Structural studies of insect cytokine-like Stress Responsive Peptide 4

by

Hannah Miller

A THESIS

submitted in partial fulfillment of the requirement for the degree

MASTER OF SCIENCE

Department of Biochemistry and Molecular Biophysics

College of Arts and Sciences

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2023

Approved by:

Major Professor

Om Prakash

Copyright

© Hannah Miller 2023

Abstract

Similar to innate immunity in vertebrates, insects rely on a variety of both humoral and cellular responses to defend themselves from pathogen invasion. A family of peptides has been identified in the hemolymph of six orders of insects that function like cytokines to regulate immune responses. This family of peptides is activated during times of stress - such as pathogen invasion, wounding, and heat or cold shock – and are rightfully named Stress Responsive Peptides (SRPs). The SRP family plays a variety of roles within the insect immune system making each of their structure-activity relationships compelling. Mature *Manduca Sexta* Stress Responsive Peptide-4 (SRP-4) is a 26-residue peptide stabilized with a disulfide bond between residues 7C and 18C. Its structure was investigated through 2D ^{1}H-^{1}H Nuclear Magnetic Resonance (NMR) Spectroscopy and Nuclear Overhauser Effect (NOE) derived constraints. We observed that SRP-4 contains two short beta strands between residues 10G-12V and 15G-18C, which are connected by a beta-turn. The N- and C-termini of SRP-4 are intrinsically disordered. SRP-4 bears structural similarity to the previously studied SRP-1 and -2. The structural studies of this peptide also revealed a resemblance to the C-terminal subdomain of the human epidermal growth factor (EGF). The solved structure of this peptide within the stress responsive peptide family aids in our understanding of their unique roles in the immune system.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Thank you, Dr. Om Prakash, for your encouragement and guidance throughout my time at Kansas State - I appreciate all your help. Thank you to the NMR lab group, Nitin Mishra, Andy Su, and Tomohiro Kimura for all assistance you have provided. A big thank you to Nitin for always being available to answer my questions and teach me new techniques.

Thank you to Dr. Ho Leung Ng for accepting me into your lab after my freshman year and allowing me to begin my research journey. Thank you to my team member and friend, Mian Huang, for sharing with me your knowledge and love for science. I will be forever grateful for all you have taught me.

Thank you to my committee member, Dr. Michal Zolkiewski, for sharing your advice on this thesis and assisting me as I complete my graduate studies. Thank you, Dr. Haobo Jiang, of Oklahoma State University, for your support in obtaining the peptide.

A final and generous thank you to my parents, Arnold and Esther Miller, who have continuously supported me. Without you, my success would not be possible.

Structural Studies of Insect Cytokine-Like Stress Responsive Peptide-4

Introduction

Cytokines are communicatory proteins that aid in cell-to-cell messaging within the immune system. They exist in both vertebrates and invertebrates and are responsible for inducing immune responses such as cell proliferation, differentiation, and chemotaxis (Schrag *et al*., 2017). While hundreds of cytokines have been discovered and characterized in mammals, the study of cytokines in insects is relatively novel. Cytokine-like peptides have been identified in the hemolymph of six orders of insects: Lepidoptera, Diptera, Hymenoptera, Coleoptera, Hemiptera, and Orthoptera (Shears *et al*., 2018). The first peptide discovered was the growth-blocking (GBP) peptide in *Mythimna separata*, the armyworm, recognized for its ability to protect the insect against parasites by delaying host larval growth (Hayakawa *et al*., 1991). Shortly after, the paralytic peptide (PP) was characterized in *Manduca sexta*, the tobacco hornworm, after it induced paralysis upon injection into the larva of the same species (Skinner *et* al., 1991; Wang *et* al., 1999). Since these discoveries, several cytokine-like peptides have been isolated that appear after stressful conditions, much like the GBP and PP. As their functions within the immune system continue to be investigated, the newly discovered homologous peptides are grouped into the family of Stress Responsive Peptides (SRPs) (Schrag *et al*., 2019).

SRPs are secreted by the fat body, into the hemolymph, as unstructured, inactive, proteins around 60-150 residues long; they are called pre-pro-peptides (Wang *et al*., 1999; Yamaguchi *et al.*, 2012; Matsumoto *et al*., 2012). The pre-region is cleaved by serine proteases within the hemolymph. The pro-peptide then awaits activation as an extracellular precursor, or it can be stored within a secretory vesicle (Nakatogawa *et al*., 2009). Upon introduction of a stressor, the C-terminal region of the pro-peptide is cleaved by a serine protease – the exact pathway from stress induction to cleavage is still under investigation (Wang *et al.*, 2022). The cleaved C-terminus gives rise to the mature, active, SRP. The active SRPs appear to take on different roles within the immune system including wound healing, antimicrobial peptide (AMP) activation, and hemocyte proliferation (Yamaguchi *et al.*, 2012).

Active SRPs are 23 to 32 residues in length, consisting of groups of conserved residues within their sequences (Schrag *et al.*, 2017; Matsumoto *et al.*, 2012) (Figure 1). Among the conserved residues are two cysteines that form a disulfide bond, giving the peptides structural constraint within their core region (residues 8-21) (Matsumoto *et al.*, 2012). This core region is more conserved in sequence (C-X(2)-G-X(4,6)-G-X(1,2)-C-[K/R]) than the N- and C-termini. Previous NMR studies found that a β-hairpin exists within this constrained core region with the N- and C-termini being intrinsically disordered (Schrag *et al.*, 2019; Schrag *et al.*, 2017). The β-hairpin motif contains alternating hydrophobic residues and is predicted to aid in binding to the SRP receptors (Clark *et al.*, 2001). The N-terminal extension, existing before the first cysteine, is important for recognition and activation by proteases; it also plays an imperative role in the function of the plasmatocyte spreading peptide. The C-terminal extension, containing acidic residues and a hydrophobic tyrosine, is thought to have varying functions between peptides. Through mutational studies, it was revealed the N-terminus may aid in stabilizing the peptide structure and maintaining biological activity (Dembele, 2021).

```
MsSRP1    FGVRVGTCPSGYVRR GTFCFPDDDY
MsSRP2    FGVKDGKCPSGRVRRLG ICVPDDDY
MsSRP4    SL KRGTCPDGRVLF GRVCVEIDSLDY
```

Figure 1 Aligned Sequences of MsSRP-1, -2, and -4

The structural studies of *Manduca sexta* SRP-1 and -2 were performed by past members in our lab (Schrag *et al.*, 2019; Schrag *et al.*, 2017). They closely investigated the β-hairpin structure of the peptides using 2D ^{1}H-^{1}H NMR spectroscopy and NOE derived distance constraints. SRP-1 is a 25-residue peptide containing β-strands from resides Y12-R15 and F18-F20 with a type-II β-turn from R15-F18 (Schrag *et al.*, 2019). The structure is stabilized by a disulfide bond between C8 and C19. SRP-2 is a 25-residue peptide containing β-strands from residues R12-R15 and I18-V20 with a type-I' β-turn from R15-I18 (Schrag *et al.*, 2017). The structure is stabilized by a disulfide bond between C8 and C19. Both peptides contain anti-parallel β-strands with four and three residues per strand, connected by a two residue turn.

As additional structural research is done on insect cytokines the degree of similarity which they share with mammalian cytokines becomes more clear. Mammalian Epidermal Growth Factor (EFG) contains a sequence motif much like that of the conserved SRPs (Matsumoto *et al.*, 2012) (Figure 2). EGF also contains a β-hairpin within its C-terminal subdomain appearing similar in structure to the previously described SRP-1 and -2. A previous study on GBP revealed that it could bind and activate EGF receptors within human keratinocytes, causing cell proliferation (Matsumoto *et al.*, 2012). The sequential and structural similarities combined with the ability of an insect peptide to activate the human receptor raise the hypothesis that SRPs and EGF may be ancestrally related.

```
Core region of insect cytokines    C-X(2)-G-X(4,6)-G-X(1,2)-C-[KR]
EGF C-terminal subdomain                -C-V-V-G-Y-I-G-E-R-C
```

Figure 2 Alignment of the conserved core region of insect cytokines and the EGF C-terminal subdomain

In this chapter, I continue to advance the study of SRPs by analyzing the structure of *Manduca Sexta* SRP-4 using 2D ^{1}H-^{1}H NMR correlation spectroscopy and NOE derived distance constraints. SRP-4 is a 26-residue peptide stabilized with a disulfide bond. SRP-4's sequence is 7.7% identical and 46.2% similar to SRP-1 and 3.8% identical and 50% similar to SRP-2. These homologies are dissimilar to the previously studied peptides, SRP-1 and -2, which are 67.5% identical and 75% similar. This raised the question of whether SRP-4 maintained the conserved structure previously found within this family of peptides.

From my studies, I found that, despite sequential differences, the structure remains conserved between the studied SRP homologs. Analysis of the NMR data revealed SRP-4 contains two anti-parallel β-strands connected by a β-turn. The solved structure of the peptides within this family will allow for further structure-activity relationship studies providing insight into their unique roles in the insect immune system.

Materials and Methods

Peptide Synthesis

The 26-residue MsSRP-4 (SLKRGTCPDGRVLFGRVCVEIDSLDY) was chemically synthesized (acquired from PepMic Ltd, Jiangsu, China) by automated protocols of step-wise solid phase synthesis using the standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry. The peptide was purified to >95% purity by reverse-phase HPLC on a C18 column. Its mass and purity were further confirmed by the matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Applied Biosystems, Foster City, CA). SRP-4 was lyophilized and stored as a dry powder until use for NMR spectroscopy experiments.

Circular Dichroism

To predict the secondary structure of SRP-4 a circular dichroism analysis was performed with a Jasco-8155 spectropolarimeter (JASCO, Tokyo, Japan) using a 1mm path length cell over 190 to 250 nm range. The spectra were acquired every 1 nm, with a 2-second averaging time per point, at 25°C. I prepared the SRP-4 samples to a concentration of 50 μM in pure water with increasing volume/volume percentages (0, 10, 20, 30, 40, 50%) of 2,2,2-trifluoroethanol (TFE).

Nuclear Magnetic Resonance Spectroscopy

To prepare SRP-4 for NMR experiments I dissolved the peptide in 10% D_2O/ 90% H_2O with a final concentration of 2.2 mM. The sample was then placed in a 3 mm Wilmad 335 NMR sample tube for NMR experimentation. The 2D ^1H-^1H NMR experiments were performed on a 14 Tesla Bruker AVANCE NEO (Bruker Biospin Corporation, Billerica, MA) NMR spectrometer, operating at 600MHz for proton, using a 5mm Pulse Field Gradient inverse detection triple resonance HCN cryogenic probe. The invariant nature of the NMR chemical shifts and line widths upon 10-fold dilution indicated that SRP-4 was monomeric in solution at the concentration used for 2D NMR analysis. To collect proton resonance information, I performed a standard suit of a 2D ^1H-^1H Correlation NMR Spectroscopy experiments including 2D ^1H-^1H Correlation NMR Spectroscopy (COSY), 2D ^1H-^1H Total Correlation NMR Spectroscopy (TOCSY), and a 2D ^1H-^1H Nuclear Overhauser Effect Spectroscopy (NOESY). A total of 512 increments of 4K data points were collected for these experiments. NOESY experiments were performed with 150, 300, and 400 ms mixing times. 2D ^1H-^1H TOCSY spectra were recorded by using MLEX-17 for isotropic mixing for 80ms at a B1 field strength of 8 KHz. Water peak suppression was obtained by low-power irradiation of the water peak during relaxation delay. 2D ^1H-^{15}N Heteronuclear Single Quantum Coherence (HSQC) and ^1H-^{13}C HSQC NMR spectra were also recorded to obtain additional information on amide groups and proton-bearing carbons, respectively. The data sets were acquired at a temperature of 25°C. To collect information on hydrogen bonds, I performed a temperature gradient experiment by collecting 2D ^1H-^1H TOCSY data at 12.5, 15, 20, 25, 30, and 35°C respectively. I processed all data using Buker Topspin 4.1.4. Analysis of all spectra and resonance assignments was done using the NMR spectrum analysis software CARA (http:://cara.nmr.ch) (Keller, 2004). The ^1H chemical shifts were referenced to the external standard 2,2-dimethyl-2-silapentane-5-sulphonic acid (DSS) at 25°C. ^{15}N and ^{13}C chemical shifts were referenced indirectly from DSS.

Chemical Shift Assignments and Structure Calculations

I first assigned the proton chemical shifts of the residue side chains using the 2D ^1H-^1H TOCSY spectra. I used the ^1H-^{15}N HSQC to identify amide proton peaks which were overlapped in 2D ^1H-^1H TOCSY spectra. This is done by identifying each residue's characteristic NH chemical shift. If any side chain proton peaks were still unclear, I analyzed their corresponding

CH chemical shift on the 2D ^{1}H-^{13}C HSQC spectrum to correctly identify which residue the proton peak belonged to. The chemical shifts of the beta carbons of 7C and 18C in 2D ^{1}H-^{13}C HSQC were used to confirm the disulfide bond. Next, I conducted the sequential residue assignment by overlaying cross peaks of the 2D ^{1}H-^{1}H COSY and NOESY spectra acquired under similar conditions. I was able to identify sequential residues by following the i residue NH/Hα COSY peak to the i+1 residue Hα/NH NOESY peak (Figure 3). All NOE cross-peaks were assigned using the 2D ^{1}H-^{1}H NOESY spectrum acquired with the 400ms mixing time. NOE cross-peaks arise when two protons are close together in space. The volume of the cross-peak is directly related to the proton-proton distance – the closer the protons, the greater the peak volume. The volume of the NOE-cross peaks was then used to determine proton-proton distance constraints.

I used all assigned chemical shifts, including those from the 2D ^{1}H-^{1}H COSY, TOCSY, NOESY, ^{1}H-^{15}N HSQC, and ^{1}H-^{13}C HSQC, for structure determination. The distance constraints obtained from the NOESY spectrum recorded at 400 ms mixing time were used for the structure calculations. The structure calculations were performed using CYANA 2.1 software. CYANA is a computer program that arranges atoms in space based on the given atom-atom distances and protein amino acid sequence. CYANA obtains the proton-proton distance values from the 2D ^{1}H-^{1}H NOESY peak volumes. The NOESY peak volumes come from assigned NOE peaks and are collected using the CARA software. In addition to NOE based distance constraints, I included the hydrogen bonds and disulfide bond distance constraints. The hydrogen bonds were identified from the temperature gradient experiments and the disulfide bond was identified from the 2D ^{1}H-^{13}C HSQC experiment. I set each round of calculations to generate 200 conformers from which the 20 lowest energy structures were selected and analyzed with PyMOL. These structures were analyzed for any restraint violations and corrected until the latest round of structure calculations generated no violations greater than 0.3Å.

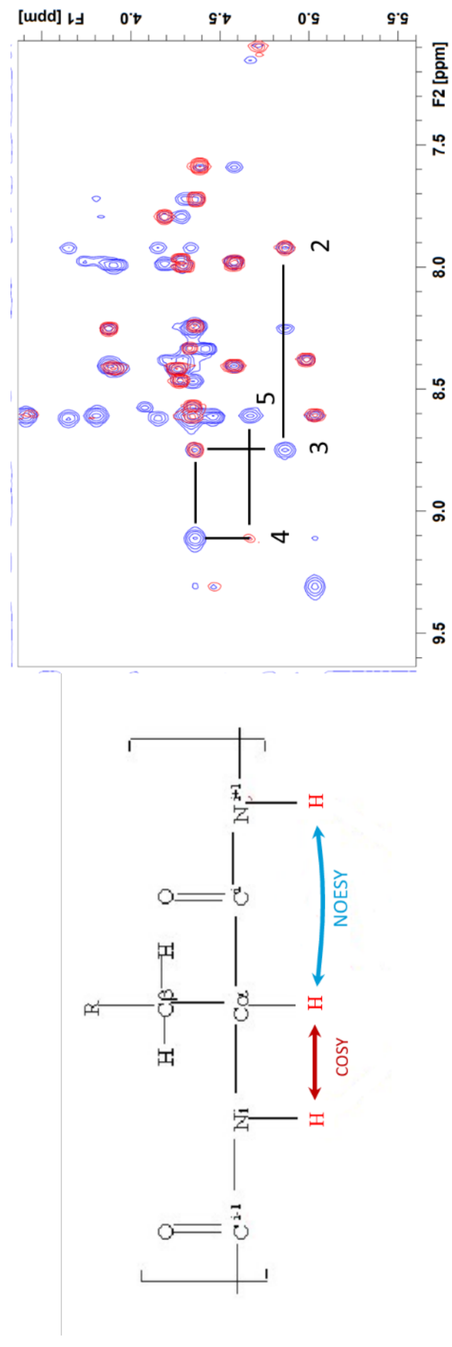

Figure 3 A pictorial representation of how the overlapping of 2D ^{1}H-^{1}H COSY and NOESY spectra are used to assign sequential residues. The COSY spectrum is in red and the NOESY spectrum is in blue. The numbers represent sequential residues.

Results and Discussion

NMR results and Chemical Shift Assignments

I identified the amino acid resonances and assigned their proton chemical shifts using the fingerprint region of the 2D ^1H-^1H TOCSY (Table 1 & Figure 4). The fingerprint region contains the NH to side chain protons with the NH protons existing between 6.5 – 10 ppm in the first dimension and the side chain protons existing between 0 – 6 ppm in the second dimension (Brinson *et al.*, 2019). All amino acid resonances can be found within this region except for proline, due to it lacking an amide proton, and the first residue, serine, due to the exchange of the amide proton with the solvent. To identify the proton chemical shifts of S1 and P8, I utilized the crosspeaks from 0 – 4.5 ppm in the first and second dimensions of the 2D ^1H-^1H TOCSY (Figure 5). This region contains the peaks of aliphatic side chains and the Hα, β, δ, and γ of Proline. Any other amino acid resonances which were unclear or overlapped within the fingerprint region were made clear by the assignment of the NH peaks in the 2D ^1H-^{15}N HSQC (Figure 6).

I determined SRP-4's amino acid sequential connectivity using the $d_{\alpha N}(i, i+1)$ proton connectivities as mentioned in the methods. Further structural information was provided by assigning short, medium, and long-range proton-proton connectivities, including $d_{\alpha N}$, $d_{\beta N}$, and d_{NN}, on the 2D ^1H-^1H NOESY spectra. I observed a number of non-sequential connectivities between residues that were characteristic of antiparallel beta-strands (Figure 7). To gain more information on the connectivity between the two beta-strands I analyzed the 2D ^1H-^1H TOCSY temperature gradient experimental data. The amino acid resonances whose proton chemical shifts were not altered by an increase in temperature indicated that they were located in a nonflexible, hydrogen-bonded, conformation. It was determined that G10, R11, V12, G15, V17, and C18 were all involved in hydrogen bonding as indicated by their nonfluctuating proton chemical shifts from 15 to 35°C (Figure 8). This observation suggested that these residues were involved in antiparallel beta-strands.

To confirm the presence of the disulfide bond predicted to exist within SRP-4, I analyzed the C_β chemical shifts of C7 and C18 within the 2D ^1H-^{13}C HSQC spectrum. The C_β shifts of oxidized (S-S) cysteines will be more down-field with chemical shift values between 33ppm and 51ppm (Sharma *et al.*, 2000). Whereas reduced (S-H) cysteines will be more up-field with chemical shift values between 24ppm and 34ppm. The C_β chemical shifts of C7 and C18 are 36.06ppm and 37.67ppm, respectively. These chemical shift values confirmed the presence of the disulfide bond, and this connectivity was included in structure calculations.

Table 1 SRP-4 Proton Chemical Shift (ppm) Assignments

Residue	HN (ppm)	HA (ppm)	HB (ppm)	HD and others (ppm)
1SER	-	4.063	3.876	-
2LEU	8.569	4.34	1.52	**Hδ** 0.807
3LYS	8.46	4.263	1.732, 1.676	**Hγ** 1.34 **Hε** 2.902
4ARG	8.413	4.26	1.726	**Hδ** 3.11 **Hγ** 1.586, 1.582 **Hε** 7.083
5GLY	8.405	3.907	-	-
6THR	7.982	4.568	4.183	**Hγ** 1.104
7CYS	7.582	4.377	2.956, 2.843	-
8PRO	-	4.408	2.265, 1.957	**Hδ** 3.313 **Hγ** 1.832, 1.737
9ASP	8.33	4.33	2.67	-
10GLY	8.607	4.314	-	-
11ARG	8.599	5.026	1.833	**Hδ** 3.41 **Hγ** 1.673
12VAL	9.304	4.454	2.061	**Hγ** 0.882, 0.771
13LEU	7.787	4.18	1.5, 1.33	**Hδ** 0.714, 0.627
14PHE	8.599	4.33	3.102, 2.979	**Hδ** 7.14, 7.24 **Hε** 7.19, 7.24
15GLY	8.599	4.145, 3.642	-	-
16ARG	7.913	4.86	1.656	**Hδ** 2.825 **Hγ** 1.470, 1.265
17VAL	8.741	4.353	1.866	**Hε** 0.747, 0.704
18CYS	9.147	4.671	3.02, 2.783	-
19VAL	7.716	4.347	2.053	**Hγ** 0.878
20GLU	8.606	3.797	2.034, 1.938	**Hγ** 2.564, 2.364
21ILE	8.241	3.866	1.654	**Hδ** 0.578 **Hγ** 1.011, 1.302, 0.781
22ASP	8.373	4.972	2.687	-
23SER	7.966	4.572	3.822, 3.728	-
24LEU	8.24	4.357	1.65	**Hδ** 0.686, 0.578 **Hγ** 1.551
25ASP	8.397	4.57	2.786	-
26TYR	7.975	4.283	2.706, 2.599	**Hδ** 6.97 **Hε** 6.72

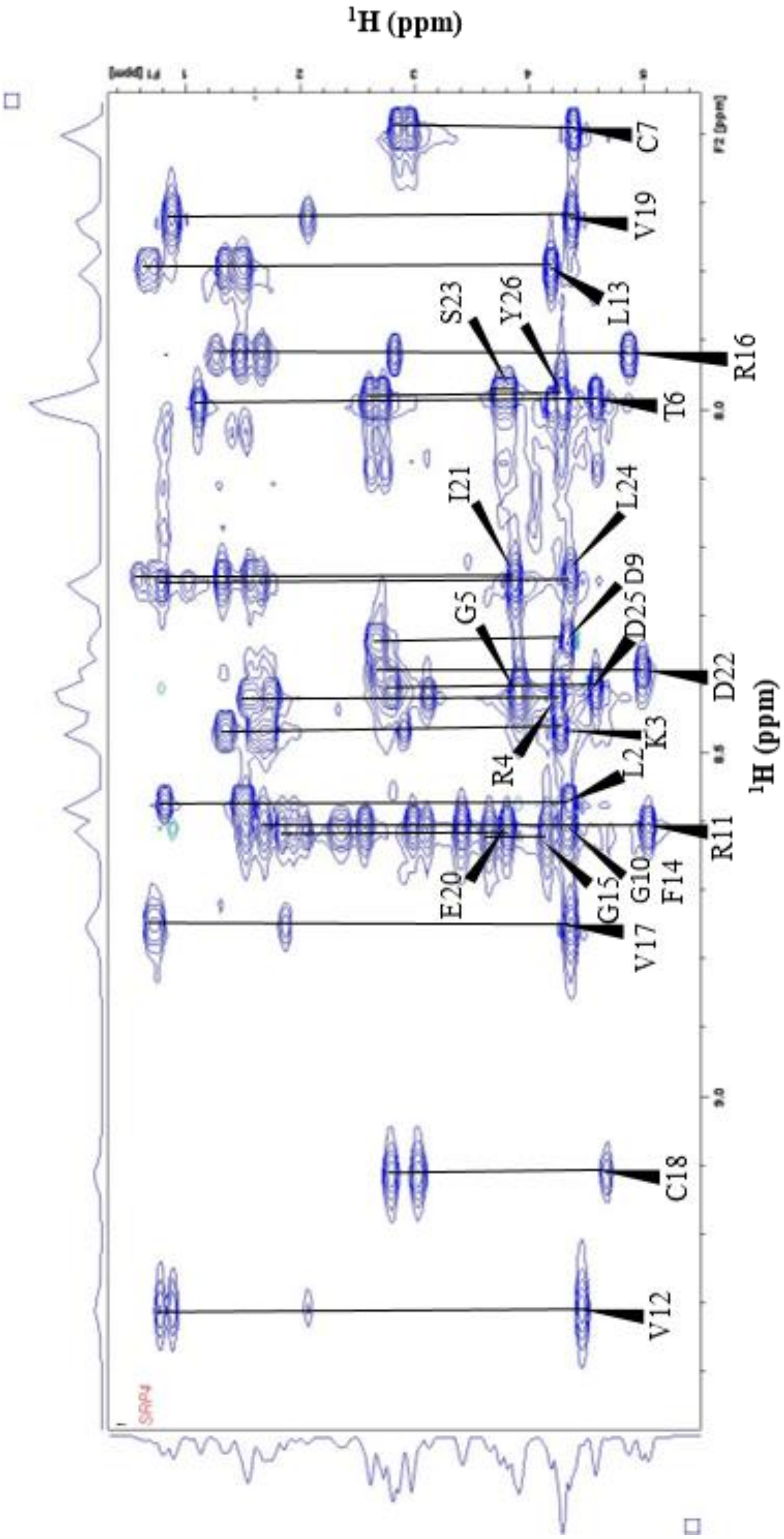

Figure 4 2D ^{1}H-^{1}H TOCSY Fingerprint region representing NH-CαH, NH-CβH, NH-CγH, NH-CδH and NH-CεH cross peaks. Arrows point to Hα-HN cross peak of the corresponding residue.

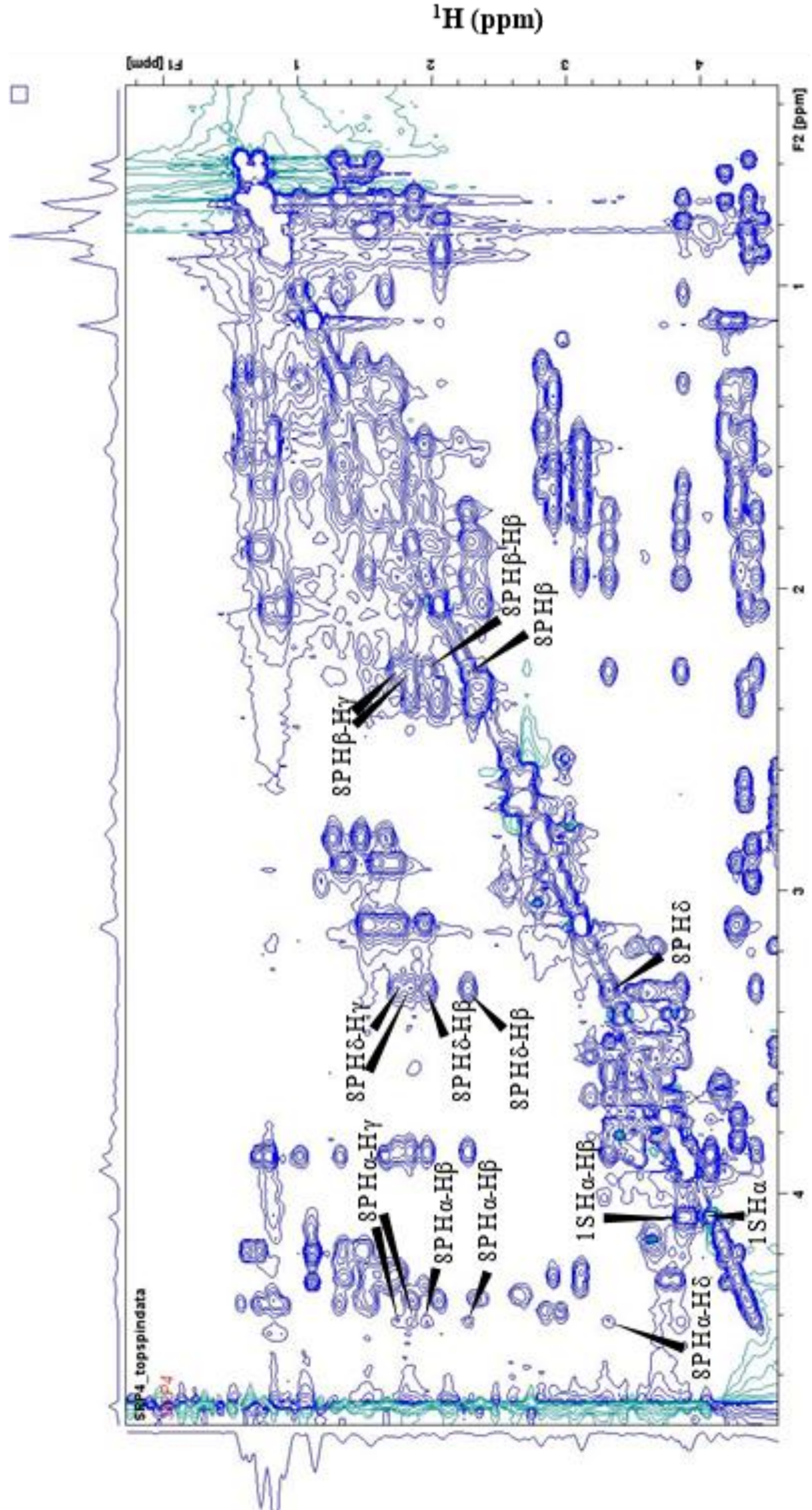

Figure 5 Aliphatic and Proline sidechain region of the 2D ^{1}H-^{1}H TOCSY. The chemical shifts of P8 and S1 are identified within this region.

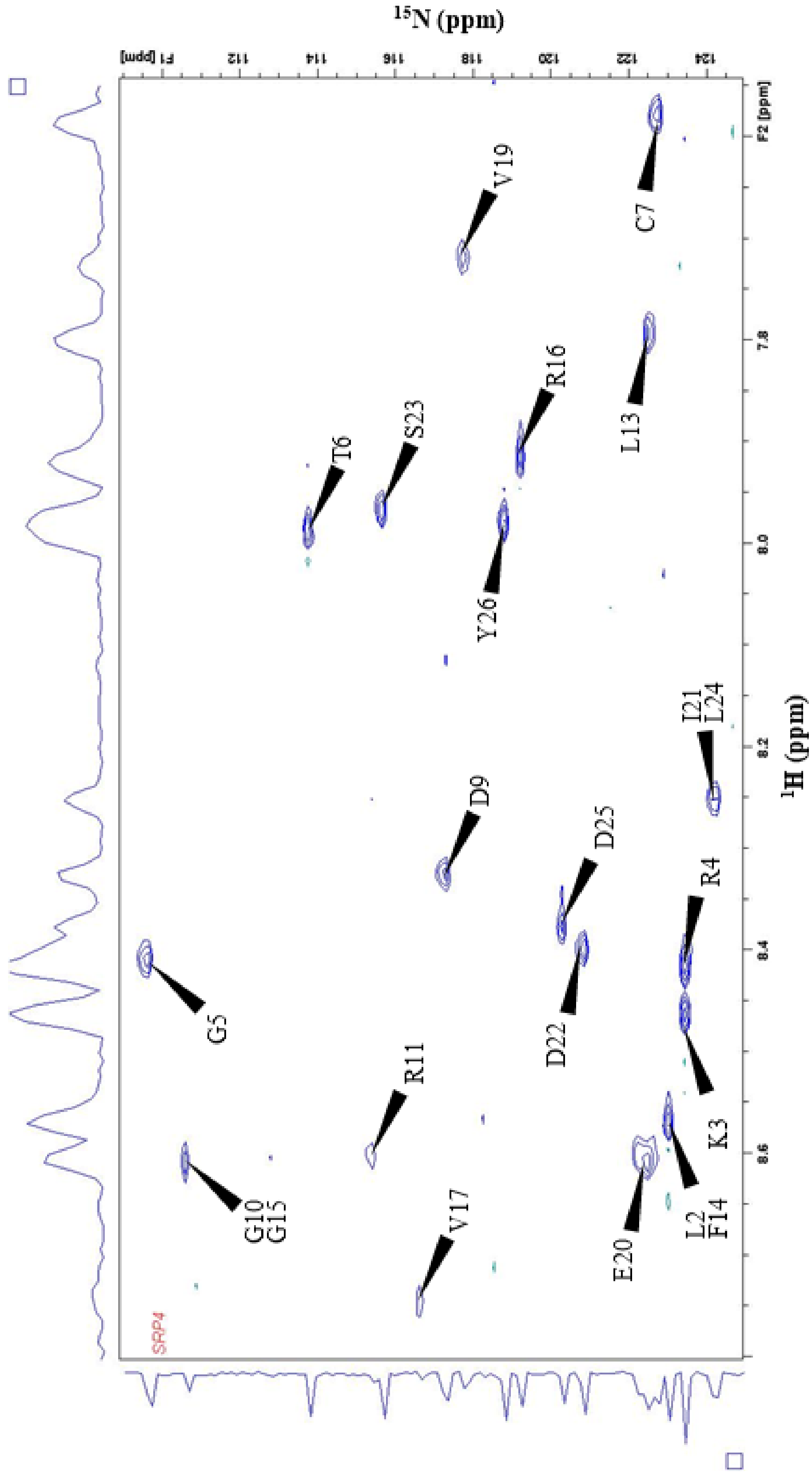

Figure 6 2D ^{1}H-^{15}N HSQC. Peaks indicate the NH of the corresponding residue.

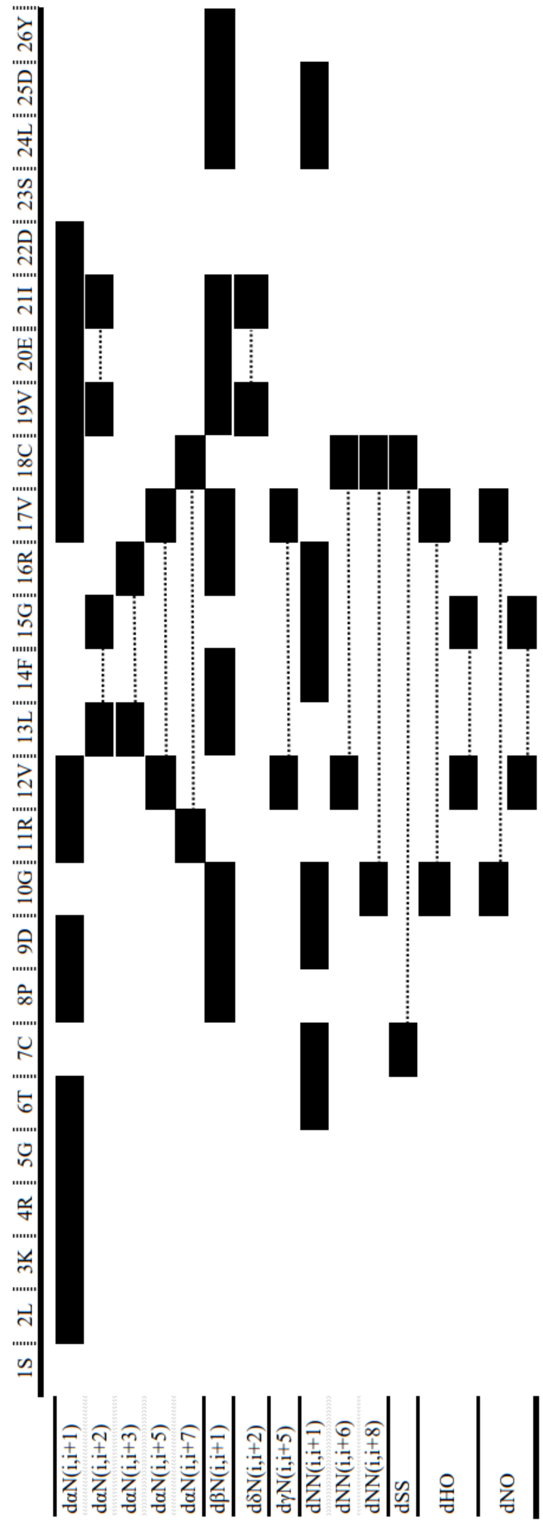

Figure 7 Summary of NOE connectivities for SRP-4. The left column represents the distance between the corresponding elements. The top row represents the amino acid sequence. Each solid bar represents residue connectivities. Any residues which have a non-sequential connection are denoted by dotted lines.

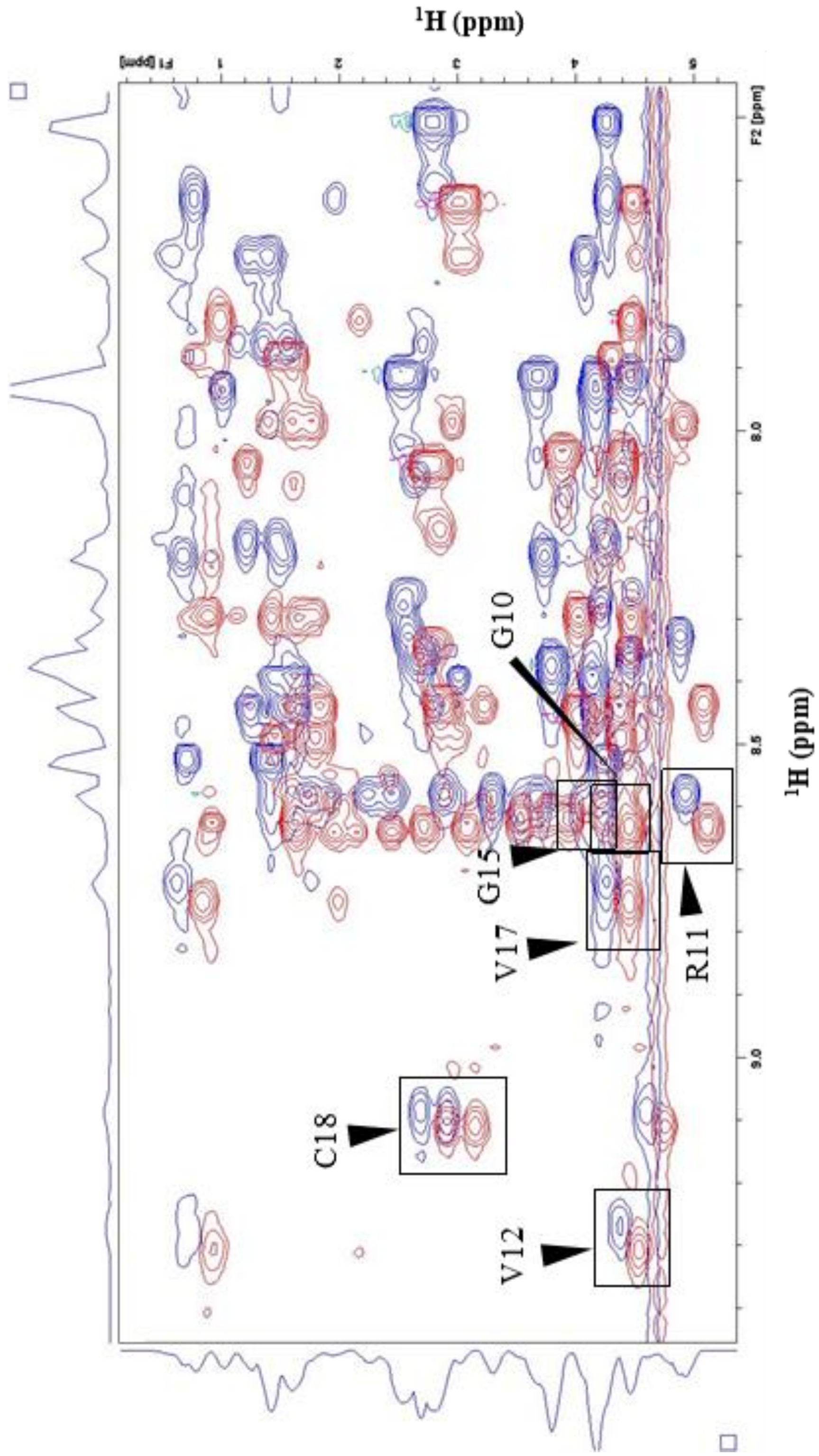

Figure 8 2D ^{1}H-^{1}H TOCSY temperature gradient experiment. The 2D ^{1}H-^{1}H TOCSY collected at 35°C is shown in red and the 2D ^{1}H-^{1}H TOCSY collected at 15°C is shown in blue. The Hα chemical shifts of the hydrogen-bonded residues are indicated within each box.

Secondary Structure Prediction for SRP-4

To preliminarily determine the secondary structure of SRP-4 I employed the chemical shift index (CSI) prediction method. The CSI method utilizes the difference between the observed Hα chemical shift and the reported Wishart "random coil" chemical shift value for the same residue type (Wishart *et al.*, 1992). Beta-strands have groupings of Hα whose chemical shifts are greater than the Wishart values and helical or turn segments have chemical shifts that are less than the Wishart values. I calculated the CSI value for each residue (Table 2) and charted the values to observe the difference (Figure 9). From the CSI, I predicted that SRP-4 contains two short beta-strands between residues 10G-12V and 16R-19V, as indicated by the two positive regions. I hypothesized that these beta-strands are connected by a short turn between residues 13L-15G, as indicated by the negative region. The rest of the peptide appeared to be disordered as indicated by the inconsistent CSI values outside of the core, beta-strand, region.

Table 2 SRP-4 CSI (ppm) calculated from the difference between each residue's observed Cα-H and its corresponding Wishart value (Wishart *et al.*, 1992).

#	Residue	Hα (ppm)	Wishart Value (ppm)	CSI (ppm)
1	SER	4.063	4.47	-0.407
2	LEU	4.34	4.32	0.02
3	LYS	4.263	4.33	-0.067
4	ARG	4.26	4.35	-0.09
5	GLY	3.907	3.96	-0.053
6	THR	4.568	4.35	0.218
7	CYS	4.377	4.54	-0.163
8	PRO	4.408	4.42	-0.012
9	ASP	4.33	4.71	-0.38
10	GLY	4.314	3.96	0.354
11	ARG	5.026	4.35	0.676
12	VAL	4.454	4.12	0.334
13	LEU	4.18	4.32	-0.14
14	PHE	4.33	4.63	-0.3
15	GLY	4.145, 3.642	3.96	-0.318
16	ARG	4.86	4.35	0.51
17	VAL	4.353	4.12	0.233
18	CYS	4.671	4.54	0.131
19	VAL	4.347	4.12	0.227
20	GLU	3.797	4.33	-0.533
21	ILE	3.866	4.17	-0.304
22	ASP	4.972	4.71	0.262
23	SER	4.572	4.47	0.102
24	LEU	4.357	4.32	0.037
25	ASP	4.57	4.71	-0.14
26	TYR	4.283	4.55	-0.267

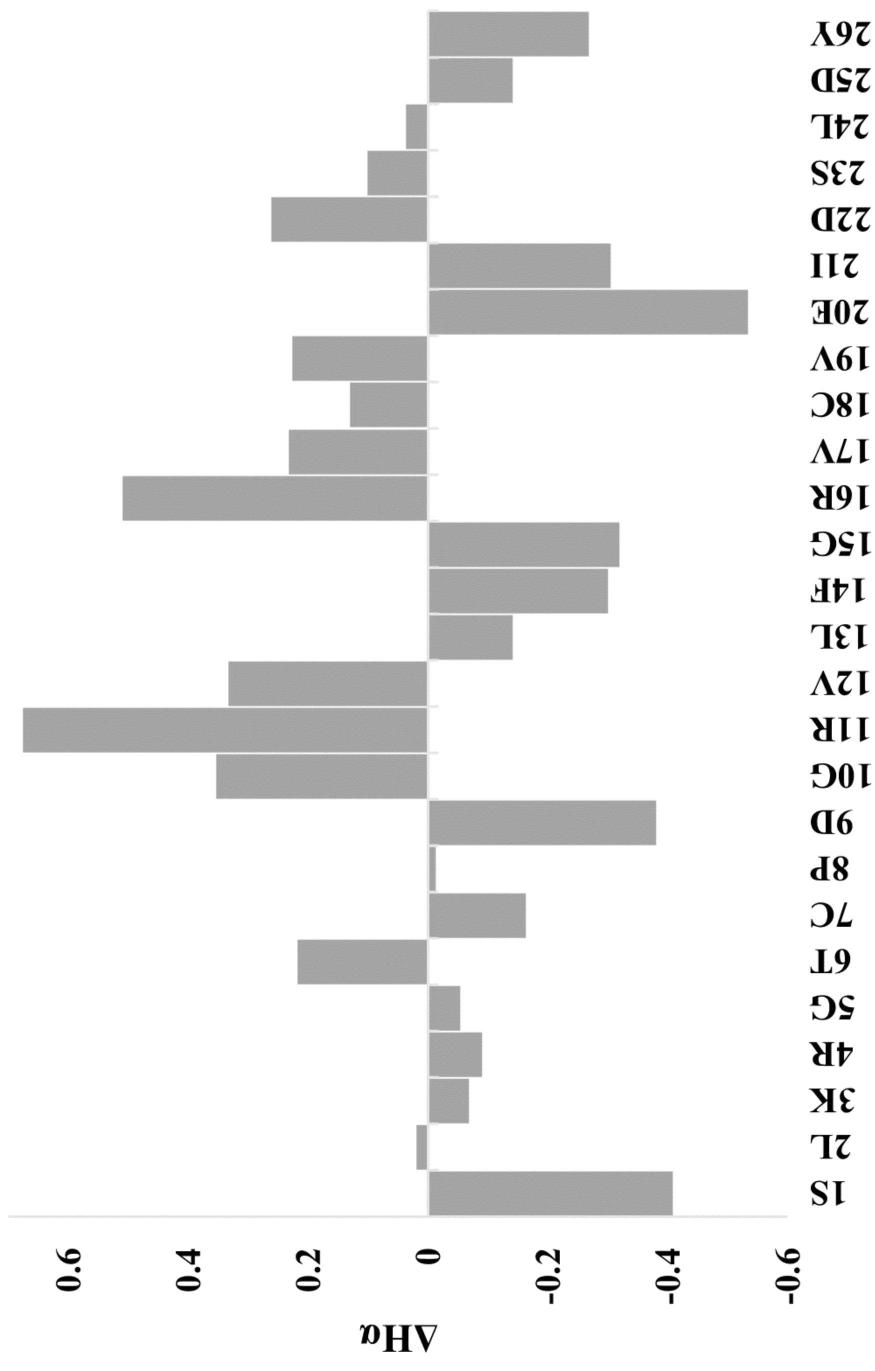

Figure 9 Cα-H Chemical Shift Index (ppm) of SRP-4. The positive regions from 10G-12V and 16R-19V indicate the presence of beta-strands.

To further verify my prediction of the secondary structure of SRP-4, I dissolved the peptide in increasing concentrations of TFE and collected circular dichroism data (Figure 10). TFE was used to stabilize the peptide's secondary structure in an aqueous solution. This is possible because TFE surrounds the peptide, displaces the water molecules, and allows for the hydrogen bonds within the peptide to remain intact (Roccatano *et al.*, 2002). In pure water, SRP-4 appears primarily disordered as represented by the minima around 190 and 200 nm, this is an indication that water may partially disrupt the intrapeptide hydrogen bonds. As the concentration of TFE increases, the ellipticity around 215 nm begins to decrease, and the ellipticity at 190 nm increases. This trend is characteristic of beta-strands. However, even as TFE increases, the minima around 200nm are still largely present indicating that SRP-4 remains partially in a random coil conformation. This supports my hypothesis that SRP-4 is largely random coil with the presence of short beta-strands.

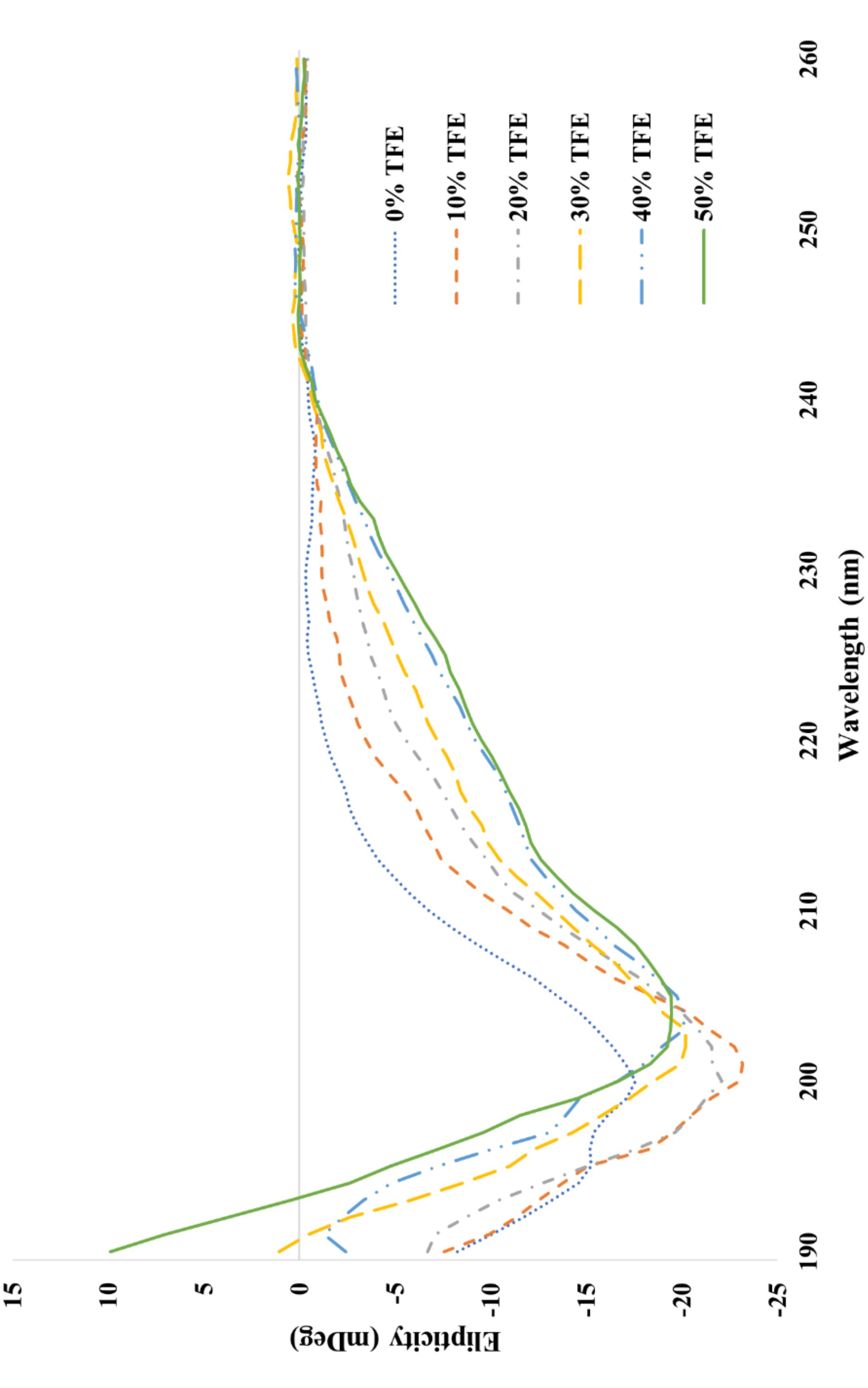

Figure 10 Circular Dichroism data of SRP-4 in v/v percent TFE concentration.

Structure Calculations and NMR Structure of SRP-4

To determine the tertiary structure of SRP-4, I used a total of 158 distance constraints including 109 intra-residue, 28 sequential, 3 medium, 8 long, 2 disulfide, and 8 hydrogen-bond constraints in structure calculations (Table 3). From the 200 conformers generated from these constraints, 20 of the lowest energy structures were used for SRP-4's structural analysis. The 20 structures were in agreement with the previous CSI prediction and contained no distance violation greater than 0.3Å. The superimposition of the 20 structures showed considerable flexibility within the N- and C-termini with a more rigid alignment of residues 10-18, forming the core region. I used the alignment of the core region of the 20 structures to determine the backbone and non-hydrogen atom root mean square deviation values which are 0.56Å and 1.38Å, respectively. The lowest energy structure out of the 20 represents the NMR solution structure for SRP-4 (Figure 11). A closer examination of the lowest energy structure revealed that SRP-4 is composed of two short antiparallel beta-strands between residues 10G-12V and 15G-18C, connected by a beta-turn between residues 13L-14F, with intrinsically disordered N- and C-termini. This structure is stabilized by a disulfide bond between residues 7C and 18C.

Table 3 Structural statistics of the 20 Lowest Energy Structures of SRP-4.

NOE Constraints	Number
Total	158
Intra-residue	109
Sequential	28
Medium Range	3
Long Range	8
Disulfide Bonds	2
Hydrogen Bonds	8
Constraints per residue	6.08
Pairwise RMSD to Mean Structure (residues 10-18)	
Backbone atoms	0.56 +/- 0.22 Å
Heavy Atoms	1.38 +/- 0.34 Å

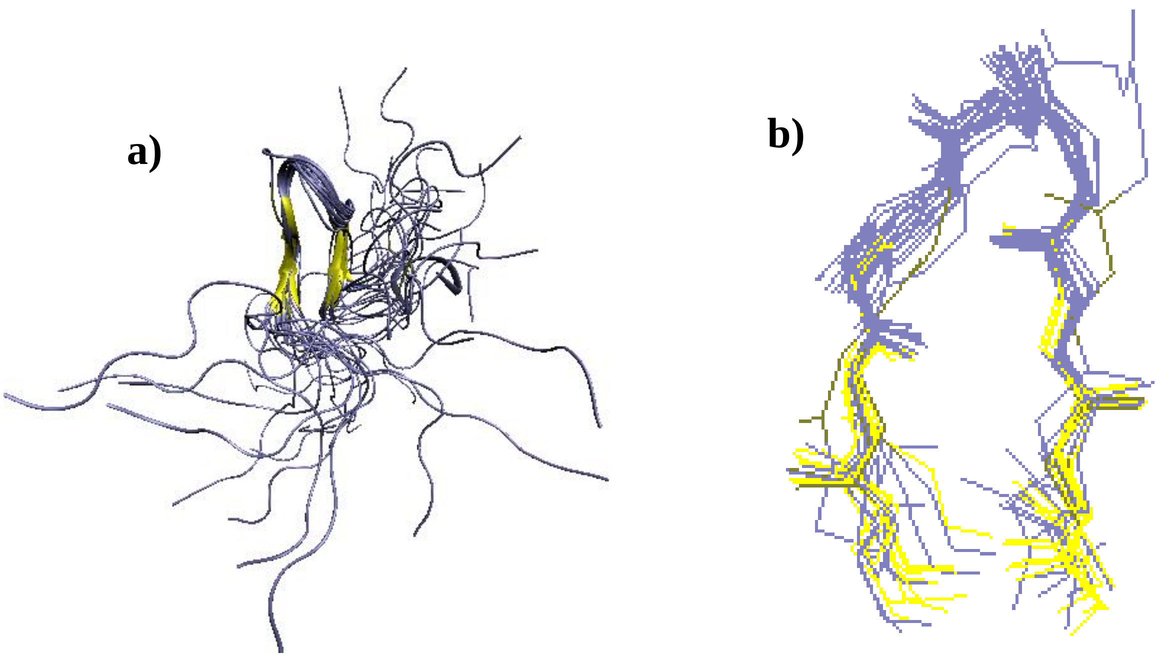

Figure 11 a) Overlay of the cartoon diagrams of the 20 lowest energy structures of SRP-4 b) Overlay of the line diagrams of residues 10 – 18 of the 20 lowest energy structures of SRP-4. The yellow residues represent the residues in beta strands.

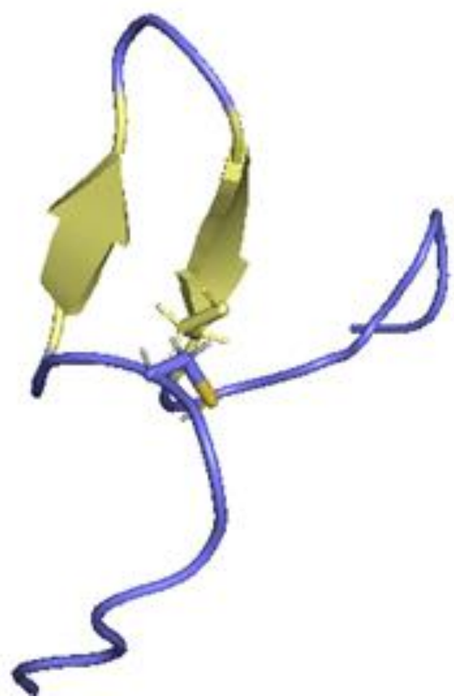

Figure 12 Cartoon diagram of the lowest energy structure of SRP-4.

Conclusions

The SRP family plays a variety of roles within the insect immune system, including wound healing, antimicrobial peptide (AMP) activation, and hemocyte proliferation, making each of their structure-activity relationships compelling. Mature *Manduca Sexta* Stress Responsive Peptide-4 (SRP-4) is a 26-residue peptide stabilized with a disulfide bond between residues 7C and 18C. From CSI and CD investigation, it was revealed early on that SRP-4 may contain beta-strands within its core region, with disordered N- and C- termini. The analysis of 2D ^{1}H-^{1}H Nuclear Magnetic Resonance (NMR) spectra and Nuclear Overhauser Effect (NOE) derived constraints revealed that SRP-4 contains two short beta strands between residues 10G-12V and 15C-18C, which are connected by a beta-turn between residues 13L-14F. Structure calculations revealed that the core, beta-strand, region of the 20 lowest energy SRP-4 conformations present a low RMSD to mean structure of 0.56 Å, indicating high agreeability and stability. This study adds to the continued investigation of the SRP family and will provide insight into their unique roles in the insect immune system.

References

Brinson, R. G., & Marino, J. P. (2019). 2D J-Correlated Proton NMR Experiments for Structural Fingerprinting of Biotherapeutics. *Journal of Magnetic Resonance*, *307*, 106581. https://doi.org/10.1016/j.jmr.2019.106581

Clark, K. D., Volkman, B. F., Thoetkiattikul, H., Hayakawa, Y., & Strand, M. R. (2001). N-terminal residues of plasmatocyte-spreading peptide possess specific determinants required for biological activity. *The Journal of biological chemistry*, *276*(40), 37431–37435. https://doi.org/10.1074/jbc.M105235200

Dembele, H. (2021). NMR and molecular dynamics structural investigations of Manduca Sexta stress response peptides and Anopheles gambiae ecdysis triggering hormone. *Thesis.*

Hayakawa Y. (1991). Structure of a growth-blocking peptide present in parasitized insect hemolymph. J. Biol. Chem. 266, 7981–7984.

Keller, R. (2004). *The computer aided resonance assignment tutorial*. Cantina Verlag: Goldau, Switzerland.

Matsumoto, H., Tsuzuki, S., Date-Ito, A., Ohnishi, A., & Hayakawa, Y. (2012). Characteristics common to a cytokine family spanning five orders of insects. *Insect Biochemistry and Molecular Biology*, *42*(6), 446-454. https://doi.org/10.1016/j.ibmb.2012.03.001

Nakatogawa, S., Oda, Y., Kamiya, M., Kamijima, T., Aizawa, T., Clark, K. D., Demura, M., Kawano, K., Strand, M. R., & Hayakawa, Y. (2009). A Novel Peptide Mediates Aggregation and Migration of Hemocytes from an Insect. *Current Biology*, *19*(9), 779-785. https://doi.org/10.1016/j.cub.2009.03.050

Roccatano, D., Colombo, G., Fioronir, M., Mark, A. E. (2002). Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proceeding of the National Academy of Sciences*, 99(19), 12179-12184. https://doi: 10.1073/pnas.182199699

Schrag, L. G., Cao, X., Dembele, H., Liu, X., Souhail, Q. A., Kanost, M. R., Chen, J., Jiang, H., & Prakash, O. (2019). Expression and Characterization of Manduca sexta Stress Responsive Peptide-1; an Inducer of Antimicrobial Peptide Synthesis. *Biochemistry and molecular biology*, *4*(3), 42. https://doi.org/10.11648/j.bmb.20190403.12

Schrag, L. G., Cao, X., Herrera, A. I., Wang, Y., Jiang, H., & Prakash, O. (2017). Solution Structure and Expression Profile of an Insect Cytokine: Manduca sexta Stress Response Peptide-2. *Protein and peptide letters*, *24*(1), 3–11. https://doi.org/10.2174/0929866524666161121142840

Schrag, L.G., Herrera, A. I., Cao, X., Prakash, O., Jiang, H. (2017). Structure and Functions of Stress-Responsive Peptides in Insects. *Peptide-based Drug Discovery: Challenges and New Therapeutics Drug Discovery Series 59,* 438-451, The Royal Society of Chemistry, London, UK.

Sharma, D., Rajarathnam, K. (2000). 13C NMR chemical shifts can predict disulfide bond formation. *J Biomol NMR* **18**, 165–171. https://doi.org/10.1023/A:1008398416292

Shears, S. B., & Hayakawa, Y. (2018). Functional Multiplicity of an Insect Cytokine Family Assists Defense Against Environmental Stress. *Frontiers in Physiology*, *10*. https://doi.org/10.3389/fphys.2019.00222

Skinner, W., Dennis, P., Li, J., Summerfelt, R., Carney, R., & Quistad, G. (1991). Isolation and identification of paralytic peptides from hemolymph of the lepidopteran insects Manduca sexta, Spodoptera exigua, and Heliothis virescens. *Journal of Biological Chemistry*, *266*(20), 12873-12877. https://doi.org/10.1016/S0021-9258(18)98775

Wang, Y., Jiang, H., & R. Kanost, M. (1999). Biological activity of Manduca sexta paralytic and plasmatocyte spreading peptide and primary structure of its hemolymph precursor. *Insect Biochemistry and Molecular Biology*, *29*(12), 1075-1086. https://doi.org/10.1016/S0965-1748(99)00086-7

Wang, Y., Kanost, M. R., & Jiang, H. (2022). A mechanistic analysis of bacterial recognition and serine protease cascade initiation in larval hemolymph of Manduca sexta. *Insect Biochemistry and Molecular Biology*, *148*, 103818. https://doi.org/10.1016/j.ibmb.2022.103818

Wishart, D.S., Sykes, B.D., Richards, F.M. (1992). The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, 31(6), 1647-1651. doi: 10.1021/bi00121a010

Yamaguchi, K., Matsumoto, H., Ochiai, M., Tsuzuki, S., & Hayakawa, Y. (2012). Enhanced expression of stress-responsive cytokine-like gene retards insect larval growth. *Insect Biochemistry and Molecular Biology*, *42*(3), 183-192. https://doi.org/10.1016/j.ibmb.2011.11.009