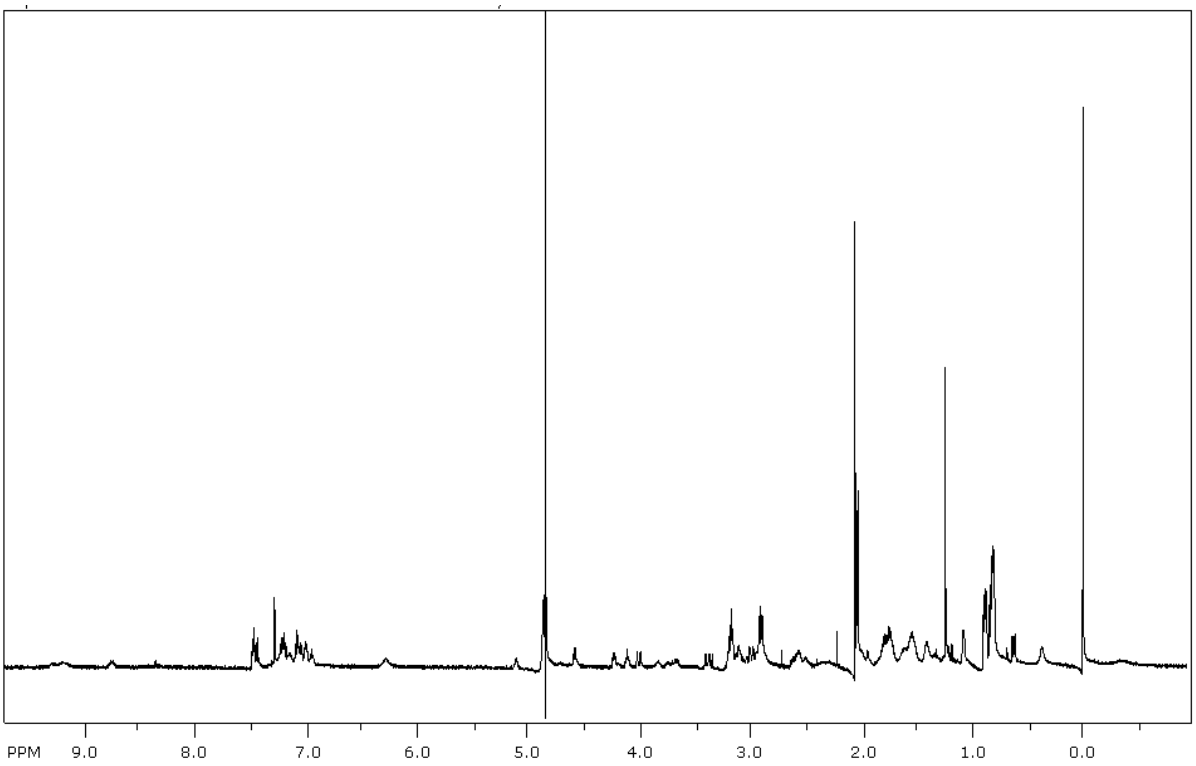


Figure 6.30. [1]HNMR of Mut1C-S12 Cyclic.

Table 6.19. Proton Chemical Shift Assignments for Peptide Mut1C-S12 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.82	2.55			
Arg	4.21	1.69	1.25	2.90	
Trp	4.82	3.08			
Thr	4.73	3.84	1.07		
Glu	5.45	2.27, 1.94			
Val	4.12	1.78	0.80		
Phe	4.99	2.93			
Val	4.64	1.98			
Asn	4.24	2.92, 2.60			
Gly	4.00, 3.38				
Arg	4.59	1.73	1.55	3.18	
Phe	5.13	2.98			
Leu	5.25	1.79		1.40	
Lys[a]	3.76	1.77	1.24		
Trp	5.13	3.14			
Leu	4.70	1.57		0.84	
Lys[a]	4.21				
Cys	5.13	2.98, 2.37			

[a] Highlighted values are those that could not be determined unambiguously.

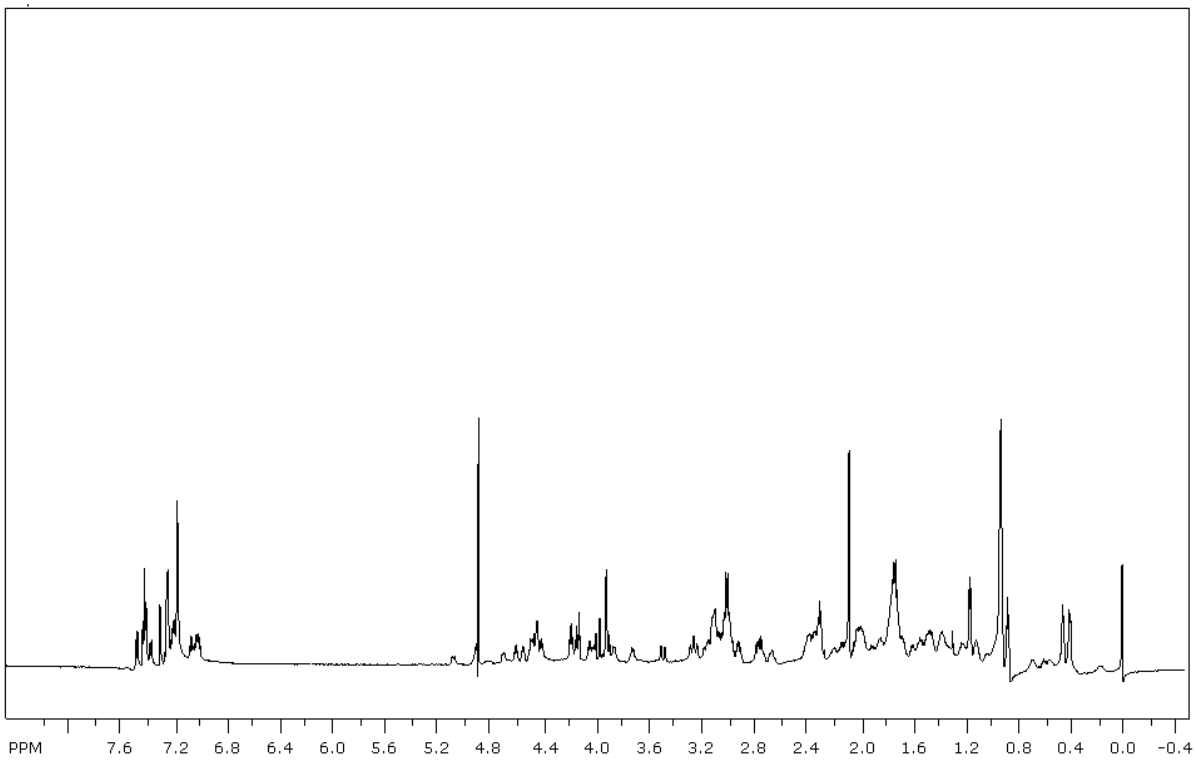


Figure 6.31. ^{1}HNMR of Mut1C-S23.

Table 6.20. Proton Chemical Shift Assignments for Peptide Mut1C-S23 (in ppm).

Residue	α	β	γ	δ	ε
Gly	3.86				
Arg	4.51	1.62	1.39	2.89	
Phe	4.85	3.09			
Leu	4.34	1.62	1.34		
Lys	4.45	1.67	1.30		2.92
Trp	4.89	3.12, 3.04			
Leu	4.76	1.50		0.86	
Lys	4.15	1.17, 0.10	0.44, 0.33	1.05, 0.83	2.30
Val	4.10	1.84	0.83		
Asn	4.39	2.98, 2.71			
Gly	4.09, 3.44				
Orn	4.66	1.88	1.72	3.04	
Trp	5.04	3.20, 3.00			
Gln	4.79	2.15, 1.94	2.31		
Lys	3.97	1.27, 0.35	0.62	1.10, 0.91	2.65
Thr	4.41	4.01	1.08		
Trp	4.79	3.10, 2.94			
Glu	4.43	1.94	2.25		
Lys	4.56	1.72	1.43	1.66	2.96
Pro	4.42	2.28, 1.96			
Gly	3.90				

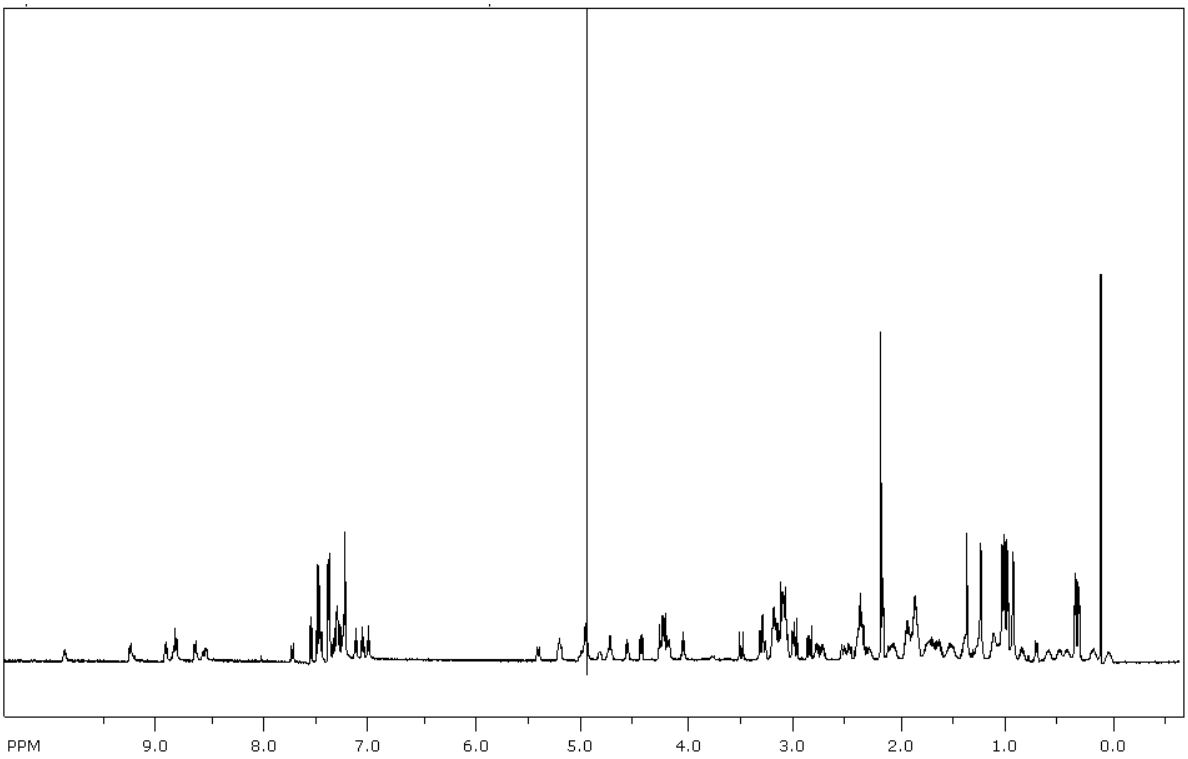


Figure 6.32. ^{1}NMR of Mut1C-S23 Cyclic.

Table 6.21. Proton Chemical Shift Assignments for Peptide Mut1C-S23 Cyclic (in ppm).

Residue	α	β	γ	δ	ε
Cys	4.94	3.17, 3.01			
Phe	4.94	3.01			
Leu	4.67	1.80	1.43		
Lys	4.68	1.81	1.44	1.64	2.99
Trp	5.12	3.20, 2.90			
Leu	4.90	1.57		0.93	
Lys	4.17	1.16, -0.057	0.39, 0.18	0.99	2.28
Val	4.15	1.92	0.86		
Asn	4.37	3.07, 2.77			
Gly	4.17, 3.40				
Orn	4.76	1.87	1.78	3.11	
Trp	5.15	3.21, 2.94			
Gln	4.92	2.20, 1.99	2.36		
Lys	4.13	1.26, 0.33	0.52	0.87	2.66
Thr	4.51	3.96	1.13		
Trp	4.95	3.14			
Glu	4.66	2.26			
Cys	5.35	3.00, 2.44			

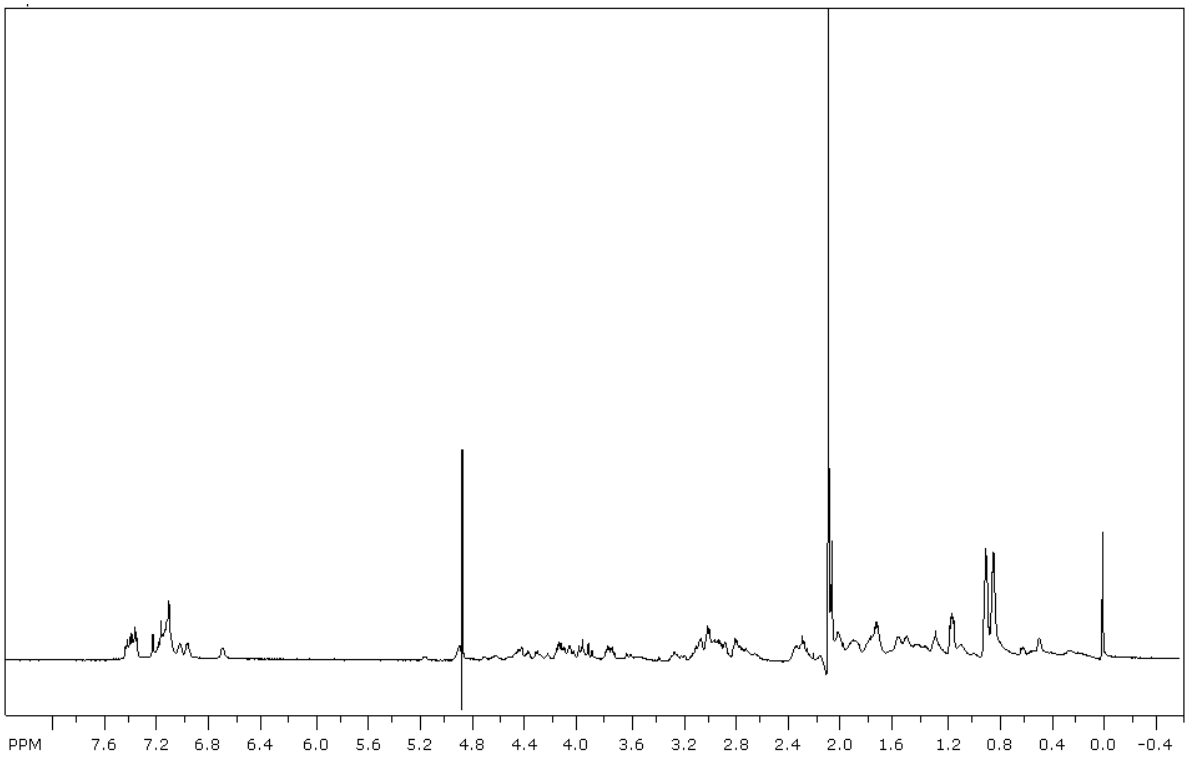


Figure 6.33. ^{1}HNMR of Mut1C.

Table 6.22. Proton Chemical Shift Assignments for Peptide Mut1C (in ppm).

Residue	α	β	γ	δ	ε
Ser	4.42	3.70			
Arg	4.60	1.84	1.48	3.01	
Trp	4.91	2.87, 2.73			
Thr	4.40	4.02	1.09		
Glu	4.69	1.73	2.26		
Val	4.10	1.84	0.81		
Phe	4.69	2.26			
Val	4.29	1.95	0.85		
Asn	4.43	2.94, 2.70			
Gly	0.43				
Arg	4.68	1.73	1.49	2.95	
Phe	4.80	3.06			
Leu	4.62	1.69			
Lys	4.21	1.29, 0.44	1.01, 0.21		
Trp	4.97	3.06			
Leu	4.77	1.51	1.37	0.78	
Lys	3.96	1.34	0.98, 0.74	1.16	2.71
Val	4.62	1.88	1.22		
Asn	4.35	2.90, 2.61			
Gly	0.57				
Orn	4.60	1.69	1.48	3.00	
Trp	5.14	3.20, 2.95			
Gln[a]	4.77				
Lys	4.14	0.49	0.12		2.29
Thr	4.45	4.10	1.11		
Trp	4.84	3.26, 3.08			
Glu	4.49	1.74	2.24		
Lys[a]	4.26	1.48	1.24		
Pro	4.45	2.30, 2.01	2.11	3.94, 3.77	
Gly	3.94				

[a] Highlighted values are those that could not be determined unambiguously.

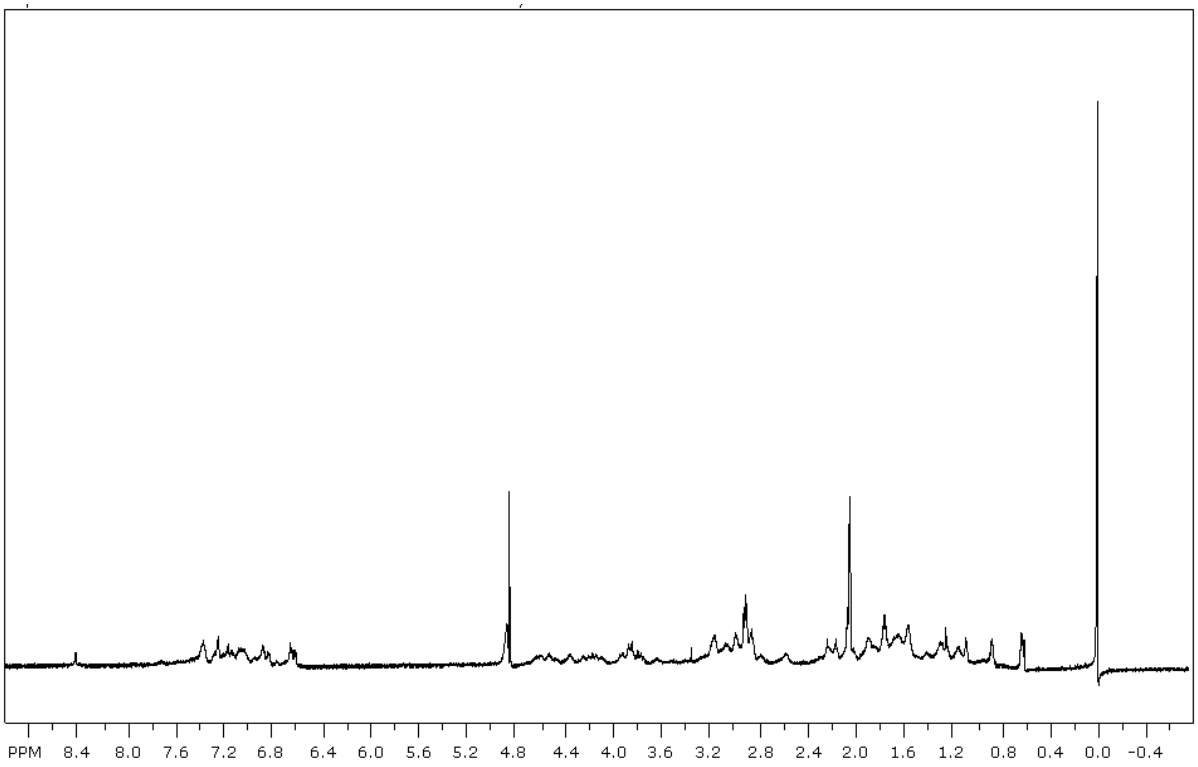


Figure 6.34. ^{1}HNMR of Mut1D.

Chapter VII

Recognition of Modified, Damaged, and Mismatched DNA Bases by β-Sheet Peptides

A. Background and Significance

OB-fold domains are thought of primarily as single-stranded DNA binding domains that recognize single strands of DNA such as those found at the replication fork[1] and in telomeric DNA.[2] While that is a primary function of OB-fold domains, the depth of their cellular roles is much greater. In fact, OB-fold domains are single-stranded DNA binding motifs that emerge in various processes in which single-stranded DNA or single, unpaired nucleotide bases are present and become exposed.[3] This occurs at the mRNA cap,[4] in

---

[1] (a) Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004**, *14*, 36-42. (b) Mer, G.; Bochkarev, A.; Gupta, R.; Bochkareva, E.; Frappier, L.; Ingles, J. C.; Edwards, A. M.; Chazin, W. J. *Cell* **2000**, *103*, 449-456. (c) Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. *Nature* **1997**, *385*, 176-181. (d) Wold, M. S. *Annu. Rev. Biochem.* **1997**, *66*, 61-92.

[2] (a) Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. *Biochemistry* **2003**, *42*, 3751-3758. (b) Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. *Science* **2002**, *296*, 145-147. (c) Croy, J. E.; Wuttke, D. S. *Trends Biochem. Sci.* **2006**, *31*, 516-525.

[3] Pestryakov, P. E.; Lavrik, O. I. *Biochemistry (Moscow)* **2008**, *73*, 1388-1404.

[4] (a) Guilligay, D.; Tarendeau, F.; Resa-Infante, P.; Coloma, R.; Crepin, T.; Sehr, P.; Lewis, J.; Ruigrok, R. W. H.; Ortin, J.; Hart, D. J.; Cusack, S. *Nat. Struct. Mol. Biol.* **2008**, *15*, 500-506. (b) Rutkowska-Wlodarczyk, I.; Stepinski, J.; Dadlez, M.; Darzynkiewicz, E.; Stolarski, R.; Niedzwiecka, A. *Biochemistry* **2008**, *47*, 2710–2720. (c) Richter, J. D.; Sonenberg, N. *Nature* **2005**, *433*, 477-479. (d) Quiocho, F. A.; Hu, G. H.; Gershon, P. D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 78-86. (e) Sonenberg, N.; Gingras, A.-G. *Curr. Opin. Cell Biol.* **1998**, *10*, 268-275.

telomeric DNA,[2] during recombination,[5] and during DNA replication and repair.[1] Single-stranded DNA binding proteins, many of which have OB-fold domains, are often involved in recognizing damaged DNA.[6, 7] The primary eukaryotic ssDNA binding protein Replication Protein A[1] has been shown to bind damaged and mismatched DNA.[1d] It is thought that these proteins use the same mechanism of recognition as they do to bind single-stranded DNA. In this way, the peptide or protein recognizes the damaged or mismatched base as it would a single-stranded nucleotide.[1d] DNA methylation, for example, is one such DNA modification often found in nature. Methylation is known to occur at the mRNA cap[4] as well as in other DNA sequences such as the promoter regions of certain genes.[8] As shown in Figure 8.1, the PB2 cap binding domain recognizes the m^7GTP cap analog via a β-sheet with specific aromatic stacking interactions stabilizing the complex (Figure 7.1b).[4a-4b]

---

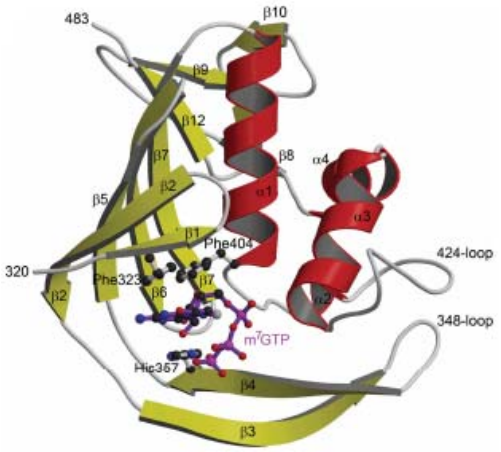
[5] (a) Min, J.; Park, P.; Ko, E.; Choi, E.; Lee, H. *Biochem. Biophys. Res. Commun.* **2007**, *362*, 958-964. (b) Nair, P. A.; Nandakumar, J.; Smith, P.; Odell, M.; Lima, C. D.; Shuman, S. *Nat. Struct. Mol. Biol.* **2007**, *14*, 770-778. (c) Sugiyama, T.; Kantake, N.; Wu, Y.; Kowalczykowski, S. C. *EMBO J.* **2006**, *25*, 5539-5548. (d) Moure, C. M.; Gimble, F. S.; Quiocho, F. A. *J. Mol. Biol.* **2003**, *334*, 685-695. (e) Stano, N. M.; Patel, S. S. *J. Mol. Biol.* **2002**, *315*, 1009-1025. (f) Cheetham, G. M. T.; Jeruzalmi, D.; Steitz, T. A. *Nature*, **1999**, *399*, 80-83. (g) Cheetham, G. M. T.; Steitz, T. A. *Science* **1999**, *286*, 2305-2309.

[6] (a) Malta, E.; Moolenaar, G. F.; Goosen, N. *J. Biol. Chem.* **2006**, *281*, 2184-2194. (b) Truglio, J. J.; Karakas, E.; Rhau, B.; Wang, H.; DellaVecchia, M. J.; Van Houten, B.; Kisker, C. *Nat. Struct. Mol. Biol.* **2006**, *13*, 360-364. (c) Skorvaga, M.; DellaVecchia, M. J.; Croteau, D. L.; Theis, K.; Truglio, J. J.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2004**, *279*, 51574-51580. (d) Skorvaga, M.; Theis, K.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2002**, *277*, 1553-1559.

[7] Min, J. H., Pavletich, N. P. *Nature* **2007**, *449*, 570-576.

[8] (a) Desmond, J. C.; Raynaud, S.; Tung, E.; Hofmann, W.-K.; Haferlach, T.; Koeffler, H. P. *Leukemia* **2007**, *21*, 1026-1034. (b) Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J. G.; Ikegami, T.; Nakao, M.; Shirakawa, M. *Cell* **2001**, *105*, 487-497. (c) Ballestar, E.; Wolffe, A. P. *Eur. J. Biochem.* **2001**, *268*, 1-6.

(a)



(b)

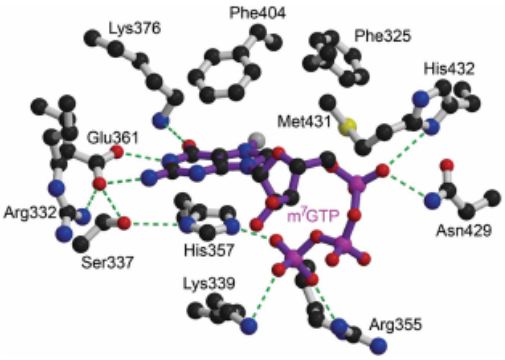


Figure 7.1. Structure of the PB2 cap binding domain bound to m^7GTP.[4a-4b] (a) Ribbon diagram of the full structure of the cap binding domain with the β-sheet portion shown in gold and the m^7GTP as well as the sidechains involved in stacking interactions shown as a ball-and-stick model. (b) Specific interactions between the PB2 cap binding domain and m^7GTP. Green dotted lines represent proposed hydrogen bonds.

There are several other examples of systems involving β-hairpin recognition of damaged DNA. The nucleotide excision repair pathway utilizes β-hairpins to recognize lesions in duplex DNA. In prokaryotes, the protein UvrB interacts with DNA by inserting a flexible β-hairpin between the two strands of the duplex DNA. This hairpin uses a cluster of conserved aromatic residues to form tight contacts with the DNA.[6] Similarly, eukaryotic protein Rad4 has been shown to insert a β-hairpin into double stranded DNA, causing

damaged bases to flip out and allowing the protein to recognize the unpaired bases.[7] In each case, regulation of the modification or repair by the appropriate repair pathways is crucial to maintain genomic stability and to prevent disease.[9] Thus, an important application of the current research is the relevance of the OB-fold properties in the designed β-sheet peptides to the study and recognition of damaged and mismatched DNA bases.

Though much of the work done in this laboratory has focused on designing peptides to bind short ssDNA sequences, this chapter delves into the application of peptides recognizing damaged or modified and mismatched nucleotide bases. In this study we endeavored to determine whether minimalistic OB-fold peptides shown to bind ssDNA with selectivity in previous work also bind damaged or mismatched bases and to understand more about the role of flipped out bases.

B. Results

i. Sequence Design

The WW domain Mutant 1 (Mut1; Figure 7.2) was applied to the selective recognition of damaged bases in ssDNA and mismatched bases in duplex DNA sequences. Because of the success with Mut1 as a peptide selective for single-stranded DNA over duplex DNA, Mut1 was the only peptide used in studies with damaged and mismatched bases. The DNA sequences were varied to study several different types of DNA damage (Table 7.1) All binding studies were conducted by either fluorescence quenching or fluorescence anisotropy as in previous chapters and described in the Experimental Section.

[9] (a) Mathers, J. C.; Coxhead, J. M.; Tyson, J. *Curr. Cancer Drug Tar.* **2007**, *7*, 425-431. (b) Sansom, O. J.; Maddison, K.; Clarke, A. R. *Nat. Clin. Pract. Oncol.* **2007**, *4*, 305-315.

(a) Mut1:

Ac-S-R-W-T-E-H-K-S-N-G-R-T-Y-Y-**W**-N-**K**-V-N-G-O-**W**-Q-**K**-T-W-E-K-P-G-NH_2

(b)

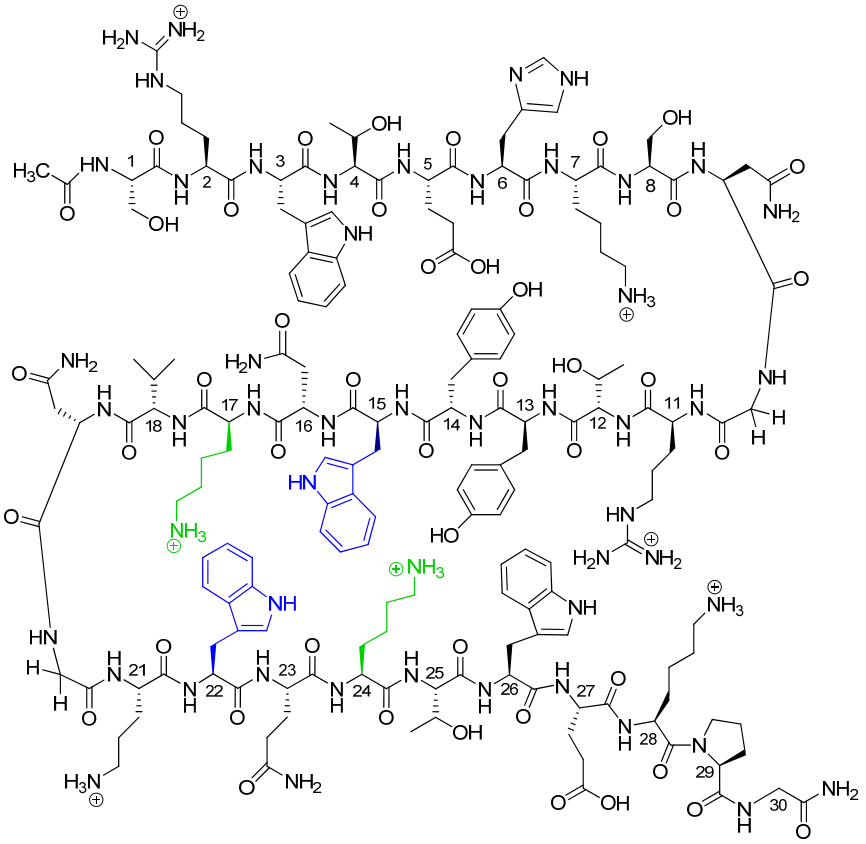


Figure 7.2. WW domain Mut1. (a) Sequence of Mut1 with residues in the binding pocket highlighted. (b) Structure of Mut1. The WKWK binding cleft is shown with Trp residues in blue and Lys residues in green.

Table 7.1. DNA sequences used as damaged or mismatched DNA.[a]

DNA Sequence Name	Sequence
(a) TTGTT	5'-TTGTT-3'
(b) TT(8oxoG)TT	5'-TT(**8oxoG**)TT-3'
(c) Duplex DNA	5'-CCATCGCTACC-3' 3'-GGTAGCGATGG-5'
(d) Duplex DNA with G-G Mismatch	5'-CCATC**G**CTACC-3' 3'-GGTAG**G**GATGG-5'
(e) Duplex DNA with C-C Mismatch	5'-CCATC**C**CTACC-3' 3'-GGTAG**C**GATGG-5'

[a]DNA bases shown in blue are either the damaged base or the mismatched base pairs.

ii. 8-oxo-Guanine Recognition Studies

The first set of experiments conducted involved studying oxidative DNA damage, and 8-oxo-guanine (8oxoG) (Figure 7.3) is known to be one of the most prevalent manifestations of oxidative damage.[10] In duplex DNA, damaged DNA sites are known to have a sharp bend, and molecular dynamics simulations have shown that in the case of 8-oxo-guanine, the 8oxoG causes distortion or kinks in the DNA backbone which assists these proteins in recognizing the damaged base.[10] Within the context of a single-stranded DNA sequence, we studied the recognition of this specific modification by Mut1. We used a short, five base pair sequence, 5'-TTGTT-3' (Table 7.1a), with 8oxoG replacing the guanine for comparison (Table 7.1b), to test the binding interaction between Mut1 and 8oxoG. Binding constants

[10] (a) Miller, J. H.; Fan-Chiang, C.-C. P.; Straatsma, T. P.; Kennedy, M. A. *J. Am. Chem. Soc.* **2003**, *125*, 6331-6336. (b) Malins, D. C.; Polissar, N. L.; Ostrander, G. K.; Vinson, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 12442-12445.

were determined by fluorescence quenching experiments. Penta-thymidine displayed the weakest affinity for $(WKWK)_2$ in previous studies,[11] so it was thought that placing 8oxoG in the middle of this sequence may cause enhanced binding or selectivity by Mut1, a similar peptide to WKWK dimer, for this specific damage.

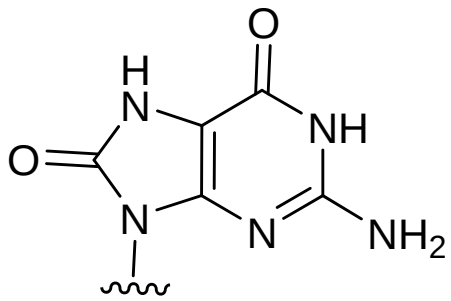


Figure 7.3. Structure of 8-oxo-guanine, a base damaged due to oxidation.

Mut1 binds the control sequence, 5'-TTGTT-3', with a K_d of about 50 μM, while it binds the analogous sequence including the 8oxoG with a dissociation constant of about 180 μM (Table 7.2, Figure 7.4). The binding affinity for both sequences is very weak, and the conditions were not optimal for binding in this case. The data suggests, however, that there is no real selectivity for the 8-oxo-guanine. This is consistent with the research indicating that the 8oxoG and other DNA damaged sites distort the DNA helix when placed in duplex DNA, and this may be the primary factor in the damaged base recognition. In that case, there would be no recognition of the base itself, but recognition would be due to the helix distortion. If this experiment were repeated in a duplex, we could determine if there is any selectivity for a mismatched base that could be flipped out into solution and more accessible for recognition by Mut1. Because of the weak binding, further studies in this area were not continued.

[11] Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24-25.

Table 7.2. Dissociation constants for Mut1 binding to two ssDNA sequences.[a]

DNA Sequence	K_d, μM (error)
5'-TTGTT-3'	50 (20)
5'-TT(8oxoG)TT-3'	180 (30)

(a) Conditions: 5 μM peptide concentrations; 10 mM Na_2HPO_4, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting.

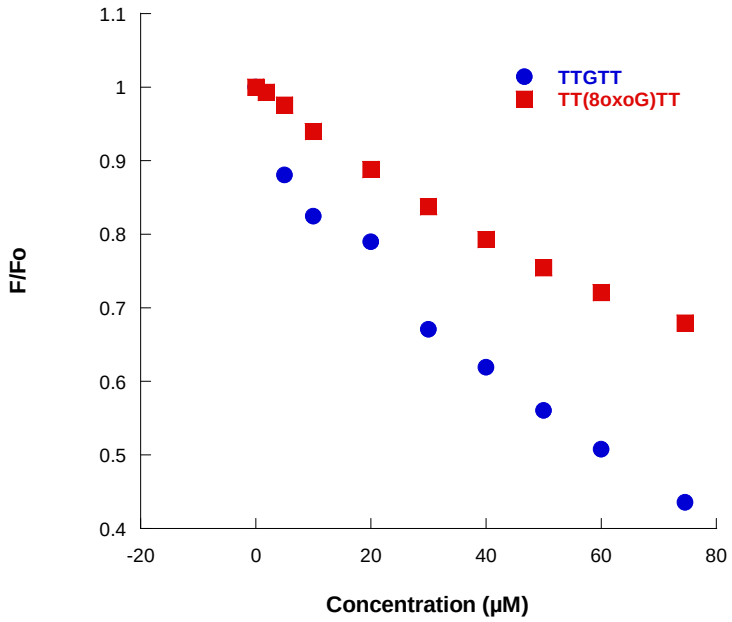


Figure 7.4. Fluorescence titrations of Mut1 with the single-stranded DNA sequences 5'-TTGTT-3' and 5'-TT(8oxoG)TT-3' (Table 7.1); 5 μM peptide concentrations; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

iii. Recognition of Duplex DNA Containing Single Mismatches

We then explored the ability of Mut1 to recognize mismatched DNA bases. Studies have shown that a conserved Phe residue stacks with the mismatched DNA in the mismatch

repair pathway.[12] Similarly, the base excision repair pathway uses an aromatic pocket for recognition of damaged bases by glycosylase enzymes.[13] We wanted to determine if the designed beta-sheet peptide Mut1 could recognize a mismatched DNA through its Trp binding pocket.

The first DNA sequence used was similar to the 11-mer duplex DNA sequence that has been used with all other experiments with this project (Table 8.1c). In this set of experiments, the fluorophore TAMRA was placed on Mut1 as in previous chapters (see Experimental Section) to attain optimal binding conditions using fluorescence anisotropy. We employed a G-G mismatch (Table 7.1d) as well as a C-C mismatch (Table 7.1e) for these experiments. Pyrimidine-pyrimidine mismatches are known to be the most destabilizing,[14] so that would indicate that the C-C mismatch would be the least stable in this case.

Fluorescence anisotropy studies showed that there is no significant selectivity for either of the mismatches by Mut1. Mut1 exhibits weak binding for each. Mut1 binds the duplex sequence with a K_d of about 390 μM. Similarly, the binding constant for the interaction between Mut1 and the duplex with a G-G mismatch is about 315 μM. Mut1 binds the double helix with the C-C mismatch with a dissociation constant of about 260 μM (Table

---

[12] (a) Kunkel, T. A.; Erie, D. A. *Annu. Rev. Biochem.* **2005**, *74*, 681-710. (b) Yamamoto, A.; Schofield, M. J.; Biswas, I.; Hsieh, P. *Nucleic Acids Res.* **2000**, *28*, 3564-3569. (c) Malkov, V. A.; Biswas, I.; Camerini-Otero, R. D.; Hsieh, P. *J. Biol. Chem.* **1997**, *272*, 23811-23817.

[13] (a) Cao, C. Y.; Kwon, K.; Jiang, Y. L.; Drohat, A. C.; Stivers, J. T. *J. Biol. Chem.* **2003**, *278*, 48012-48020. (b) Eichman, B. F.; O'Rourke, E. J.; Radicella, J. P.; Ellenberger, T. *EMBO J.* **2003**, *22*, 4898-4909. (c) Asaeda, A.; Ide, H.; Asagoshi, K.; Matsuyama, S.; Tano, K.; Murakami, A.; Takamori, Y.; Kubo; K. *Biochemistry*, **2000**, *39*, 1959-1965.

[14] (a) Cordier, C.; Pierre, V. C.; Barton, J. K. *J. Am. Chem. Soc.* **2007**, *129*, 12287-12295. (b) Jackson, B. A., Barton, J. K. *Biochemistry* **2000**, *39*, 6176-6182. (c) Tikhomirova, A.; Beletskaya, I. V.; Chalikian, T. V. *Biochemistry* **2006**, *45*, 10563-10571. (d) Aboul-ela, F.; Koh, D.; Martin, F. H.; Tinoco, I. Jr. *Nucleic Acids Res.* **1985**, *13*, 4811-4824.

7.3; Figure 7.5). Thermal UV experiments were attempted to gain more knowledge about these interactions, but those studies were unsuccessful due to instrument problems.

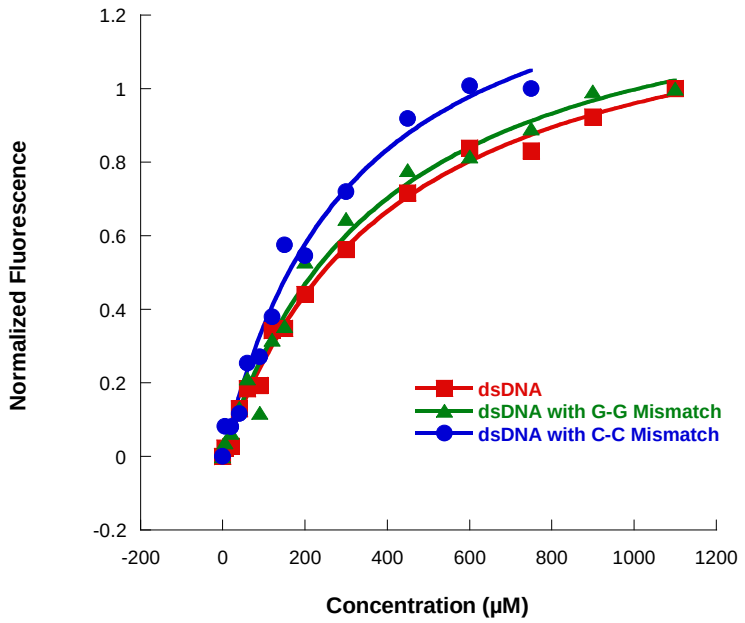


Figure 7.5. Fluorescence titrations of Mut1 with the double-stranded DNA sequence and the sequence with a G-G mismatch and a C-C mismatch (Table 7.1c-e); 20 μM peptide concentrations; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

Table 7.3. Dissociation constants for binding of Mut1 to dsDNA and the corresponding mismatched DNA sequences.[a]

DNA Sequence	K_d, μM (error)
Duplex DNA	390 (45)
Duplex DNA with G-G Mismatch	315 (40)
Duplex DNA with C-C Mismatch	260 (55)

(a) Conditions: 20 μM peptide concentrations; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Each value is the average of at least two measurements. The error is from the fitting.

According to these binding affinities, the conditions for binding are still not optimized because of the high DNA concentrations required. The design for Mut1 may have fulfilled its original purpose of gaining selectivity for ssDNA at the expense of weak binding for single base mismatches. By having the first strand inhibit duplex DNA binding, the additional strand may make the full peptide too large for the single mismatch to be accessible for favorable binding interactions. Strands 2&3 of Mut1 (Mut1-S23) may be better equipped to recognize a single mismatch because that peptide includes the full binding pocket but excludes Strand 1 which may inhibit the peptide from closely contacting the mismatched base. Some small preference may exist for the C-C mismatch, but the conditions are not optimal to determine this due to the weak binding. Experiments with more ideal conditions have not been attempted.

C. Discussion, Conclusions, and Future Directions

Initial attempts at recognizing damaged and mismatched DNA were not successful due to conditions that were not ideal. Data suggests that the original purpose for the design of Mut1 was realized and that duplex DNA binding has been inhibited due to the first strand. Previous data indicates that the peptide WKWK binds duplex DNA via some type of groove binding, and the first strand included in this peptide, Mut1, seems to make the peptide too large to form the same favorable contacts as it does in recognizing ssDNA. The idea for applying this peptide to damaged and mismatched bases is that the binding pocket would recognize the damaged or mismatched base like it would a single-stranded sequence[11] or ATP.[15] A truncated version of Mut1 composed of Strands 2&3 (Mut1-S23) may be the best binder of a damaged or mismatched DNA base. Many of the proteins which bind DNA

[15] (a) Butterfield, S. M.; Waters, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581. (b) Butterfield, S. M.; Sweeney, M. M.; Waters, M. L. *J. Org. Chem.* **2005**, *70*, 1105-1114.

damage utilize β-hairpins in the recognition event,[6, 7] so the hairpin forming Strands 2&3 may display the tightest binding of a single base in the context of a duplex sequence such as in DNA damage. Several additional studies could be conducted to better understand DNA recognition by different types of proteins. These include comparing Mut1, Mut1-S23, and related peptides to binding different DNA modifications such as oxidative damage, methylation, and various mismatches. New designs could be completed to improve binding affinity such that these types of studies could be more useful. These experiments have not been attempted but could shed light on the mechanisms by which proteins recognize the various DNA modifications that occur within the cell.

D. Experimental Section

i. Peptide Synthesis and Purification

Peptides were synthesized via automated solid phase peptide synthesis using an Applied Biosystems Pioneer Peptide Synthesizer. Fmoc protected amino acids were used with a PEG-PAL-PS resin. Amino acid residues were activated with HBTU (O-benzotriazole-N,N,N',N',-tetramethyluronium hexafluorophosphate) and HOBT (N-hydroxybenzotriazole) along with DIPEA (diisopropylethylamine) in DMF (N,N-dimethylformamide). Amino acids were deprotected with 2% DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) and 2% piperidine in DMF for approximately 10 minutes. Each amino acid was coupled on an extended cycle of 75 minutes to improve coupling. The N-terminus of each peptide was acetylated using 5% acetic anhydride and 6% lutidine in DMF for 30 minutes. Cleavage of the peptides from the resin as well as sidechain deprotection was performed in 95% trifluoroacetic acid (TFA), 2.5% H_2O, and 2.5% triisopropylsilane (TIPS) for three hours. TFA was evaporated by bubbling with nitrogen,

and ether was added to the resulting product. The peptide was then extracted with water and lyophilized to a powder.

Peptides were purified by reversed-phase HPLC. A Vydac C-18 semi-preparative column was used for separation with a gradient of 5-35% solvent B over 25 minutes with solvent A 95:5 water:acetonitrile, 0.1% TFA and solvent B 95:5 acetonitrile:water, 0.1% TFA. Peptides were then lyophilized and peptide sequence was confirmed by MALDI mass spectrometry. After purification, all peptides were desalted with a Pierce D-Salt Polyacrylamide 1800 desalting column.

ii. DNA Sample Preparation

DNA sequences were purchased from IDT (Integrated DNA Technologies, Inc.). All DNA samples were dissolved in 10mM Na_2HPO_4, 100 mM NaCl, adjusted to pH 7.0. Concentrations of both DNA strands were determined using a Perkin Elmer Lambda 35 UV/Vis Spectrometer. Absorbance values were determined at 260 nm, and concentrations were calculated using the extinction coefficients of the two DNA strands ($\varepsilon_{260,\ ssDNA} = 95500$ $M^{-1}\cdot cm^{-1}$ and $\varepsilon_{260,\ dsDNA} = 112600$ $M^{-1}\cdot cm^{-1}$). Equal concentrations of the two strands (in sodium phosphate buffer, pH 7.0) were pooled in a final concentration of 100 mM NaCl. The solution was heated at 95 ºC for 5 minutes to anneal the strands and was then allowed to cool to room temperature before storing at -20 ºC.

iii. Fluorescence Titrations

To determine the recognition of single-stranded and double-stranded oligonucleotides by the peptides, fluorescence titrations were performed which followed the Trp quenching with increasing oligonucleotide concentration. Peptide and nucleotide samples were prepared in 10 mM sodium phosphate buffer, pH 7.0. Peptide concentrations were

determined in 5 M guanidine hydrochloride by recording the absorbance of the Trp residues at 280 nm ($\varepsilon = 5690\ M^{-1}cm^{-1}$) by UV/vis spectroscopy. Concentrations of nucleotides were determined by UV/vis spectroscopy by observing the absorbance at 260 nm. Fluorescence scans were obtained on a Cary Eclipse Fluorescence Spectrophotometer from Varian. The experiments were performed at 298 K using an excitation wavelength of 297 nm. Fluorescence emission intensities of the Trp residues at 348 nm were fit as a function of nucleotide concentration to the binding equation (Equation 7.1) on Kaleidagraph using non-linear least squares fitting.[16]

Equation 7.1. $I = [I_o + I_\infty([L]/K_d)]/[1 + ([L]/K_d)]$

where I is the observed fluorescence intensity, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, and K_d is the dissociation constant. Oligonucleotides have an observable absorbance at the excitation wavelength of Trp (297 nm), and therefore there is an inner filter effect for which one must take account. The absorbance of the oligonucleotides at 297 nm was monitored at known concentrations and the extinction coefficient was determined. New absorbance values were determined for each oligonucleotide concentration. Corrected fluorescence values were determined from the following equations (Equation 7.2 and Equation 7.3).[17]

Equation 7.2. $F_c = F_o/C_i$

Equation 7.3. $C_i = (1\text{-}10^{-Ai})/(2.303)(A_i)$

[16] Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994**, *3*, 1261-1266.

[17] Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992**, *212*, 424-458.

where F_c is the corrected fluorescence, F_o is the fluorescence observed, and C_i is the correction factor for each absorbance value (i). A_i is the new absorbance value for each concentration determined by the extinction coefficient.

iv. Fluorescence Anisotropy

As in previous chapters, a fluorophore, 5-(and -6)-Carboxytetramethylrhodamine, mixed isomers (TAMRA), was purchased from Biotium, Inc., and was employed to provide an alternative to fluorescence quenching experiments using Mut1. Because of the inner filter effect, many of the binding interactions cannot be measured accurately using the fluorescence quenching method. TAMRA was coupled onto the peptide Mut1 at Orn 21. The synthesis was completed by coupling Lys(ivDde) in the original ornithine position (Orn21Lys). The ivDde protecting group was orthogonally deprotected by treatment with 2% hydrazine in DMF. Manual coupling of TAMRA was performed with two equivalents of TAMRA (100 mg bottle used for 0.1 mmol peptide) and four equivalents of HOBT, HBTU, and DIPEA in DMF. Cleavage from the resin and sidechain deprotection was completed as with all other peptides. The resulting peptide was purified by HPLC and its sequence and purity determined by mass spectrometry. Peptide concentrations were determined by UV-Vis using TAMRA's extinction coefficient of 91000 $M^{-1}\cdot cm^{-1}$ at wavelength 559 nm using the same conditions as with unlabeled peptides. This extinction coefficient was supplied by Integrated DNA Technologies, www.idtdna.com. The excitation wavelength used in the experiments is 559 nm, and the observed emission wavelength is 583 nm. Anisotropy was determined by the software that came with the instrument. The anisotropy was fit to the following equation (Equation 7.4) using Kaleidagraph to determine the binding constant

Equation 7.4. $$F = \frac{(((-(-K_d - [L] - [P]) - (\sqrt{((-K_d - [L] - [P])^2 - (4 \times [L] \times [P]))})))}{(2 \times [P])) \times (I_\infty - I_o)) + I_o}$$

where F is the fluorescence anisotropy, I_o is the initial fluorescence intensity of the peptide, I_∞ is the fluorescence intensity at binding saturation, [L] is the concentration of added nucleotide, [P] is the peptide concentration for each fraction, and K_d is the dissociation constant. Equation 7.4 was derived from equations given by Wang and coworkers.[18]

Overlay fluorescence plots are provided in the text of the chapter, but fits of each individual binding experiment are included in the Experimental Section below for completion.

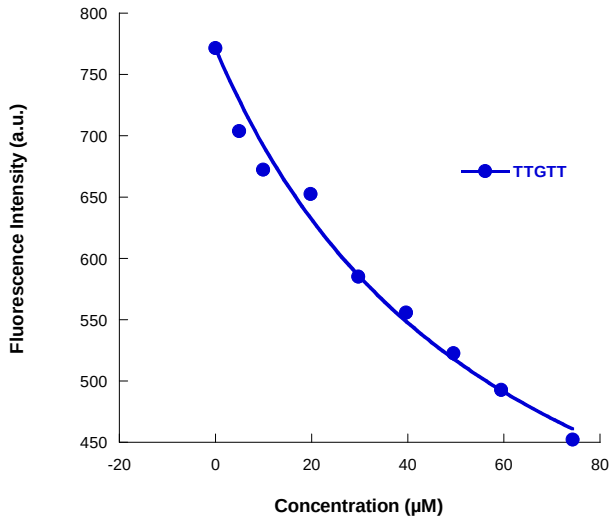


Figure 7.6. Fluorescence titrations of Mut1 with the single-stranded DNA sequence 5'-TTGTT-3' (Table 7.1a); 5 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

[18] Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768-779.

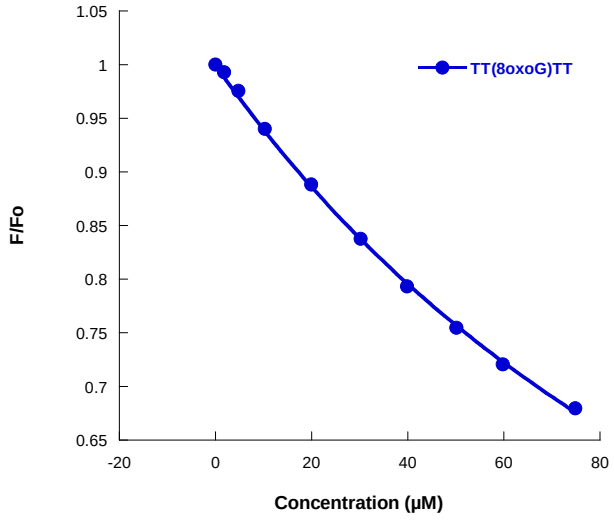


Figure 7.7. Fluorescence titrations of Mut1 with the single-stranded DNA sequence 5'-TT(8oxoG)TT-3' (Table 7.1b); 5 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. This data is normalized data.

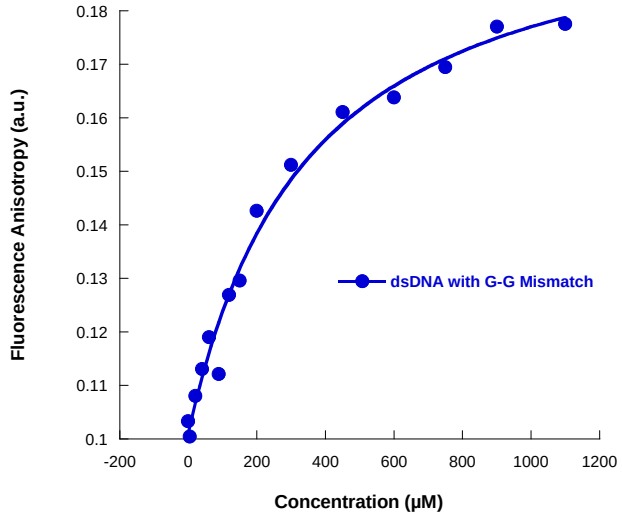


Figure 7.8. Fluorescence titration of Mut1 with the double-stranded DNA sequence containing a G-G mismatch (Table 7.1d); 20 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K. Experiment performed by fluorescence anisotropy.

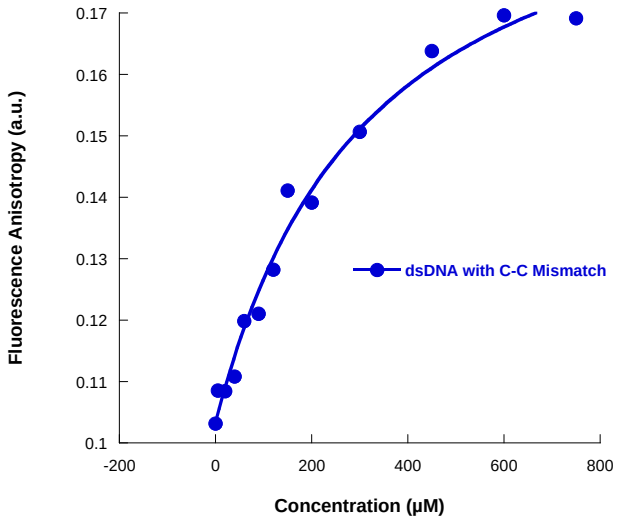


Figure 7.9. Fluorescence titration of Mut1 with the double-stranded DNA sequence with a C-C mismatch (Table 7.1e); 20 μM peptide concentration; 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 298 K.

Bibliography

Aboul-ela, F.; Koh, D.; Martin, F. H.; Tinoco, I. Jr. *Nucleic Acids Res.* **1985**, *13*, 4811-4824.

Aggarwal, S.; Harden, J. L.; Denmeade, S. R. *Bioconjugate Chemistry* **2006,** *17*, 335-340.

Ahn, D. R.; Yu, J. *Bioorg. Med. Chem.* **2005**, *13*, 1177-1183.

Allen, M. D.; Yamasaki, K.; Ohme-Takagi, M.; Tateno, M.; Suzuki, M. *EMBO J.* **1998,** *17*, 5484-5496.

Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. *Biochemistry* **2003**, *42*, 3751-3758.

Asaeda, A.; Ide, H.; Asagoshi, K.; Matsuyama, S.; Tano, K.; Murakami, A.; Takamori, Y.; Kubo; K. *Biochemistry* **2000**, *39*, 1959-1965.

Athanassiou, Z.; Dias, R. L.; Moehle, K.; Dobson, N.; Varani, G.; Robinson, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 6906-6913.

Athanassiou, Z.; Patora, K.; Dias, R. L.; Moehle, K.; Robinson, J. A.; Varani, G. *Biochemistry* **2007**, *46*, 741-751.

Aza-Blanc, P.; Cooper, C. L.; Wagner, K.; Batalov, S.; Deveraux, Q. L.; Cooke, M. P. *Molecular Cell* **2003,** *12*, 627-637.

Bahr, M.; Valis, L.; Wagenknecht, H.-A.; Weinhold, E. *Nucleosides, Nucleotides, and Nucleic Acids* **2007**, *26*, 1581-1584.

Ball, L. J.; Kuhne, R.; Schneider-Mergener, J.; Oschkinat, H. *Angew. Chem. Int. Ed.* **2005**, *44*, 2852-2869.

Ballestar, E.; Wolffe, A. P. *Eur. J. Biochem.* **2001**, *268*, 1-6.

Barbas, C. F., 3rd; Bain, J. D.; Hoekstra, D. M.; Lerner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89*, 4457-4461.

Bedford, M. T.; Chan, D. C.; Leder, P. *EMBO J.* **1997**, *16*, 2376-2383.

Blackwell, T. K..; Weintraub, H. *Science* **1990**, *250*, 1104-1110.

Blanco, F. J.; Jimenez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887.

Bochkarev, A.; Bochkareva, E. *Curr. Opin. Struct. Biol.* **2004,** *14*, 36-42.

Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. *Nature* **1997**, *385*, 176-181.

Burgess, K.; Martinez, C. I.; Russell, D. H.; Shin, H.; Zhang, A. J. *J. Org. Chem.* **1997**, *62*, 5662-5663.

Bussow, K.; Cahill, D.; Nietfeld, W.; Bancroft, D.; Scherzinger, E.; Lehrach, H.; Walter, G. *Nucleic Acids Res.* **1998**, *26*, 5007-5008.

Butterfield, S. M.; Waters, M. L., *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.

Butterfield, S. M.; Cooper, W. J.; Waters, M. L., *J. Am. Chem. Soc.* **2005**, *127*, 24-25.

Butterfield, S. M.; Sweeney, M. M.; Waters, M. L., *J. Org. Chem.* **2005**, *70*, 1105-1114.

Camperi, S. A.; Marani, M. M.; Iannucci, N. B.; Cote, S.; Albericio, F.; Cascone, O. *Tettrahedron Lett.* **2005**, *46*, 1561-1564.

Cao, C. Y.; Kwon, K.; Jiang, Y. L.; Drohat, A. C.; Stivers, J. T. *J. Biol. Chem.* **2003**, *278*, 48012-48020.

Chan, D. C.; Bedford, M. T.; Leder, P. *EMBO J.* **1996**, *15*, 1045-1054.

Chan, K.-W.; Lee, Y.-J.; Wang, C.-H.; Huang, H.; Sun, Y.-J. *J. Mol. Biol.* **2009**, *388*, 508-519.

Chandra, A.; Hughes, T. R.; Nugent, C. I.; Lundblad, V. *Genes Dev.* **2001**, *15*, 404-414.

Cheetham, G. M. T.; Jeruzalmi, D.; Steitz, T. A. *Nature* **1999**, *399*, 80-83.

Cheetham, G. M. T.; Steitz, T. A. *Science* **1999**, *286*, 2305-2309.

Chen, H. I.; Sudol, M. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7819-7823.

Cheng, A. C.; Chen, W. W.; Fuhrmann, C. N.; Frankel, A. D. *J. Mol. Biol.* **2003**, *327*, 781-796.

Cline, L. L.; Waters, M. L. *Org. Biomol. Chem.* **2009**, *7*, 4622-4630.

Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A., *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98*, 5578-5583.

Combs, A. P.; Kapoor, T. A.; Feng, S.; Chen, J. K.; Daude-Snow, L. F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1996**, *118*, 287-288.

Cooper, W. J.; Waters, M. L. *Org. Lett.* **2005**, *7*, 3825-3828.

Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652-3711.

Cordier, C., Pierre, V. C., Barton, J. K. *J. Am. Chem. Soc.* **2007**, *129*, 12287-12295.

Cowell, S. M.; Gu, X.; Vagner, J.; Hruby, V. J. *Methods Enzymol.* **2003**, *369*, 288-297.

Croy, J. E.; Wuttke, D. S. *Trends Biochem. Sci.* **2006**, *31*, 516-525.

Dalby, P. A.; Hoess, R. H.; DeGrado, W. F. *Protein Sci.* **2000**, *9*, 2366-2376.

Danek, S. C.; Queffelec, J.; Morken, J. P. *Chem. Commun.* **2002**, 528-529.

Desmond, J. C.; Raynaud, S.; Tung, E.; Hofmann, W.-K.; Haferlach, T.; Koeffler, H. P. *Leukemia* **2007**, *21*, 1026-1034.

Eichman, B. F.; O'Rourke, E. J.; Radicella, J. P.; Ellenberger, T. *EMBO J.* **2003**, *22*, 4898-4909.

Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346*, 818-822.

Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Peptide Sci.* **2005**, *80*, 303-311.

Famulok, M.; Jenne, A. *Curr. Opin. Chem. Biol.* **1998**, *2*, 320-327.

Fasan, R.; Dias, R. L.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *Angew. Chem. Int. Ed. Engl.* **2004,** *43*, 2109-2112.

Fasan, R.; Dias, R. L.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R.; Grutter, M. G.; Robinson, J. A. *ChemBioChem* **2006,** *7*, 515-526.

Fernandez-Escamilla, A. M.; Ventura, S.; Serrano, L.; Jimenez, M. A. *Protein Sci.* **2006**, *15*, 2278-2289.

Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233-1251.

Gotfredsen, C. H.; Grotli, M.; Willert, M.; Meldal, M.; Duus, J. O. *J. Chem. Soc., Perkin Trans. 1* **2000**, 1167-1171.

Griffiths-Jones, S. R.; Searle, M. S. *J. Am. Chem. Soc*. **2000**, *122*, 8350-8356.

Guilligay, D.; Tarendeau, F.; Resa-Infante, P.; Coloma, R.; Crepin, T.; Sehr, P.; Lewis, J.; Ruigrok, R. W. H.; Ortin, J.; Hart, D. J.; Cusack, S. *Nat. Struct. Mol. Biol.* **2008**, *15*, 500-506.

Gunawardane, R. N.; Sgroi, D. C.; Wrobel, C. N.; Koh, E.; Daley, G. Q.; Brugge, J. S. *Cancer Res.* **2005,** *65*, 11572-11580.

Hahn, W. C.; Dunn, I. F.; Kim, S. Y.; Schinzel, A. C.; Firestein, R.; Guney, I.; Boehm, J. S. *Biochim. Biophys. Acta.* **2009,** *1790*, 478-484.

Halkes, K. M.; Gotfredsen, C. H.; Grotli, M.; Miranda, L. P.; Duus, J. O.; Meldal, M. *Chemistry* **2001**, *7*, 3584-3591.

Hamamoto, K.; Kida, Y.; Zhang, Y.; Shimizu, T.; Kuwano, K., *Microbiol Immunol.* **2002,** *46*, 741-749.

Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 3548-3553.

Harrison, S. C. *Nature* **1991**, *353*, 715-719.

Hiemstra, H. S.; Duinkerken, G.; Benckhuijsen, W. E.; Amons, R.; de Vries, R. R.; Roep, B. O.; Drijfhout, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94*, 10313-10318.

Hillier, B. J.; Rodriguez, H. M.; Gregoret, L. M. *Fold. Des.* **1998,** *3*, 87-93.

Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86.

Horwitz, M. S. Z.; Loeb, L. A. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 7405-7409.

Hu, L. A.; Tang, P. M.; Eslahi, N. K.; Zhou, T.; Barbosa, J.; Liu, Q. *J. Biomol. Screen.* **2009,** *14*, 789-797.

Huang, P. Y.; Carbonell, R. G. *Biotechnol. Bioeng.* **1999**, *63*, 633-641.

Hughes, M. D.; Zhang, Z. R.; Sutherland, A. J.; Santos, A. F.; Hine, A. V. *Nucleic Acids Res.* **2005,** *33*, e32.

Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 6518-6519.

Hughes, R. M.; Benshoff, M. L.; Waters, M. L. *Chem. Eur. J.* **2007**, *13*, 5753-5764.

Jackson, B. A., Barton, J. K. *Biochemistry* **2000**, *39*, 6176-6182.

Jager, M.; Dendle, M.; Fuller, A. A.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 2306-2313.

Jager, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10648-10653.

Jager, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373-393.

Jiang, X.; Kowalski, J.; Kelly, J. W. *Protein Sci.* **2001**, *10*, 1454-1465.

Joyce, G. F. *Gene* **1989,** *82*, 83-87.

Kaptein, R. *Curr. Opin. Struct. Biol.* **1993**, *3*, 50-56.

Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3954-3959.

Kato, Y.; Nagata, K.; Takahashi, M.; Lian, L.; Herrero, J. J.; Sudol, M.; Tanokura, M. *J. Biol. Chem.* **2004**, *279*, 31833-31841.

Kato, Y.; Miyakawa, T.; Kurita, J.; Tanokura, M. *J. Biol. Chem.* **2006**, *281*, 40321-40329.

Kaul, R.; Angeles, A. R.; Jager, M.; Powers, E. T.; Kelly, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 5206-5212.

Kiehna, S. E.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2657-2667.

Kijanka, G.; Murphy, D. *J. Proteomics* **2009**, *72*, 936-944.

Kloks, C. P.; Spronk, C. A.; Lasonder, E.; Hoffmann, A.; Vuister, G. W.; Grzesiek, S.; Hilbers, C. W. *J. Mol. Biol.* **2002**, *316*, 317-326.

Koepf, E. K.; Petrassi, M.; Ratnaswamy, G.; Huff, M. E.; Sudol, M.; Kelly, J. W. *Biochemistry* **1999**, *38*, 14338-14351.

Koh, E. Y.; Chen, T.; Daley, G. Q. *Nucleic Acids Res.* **2002**, *30*, e142.

Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253-256.

Kraemer-Pecore, C. M.; Lecomte, J. T. J.; Desjarlais, J. R. *Protein Sci.* **2003**, *12*, 2194-2205.

Krajewska, W. M. *Int. J. Biochem.* **1992**, *24*, 1885-1898.

Kramer, A.; Volkmer-Engert, R.; Malin, R.; Reineke, U.; Schneider-Mergener, J. *Pept. Res.* **1993,** *6*, 314-319.

Krook, M.; Mosbach, K.; Lindbladh, C. *Biochem. Biophys. Res. Commun.* **1994**, *204*, 849-854.

Kunkel, T. A.; Erie, D. A. *Annu. Rev. Biochem.* **2005**, *74*, 681-710.

Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84.

Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997,** *97*, 411-448.

Lam, K. S.; Lehman, A. L.; Song, A.; Doan, N.; Enstrom, A. M.; Maxwell, J.; Liu, R. *Methods Enzymol.* **2003,** *369*, 298-322.

Latchman, D. S. *Int. J. Biochem. Cell. Biol.* **1997**, *29*, 1305-1312.

Leeper, T. C.; Athanassiou, Z.; Dias, R. L.; Robinson, J. A.; Varani, G. *Biochemistry* **2005,** *44*, 12362-12372.

Lim, W. A.; Fox, R. O.; Richards, F. M. *Protein Sci.* **1994**, *3*, 1261-1266.

Liu, D. R.; Schultz, P. G. *Angew. Chem. Int. Ed.* **1999**, *38*, 36-54.

Liu, R.; Marik, J.; Lam, K. S. *Methods Enzymol.* **2003,** *369*, 271-287.

Lohman, T. M.; Mascotti, D. P. *Methods Enzymol.* **1992,** *212*, 424-458.

Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860-2874.

Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303-1324.

Macias, M. J.; Hyvonen, M.; Baraldi, E.; Schultz, J.; Sudol, M.; Saraste, M.; Oschkinat, H. *Nature* **1996**, *382*, 646-649.

Macias, M. J.; Gervais, V.; Civera, C.; Oschkinat, H. *Nat. Struct. Biol.* **2000**, *7*, 375-379.

Malins, D. C.; Polissar, N. L.; Ostrander, G. K.; Vinson, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 12442-12445.

Malkov, V. A.; Biswas, I.; Camerini-Otero, R. D.; Hsieh, P. *J. Biol. Chem.* **1997**, *272*, 23811-23817.

Malta, E.; Moolenaar, G. F.; Goosen, N. *J. Biol. Chem.* **2006**, *281*, 2184-2194.

Mascotti, D. P.; Lohman, T. M., *Biochemistry* **1992**, *31*, 8932-8946.

Mascotti, D. P.; Lohman, T. M., *Biochemistry* **1993**, *32*, 10568-10579.

Mathers, J. C.; Coxhead, J. M.; Tyson, J. *Curr. Cancer Drug Tar.* **2007**, *7*, 425-431.

Max, K. E.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *J. Mol. Biol.* **2006**, *360*, 702-714.

Max, K. E.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. *FEBS J.* **2007,** *274*, 1265-1279.

Maynard, A. J.; Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 1996-2007.

Mer, G.; Bochkarev, A.; Gupta, R.; Bochkareva, E.; Frappier, L.; Ingles, J. C.; Edwards, A. M.; Chazin, W. J. *Cell* **2000**, *103*, 449-456.

Miller, J. H.; Fan-Chiang, C.-C. P.; Straatsma, T. P.; Kennedy, M. A. *J. Am. Chem. Soc.* **2003**, 125, 6331-6336.

Min, J.; Park, P.; Ko, E.; Choi, E.; Lee, H. *Biochem. Biophys. Res. Commun.* **2007**, *362*, 958-964.

Min, J. H.; Pavletich, N. P. *Nature* **2007**, *449*, 570-576.

Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. *Science* **2002,** *296*, 145-147.

Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew. Chem. Int. Ed. Engl.* **2007,** *46*, 9101-9104.

Moodie, S. L.; Mitchell, J. B. O.; Thornton, J. M. *J. Mol. Biol.* **1996**, *263*, 486-500.

Moreno, R.; Jiang, L.; Moehle, K.; Zurbriggen, R.; Gluck, R.; Robinson, J. A.; Pluschke, G. *ChemBioChem* **2001**, *2*, 838-843.

Morken, J. P.; Kapoor, T. M.; Feng, S.; Shirai, F.; Schreiber, S. L. *J. Am. Chem. Soc.* **1998**, *120*, 30-36.

Moure, C. M., Gimble, F. S., Quiocho, F. A. *J. Mol. Biol.* **2003**, *334*, 685-695.

Murray, J. K.; Farooqi, B.; Sadowsky, J. D.; Scalf, M.; Freund, W. A.; Smith, L. M.; Chen, J.; Gellman, S. H. *J. Am. Chem. Soc.* **2005,** *127*, 13271-13280.

Murzin, A. G. *EMBO J.* **1993**, *12*, 861-867.

Nair, P. A.; Nandakumar, J.; Smith, P.; Odell, M.; Lima, C. D.; Shuman, S. *Nat. Struct. Mol. Biol.* **2007**, *14*, 770-778.

Neidle, S.; Parkinson, G. N. *Curr. Opin. Struct. Biol.* **2003**, *13*, 275-283.

Newkirk, K.; Feng, W.; Jiang, W.; Tejero, R.; Emerson, S. D.; Inouye, M.; Montelione, G. T. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 5114-5118.

Nguyen, H.; Jager, M.; Moretto, A.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3948-3953.

Nguyen, H.; Jager, M.; Kelly, J. W.; Gruebele, M. *J. Phys. Chem. B* **2005**, *109*, 15182-15186.

Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J. G.; Ikegami, T.; Nakao, M.; Shirakawa, M. *Cell* **2001**, *105*, 487-497.

Oliphant, A. R.; Nussbaum, A. L.; Struhl, K. *Gene* **1986**, *44*, 177-183.

Otte, L.; Wiedemann, U.; Schlegel, B.; Pires, J. R.; Beyermann, M.; Schmieder, P.; Krause, G.; Volkmer-Engert, R.; Schneider-Mergener, J.; Oschkinat, H. *Protein Sci.* **2003**, *12*, 491-500.

Paulick, M. G.; Hart, K. M.; Brinner, K. M.; Tjandra, M.; Charych, D. H.; Zuckermann, R. N. *J. Comb. Chem.* **2006**, *8*, 417-426.

Perrin, D. M. *Comb. Chem. High Throughput Screen.* **2000,** *3*, 243-269.

Pestryakov, P. E.; Lavrik, O. I. *Biochemistry (Moscow)* **2008**, *73*, 1388-1404.

Pires, J. R.; Parthier, C.; do Aido-Machado, R.; Wiedemann, U.; Otte, L.; Bohm, G.; Rudolph, R.; Oschkinat, H. *J. Mol. Biol.* **2005**, *348*, 399-408.

Pollock, R.; Treisman, R. *Nucleic Acids Res.* **1990,** *18*, 6197-6204.

Quiocho, F. A.; Hu, G. H.; Gershon, P. D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 78-86.

Ramirez-Alvarado, M.; Blanco, F. J.; Niemann, H.; Serrano, L. *J. Mol. Biol.* **1997**, *273*, 898-912.

Richter, J. D.; Sonenberg, N. *Nature* **2005**, *433*, 477-479.

Riemen, A. J.; Waters, M. L. *Biochemistry* **2009**, *48*, 1525-1531.

Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 1195-1195.

Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 388-395.

Rutkowska-Wlodarczyk, I.; Stepinski, J.; Dadlez, M.; Darzynkiewicz, E.; Stolarski, R.; Niedzwiecka, A. *Biochemistry* **2008**, *47*, 2710–2720.

Samson, I.; Rozenski, J.; Samyn, B.; Van Aerschot, A.; Van Beeumen, J.; Herdewijn, P. *J. Biol. Chem.* **1997**, *272*, 11378-11383.

Sansom, O. J.; Maddison, K.; Clarke, A. R. *Nat. Clin. Pract. Oncol.* **2007**, *4*, 305-315.

Schindelin, H.; Marahiel, M. A.; Heinemann, U. *Nature* **1993,** *364*, 164-168.

Schnuchel, A.; Wiltscheck, R.; Czisch, M.; Herrler, M.; Willimsky, G.; Graumann, P.; Marahiel, M. A.; Holak, T. A. *Nature* **1993,** *364*, 169-171.

Schouten, J. A.; Ladame, S.; Mason, S. J.; Cooper, M. A.; Balasubramanian, S. *J. Am. Chem. Soc.* **2003**, *125*, 5594-5595.

Schroder, K.; Graumann, P.; Schnuchel, A.; Holak, T. A.; Marahiel, M. A. *Mol. Microbiol.* **1995,** *16*, 699-708.

Scott, J. K.; Smith, G. P. *Science* **1990,** *249*, 386-390.

Shamoo, Y.; Friedman, A. M.; Parsons, M. R.; Konigsberg, W. H.; Steitz, T. A. *Nature* **1995**, *376*, 362-366.

Sharman, G. J.; Searle, M. S. *Chem. Commun*. **1997**, 1955-1956.

Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc*. **1998**, *120*, 5291-5300.

Shen, J.; Gai, D.; Patrick, A.; Greenleaf, W. B.; Chen, X. S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 11248-11253.

Sillerud, L. O.; Larson, R. S. *Current Protein and Peptide Science* **2005**, *6*, 151-169.

Simonsson, T., *Biol. Chem.* **2001**, *382*, 621-628.

Skorvaga, M.; Theis, K.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2002**, *277*, 1553-1559.

Skorvaga, M.; DellaVecchia, M. J.; Croteau, D. L.; Theis, K.; Truglio, J. J.; Mandavilli, B. S.; Kisker, C.; Van Houten, B. *J. Biol. Chem.* **2004**, *279*, 51574-51580.

Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391-410.

Sonenberg, N.; Gingras, A.-G. *Curr. Opin. Cell Biol.* **1998**, *10*, 268-275.

Stanger, H. E.; Gellman, S. H. *Journal of the American Chemical Society* **1998,** *120*, 4236-4237.

Stano, N. M.; Patel, S. S. *J. Mol. Biol.* **2002**, *315*, 1009-1025.

Stewart, A. L.; Waters, M. L. *ChemBioChem* **2009,** *10*, 539-544.

Sugiyama, T.; Kantake, N.; Wu, Y.; Kowalczykowski, S. C. *EMBO J.* **2006**, *25*, 5539-5548.

Sundquist, W. I.; Klug, A. *Nature* **1989**, *342*, 825-829.

Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126*, 2028-2034.

Taylor, S. J.; Morken, J. P. *Science* **1998**, *280*, 267-270.

Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 115-133.

Tikhomirova, A.; Beletskaya, I. V.; Chalikian, T. V. *Biochemistry* **2006**, *45*, 10563-10571.

Tisne, C.; Dardel, F. *Comb. Chem. High Throughput Screen.* **2002**, *5*, 523-529.

Travers, A. *DNA—Protein Interactions.* Chapman and Hall: London, 1993.

Truglio, J. J.; Karakas, E.; Rhau, B.; Wang, H.; DellaVecchia, M. J.; Van Houten, B.; Kisker, C. *Nat. Struct. Mol. Biol.* **2006**, *13*, 360-364.

Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.

Tugyi, R.; Uray, K.; Ivan, D.; Fellinger, E.; Perkins, A.; Hudecz, F. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 413-418.

Vetter, D.; Thamm, A.; Schlingloff, G.; Schober, A. *Mol. Divers.* **2000**, *5*, 111-116.

Voorhoeve, P. M.; le Sage, C.; Schrier, M.; Gillis, A. J. M.; Stoop, H.; Nagel, R.; Liu, Y. P.; van Duijse, J.; Drost, J.; Griekspoor, A.; Zlotorynski, E.; Yabuta, N.; De Vita, G.; Nojima, H.; Looijenga, L. H. J.; Agami, R. *Cell* **2006**, *124*, 1169-1181.

Wang, B.; Dickinson, L. A.; Koivunen, E.; Ruoslahti, E.; Kohwi-Shigematsu, T. *J. Biol. Chem.* **1995**, *270*, 23239-23242.

Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768-779.

Warren, E. M.; Huang, H.; Fanning, E.; Chazin, W. J.; Eichman, B. F. *J. Biol. Chem.* **2009**, *284*, 24662-24672.

Whitehurst, A. W.; Bodemann, B. O.; Cardenas, J.; Ferguson, D.; Girard, L.; Peyton, M.; Minna, J. D.; Michnoff, C.; Hao, W. H.; Roth, M. G.; Xie, X. J.; White, M. A. *Nature* **2007**, *446*, 815-819.

Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871-880.

Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333.

Wold, M. S. *Annu. Rev. Biochem.* **1997**, *66*, 61-92.

Wuthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley-Interscience: New York, 1986.

Yamamoto, A.; Schofield, M. J.; Biswas, I.; Hsieh, P. *Nucleic Acids Res.* **2000**, *28*, 3564-3569.

Yang, L.; Schepartz, A. *Biochemistry* **2005**, *44*, 7469-7478.

Yu, P.; Liu, B.; Kodadek, T. *Nat. Biotechnol.* **2005**, *23*, 746-751.

Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991,** *350*, 718-720.

Zarrinpar, A.; Lim, W. A. *Nat. Struct. Biol.* **2000**, *7*, 611-613.