5-20-2020

Designing Peptides to Bind kappa B DNA and Mimic NF-kappa B Protein Complexes

Nicholas A. Markowich
Georgia Southern University

Designing Peptides to Bind kappa B DNA and Mimic NF-kappa B Protein Complexes

An Honors Thesis submitted in partial fulfillment of the requirements for Honors in Department of Chemistry and Biochemistry

By
Nicholas Markowich

Under the mentorship of Dr. Amanda Stewart White

Transcription factors are important proteins that regulate gene expression and protein synthesis. Transcription factors can either boost the gene's transcription rate by helping RNA polymerase activate transcription or restrict it by interfering with RNA polymerase, thereby repressing transcription. Nuclear factor-kappa-B (NF-κB) transcription factors are a family of proteins that control the synthesis of proteins involved in many cellular processes such as inflammatory and immune responses, cell growth, and apoptosis. However, the overexpression and activation of these transcription factors is linked to deadly conditions such as cancer and neurodegenerative diseases which currently have few safe cures. The goal of this research is to design and synthesize peptides to mimic the size, shape, structure and function of NF-κB. This will allow peptides to compete with and inhibit NF-κB from binding DNA, thus preventing the overexpression of proteins. To explore such possibilities, one peptide was analyzed. The binding affinity for between the peptide and κB DNA was determined using Isothermal Titration Calorimetry (ITC). The secondary structure of the peptide was determined using circular dichroism (CD). Initial data will direct the synthesis of new peptides that will show improved DNA binding affinity and will be screened for inhibition studies. A review of other β-sheet peptides is included to provide more context of their importance in supporting life and how they bind to DNA.

Thesis Mentor:_____

Dr. Amanda Stewart White

Honors Director:_____

Dr. Steven Engel

May 2020
Chemistry and Biochemistry
University Honors Program
Georgia Southern University

Acknowledgements

I would like to take a moment to thank the aid of Dr. Amanda Stewart White, who has supported my Honors research to the best of her ability throughout my time in the program. I would also like to thank the members of the Honors Program for allowing me to appreciate the benefits of membership during my time at Georgia Southern University.

Summary/Review of β sheet Peptides Section

There are many different types of β-sheet DNA-binding proteins that play a vital role in the function of the body. NF-κB, for example, has the important role of regulating DNA transcription factors that in turn control gene expression and various cell processes vital to survival. Other important β-sheet DNA-binding proteins include replication protein A (RPA), xeroderma pigmentosum group A (XPA), X-ray repair cross-complementing protein 4 (XRCC4), SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A like 1 (SMARCAL1) protein, zinc finger RAN-binding domain containing 3 (ZRANB3) protein, helicase-like transcrpition factor (HLTF) protein, and the proliferating cell nuclear antigen (PCNA) sliding clamp.

RPA is a protein complex consisting of three subunits of approximately 70, 32, and 14 kDA each (heterotrimer) that each consist of α-helices (meaning a tendency towards methionine, alanine, leucine, glutamate, and lysine), a large number of β-sheets (meaning a tendency towards tyrosine, phenylalanine, tryptophan, threonine, valine, isoleucine, and proline), and loop regions (meaning a tendency towards tryptophan, phenylalanine, histidine, aspartate, tyrosine, leucine, glutamate, isoleucine, and valine). Using its six OB (Oligonucleotide/oligosaccharide-binding) folds, it mainly binds to single-stranded DNA (ssDNA) in a non-specific sequence manner in order to function in the role of DNA repair and is a major player in binding with other proteins involved in nucleotide excision repair (NER) (Flynn & Zou 267). Four of these OB folds are DNA-binding domains on its β-sheets (three on RPA70 and one on RPA32) where the ssDNA binds to, with the amino acid tryptophan (W-101 in the RPA32 subunit DBD-D, for example) being the specific binding agent for the ssDNA to a loop region (Pohkrel 9414-

9415). If the first binding domain on RPA70 (DBD-A) fails to bind the ssDNA due to disruption, it destroys RPA's binding affinity to ssDNA, while point mutations in DBD-A or DBD-B or deletion of DBD-C just reduce said binding affinity. None of these disruptions affect its functionality in DNA replication, repair and checkpoint response, however, suggesting that RPA's functions are not reliant on each other (Flynn & Zou 267).

XPA is a protein consisting of α-helices (meaning a tendency towards methionine, alanine, leucine, glutamate, and lysine), β-sheets (meaning a tendency towards tyrosine, phenylalanine, tryptophan, threonine, valine, isoleucine, and proline), and loop regions (meaning a tendency towards tryptophan, phenylalanine, histidine, aspartate, tyrosine, leucine, glutamate, isoleucine, and valine). XPA shares several visual characteristics with RPA aside from its much smaller size, but it prefers binding to branched DNA over linear DNA (Krasikova 2). Unlike RPA, XPA only has one DNA-binding domain in its central region, where it binds to DNA through the β-hairpin formed by β4-β5 specifically on residue Trp175, the antiparallel β-sheet (β3-β5) and the helix α2. In fact, two XPA molecules will work together to bind to a DNA duplex on either side (Lian 467-468). RPA and XPA are both close partners and important proteins in the NER process, where their binding to damaged DNA and other proteins trigger the activation of other proteins that repair DNA (Krasikova 2).

XRCC4 is an important protein in the process of specifically repairing DNA breaks. It mainly functions as a scaffolding protein for multiple non-homologous end-joining (NHEJ) proteins such as XLF, DNA Ligase IV, APTX, and PNK. It is comprised of an N-terminal head domain (NTD), made up of a 7-stranded anti-parallel β sandwich

(meaning a tendency towards tyrosine, phenylalanine, tryptophan, threonine, valine, isoleucine, and proline), and a helix-turn-helix motif, which consists of two α helices (meaning a tendency towards methionine, alanine, leucine, glutamate, and lysine) joined by a short strand of amino acids. It also has an elongated, flexible α-helical coiled-coil region, and carboxyl termini tails which are phosphorylated by casein kinase II (CK2) that binds the necessary other proteins (Andres 13). However, for the interests of the protein-DNA interactions, only the NTD is of true interest, as this part of the protein is what binds to the structurally similar protein XRCC4-like factor (XLF). The bonding results in forming alternating chains of superhelical filaments that can bind to long double-stranded DNA substrates of over 100 base pairs via either grooves on the outside of the filament or through the central channel of the helices. The filament can then align broken or mismatched DNA end breaks for more efficient end processing and ligation (Andres 14). Overall, XRCC4 acts as an intermediary to the process of repairing DNA breaks via bonding to multiple proteins and the DNA rather than being a true initiator.

SMARCAL1 is a protein consisting of 954 amino acids and containing an RPA binding domain at the N-terminus followed by two HARP domains, and two lobed ATPase domains on the C-terminal half (Poole and Cortez 697). Its major purpose is to reanneal the complimentary strands of ssDNA bound by RPA and evict the latter protein, thus producing dsDNA. It does this in short bursts in order to in theory prevent excessive reannealing of the strands. Excessive reannealing would also interfere with SMARCAL1's role in fork reversal, which is the process by which the replication "fork" created as the ssDNA strands combine back together undergoes in order to prevent replication stress from stalling production. This involves reannealing the parental and

nascent DNA strands to reverse the direction of the replication fork, which creates a "chicken foot" appearance when viewed (Poole and Cortez 698-699). The RPA binding domain on amino acid units 2-30 is necessary to interact with RPA subunit 32, which allows RPA to localize SMARCAL1 to the areas of replication stress and to either activate or inhibit its function as necessary (Poole and Cortez 701). The HARP domains on amino acid units 226-303 and 327-398 are the particularly important parts of the protein as they consist of a four-stranded anti-parallel β sheet (meaning a tendency towards tyrosine, phenylalanine, tryptophan, threonine, valine, isoleucine, and proline) decorated with two α helices (meaning a tendency towards methionine, alanine, leucine, glutamate, and lysine). Their purpose is to act as substrate recognition domains (SRDs) that confer DNA binding preference for junction DNA (Poole and Cortez 700). Without these, it would be impossible for them to do the job of reannealing the DNA. SMARCAL1 also bears a role in telomere replication through a similar manner (Poole and Cortez 704). Notably, however, SMARCAL1 does not bind DNA substrates composed entirely of ssDNA or dsDNA; however, it displays a strong preference for binding DNA substrates with a ssDNA/dsDNA junction (Poole and Cortez 699).

ZRANB3 is a 1079-amino acid protein and, based on sequence homology, is the most closely related SNF2 protein to SMARCAL1 (Poole and Cortez 705). It contains two lobed ATPase domains following each other on the N-terminus, followed by two protein interaction sites for PIP and NZF respectively, a SRD, a nuclease for HNH, and another protein interaction site for APIM on the C-terminus (Poole and Cortez 697). Like SMARCAL1 it has a strong preference for binding DNA substrates with a ssDNA/dsDNA junction while not binding to substrates composed entirely of ssDNA or

dsDNA, and it has an SRD like SMARCAL1's HARP domains (Poole and Cortez 706). However, not only is this SRD on a different section of the protein at amino acids units 621-650, it is heavily implied it has a similar but not identical structure to the HARP domains since while ZRANB3 functions in replication stress responses and fork remodeling just like SMARCAL1, it lacks SMARCAL1's role in telomere replication and possesses endonuclease activity that depends on ATP hydrolysis by the intact motor domain and a C-terminal nuclease domain (Poole and Cortez 705-706). It has also been noted that ZRANB3 is localized to the areas of replication stress by its interactions with PCNA rather than RPA since it lacks a binding domain to the latter but does have them to the former via the PIP (PCNA-interacting protein), but the presence and absence of RPA can still affect whether or not fork reversal is inhibited (Poole and Cortez 707).

HLTF is a 1009-amino acid protein that contains a HIRAN domain that acts as its SRD at the N terminus, followed by two ATPase domains, and a RING domain between the two ATPase domains that acts as an E3 ubiquitin ligase (Poole and Cortez 697). The HIRAN domain consists of an arrangement of β sheets and α helices similar to the HARP domains of SMARCAL1 and the SRD of ZRANB3, but they are longer than the former in length (Poole and Cortez 700). Like SMARCAL1 and ZRANB3, HLTF binds to dsDNA with low affinity, but unlike them it will bind ssDNA and prefers binding to a replication fork-like structure due to how the HIRAN domain recognizes ssDNA with a preference for 3' overhangs possessing a hydroxide group (3'-OH) (Poole and Cortez 709). While it works in fork replication, it can only act in fork reversal rather than fork restoration, and as it lacks any RPA binding domain RPA seemingly has no effect on its regulation while PCNA presence does catalyze fork reversal (Poole and Cortez 710).

PCNA is a ring-shaped homotrimer encircling the DNA helix whose purpose is to provide a moving platform for DNA polymerases and other proteins to dock for DNA replication and repair, and it can be considered the eukaryotic DNA sliding clamp equivalent of the bacterial β-clamp (March and Biasio 663). The clamp consists of two layers: an outer layer of six β-sheets and an inner layer of 12 α-helices lining the central channel, which is rich in lysine and arginine residues. The outer layer contains the protein-protein interaction sites, while the inner layer contains key basic residues that match to areas of the DNA's double helix that allow it to recognize where on the DNA structure it is and in turn allow for rotational DNA backbone tracking and the clamp sliding along the helix as necessary, and these key residues cause the DNA to tilt and bend as it runs through the clamp to match with them (March and Biasio 663-665). The PCNA moves along DNA through two different modes of travel: transitional diffusion where PCNA travels along the DNA rapidly without making contact, and cogwheel diffusion where PCNA rotationally tracks the helical pitch of the DNA duplex through spiral motion, creating short-lived DNA interactions at a 30º tilt to the clamp central axis that keep the clamp oriented. Cogwheel diffusion in particular is the most likely mode of PCNA sliding used during DNA replication, as it keeps the PCNA in the proper orientation to bind to the polymerase pol δ, which duplicates DNA, and also means there will always be an "active" spot ready to continue synthesis uninterrupted (March and Biasio 666-667).

These various proteins share several characteristics with each other in some form or another yet also bear differences. RPA and XPA obviously bear the most similarities to each other, since as previously mentioned XPA is structurally like RPA if smaller and

with only one DNA binding domain to RPA's four. They both can play roles in many processes such as DNA replication, repair, and checkpoint response. The obvious difference is the fact that XPA has different binding preferences to RPA in that it prefers branched DNA to linear DNA, likely due to how it has fewer binding domains and its smaller size. XPA noticeably doesn't rely on RPA for activation and instead works as a partner alongside it as necessary.

XRCC4, meanwhile, is notably much different to RPA in terms of structure and function, particularly since its main role is as a scaffolding protein that acts solely in the process of mainly repairing DNA breaks rather than the many functions RPA and XPA do, and as a result it is much more specialized. After all, it needs to bind to the structurally similar XLF protein on its NTD with its anti-parallel β sandwich and helix turn helix motif first in order to make the filament able bind to DNA while RPA and XPA have no such issue, and even then it prefers binding to large sections of DNA while RPA and XPA have no such size restrictions. XRCC4 is not so much the initiator like RPA and XPA but rather the hub that serves to collect together the necessary proteins to prepare them for DNA binding.

SMARCAL1 is also more specialized like XRCC4 but noticeable in how it works in conjecture with RPA to fulfill its functions of reannealing DNA, even having an RPA-binding domain to react to the presence of the protein in question. Thus, SMARCAL1 unlike the others could be considered something of a "follower" protein that responds to the process already in motion to fulfill its duty and thus can activated and stopped by RPA as required. The HARP domains of SMARCAL1 meanwhile are what bind to DNA as necessary, but as previously mentioned due to the different structure they have a

different preference for the kind of DNA they bind by preferring a combination of ssDNA and dsDNA rather than the branched DNA of XPA or the linear ssDNA preference of RPA. ZRANB3 meanwhile is very similar in structure to SMARCAL1 and bears many of the same functions and characteristics. One of the biggest differences, of course, is that since ZRANB3 lacks an RPA binding domain it's regulated less directly by the protein than SMARCAL1 and in fact can act alongside SMARCAL1 in a manner like how RPA and XPA can work together where a deficiency of one doesn't have a negative effect on the other and they can cover each other to an extent. HLTF meanwhile lacks an RPA binding domain and seemingly uninfluenced by RPA's presence while like ZRANB3 it is still affected by PCNA, and the similar-yet-longer structure of its HIRAN domain also causes HLTF to only act in fork reversal rather than also fork restoration. Thus, the highly-similar structures with key differences allow them to have similar functions, but also their own unique roles that allow them to work mutually yet independently as necessary for the body.

PCNA is like XRCC4 since binds DNA and other proteins on two different parts of its total structure and like XRCC4 acts as a "scaffolding" to allow proteins and DNA to properly bind. Where they differ is that PCNA binds to other proteins on its β sheet outer layer while it binds to DNA on its α helix inner layer before letting the two interact with each other. PCNA is also much more active than XRCC4 since it effectively acts as the major binding area for proteins and DNA. PCNA also aids in DNA replication and repair via being modulated by two different processes. The first is the acetylation of lysine residues on the inner layer of the sliding clamp, with certain residues degraded by certain proteins promoting the removal of chromatin-bound PCNA and its degradation in

NER, while the acetylation of other residues in response to DNA damage triggers homologous recombination suppresses cell sensitivity to the damaging agents. The second is interacting with the protein factor $p15^{PAF}$, which bears a role in DNA replication and repair by binding to both the PCNA and directly to the DNA. This allows it to modulate PCNA sliding by forcing cogwheel diffusion over transitional diffusion due to binding DNA in the clamp to PCNA (March and Biasio 668-670).

It's noticeable that despite possessing many of the same secondary structures of β sheets and α helices, these proteins have much different roles in how they bind DNA and how that binding affects the necessary function they are carrying out. This in turn cuts to the heart of the matter that it is not merely the presence of secondary structures created by the specific amino acids of a protein that determines protein functionality and binding, but how those secondary structures are arranged in the protein to carry those out. Scientists in their attempts to understand and unlock the secrets of proper DNA-binding have had to observe the placement of these secondary structures and how they interact with each other and specific DNA sequences in order synthesize smaller proteins that can mimic just enough of the necessary structures to still bind the specific sequences without causing any resulting expression. In learning how to do this, scientists open up the possibility of being able to control the expression of proteins in a safe way that does not involve throwing the body's natural balance of proteins out of order, but rather than be used to prevent overexpression and under-expression of certain proteins that results in otherwise incurable diseases by traditional methods.

Works Cited

Andres, Sara N.; et al.; "Recognition and Repair of Chemically Heterogeneous Structures at DNA Ends"; *Environmental and Molecular Mutagenesis*; 56; 1-21; 2015; https://drive.google.com/file/d/1DjynbmtQ2JrBpfR2EYJNd7OZ8vhbJEB2/view; Accessed April 9, 2020.

Flynn, Rachel Litman; Zou, Lee; "Oligonucleotide/oligosaccharide-binding fold proteins: a growing family of genome guardians"; *Critical Reviews in Biochemistry and Molecular Biology*; 45(4); 266–275; 2010; https://drive.google.com/file/d/1vijTnZcy2bDMvFHphXVHRrjA6Hj72iWw/view ; Accessed February 6, 2020.

Krasikova, Yuliya S.; et al.; "RPA and XPA interaction with DNA structures mimicking intermediates of the late stages in nucleotide excision repair"; *PLOS One*; 13; 1-20; January 10, 2018; https://doi.org/10.1371/journal; Accessed February 6, 2020.

Lian, Fu-Ming, et al; "New Structural Insights into the Recognition of Undamaged Splayed-Arm DNA with a Single Pair of Non-Complementary Nucleotides by Human Nucleotide Excision Repair Protein XPA"; *International Journal of Biological Macromolecules*; 148; 466–474; April 2018; https://www.sciencedirect.com/science/article/pii/S0141813019393109; Accessed February 13, 2020.

March, Matteo De; Biasio, Alfredo De; "The dark side of the ring: role of the DNA sliding surface of PCNA"; *Critical Reviews in Biochemistry and Molecular Biology*; 52:6; 663-673;

https://drive.google.com/file/d/1YEGL9_AwAQUnTb7soro8AmI29XgetFy4/view; Accessed February 6, 2020.

Pokhrel, Nilisha; "Monitoring Replication Protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids"; *Nucleic Acids Research*; 45; 16; 9413–9426; 2017; https://drive.google.com/file/d/1FamknwJQtkaEwEouf8Af40eARyfwnrYa/view; Accessed February 6, 2020.

Poole, Lisa A.; Cortez, David; "Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability"; *Critical Reviews in Biochemistry and Molecular Biology*; 52(6); 696-714; 2017; http://dx.doi.org/10.1080/10409238.2017.1380597; Accessed April 9, 2020.

Experimental Section

Introduction

In order to properly analyze the sample peptide NF-κB WKWK (Ac-QRFRWVRVNGKYIKVQLE-NH_2) and how it interacts with annealed sample κB DNA combined with its complement (both 5'-TGG-GAA-TTC-CCA-3'), the processes of circular dichroism (CD) and isothermal titration calorimetry (ITC) were used. Circular dichroism is a method used to determine the secondary structures of peptides and protein, with noticeable dips and shifts at certain points on the wavelength indicating the presence of alpha helices, beta sheets, and random coils. Isothermal titration calorimetry is a technique used to study how small molecules bind to larger molecules by measuring the absorbance and release of heat during the binding process to determine the favorability of the reaction. For this research, CD was performed on the sample peptide NF-κB WKWK under various concentrations for each run in order to determine its potential secondary structures, and ITC was performed on the sample NF-κB WKWK and the annealed sample κB DNA combined with its complement over the course of three runs under various conditions.

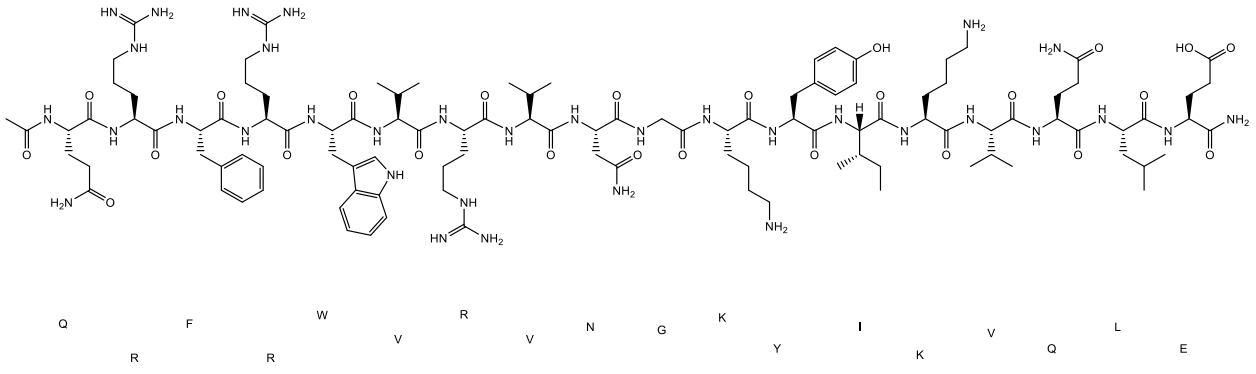

Figure 1: Sample peptide NF-κB WKWK (Ac-QRFRWVRVNGKYIKVQLE-NH_2)

Circular Dichroism of NF-κB WKWK

Procedure

To begin the process of circular dichroism (CD), the researcher first opened the N_2 gas cylinder via the regulator and let it run for 10 minutes. Once the time was up, they then turned on the "isotemp" device and set it to 25.00 ºC, adjusted the N_2 cylinder air flow until it was approximately 120 kPa and 7-8 psi on the gauge, turned on the "Jasco MPTC-490S," and finally turned on J-815 spectrometer itself. From the computer, they opened the "Spectra Manager" app before clicking on "Temperature Interval Measurement" in the menu. They connected the spectrometer via selecting the "Control" tab, then "Select Accessory" and "Auto-Peltier 6-cell Changer," and then they selected the "Measure" tab and clicked "Baseline" to adjust all usable cells to the desired temperature of 25 ºC via selecting them all.

Between the processes of letting the N_2 run and allowing the cells to warm up, the researcher prepared the samples to run. Taking two cuvettes previously soaked in methanol, the researcher cleaned them out via at least six rinses with deionized water. One cuvette was filled with 400 µL of filtered 10 mM Na_2HPO_4 buffer (pH 7.02) to serve as the blank, while the other was filled with the sample peptide NF-κB WKWK (Ac-QRFRWVRVNGKYIKVQLE-NH_2). Said peptide was prepared via approximately 1.5 mg peptide being dissolved in 1000 µL of buffer, with 400 µL transferred to another cuvette. The blank and sample were placed in Cells 1 and 2 respectively of the spectrometer once the cells were at the desired baseline temperature before being run through, while at the same time the researcher determined the concentration of the peptide through a UV-vis spectrometer (Genesys 50) with 20 µL of the remaining

sample. Once the data was run through, it was determined the concentration was too low and possibly even under 1.5 mg due to an insignificant dip at 210 nm.

As there was very little of the peptide left in the container, the extra 580 µL was dumped back into the container and mixed to obtain a higher concentration determined through the UV-vis spectrometer once more with 10 µL of sample before repeating the CD run with 400 µL of the new sample. As the run determined the concentration was too high, the sample was divided in half with 200 µL of buffer being added to half of the sample to cut the concentration. When the next run still determined the concentration too high, 300 µL of peptide was taken and added with 100 µL of buffer to run again, and finally that result was divided in half and had 200 µL of buffer added to receive a final concentration. As the determined concentrations were subject to error due to the spectrometer and formula originally used, the concentration of the new peptide was determined via assuming the total amount of peptide in the container equaled 1.5 mg and worked from there, which was extrapolated to the following peptides to obtain their potential concentrations. As the curves proved binding, such concentrations could be guessed at since the experiment could be repeated with better measured concentrations.

Calculations

Calculation of Concentration for CD from UV-vis spectrometer (Genesys 50)

w/Parameters: 200-400 nm; 2.0 nm; Fast

$$A = \in * l * c \; with \; \in = 230900 \frac{L}{mol * cm} \; and \; l = 1 \; cm$$

Original Peptide Conc: 20 µL peptide, 380 µL buffer with an Abs at 250 nm of 0.012

$$c = \frac{0.012}{230900 * 1} = 5.197 * 10^{-8} M * 10^{6} = 0.05197 \; \mu M * \frac{400 \; \mu L}{20 \; \mu L} = 1.0394 \; \mu M$$

New Peptide Conc: 10 µL peptide, 390 µL buffer with an Abs at 250 nm of 0.056

$$c = \frac{0.056}{230900 * 1} = 2.425 * 10^{-7} M * 10^{6} = 0.2425\ \mu M * \frac{400\ \mu L}{10\ \mu L} = 9.701\ \mu M$$

Calculation of Concentrations for CD from assumed amounts:

New Peptide (assuming it contained 1.5 mg WKWK with Molar Weight of 2360.77 g):

$$1.5\ mg\ WKWK * \frac{0.001\ g}{1\ mg} = 0.0015\ g\ WKWK * \frac{1\ mol}{2360.77\ g} = 6.35 * 10^{-7} mol$$

* If assuming the volume is the 580 µL (0.00058 L):

$$\frac{6.35 * 10^{-7} mol}{0.00058\ L} = 0.001095493\ M * 10^{6} = 1095.493\ \mu M$$

* If assuming the volume is 1000 µL (0.001 L):

$$\frac{6.35 * 10^{-7} mol}{0.001\ L} = 6.35 * 10^{-4}\ M * 10^{6} = 635.386\ \mu M$$

The assumed range of Concentration for the New Peptide was 635 to 1095 µM

Old Peptide: Using the ratio between the values found on the UV-vis spectrometer…

$$(635\ to\ 1095)\mu M * \frac{1.0394\ \mu M}{9.701\ \mu M} = (66\ to\ 113)\mu M$$

New Peptide #2

$$(635\ to\ 1095)\mu M * \frac{200\ \mu L}{400\ \mu L} = (318\ to\ 548)\mu M$$

New Peptide #3

$$(318\ to\ 548)\mu M * \frac{300\ \mu L}{400\ \mu L} = (238\ to\ 411)\mu M$$

New Peptide #4

$$(238\ to\ 411)\mu M * \frac{200\ \mu L}{400\ \mu L} = (119\ to\ 205)\mu M$$

Results

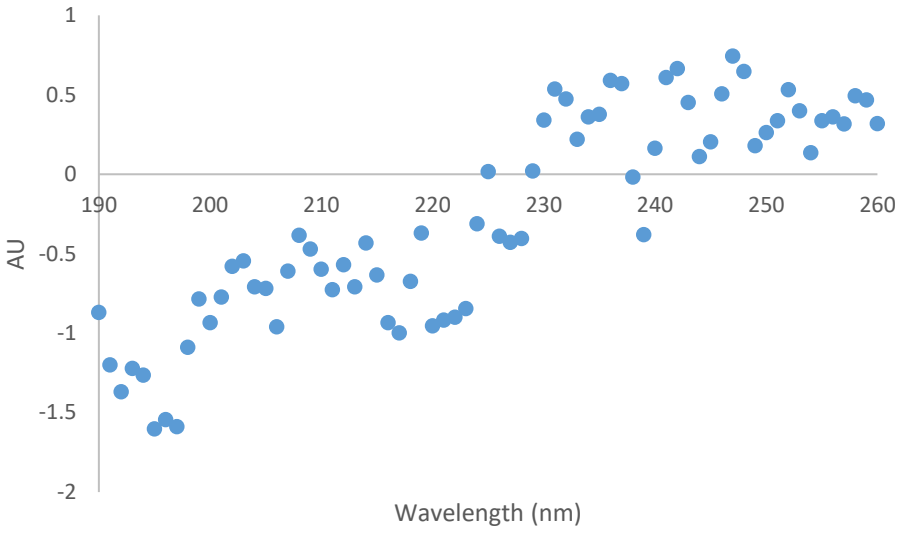

Figure 2: CD spectrometer scan of Cell 1 "Buffer" containing 400 µL of 10 mM Na_2HPO_4, pH 7.02 (filtered) in first scan.

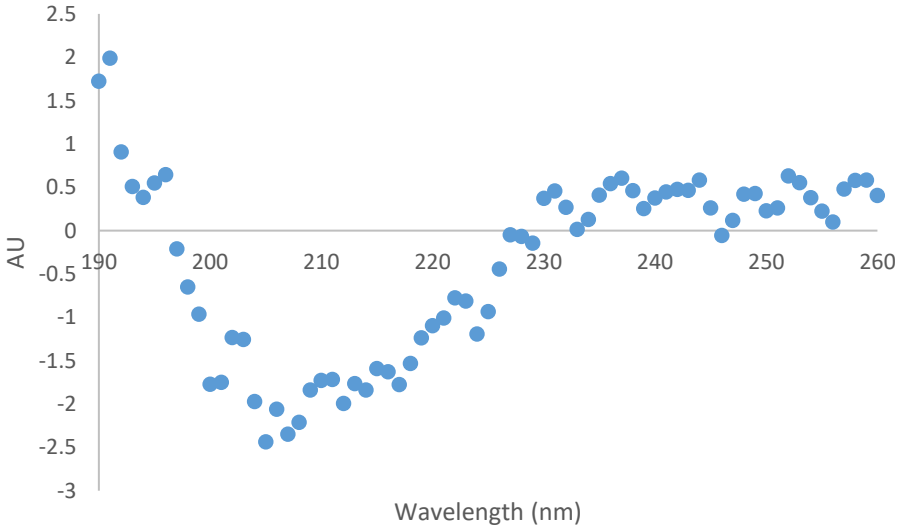

Figure 3: CD spectrometer scan of Cell 2 "WKWK Hairpin" containing 400 µL of 66-113 µM WKWK peptide in first scan.

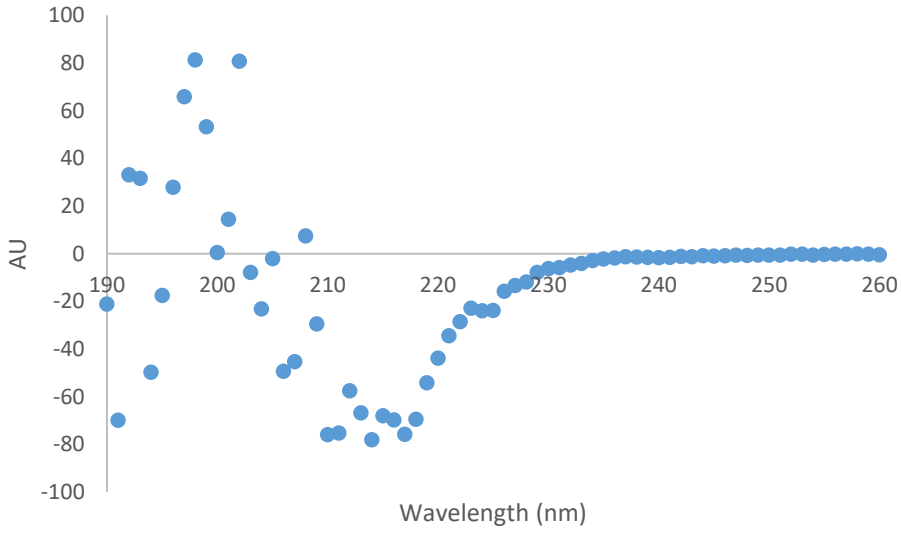

Figure 4: CD spectrometer scan of Cell 2 "WKWK Hairpin" containing 400 µL of 635-1095 µM WKWK peptide in second scan.

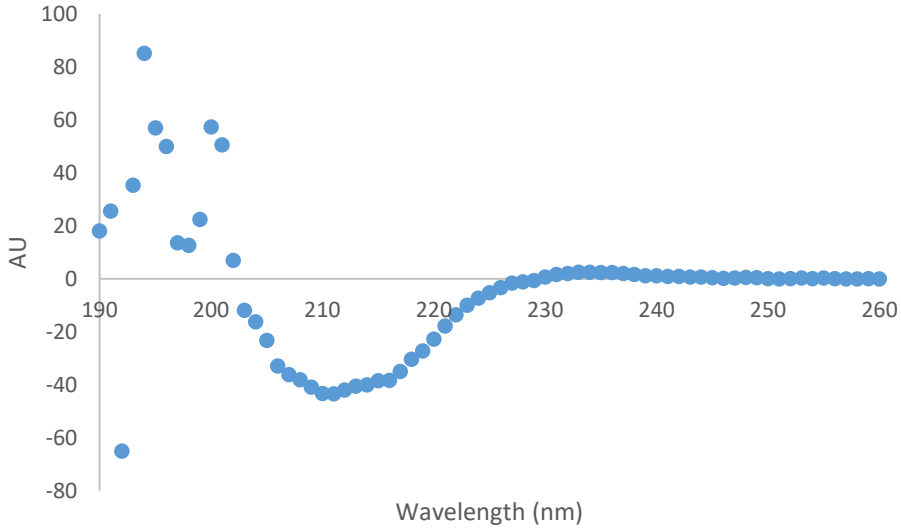

Figure 5: CD spectrometer scan of Cell 2 "WKWK Hairpin" containing 400 µL of 318-548 µM WKWK peptide in third scan.

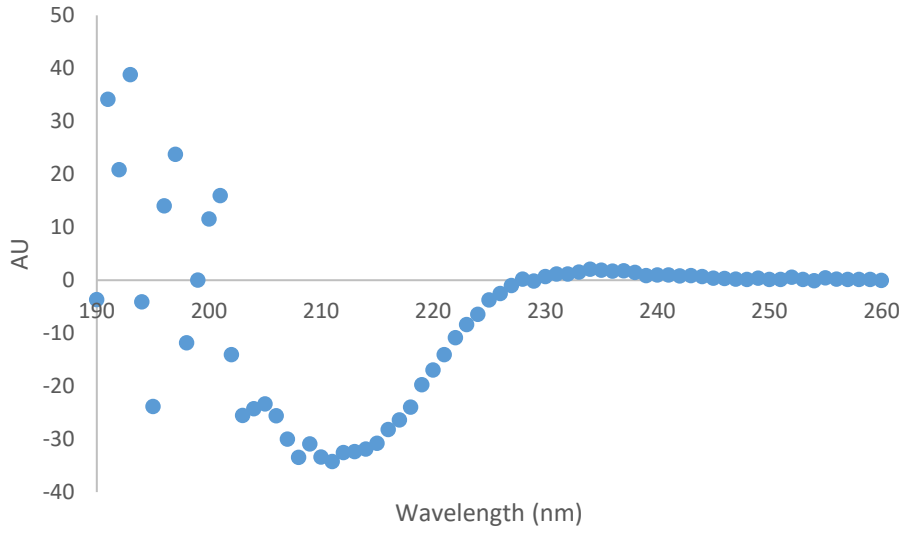

Figure 6: CD spectrometer scan of Cell 2 "WKWK Hairpin" containing 400 µL of 238-411 µM WKWK peptide in fourth scan.

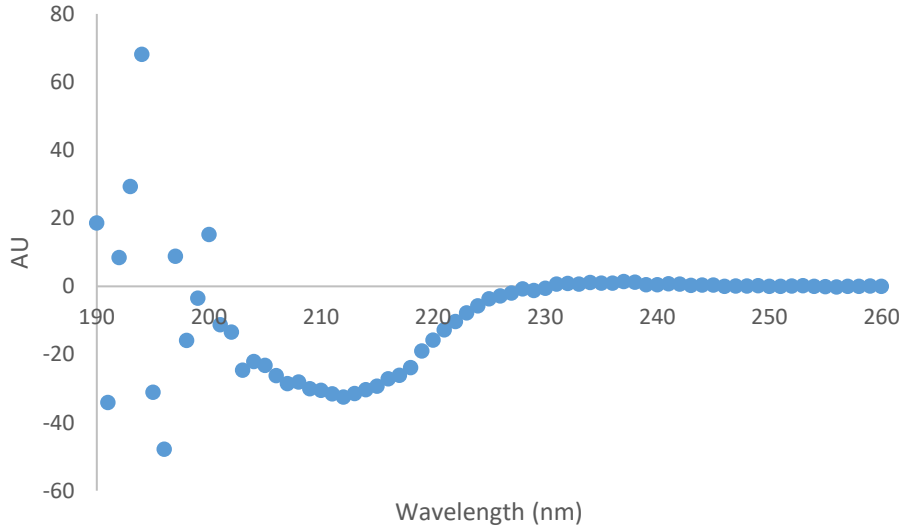

Figure 7: CD spectrometer scan of Cell 2 "WKWK Hairpin" containing 400 µL of 119-205 µM WKWK peptide in fifth scan.

Discussion

For the CD runs, there was a dip in the graphs at 210-215 nm indicating there was a β-sheet in the peptide featured in Figure 1 even in the earliest stages as shown by Figure 3, but it was too insignificant to be truly noteworthy. Increasing the concentration showed a much more noticeable dip in the same area as shown in Figure 4 and cutting the concentration down from there showed a smoother yet still obvious dip as shown by Figures 4-7. While exact concentrations of the peptide were subjective due to a failure to accurately log it before running the experiment, the obvious dip at the necessary wavelength at all runs indicated that the peptide did indeed have β-sheets and the process could be repeated effectively, so performing the experiment again with the same peptide after obtaining a proper concentration would indeed be possible.

Isothermal titration calorimetry of NF-κB WKWK with annealed κB DNA and compliment

Procedure

To begin the process of an isothermal titration calorimetry (ITC) run, the researcher first turned on the instrument and cleaned it. As each of the three runs was the first of the day, they used a Hamilton Syringe to put 300 mL of CONTAD 70 into the instrument's cell before having the instrument run "Detergent Soak and Rinse (Long)" and following the needed prompts, which took approximately 40 minutes to run to completion before extracting and throwing the waste detergent away into the sink. While the instrument was being cleaned, the researcher prepared the sample peptide and DNA.

The sample peptide was already pre-prepared in filtered ITC buffer 10 mM Na_2HPO_4 / 100 mM NaCl (pH 7.10) and its concentration found courtesy of the

NanoDrop in Protein A280 Mode (552.7 µM), which was then made into 500 µL of 5 µM peptide in a 1500 µL micro-centrifuge tube. The sample double-stranded DNA κB DNA 2019 and its complement (both 5'-TGG-GAA-TTC-CCA-3') were pre-prepared and annealed before their concentration was found in the NanoDrop in Nucleic Acid Mode (668.7 µM), which was then made into 60 µL of 500 µM DNA in a 100 µL micro-centrifuge tube.

Once the samples were ready and the instrument cleaned, the researcher went to the "Advanced Experimental Design" tab of the ITC, named the file ("WKWKtoDNA111419"), and set the Experimental Parameters (Total injections 20, Cell temperature (C) 25, Reference power (ucal/sec.)(0-12.65) 8, Initial delay (sec.) 60, Syringe Concentration (mM) 0.5, Cell Concentration (mM) 0.005, and Stirring speed (RPM) 750), and Injection Parameters (The first injection: Volume 0.4, duration 0.8, spacing 300, filter 5; all subsequent injections: Volume 2.0, duration 4.0, spacing 300, filter 5). They then loaded the DNA tube into the "load" area on the instrument, set the syringe to the rest position with the tube hooked up and selected "Syringe Fill" under the "Instrument Controls" tab to follow the directions of loading the DNA into the syringe.

Once completed, they took the Hamilton syringe and cleaned it once with methanol and twice with ITC buffer before drawing up 350 µL of the peptide sample and removing bubbles from the mixture before injecting it to fill the instrument cell. Once finished, they tapped around inside the cell to loosen and pop any spare bubbles before cleaning the Hamilton syringe with methanol and moving the filled DNA syringe into its rest position. They then removed the tube attached to the syringe and connected it to the

back of the instrument before firmly but carefully inserting the syringe into the cell and clicking “Start” to run the experiment through its baseline and the run proper.

After approximately two hours, the run was completed, and the file was analyzed in “MicroCal Analysis Launcher ITC200.” The run was then repeated but with a reduced DNA concentration from 500 µM to 350 µM (“WKWKtoDNA120419”), and then that run was repeated with the same changed concentration but with the temperature increased from 25 ºC to 30 ºC and a different filtered ITC buffer (“WKWKtoDNA120519”).

Calculations

Calculation of Concentrations for ITC Runs

Calculation of Peptide Concentration from Average NanoDrop Protein A280 w/ “5M GuHCl & Peptide” (83 µL of 2.5 mg/mL peptide in 500 µL of solution) and Molar Weight (2360.77 g/mol)

$$\frac{0.2166\ mg}{2360.77\ g} * \frac{1\ g}{1000\ mg} * \frac{1000\ mL}{1\ L} * 10^6 * \frac{500}{83} = 552.7\ \mu M$$

Calculation of DNA Concentration of both KB DNA and its complement (814.5 + 893.9 nmol mixed together with 815 + 894 µL of 10mM Na_2HPO_4 / 100 mM NaCl) from Average ng/µL NanoDrop Nucleic Acid Mode (Conc Factor 62.80) and Molar Weight (both 3645.4 g/mol)

$$\left(\frac{4874.92 \frac{ng}{\mu L} * 10^{-9} g/ng}{\frac{2 * 3645.4\ g}{10^{-6}\ \mu L/L}} \right) * \frac{10^6\ \mu M}{1\ M} = 668.7\ \mu M$$

Peptide and DNA Concentrations for “WKWKtoDNA111419”

- Peptide: $(552.7\ \mu M)(x\ \mu L) = (5\ \mu M)(500\ \mu L) \rightarrow x\ \mu L = 4.5\ \mu L\ \rightarrow$ $5\ \mu L\ peptide, 495\ \mu L\ ITC\ buffer$

- DNA: $(668.7\ \mu M)(x\ \mu L) = (500\ \mu M)(600\ \mu L) \rightarrow x\ \mu L = 44.86\ \mu L\ \rightarrow$ $45\ \mu L\ DNA, 15\ \mu L\ ITC\ buffer$

Peptide and DNA Concentrations for "WKWKtoDNA120419"

- Peptide: $(552.7\ \mu M)(x\ \mu L) = (5\ \mu M)(500\ \mu L) \rightarrow x\ \mu L = 4.5\ \mu L\ \rightarrow$ $5\ \mu L\ peptide, 495\ \mu L\ ITC\ buffer$
- DNA: $(668.7\ \mu M)(x\ \mu L) = (350\ \mu M)(600\ \mu L) \rightarrow x\ \mu L = 31.40\ \mu L\ \rightarrow$ $31\ \mu L\ DNA, 29\ \mu L\ ITC\ buffer$

Peptide and DNA Concentrations for "WKWKtoDNA120519"

- Peptide: $(552.7\ \mu M)(x\ \mu L) = (5\ \mu M)(500\ \mu L) \rightarrow x\ \mu L = 4.5\ \mu L\ \rightarrow$ $5\ \mu L\ peptide, 495\ \mu L\ ITC\ buffer$
- DNA: $(668.7\ \mu M)(x\ \mu L) = (500\ \mu M)(600\ \mu L) \rightarrow x\ \mu L = 31.40\ \mu L\ \rightarrow$ $31\ \mu L\ DNA, 29\ \mu L\ ITC\ buffer$

Results

Table 1: Comparison of the equilibrium association constants (K_a) and equilibrium dissociation constants (K_d) across the ITC runs. The number of data points removed from the original runs for being assumed as "bad data" to generate the altered runs are listed.

File Run	K_a (M^{-1})	K_d ($1/K_a$) (M)
11/14/19 Original	5.03E4 ± 6.44E4	1.99E-5 ± 1.55E-5
11/14/19 Altered (1 Pt)	1.18E4 ± 3.50E4	8.47E-5 ± 2.86E-5
12/04/19 Original	6.22E5 ± 1.86E6	1.61E-6 ± 5.38E-7
12/04/19 Altered (3 Pts)	1.25E5 ± 1.45E5	8.00E-6 ± 6.90E-6
12/05/19 Original	5.14E3 ± 4.92E4	1.94E-4 ± 2.03E-5
12/05/19 Altered (2 Pts)	5.81E3 ± 3.41E4	1.72E-4 ± 2.93E-5

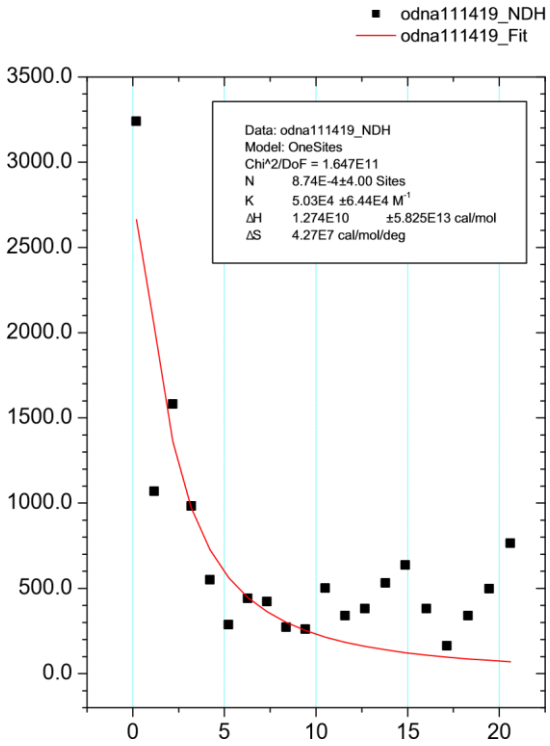

Figure 8: Original ITC run "WKWKtoDNA111419."

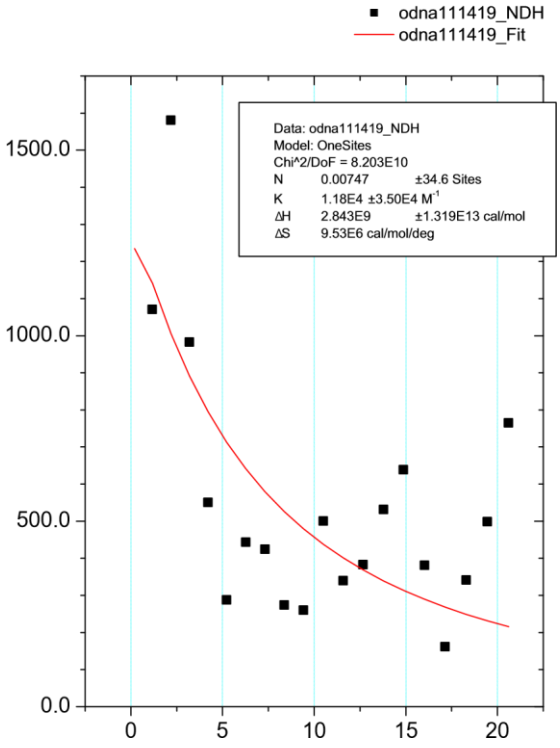

Figure 9: Altered ITC run "WKWKtoDNA111419" with one data point removed.

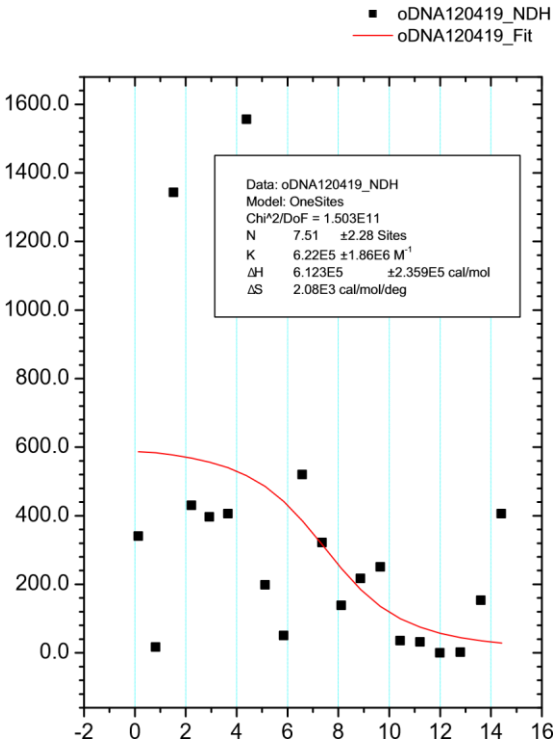

Figure 10: Original ITC run "WKWKtoDNA120419."

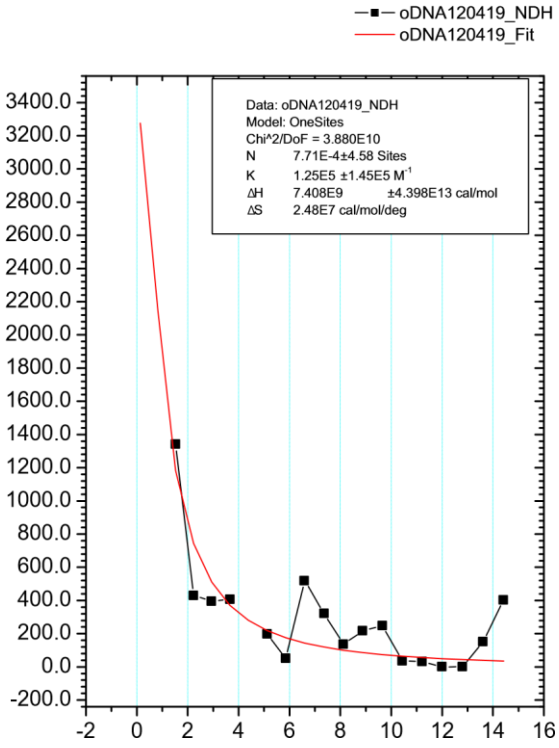

Figure 11: Altered ITC run "WKWKtoDNA120419" with three data points removed.

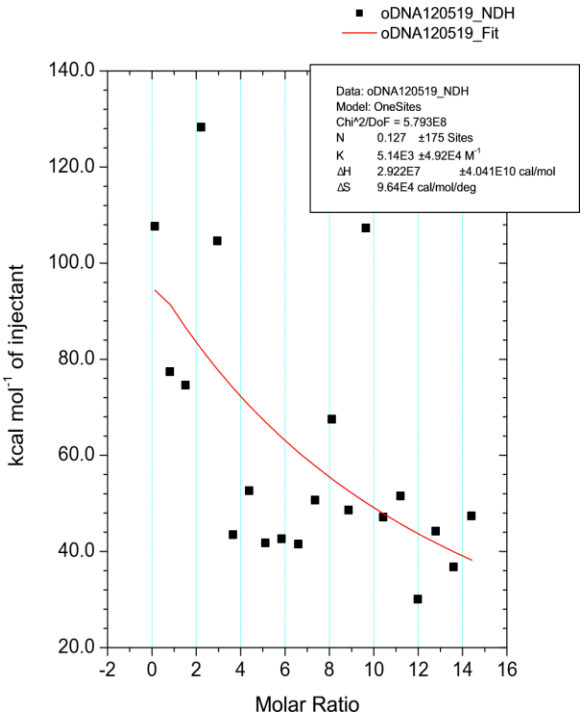

Figure 12: Original ITC run "WKWKtoDNA120519."

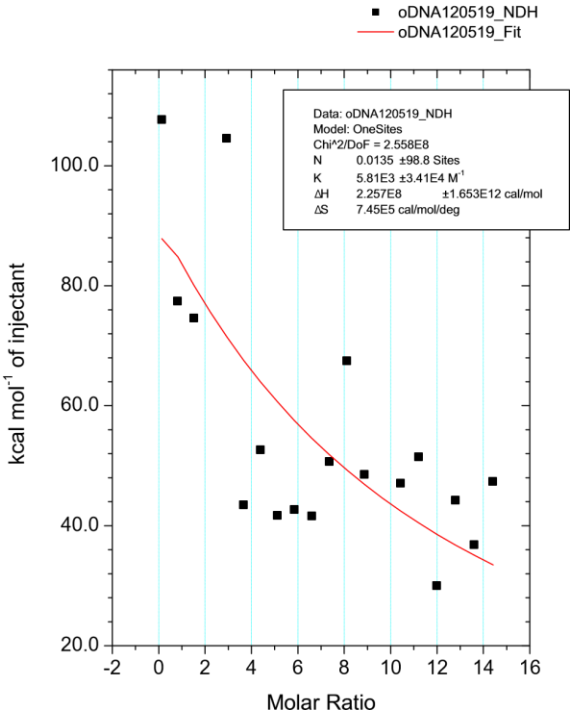

Figure 13: Altered ITC run "WKWKtoDNA120519" with two data points removed.

Discussion

For the ITC runs, the major goals were to have the points close together, to have the trend line possess a noticeable exponential decay trend, and to obtain a low-enough K_d on the line to indicate proper binding of the peptide with the DNA, with that range being preferably in the low micro-molar scale. As shown in Figures 8 and 9, the runs on 12/04/19 did possess the desired exponential decay, but the points were further apart from each other and the K_d values were larger than desired as can be seen in Table 1. Lowering the DNA concentration, as shown in Figures 10 and 11, showed marked improvement in all desired areas with closer-together points, a more obvious decay, and even pulled the K_d values into the desired micro-molar range as shown in Table 1. The runs on 12/05/19, as shown in Figures 12 and 13 and Table 1, pulled back on that somewhat, but they were still better than the runs on 11/14/19 with more obvious decay and closer-together points. While some of the changes might be attributed to the increase in temperature, there was also the possibility of the different filtered ITC buffer being a factor since the original was frozen and could not be used. Improvements on the process could still be made possible, however, with the standardization of the types of buffer to be used across runs and better accustomization to the process decreasing the chances of human error.

Overall Conclusions

As shown in both the CD and ITC, the sample peptide NF-κB WKWK possesses many of the desired traits wanted in a protein that binds NF-κB DNA. The CD shows the dip at the necessary wavelength to indicate presence of the β sheet secondary structures required. The ITC meanwhile shows there is definite bonding between the WKWK peptide and the κB DNA due to the K_d values on display, though more testing and better

conditions would certainly be required to get better answers. From this data, it would be prudent to create other peptides with similar structures to NF-κB WKWK to compare their binding affinity to κB DNA.