# Understanding the Interactions between the Nitrate Reductase Signal Peptide and its Maturation Chaperone NapD

Martin Milad

A Thesis

in

the Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science
Concordia University
Montreal, Quebec, Canada

2006

*Your file Votre référence*
*ISBN: 0-494-14229-4*
*Our file Notre référence*
*ISBN: 0-494-14229-4*

## ABSTRACT

### Understanding the Interactions between the Nitrate Reductase Signal Peptide and its Maturation Chaperone NapD

**Martin Milad**

Periplasmic Nitrate Reductase (NapA/NapB) belongs to a class of proteins that are exported in a folded state by the twin arginine translocon (Tat) machinery. For a long time it was known that NapD is essential for the proper formation of the NapA/NapB enzyme. However, the mechanism by which this is accomplished is still unknown. The focus of this study is to understand the interaction between NapA and its maturation chaperone, NapD. For the first time we will show that the N-terminal consensus leader sequence of NapA that is involved in Tat export is also the binding site for the chaperone. To characterize the mechanism co-purification experiments, biophysical techniques and Nuclear Magnetic Resonance (NMR) were used. Co-purification experiments indicated that the N-terminal fragments; NapA114, NapA35 and NapA27 were able to bind NapD. Circular Dichroism (CD) analysis of NapD in complex with NapA35 illustrated a small increase in secondary structure. NMR confirmed that the small changes in CD spectra were due to conformational changes within NapD and the formation of an alpha-helix for NapA35. Chemical shift changes from NMR demonstrated that the entire leader peptide participates in binding with NapD. The residues that underwent the greatest changes in chemical shift are Arg6, Phe8, Met9, Lys10, Val14, Ala16-21, Gly22 and Leu23. NMR also showed that the NapA35 peptide is mobile when unbound. However when NapA35 is complexed with NapD, drastic changes in mobility occurred. For the first time the structure of the Tat leader peptide, NapA35 is determined when in complex with the chaperone NapD.

**Acknowledgements**

I would like to thank Concordia University and the Department of Biochemistry for giving me the opportunity to study in this ever changing field of Biochemistry. I would like to thank my supervisor Dr. Irena Ekiel, my co-supervisor Dr. Judith Kornblatt and committee members Dr. Christine DeWolf and Dr. Joanne Turnbull.

I would also like to thank Dr. Ovidiu Minailiuc for help with NMR, particularly for structure calculations in Chapter 5.5 as well as Jing Cheng for help with cloning (Chapter 2.3.1). Also I would like to thank Carine Lievre for providing some of the constructs, NapD, (his)-NapD/NapA114 and (his)-NapD/NapA35. All other cloning, protein expression, purification, biophysical techniques and NMR sample preparations were prepared by myself. I also assigned the homonuclear and heteronuclear signals of NapA35 free and in complex with NapD, for TALOS and CYANA calculations. I would like to thank McGill University and the Québec/Eastern Canada High Field NMR Facility for the 3D $^{13}$C-NOESY HSQC and $^{13}$C-HSQC (800MHz) experiments.

# Table of Contents

Page

# List of Figures

Page

# List of Tables

Page

## Abbreviations

| | |
|---|---|
| ACN | Acetonitrile |
| CD | Circular Dichroism |
| CYANA | Combined assignment and dynamics algorithm for NMR applications |
| $D_2O$ | Deuterium Oxide |
| DTT | 1,4-dithiothreitol |
| HPLC | High Performance Liquid Chromatography |
| HSQC | Heteronuclear Single Quantum Correlation |
| MS | Mass Spectrometry |
| m/z | mass-to-charge |
| NMR | Nuclear Magnetic Resonance |
| NOESY | Nuclear Overhauser Effect |
| PDB | Protein Data Bank |
| SDS-PAGE | Sodium dodecyl sulphate Polyacrylamide Gel Electrophoresis |
| TALOS | Torsion Angle Likelihood Obtained from (chemical) Shift and Sequence Similarity |
| TFA | Trifluoroacetic acid |
| TOCSY | Total Correlated Spectroscopy |
| UV/vis | Ultraviolet/visible Spectroscopy |

# Chapter I Introduction

## 1.1 Protein Interactions

An important characteristic of proteins is their ability to specifically bind to other proteins, peptides and small molecules. The interactions of proteins with other biomolecules can result from both covalent and non-covalent bonding. Covalent interactions usually involve the formation of disulfide bonds between two thiol groups. On the other hand, non-covalent interactions consist of hydrogen bonding, van der Waals interactions and salt bridges (Dill, 1990). Regardless of the type of binding, formation of complexes is essential for proper cellular functioning.

Three decades ago Christian Anfinsen won the Nobel Prize for demonstrating that protein folding is regulated by the protein itself (Anfinsen, 1972). Since then other scientists have discovered that some proteins known as chaperones have helped in the process (Walter & Buchner, 2002). Chaperones were shown to associate with the target protein during part of its folding process, however once folding is complete (or even before) the chaperone will leave its current protein molecule to aid another. Another class of chaperones includes maturation chaperones that serve to assist in cofactor loading and retain proteins in the cytoplasm until mature multi-subunit complexes are formed.

## 1.2 Protein Export in Bacteria

The targeting and transport of proteins to and across biological membranes is an essential aspect of cell life. Approximately one quarter of the genes expressed in *Escherichia coli* encode proteins that end up being exported (Palmer *et al.,* 2005). There are two major export pathways in *E. coli*, the Secretory (Sec) pathway and the Twin-arginine translocation (Tat) pathway (Weiner *et al.,* 1998). The Sec pathway, which is the more studied of the two, exports proteins that are unstructured. It also uses a combination of ATP hydrolysis and an electrochemical membrane gradient for energy (de Keyzer *et al.,* 2003). The Tat pathway on the other hand is known to be essential in exporting folded proteins. This particular system only uses energy from the electrochemical membrane gradient (Palmer *et al.,* 2005). Little is known about the mechanism by which Tat machinery functions. However, to be exported by the Tat system proteins need a leader sequence.

The Tat machinery is involved in exporting proteins with the **SRRxFLK** consensus sequence at the N-terminal region (Alami *et al.,* 2003). The twin arginines in the sequence are the most conserved and are crucial for transport since a conservative substitution of either arginine for lysine usually blocks transport (Halbig *et al.,* 1999; Buchanan *et al.,* 2001). This leader sequence is followed by a hydrophobic region of approximately 15 amino acids, then by a more polar region for proteolytic cleavage (Stanley *et al.,* 2000). This leader peptide is often removed in the mature protein.

There has only been a few published structural studies on proteins that contain the Tat signal sequence. The crystal structures of formate dehyfrogenase H from *E. coli* (Boyington *et al.,* 1997) (PDB accession code 1AA6) and the molybdenum enzyme

dimethylsulphoxide reductase (DMSOR) from *Rhodobacter Capsulatus* have been determined (McAlpine *et al.,* 1998) (PDB accession code 4DMR). The structures of two nitrate reductase enzymes have also been solved, periplasmic nitrate reductase (Nap) from the sulphate reducing bacterium *Desulfovibrio desulfuricans* (Dias *et al.,* 1999) (PDB accession code 2NAP) and respiratory nitrate reductase (NapAB) from Rhodobacter spharoides (Arnoux *et al.,* 2003) (PDB accession code 1OGY). For all these proteins the structure was solved only for the mature protein, which begins after the cleavage of the N-terminal leader peptide. In other words the leader peptide was not included as part of the sequence. The only study that included the leader peptide as part of the structure was on the glucose:fructose oxidoreductase precursor (preGFOR) from *Zymomonas mobilis* which did not result in insight into the signal sequence structure (Nurizzo *et al.,* 2001) (PDB accession code 1H6A). The lack of density for the signal peptide in the crystal structure of preGFOR, implies that it is unstructured or disordered in the crystal. The general notion for the determination of the X-ray structure of proteins that contain the Tat signal peptide has been the deletion of the leader sequence and the focus on the catalytic domain.

Structural analysis of the Tat leader sequence was reported for the high potential iron-sulfur protein (HiPIP) from *Allochromatium vinosum* (Kipping *et al.,* 2003). Nuclear magnetic resonance (NMR) analysis of the amide proton resonance and H/D exchange matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry indicated the absence of secondary structure in the signal peptide of HiPIP. Circular dichroism studies of the Tat signal peptide of the *E. coli* SufI protein indicated an unstructured peptide in aqueous solution, with no stable secondary structure (San

Miguel *et al.,* 2003). However when the synthetic Tat leader peptide was placed in a more hydrophobic environment it was shown to adopt an alpha-helical conformation. The crystal structure of the cytochrome $b_6f$ complex from the cyanobacterium *Mastigocladus laminosus* demonstrated that membrane embedded Tat signal peptides form an alpha-helical conformation (Kurisu et al., 2003). Although secondary structure prediction programs suggest an alpha-helical structure for the twin-arginine motif and parts of the hydrophobic core the results have indicated that the peptide has an unfolded conformation in the cytoplasm.

Since Tat transport is important in periplasmic enzyme transport, it is critical that proteins are not exported before the multi-subunit enzyme is formed. This is accomplished by soluble chaperones that are localized in the cytoplasm (Turner *et al.,* 2004). These chaperones prevent pre-mature targeting of proteins before cofactor loading, correct folding, or docking of partner proteins have occurred (Hatzixanthis *et al.,* 2005). The first Tat signal-peptide-binding protein was isolated by affinity chromatography (Oresnik *et al.,* 2001). Using a glutathione-S-transferase fusion to the signal peptide of the molybterin co-factor-binding dimethyl sulphoxide (DMSO) reductase subunit DmsA, allowed for the isolation of DmsD. This led to the idea that DmsD prevents the signal peptide from interacting with the Tat machinery until complete cofactor loading of DmsA.

More recently other chaperones specific for Tat dependent precursor proteins have been identified. In *E. coli* trimethylamine N-oxide reductase (TorA) is a Tat-dependent periplasmic redox enzyme that binds a molybdenum cofactor. The twin-arginine signal peptide of TorA is the most heavily studied in the field. The TorD protein has been

identified as a cytoplasmic accessory protein required for TorA cofactor loading (Ilbert *et al.*, 2004; Pommier *et al.*, 1998). The 3D structure of the TorD homolog from *Schewanella massila* indicated that this protein consists of two alpha-helical domains connected by a short hinge region (Tranier et al., 2003) (PDB accession code 1N1C). Furthermore, by using *in vitro* molecular techniques, tight binding was demonstrated between TorD and the TorA signal peptide (Hatzixanthis *et al.*, 2005).

Most E. coli Tat substrates are not characterized. The complete list of the thirty-six polypeptides in *E. coli* known or predicted to bear the N-terminal Tat leader sequence have been summarized (Palmer et al., 2005). For almost all cases it is unknown whether a signal chaperone is needed. Here we will report the interaction between the leader peptide and its potential chaperone for the Periplasmic Nitrate Reductase enzyme in *E. coli*.

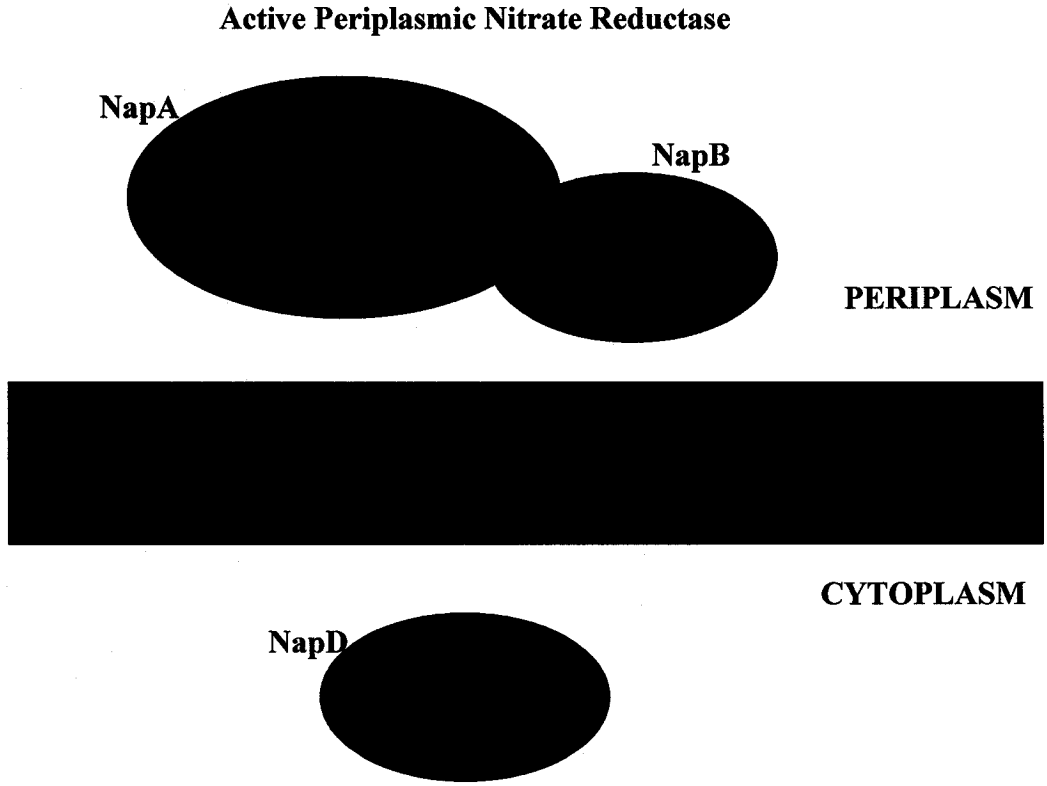
### 1.3 Periplasmic Nitrate Reductase (Nap)

When bacteria are grown in the absence of oxygen, nitrate becomes the preferred electron acceptor (Brondijk *et. al.*, 2004). Bacteria reduce nitrate ($NO_3^-$) to nitrogen gas ($N_2$), via nitrite ($NO_2^-$), nitric oxide (NO) and nitrous oxide ($N_2O$). Nitrate reductase reduces $NO_3^-$ to $NO_2^-$. Three types of nitrate-reducing enzymes have been discovered in bacteria: cytoplasmic assimilatory nitrate reductase (Nas), membrane-bound respiratory nitrate reductase (Nar) and periplasmic nitrate reductase (Nap) (Berks *et al.*, 1995). In *E. coli*, Nap is essential for anaerobic respiration at extremely low nitrate concentrations. Periplasmic nitrate reductase is encoded by the napFDAGHBC operon, (Figure 1.1) where napABCD genes have been shown to be essential (Grove et al. 1996; Potter & Cole 1999; Turner *et al.*, 2004).

Figure 1.1 Organization of nap operon in *E. coli* K12. As can be seen *NapD* is found upstream to *NapA*, perhaps serving as a means of regulating the expression of NapA. However the exact role of NapD in the post-translational assembly of NapA is unknown (Cole, 1996).


Periplasmic respiratory nitrate reductases (Figure1.2) are heterodimers composed of two subunits, a catalytic NapA subunit (90 kDa) that contains a molybdopterine guanine dinucleotide cofactor, an iron-sulfur (4Fe-4S) centre and an electron transfer unit, NapB (17 kDa) (Richardson *et al.,* 1999).

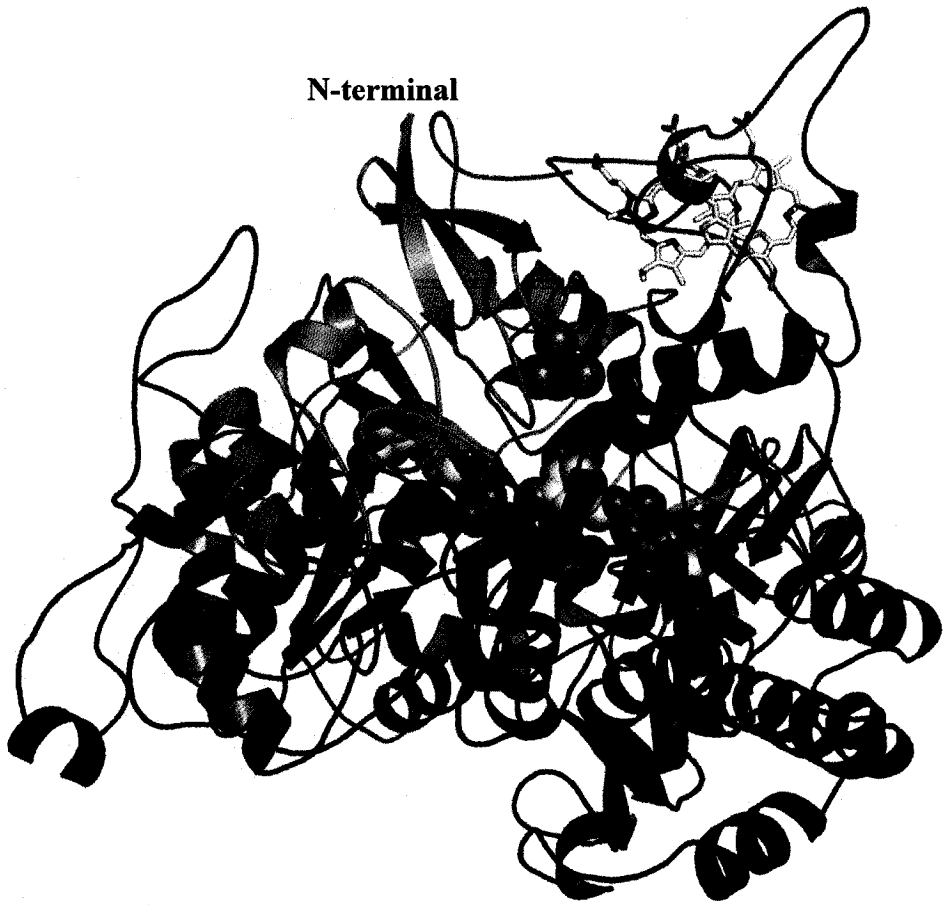
Figure 1.2 Schematic showing the active periplasmic nitrate reductase enzyme (NapA/NapB), and NapD, which acts as the chaperone involved in the maturation of NapA. NapD remains in the cytoplasm where it binds to NapA prior to export. The N-terminal sequence of NapA contains a motif, KL**SRRSFMK**, corresponding to the consensus Tat leader sequence responsible for export. This leader peptide is proteolytically cleaved in the periplasm.



Another important component of this system is NapD (10 kDa), a soluble cytoplasmic protein, which has been hypothesized to be the private chaperone involved in the maturation of NapA (Berks *et al.,* 1995). Although NapD does not cross the membrane and move into the periplasm it has been shown to be essential in the formation of the active NapA/NapB enzyme in the periplasm (Reyes *et al.*, 1998; Potter & Cole, 1999). The first experimental characterization was performed for *Rhodobacter spharoides* DSM158, by mutational analysis of the *napD* gene (Reyes *et al.,* 1998). The absence of NapD significantly decreased both the nitrite production and the nitrate reductase activity by this bacterium. In *E. coli* an in-frame deletion of *napD* was used to determine whether NapD is essential for Nap activity (Potter & Cole, 1999). The results confirmed that NapD is essential since no nitrate reduction was observed. This maturation chaperone, is unlikely to function as an electron transfer unit since it lacks conserved residues likely required for binding ligand redox cofactors (Potter *et al.,* 1999). The role of NapD in periplasmic nitrate reduction is still unknown due to the lack of structural and functional characterization.

Although the structure of periplasmic nitrate reductase has not been solved in *E. coli*, the structure for NapA/NapB has been determined by X-ray diffraction for *Rhodobacter sphaeroides* (Arnoux *et al.,* 2003) (Figure 1.3). The Tat signal peptide was not included as part of the structure, as it is believed to be unstructured. As previously mentioned, this leader sequence is cleaved in the mature protein.

Figure 1.3 Crystal structure of NapA/NapB complex in *Rhodobacter sphaeroides* (Arnoux *et al.*, 2003) (PDB accession code 1OGY). The N-terminal domain of NapA is shown in cyan. The other three domains of NapA are shown in blue. NapB is red. The Tat leader sequence is not shown, since it is cleaved and not part of mature NapA. Co-factors are shown in yellow and orange.




Furthermore, NapA consists of four domains; the N-terminal domain contains a cysteine rich motif, which binds the [4Fe-4S] cluster. The remaining three domains are structured around the molybdopterine guanine dinucleotide cofactor (Figure 1.3).

### 1.4 Tat signal peptide-chaperone NapA/NapD complex

Overall the characteristics for the Tat leader sequence are found in *E.coli* NapA (Figure 1.4). The N-terminal sequence of NapA is KL**SRRSFMK**, which corresponds to

the leader peptide. Another characteristic feature of the N-terminal region of NapA is the presence of a hydrophobic region, where we note a stretch of nine alanines from A11 to A21. There are two potential implications for the role of NapD as the chaperone. First, perhaps NapD binds to NapA, and retains NapA in the cytoplasm until the mature NapA/NapB complex is ready to be exported to the periplasm and second, NapD may bind to the multi-subunit enzyme after it is formed and help recruit it to the membrane. This hypothesis is supported by the finding that the DmsA signal-peptide-binding protein DmsD, interacts with the Tat machinery in the membrane (Papish *et al.*, 2003). NapD may also aid in the proper formation of the NapA/NapB complex. The objective of this thesis is to assess and characterize the interaction between NapD and NapA. We will show for the first time that the leader sequence of NapA, which is involved in export, is also the site to which NapD binds.

Figure 1.4 Sequence of the N-terminal of NapA. The consensus leader sequence found in Tat translocon is in bold. The hydrophobic core consists of a poly-alanine stretch from Ala15 to Ala21.

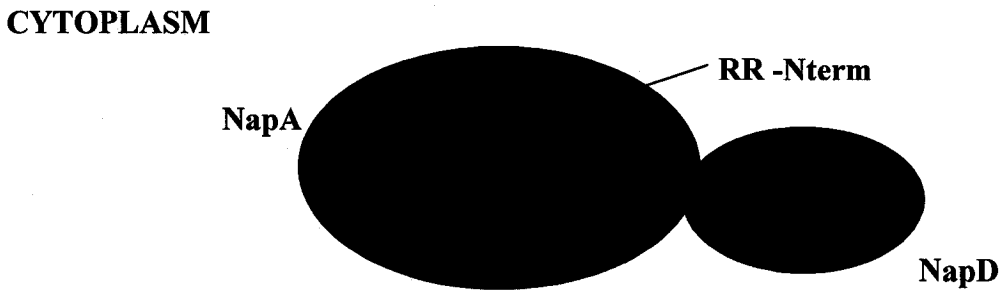
*N-term-* $^{1}$MKL**SRRSFMK** $^{11}$ANAVAAAAAA A$^{22}$GLSVPGVAR AVVG$^{35}$Q *–C-term*

## 1.5 Outline of Thesis and Statement of Purpose

Overall the goal is to structurally characterize the Tat signal peptide-chaperone interactions, so as to gain information on the mechanism of export of folded proteins. This will be accomplished by characterizing the structure of the NapD/NapA complex (Figure 1.5). Furthermore, the amino acid residues of NapA and NapD that are involved

in binding are determined. Reported for the first time will be the structure of the N-terminal Tat leader sequence when it is bound to its chaperone, NapD.

Figure 1.5 The objective of this project is to understand the interaction that occurs in the cytoplasm between NapD (chaperone) and NapA, by looking at the structural aspects of the complex.




This thesis is divided into six chapters. The first chapter established the importance of protein structure and complex formation for proper cell functioning. It introduced the Tat system as well as the periplasmic nitrate reductase enzyme. The second chapter outlines the experimental procedures used to obtain the desired protein and peptide complex. The next three chapters apply different techniques to study the complex. Chapter three is a composition of the biophysical methods applied to characterize the system. Chapters four and five utilize NMR spectroscopy to understand complex formation and characterize its structure. Finally in chapter six all the results are put together for general conclusions and biological implications.

# Chapter II Protein Production and Characterization

## 2.1 Foreword

Proteins that bind tightly can often be characterized by simple co-purification where the binding partner is co-expressed so that both components are present during purification. Complex formation can later be demonstrated by SDS-PAGE.

NapD and NapA fragments were expressed in *Escherichia coli* K-12, where only one protein was His tagged. Complex formation was probed by column chromatography using NiNTA resin, so that only the peptides that were bound to the protein were co-purified.

## 2.2 Media and Buffers

Luria-Bertani (LB) medium (Fisher Scientific): 10g/L tryptone, 5g/L yeast-extract and 5g/L NaCl.

Lysis Buffer: 50mM $NaH_2PO_4$, 300mM NaCl, and 10mM imidazole, pH 8.

Wash Buffer: 50mM $NaH_2PO_4$, 300mM NaCl, and 20mM imidazole, pH 8.

Elution Buffer: 50mM $NaH_2PO_4$, 300mM NaCl, and 250mM imidazole, pH 8.

Cleavage Buffer: 20mM Tris-HCl, 150mM NaCl and 2.5mM $CaCl_2$, pH 8

NMR Buffer: 50mM Phosphate Buffer, 150mM NaCl and 10mM DTT, pH 6.8.

## 2.3 Protein Production

### 2.3.1 Cloning

Plasmid isolations, restriction enzyme digestions, agarose gel electrophoresis, ligations, and *E.coli* transformations were performed according to standard protocols

(Sambrook *et al.,* 1989) unless otherwise stated. The genomic cDNA of *E. coli* K-12 was obtained from the ATCC Corporation (Maryland, USA). NapD was amplified by Polymerase Chain Reaction (PCR) using *Taq* polymerase and synthetic primers NapD-F and NapD-R (Table 2.1). The PCR product was then gel purified and ligated into the Nde1 and BamH1 (NEB) restriction sites of the pET-15b expression vector (Amersham Pharmacia Biotech, Piscataway, NJ). The construct was verified by DNA sequencing. The fusion protein contains a Thrombin cleavable His tag. The cloned plasmids of His-tagged NapD fusion proteins were transformed by heatshock into competent cells of the host BL21 Gold Magic (Stratagene) strain (Brinkmann *et. al.,* 1989*)*. Kanamycin 50μg/ml (Fisher) was used to select for the expression host. The transformed cells were plated on agar plates (Becton Dickinson and Co.) and were grown overnight at 37 °C. The protein was verified to be totally soluble and highly expressed before production scale growths.

Table 2.1 Primers used for cloning and their sequence. Restriction sites are colored; Nde1 is red, BamH1 is blue and Nco1 is green. Since Translational Coupling was used only the reverse primer is needed for NapA.

| **Primer** | **Sequence** |
|---|---|
| NapD-F[1] | 5'-GGA TTC **CAT ATG** CAC ACT AAC TCG-3' |
| NapD-R[2] | 5'-CG **GGA TCC** TCA TGG TGT TTC CTC-3' |
| NapD-TF[1] | 5'-GGA ATT C**CC ATG G**CG CAC ACT AAC TGG-3' |
| NapD-TR[2] | 5'-ATG GTG ATG GTG ATG CAT CAT GGT GTT TCC TCA CCT TG-3' |
| X[3]-F | 5'-GAG GAA ACA CCA TGA TGC ATC ACC ATC ACC ATC ACA G-3' |
| NapA35-R[2] | 5'-CG **GGA TCC** TTA ACC AAC AAC GGC GCG -3' |
| NapA27-R[2] | 5'-CG**G GAT CC**T TAC GGC ACG CTG AGA CCG GC-3' |
| NapA18-R[2] | 5'-CG**G GAT CC**T TAC GCA GCG GCC GCA ACG GC-3' |

[1] F- forward primer, [2]R- reverse primer, [3]X- His tag primer,

### 2.3.2 Protein Expression

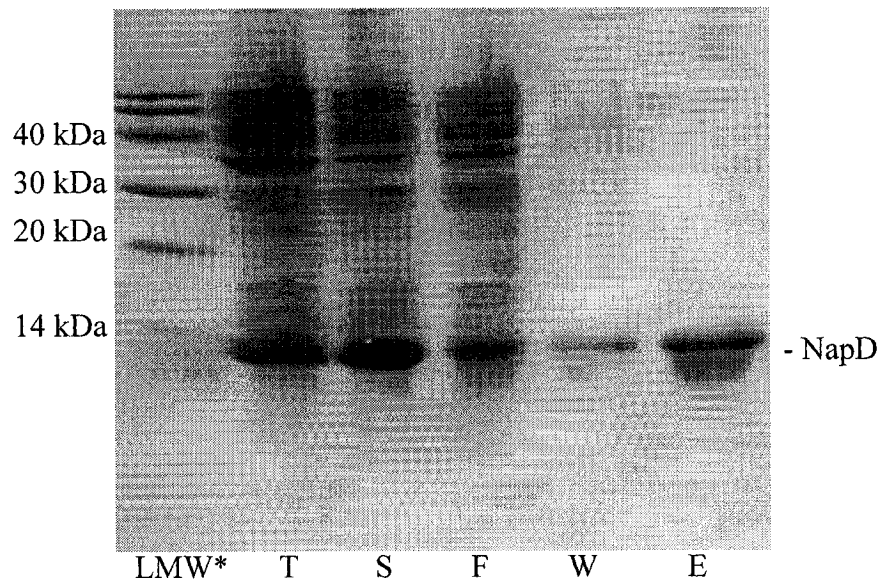
A 100ml Luria Broth (LB) culture of transformed BL21-GM was supplemented with 100μg/ml of ampicillin and 50μg/ml kanamycin and grown overnight at 37 °C. The next day, 1 L of LB was inoculated with the overnight culture. The culture was then incubated at 37 °C until the $OD_{600nm}$ was approximately 0.8. At that point, 1mM of isopropyl-1-thio-B-D-galactopyranoside (IPTG) was added to induce protein expression. The culture was then incubated at 24 °C for 16 hours to allow expression of the fusion protein. The cells were harvested the next day by centrifugation at 4 °C.

### 2.3.3 Protein Purification

The resulting cell pellet was resuspended in 35ml lysis buffer. Three milligrams of DNAse 1 was added and the suspension, on ice, was lysed by sonication (Mandel; Misonix Sonicator 3000, Power setting 5; 3 cycles of 90 sec). The extract was centrifuged for 30 minutes in a Beckman preparative centrifuge at 3000g to remove the insoluble fraction. The protein was purified by column chromatography using NiNTA resin as an affinity matrix. NiNTA resin slurry was equilibrated with lysis buffer in the gravity column prior to addition of the soluble protein. The soluble cell lysate was filtered (0.22μm Millex GP Filter unit, Millipore Corporation) and mixed with the equilibrated resin, then incubated by gentle mixing at 4 °C for 1 hour. The mixture was then added back into the column and washed with 3 volumes of wash buffer. If the His tag on the protein did not have to be cleaved the protein was eluted directly by elution buffer. If cleavage of the His tag was needed the resin was equilibrated with cleavage buffer. The N-terminal His tag was cleaved by treatment of the sample overnight at 4 °C with 4μl Thrombin (Amersham Pharmacia), to produce NapD with a three residue (Gly-Ser-His)

N-terminal extension. Benzamidine Sepharose (0.5ml) was added to the flow-through to remove thrombin. The sample was then filtered and characterized by SDS-PAGE (Figure 2.1). Purified NapD was concentrated to 500µl using an Amicon centrifugal concentrator (Millipore Corporation) with a membrane cut-off of 5kDa. The resultant concentrated protein was then dialyzed against NMR buffer and stored at 4 °C. 10mg of NapD were purified from one liter cultures.

Figure 2.1 SDS-PAGE showing the purification steps of NapD. The fractions are total (T), soluble (S), flow-through (F), wash (W) and elute (E).




*Low Molecular Weight Marker

### 2.3.4 Gel Electrophoresis and Protein Quantitation

SDS-PAGE was performed using Min Protean 3 (Biorad) (Laemmli, 1970). An 18 cm square slab apparatus was filled with 20% (w/v) resolving and a 6% (w/v) stacking polyacrylamide gel. Samples were resuspended in loading buffer and denatured by heating at 100 °C for 5 min before loading. After electrophoresis for 2 hours at 200 V, the

gel was stained by Coomassie blue (Matsudaira, 1987) for 2 hours and destained overnight.

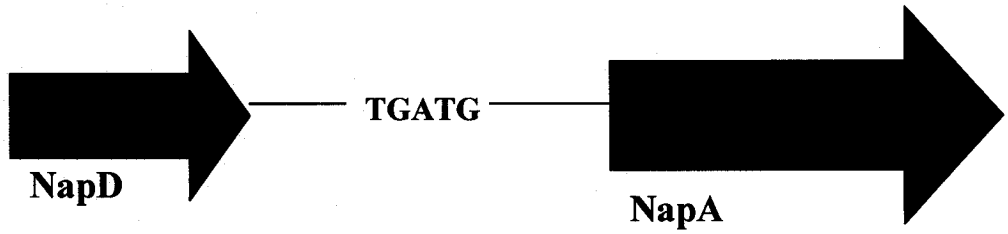
Protein concentrations were determined using the Bradford method (Biorad) and the BCA (Pierce) method, relative to BSA protein standards, following the protocols supplied by the manufacturer. The absorbance at 280 nm and absorbance at 215 and 225 nm, peptide bond method were also used to determine the concentration (Waddell, 1956). The final concentrations were taken as the average of all methods. Relative amounts for co-purified samples were obtained by scanning the SDS-gels and integrating each peak based on intensity using the program Gene Tools (data not shown).

## 2.4 Probing binding between NapD and NapA

### 2.4.1 Cloning and expression of different fragments of NapA

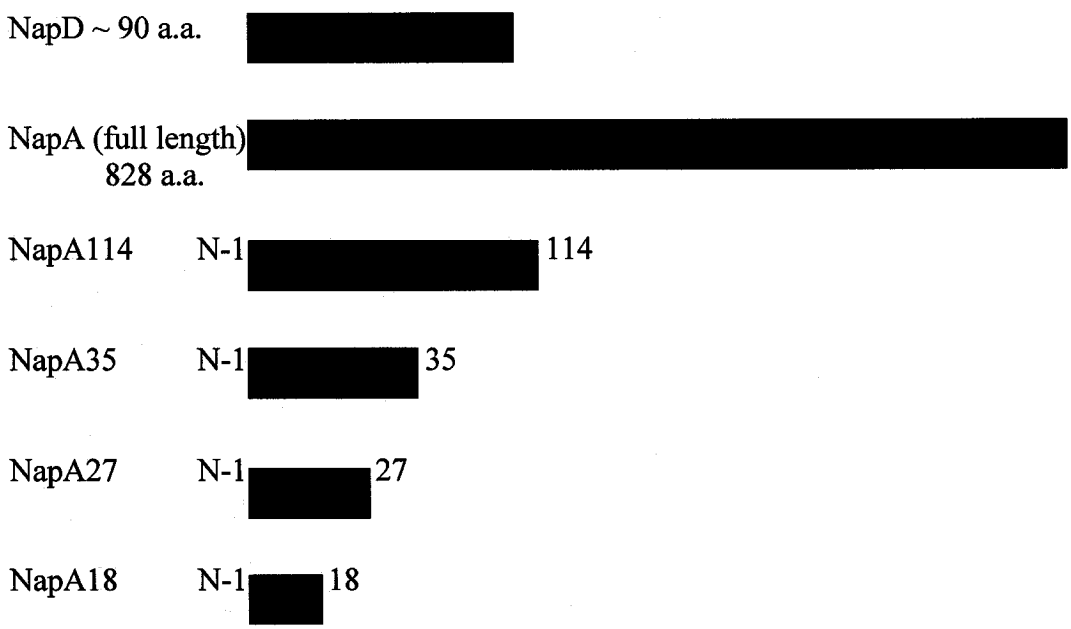
To probe the binding interactions between NapD and NapA, different constructs of NapA were evaluated. As seen in Figure 1.1 the gene encoding NapD is found upstream of the gene encoding NapA, in such a way that the stop codon of *NapD* actually overlaps with the start codon of *NapA*. This is the basis of translational coupling (Draper *et al.*, 1996; Richardson *et al.*, 2001) and can result in the stoichiometric synthesis of NapD and NapA (Berks *et al.*, 1995). Although the overlap is only one nucleotide as illustrated in Figure 2.2, it is nonetheless sufficient to promote the expression of NapA.

Figure 2.2 Translational Coupling. In *E. coli*, the stop codon of *NapD* (TGA) overlaps with the start codon of *NapA* (ATG) found downstream. This may be the basis by which NapD regulates the expression of NapA.




The primers used are tabulated in Table 2.1. NapD was cloned such that NapA will also be expressed downstream in pET-15b. The different constructs used are outlined in Figure 2.3. It is important to note however, that to probe for binding the gene products must be co-purified.

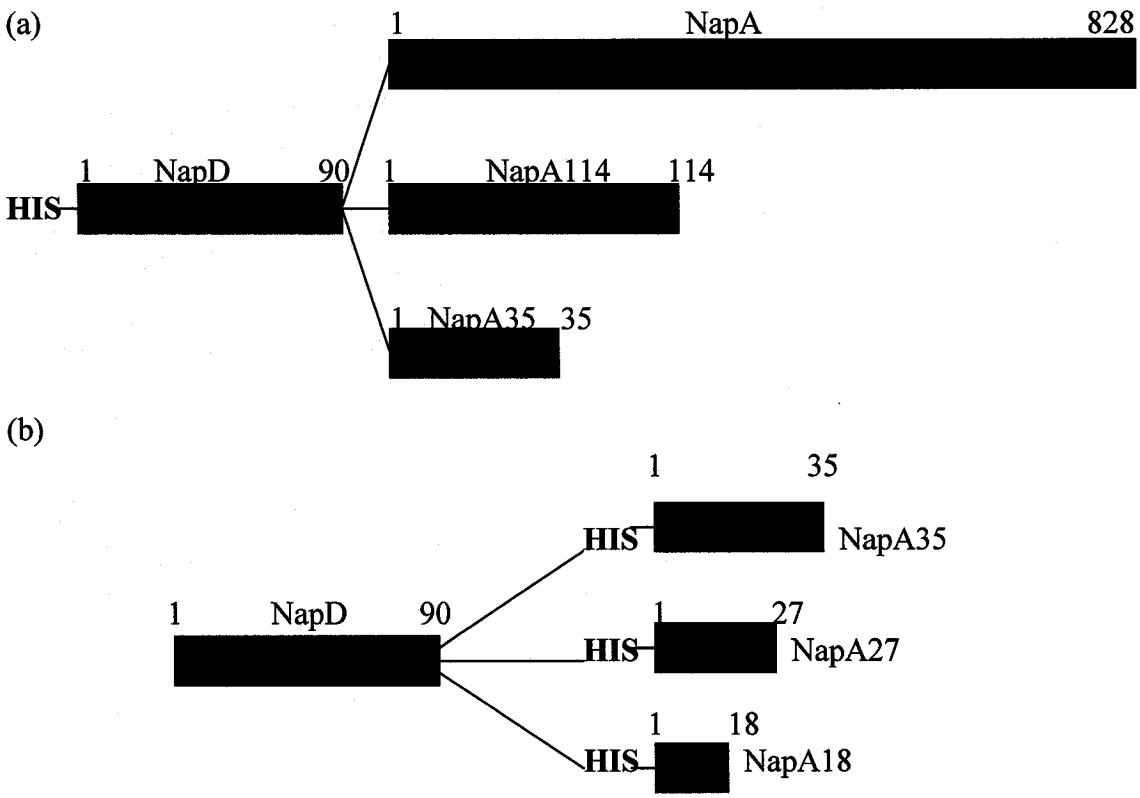
Figure 2.3 Legend showing the length of NapD and the different constructs of NapA (not drawn to scale).



For the first set of experiments, the His-tag was fused to NapD and not NapA (Figure 2.4a). For these set of experiments the NapD-F primer and the reverse primer of NapA are needed for cloning (Table 2.1). Since the stop codon of NapD overlaps with the start codon of NapA, once the ribosome has completed the translation of NapD, it will go on to translate NapA. Therefore, both NapD and NapA will be expressed in the same vector, pET-15b. The expression and purification of these constructs are the same as previously stated for NapD.

Binding was probed by co-purification experiments, such that if NapA binds to NapD it was co-purified. The first construct prepared using translational coupling was NapD with NapA114, which corresponds to a fragment of the first domain containing the iron-sulfur cluster (Arnoux *et al.,* 2003). Once it was shown to bind, full length NapA (828 amino acids) and NapA35 were also co-expressed. (Figure 2.4a) NapA35, which corresponds to the Tat signal peptide, was prepared to probe whether this fragment of NapA participates in NapD binding. As previously stated, this leader peptide is cleaved in the mature protein.

Figure 2.4 The different constructs engineered to probe for binding by co-expression. (a) The first set of experiments were carried out with the His tag on NapD, so as to survey binding based on the presence of NapA fragments. (b) Here NapA fragments contain the His tag, therefore binding is probed by looking for the presence of NapD.




An alternative approach needed to be taken for the cloning of NapD and His tagged NapA fragments (Figure 2.4b) in the pET-15b vector. The first step required the removal of the His tag from the available NapD construct (Figure 2.5). To accomplish this NapD must be ligated into Nco1 and BamH1 restriction sites of the pET-15b vector (Figure 2.5b). The construct used as the template for the first PCR was the one in which His tagged NapD was co-expressed with full length NapA. An outline is illustrated in Figure 2.6. Synthetic primers NapD-TF and NapD-TR (Table2.1) were used to amplify non-His tagged NapD. The second PCR used His-tagged NapA as the template, primer

X-F and a reverse primer depending on the desired fragment of NapA (NapA35-R, NapA27-R or NapA18-R) (Table 2.1). Primer X-F incorporates the His tag to NapA. The thrombin cleavage site comes from the initial His-tagged NapA construct. The reverse primer for NapD, NapD-TR and the forward primer for NapA, X-F, were designed so that they are complementary strands (Figure 2.5). Therefore the PCR products from the first two reactions can be joined together in the third PCR. Once NapD is cloned with His tagged NapA, the PCR product is gel purified and ligated into NcoI and BamHI restriction sites of the pET-15b vector. The sequences of all these constructs were verified by DNA sequencing.

Figure 2.5 Schematic showing the two strategies for the cloning of NapD. (a) For the incorporation of the His tag at the N-terminal of NapD, the pET15b vector is double digested with NdeI and BamHI so that NapD can be incorporated into the NdeI and BamHI restriction sites. (b) For the cloning of NapD without His tag, the pET15b vector is double digested with NcoI and BamHI so that NapD can be incorporated into the NcoI and BamHI restriction sites.

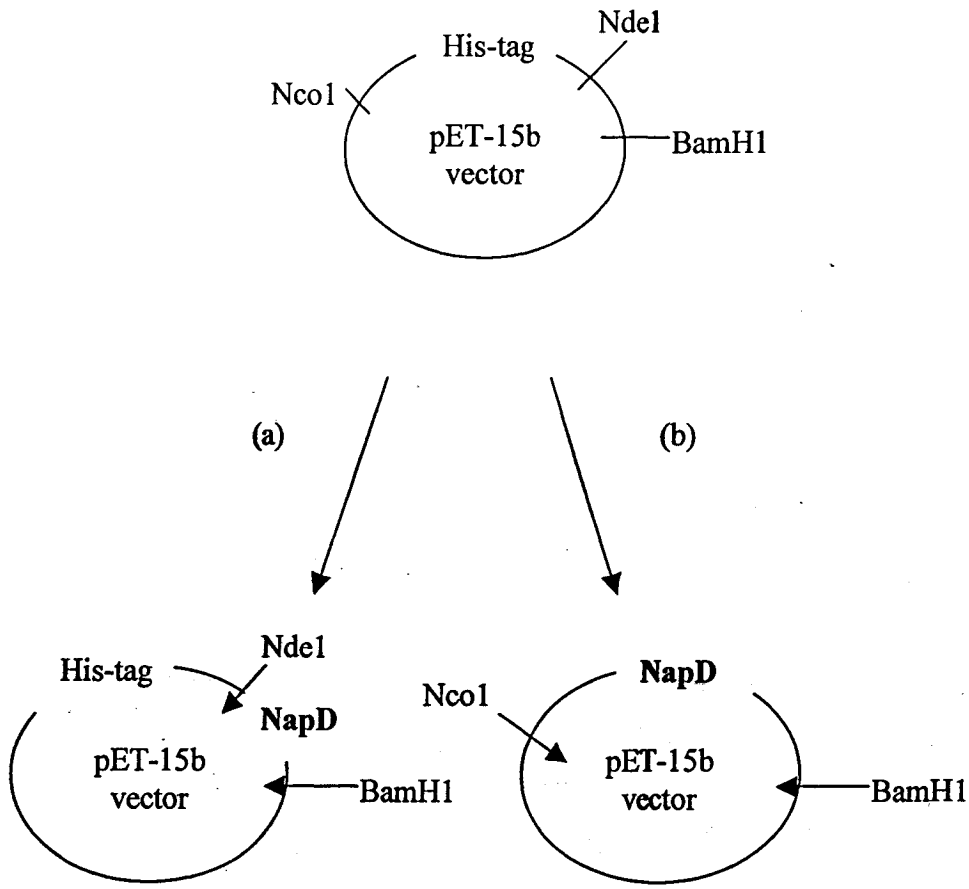
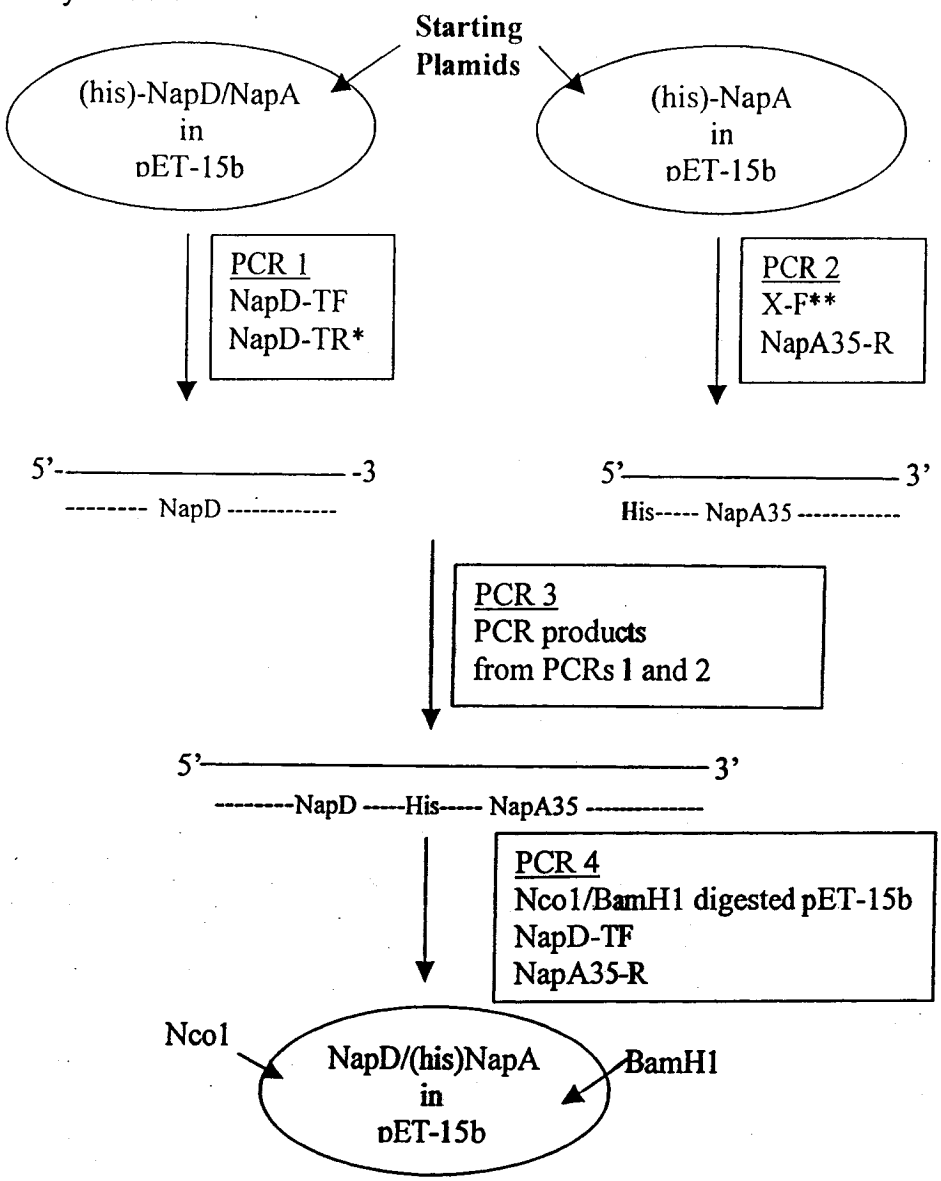


Figure 2.6 Outline of the cloning procedure for His tagged NapA35 cloned with NapD. The design for the important primers are labeled below. Primers NapD-TR and X-F are complementary strands.




NapD-TR* 5'-ATG GTG ATG GTG ATG CAT CA TGG TGT TTC CTC ACC TTG-3'
complementary base pairs of His-tag complementary base pairs of NapD

X-F** 5'-GAG GAA ACA CCA TG A TG CAT CAC CAT CAC CAT CAC AG-3'
base pairs of NapD

NapD Stop Codon | NapA Start Codon | His | His | His | His | His | His — $His_6$-Tag —

### 2.4.2 Co-Purification of NapD/NapA

SDS-PAGE was used to probe binding of co-purified fragments. Full length NapA, NapA114 (Figure 2.7) and NapA35 (Figure 2.8) were able to bind NapD. His-tagged NapD has approximately the same molecular weight as NapA114, therefore it is difficult to discern the two on a gel. However electrospray mass spectrometry was used to verify the presence of NapA114 (data not shown). The results suggest that the binding site of NapA is found within the first 35 amino acids. Another important thing to note is that the binding that occurred between NapD and NapA35 is quite strong to withstand the extensive washing done during purification.

Figure 2.7 SDS-PAGE showing NapD co-expressed with different fragments of NapA. Relevant lanes are numerically labeled. Lane 1 is NapD co-purified with NapA (full-length). Lane 2 is NapD co-purified with NapA114.

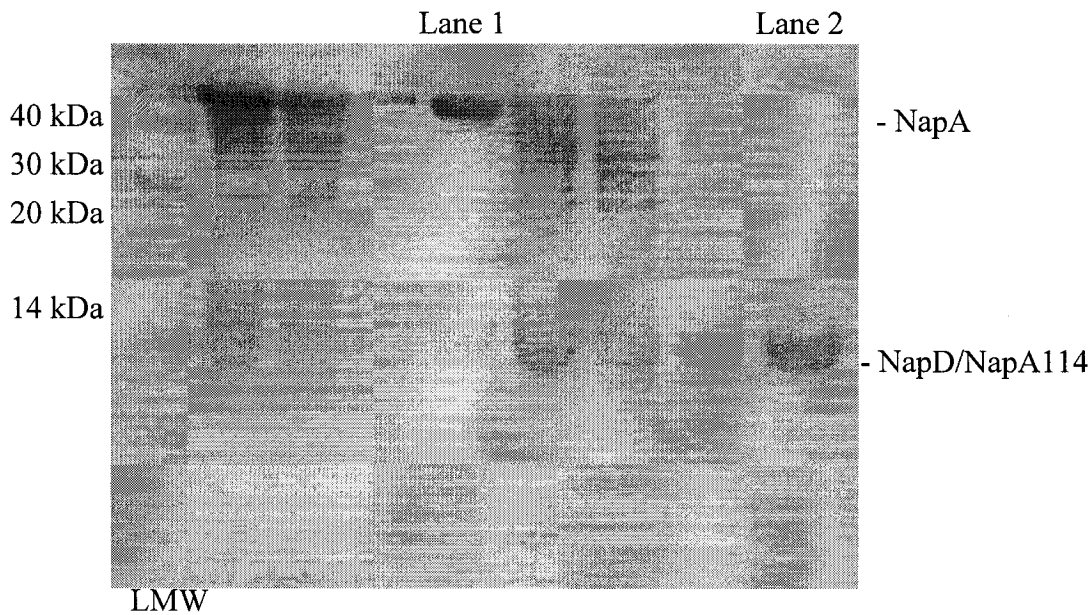
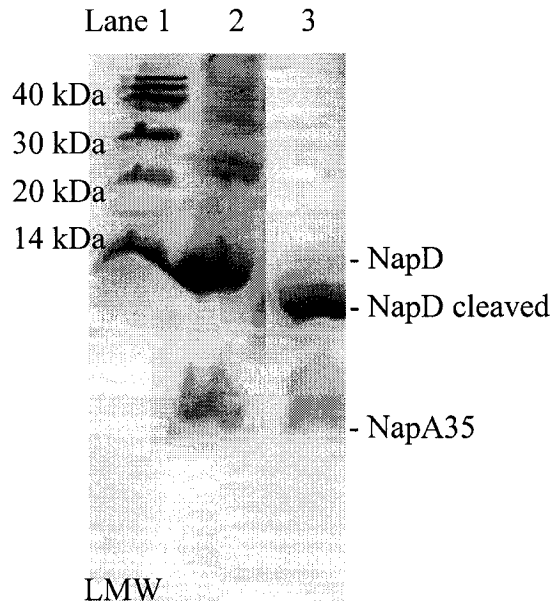
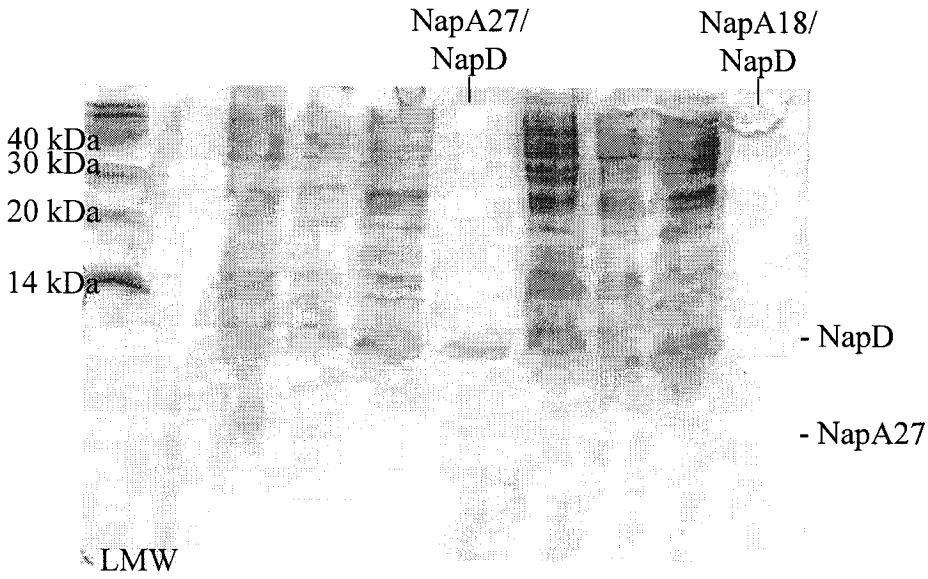


Figure 2.8 SDS-PAGE showing NapD and NapA35 co-purified. The first lane is low molecular weight marker, the second NapD/NapA35 before cleavage, whereas the third lane shows NapD/NapA35 after cleavage of His tag on NapD.




To identify a minimal fragment of NapA which will still bind NapD a different approach needed to be taken. The reason for this was that the detection limit of the SDS-PAGE gels and the level of staining with Coomasie Blue were impending. As an alternative a new set of constructs were engineered such that the His tag would be found on NapA (Figure 2.4b). As a means of validating the method NapA35 was also co-expressed in that fashion (data not shown). Two more constructs were prepared using the same approach, NapA27 and NapA18 (Figure 2.4b). All other expression and purification steps are the same as previously indicated. Binding is probed by detecting the presence of NapD since NapA is His tagged.

Figure 2.9 Co-purification results for the two constructs, where the His-tag was placed on NapA27 and NapA18.




As can be seen from the gel in Figure 2.9, NapD was present in the fraction that was co-expressed with NapA27, but not with NapA18. Although binding was confirmed for the NapA27 fragment, it does not mean that the NapA18 fragment does not bind, but it does suggest that the region between Ala18 and Gly27 is rather important in binding because once that region is removed there is definitely a loss in the binding affinity between NapD and NapA.

## 2.5 Purification and Characterization of NapA35

The results obtained from co-purification allowed for a preliminary assessment of the interaction between NapD and NapA, however the isolation of the desired fragment of NapA is essential for a more complete characterization of the system by NMR. The peptide of choice for further studies was NapA35. This fragment was chosen not only because its size is suitable for the planned NMR work but also because it is expressed and purified to reasonable amounts. NapA35 was also shown to be in a 1:1 molar ratio with NapD by size exclusion chromatography (data not shown).

Whether the construct with the His tag on NapD or NapA35 was used the initial procedures for purification were the same as previously discussed. Once NapD and NapA35 were co-purified, NapA35 needed to be isolated for studies by NMR.

The NapA35 peptide was purified by reverse phase HPLC (Waters Corporation; Vydac C18 semi-preparative column, 10 x 250 mm), using a linear gradient of 0-80% acetonitrile (ACN) in 0.1% TFA. The peaks were collected manually by monitoring the absorbance at 215nm. As can be seen from Figure 2.10, NapA and NapD were readily separated. NapA35 eluted at 48% ACN-0.1%TFA, whereas NapD, eluted at 60% ACN-0.1%TFA. For a 35 amino acid peptide NapA35 eluted at a rather high percentage of acetonitrile. This is probably due to the fact that it is quite hydrophobic, requiring a higher concentration of solvent to be eluted. To identify the main peaks, the samples were separated by SDS-PAGE (Figure 2.11).

Figure 2.10 HPLC chromatogram of the separation of NapA35 from NapD.

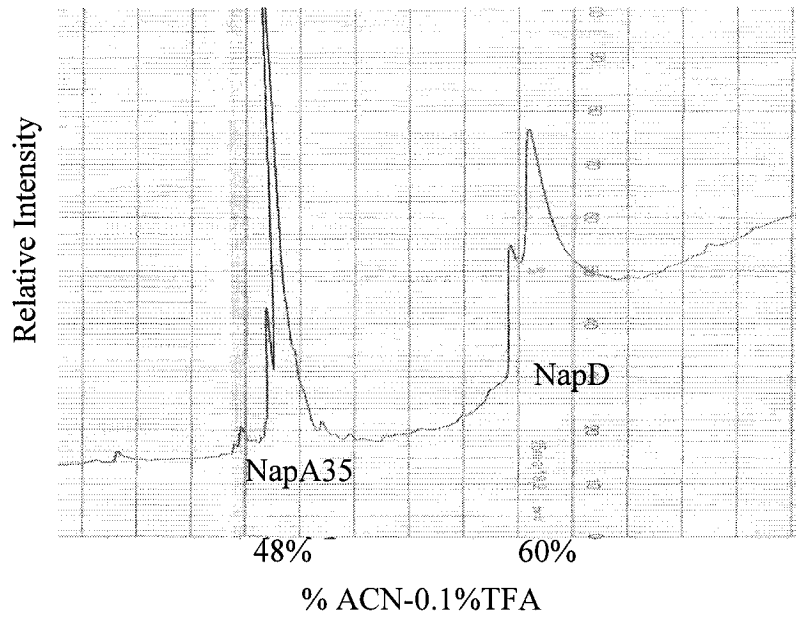
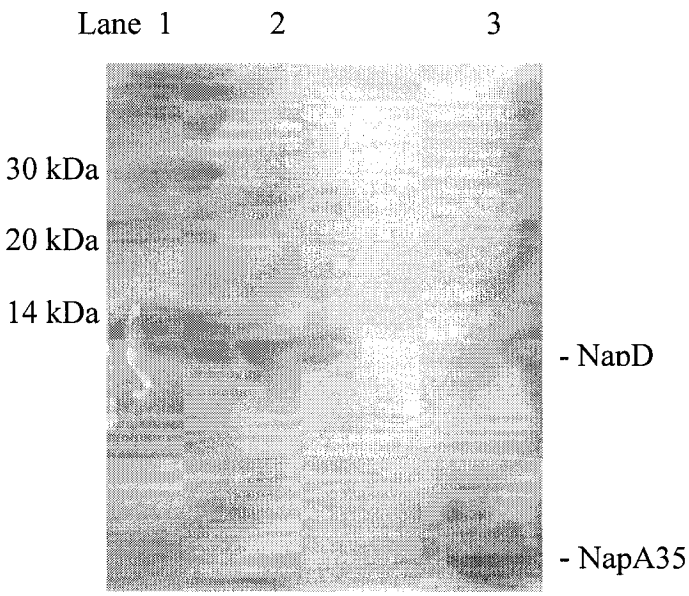


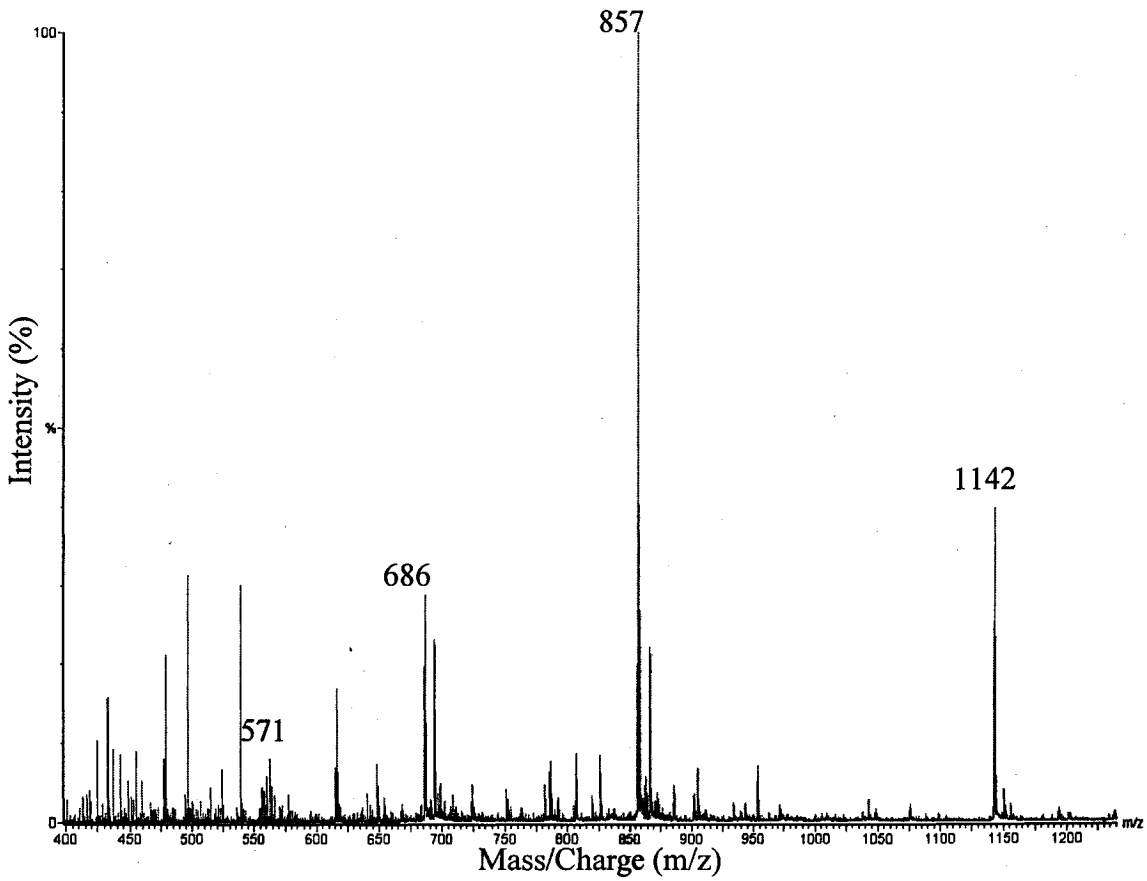

Figure 2.11 SDS-PAGE gel of the two peaks obtained from HPLC for verification of complete separation of complex. Lane 1 is low molecular weight marker. Lane 2 is NapD obtained from HPLC (2nd peak on HPLC chromatogram). Lane 3 is NapA35 obtained from HPLC (1st peak on HPLC chromatogram)




The final products were lyophilized and characterized using electrospray mass spectroscopy (SCIEX AP1 II mass spectrometer). As expected the mass of unlabeled

NapA35 was 3426 Da. (Figure 2.12) Furthermore the sequence of the peptide was verified using MS/MS.

Figure 2.12 Electrospray Mass spectrum used to determine the molecular weight of NapA35. As expected the molecular weight of NapA35 (unlabeled) is 3426 Da. The table below lists the charge and intensity of the labeled peaks. Each peak corresponds to the mass of the peptide divided by its relative charge. For example peak 857, corresponds to 3426/4, where 4 represents the charge.




| Peak | Charge | Intensity (%) |
|---|---|---|
| 1142 | 3 | 40 |
| 857 | 4 | 100 |
| 686 | 5 | 30 |
| 571 | 6 | 10 |

## 2.6 Summary

In this chapter we noted that *NapD* is located upstream to *NapA* and that the stop codon of NapD overlaps the start codon of NapA. NapD was expressed with a His tag, where the peptide fragments having no tag were co-expressed using a translational coupling system. Since NapA fragments co-purified with NapD we were able to prove that NapD does in fact bind to NapA. From the different constructs of NapA we managed to demonstrate that the binding site is found within the N-terminal region of NapA. The binding site overlaps with the Tat leader sequence, which is involved in transporting NapA out into the periplasm. Thus this signal peptide has a dual function; it triggers NapA out of the cytoplasm and binds the chaperone NapD. We also managed to show that NapA27 is actually enough to bind NapD, however it was difficult to asses from SDS-PAGE analysis whether it binds NapD with the same affinity. The fact that the NapA18 fragment did not give a positive result from the SDS-PAGE analysis suggests that the sequence from NapA18 to NapA27 is important. Although NapA18 cannot be co-purified it may still bind NapD, but with a much lower affinity.

# Chapter III Structural Characterization of NapD, NapA35 and NapA114 by Circular Dichroism, Fluorescence and Ultra-Violet Spectroscopy

## 3.1 Circular Dichroism Spectroscopy

### 3.1.1 Foreword

Circular dichroism (CD) measures the difference in absorption of left-handed and right-handed circularly polarized light in a sample solution (Creighton, 1993). CD detects optically active molecules making it sensitive to the 3D conformation of the amide backbone. Since the peptide bond angles control the magnitude and direction of electronic transitions, the ordered geometric arrangements of an alpha-helix results in CD spectra that are distinct from beta-sheet structures (Sarver and Krueger, 1991). Although CD does have its limitations for accurate predictions of secondary structure, it does present a method for probing structural changes that may occur in complexes. For preliminary characterizations, CD will be used to probe the secondary structure of NapD. An effort to understand the mechanism of interaction will begin by determining whether NapA35 and NapA114 form some sort of secondary structure upon binding to NapD.

### 3.1.2 Materials and Methods

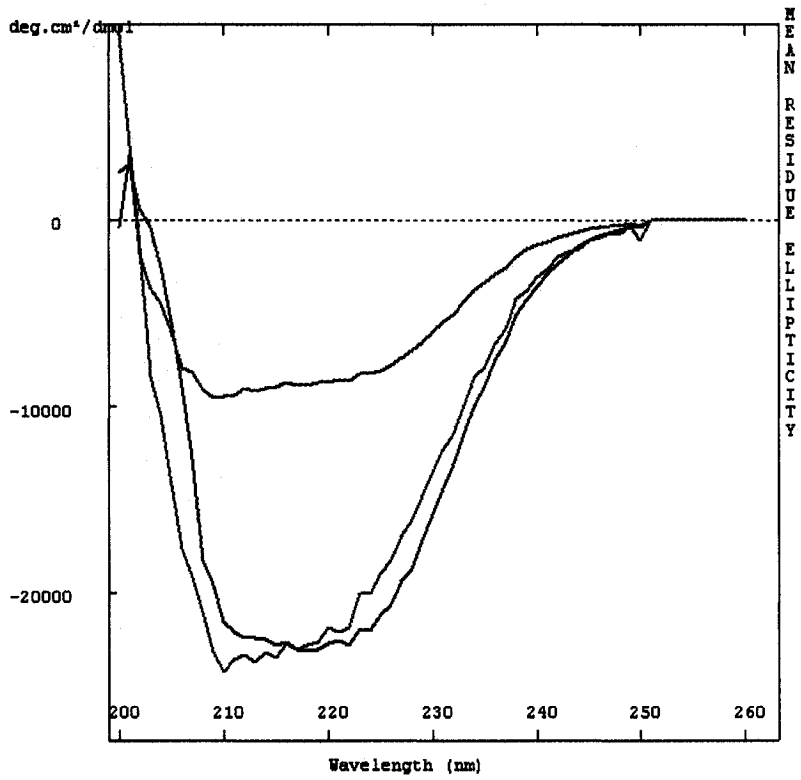
Circular Dichroism spectra were collected on a Jasco J-710 CD spectrometer. These experiments were performed with a protein concentration of 130μM for NapD in 50mM $NaH_2PO_4$-$H_2O$ Buffer, with 150mM NaCl and 10mM DTT, pH 6.8. The co-expressed peptide fragments NapA35 and NapA114 also had concentrations of 130μM. Protein concentrations were determined as previously indicated. Spectra were acquired in

the region from 200-250nm with a 0.01cm path length cell (~200μL) and a scan speed of 100nm/min with a response time of 0.25s. Baseline subtraction and smoothing were done using the Jasco software. The spectra are expressed as molar ellipticity per residue where they are the average of 5 scans. PHD (Predict Secondary Structure) is an algorithm, which predicts the secondary structure of proteins based on their amino acid propensities using information from multiple sequence alignments (Rost & Sander, 1993). PHD was used to compare the results obtained from NMR.

### 3.1.3 Results and Discussion

In an attempt to characterize the mechanism of interaction between NapD and NapA, different fragments of NapA were co-expressed with NapD. In this chapter NapA114 and NapA35 were used to study binding with NapD. When CD was initially performed the structure of NapD was unavailable, however, it was later solved by Ovidiu Minailiuc using NMR (unpublished results). The CD spectra obtained were compared with the NMR structure and PHD predictions. The spectra are expressed as mean residue ellipticity for all residues (Figure 3.1). Consequently, samples for the complex were calculated as a function of both components. The total concentration was taken into account when plotting the spectrum and analyzing it.

Figure 3.1 The CD spectra of NapD (red), NapD/NapA35 (blue) and NapD/NapA114 (green), expressed as mean residue ellipticity (deg $cm^2$/dmol).




Three methods were used to assess the secondary structure of NapD/NapA. PHD was advantageous since it allowed us to determine the secondary structure of the protein or peptide fragments when they are free in solution. However, this method was not feasible for studying changes in the complex. The accuracy of this method is approximately 70 %. The other two methods used, circular dichroism and NMR are experimental techniques that allowed for the study of the complex. Results from NMR and PHD are tabulated in Table 3.1.

Looking at the CD spectrum of NapD indicates that it is composed of both alpha-helices and beta-sheet structures. The results obtained for NapD are in agreement with the recent structure solved by NMR in our group (Ovidiu Minailiuc, unpublished results) and

those obtained by PHD. Looking at NapD/NapA35 we can see that the CD spectrum has slightly shifted in comparison to the spectra for NapD (Table 3.1). When compared with the results from the structure of NapD from NMR, we can see that NapD has a greater percentage of beta-sheets when in complex. Therefore the small changes in CD are consistent with those from NMR.

Table 3.1 Comparison between the experimental results obtained from NMR versus those predicted by PHD based on amino acid sequence. Structure of NapD obtained from NMR is compared for both the free and complexed form with NapA35. H: alpha-helical content, E: beta-sheet content, Other: not determined.

| **Protein/Complex** | **Secondary Structure (%)** | | | | | |
|---|---|---|---|---|---|---|
| | **Predicted by PHD** | | | **Obtained from NMR of NapD** | | |
| | H | E | Other | H | E | Other |
| **NapD** | 29 | 38 | 33 | 31 | 34 | 35 |
| **NapA35** | 68 | 0 | 32 | - | - | - |
| **NapD/NapA35** | 40 | 27 | 33 | *31 | *38 | *31 |
| **NapA114** | 21 | 20 | 59 | - | - | - |
| **NapD/NapA114** | 24 | 28 | 48 | - | - | - |

* Note that for the NapD/NapA35 complex the results from PHD are for both components, whereas results from NMR are only for NapD (in complex).

On the other hand for the NapD/NapA114 complex there is a decrease in CD signal suggesting the presence of a mostly unstructured peptide. This result can be interpreted using the structure of NapA/NapB (Figure 1.3) solved by X-ray (Arnoux *et al.*, 2003). NapA114 is a sub-domain of domain one. Likely the NapA114 fragment corresponding to the iron-sulfur cluster is not structured in the absence of the rest of the domain. Since the results were calculated according to the mean residue ellipticity there is a decrease in signal, however if the results had been plotted relative to NapD, the spectra would have been very similar to that of NapD alone.

Although CD is a technique that is very useful in predicting secondary structures of unknown proteins care must be taken into consideration before speculating on exact percentages. To obtain exact percentages by CD a more accurate determination of the concentration of each component in the complex would be needed. Furthermore spectra would have to be acquired in the region from 184-250nm to use deconvolution programs more accurately. This could not be done because NapD was in 50mM $NaH_2PO_4$-$H_2O$ buffer, which is known to absorb in the range where structural features display different absorptions of circularly polarized light. An alternative buffer is 10 mM sodium or potassium phosphate. The pH was adjusted using HCl, however chloride ions are known to interfere with CD at lower wavelengths. To establish the appropriate pH range for CD experiments, borate and ammonium salts may be used. 150mM NaCl was added to the buffer, however an alternative would be $Na_2SO_4$ or NaF.

Interpretation of CD would be a simple task if the analyzed protein contained only one chromophore and a limited number of regular-shaped structures (Woody *et al.*, 1996). However, the fact is that not all elements are independent electronic systems. There are several assumptions in secondary structure analysis that can lead to errors. Besides setting the experimental conditions with extreme accuracy the effect of tertiary structure is negligible. Furthermore, aromatic residues and disulfide bonds can contribute significantly to a far-UV CD spectrum (Anderson *et al.*, 2001).

In the end CD was successful in demonstrating that NapD is composed of both beta-sheet and alpha helical structures. Furthermore the spectra of NapD with NapA35 does undergo some change, however more studies are needed for precise measurements of secondary structure. Never the less, the NapA114 fragment seems relatively unstructured.

## 3.2 Fluorescence Spectroscopy

### 3.2.1 Foreword

To probe the tertiary environment of the aromatic groups of NapD in the presence of NapA114 and NapA35 fluorescence spectroscopy was performed. The fluorescence of proteins is due to the aromatic residues tyrosine and tryptophan (Teale, 1960,). For the majority of systems the fluorescence for tyrosine is lower than that for tryptophan, which dominates the emission spectra (Teale, 1960). The response of tyrosine and tryptophan fluorescence to their environment will serve to probe the conformation of NapD in complex.

### 3.2.2 Materials and Methods

Fluorescence measurements were carried out on an Aminco Bowman series 2 spectrofluorometer. Experiments were performed at 25 °C with a protein concentration of 21μM for NapD in 50mM Phosphate Buffer with 150mM NaCl and 10mM DTT, pH 6.8. Peptide concentrations were also 21μM. Fluorescence emission spectra were recorded with excitation 280nm and 295nm, with bandwidth 4nm. Spectra were collected from 300 to 400nm in 1cm quartz cuvettes.

### 3.2.3 Results and Discussion

Evaluation of the tertiary fold of NapD was accomplished by fluorescence spectroscopy. Samples were run at an excitation wavelength of 280nm and 295nm, to determine whether the aromatic residues in NapD are involved in the binding of NapA.

The simplification of the analysis at an excitation of 295nm is due to the fact that there are no Trp or Tyr residues found in the first 35 amino acid residues of NapA. (Table 3.2)

Table 3.2 Aromatic residues present in NapD and NapA fragments.

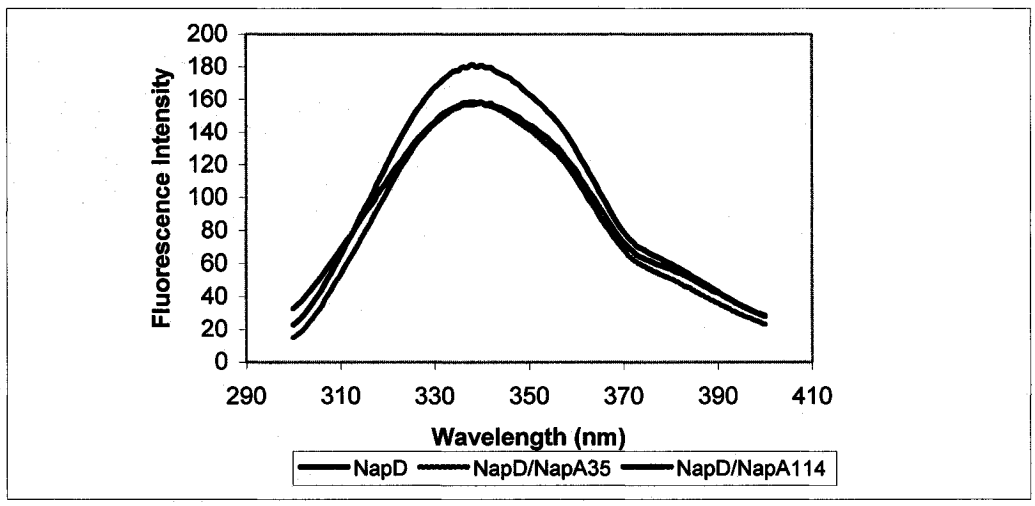
| Protein/Peptide | Aromatic Residues (not including Phe) |
|---|---|
| NapD | Trp5 and Tyr76 |
| NapA35 | No Trp or Tyr |
| NapA114 | Trp41, Tyr85,92 and 109 |

The fluorescence intensity of Trp5 in NapD was used as a spectroscopic probe of the tertiary structure. The Lambda-max of the emission spectrum is identical in NapD and NapD co-expressed with NapA35 or NapA114 for both emission spectrums at an excitation of 280nm and 295nm (Figure 3.2).

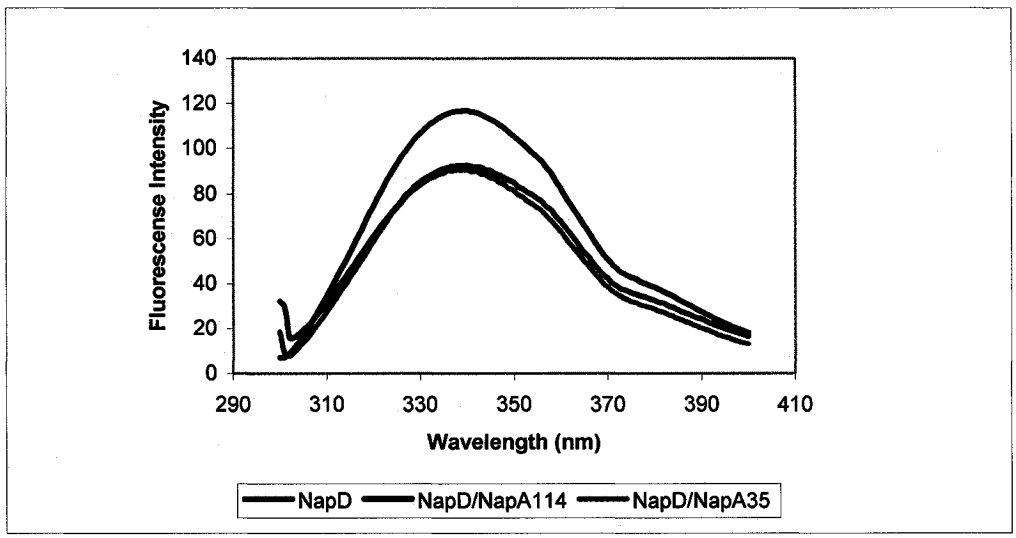
From the emission spectra excited at 295nm we can see that the Trp residue of NapD remains in the same environment in the presence of the peptides (Figure 3.2B). Furthermore, we can also conclude from the consistency of the Lambda-max in the spectra excited at 280nm that the environment of both Trp5 and Tyr76 in NapD are also unchanged in the presence of the peptides (Figure 3.2A). These experiments demonstrated that the aromatic groups of NapD are not at the binding interface.

Figure 3.2 Fluorescence Emission Spectra of NapD, NapD/NapA35 and NapD/NapA114 at an excitation of (a) 280nm and (b) 295nm.

(a) Fluorescence emission spectra at an excitation of 280nm.




(b) Fluorescence emission spectra at an excitation of 295nm.



## 3.3 Ultra-Violet Spectroscopy

### 3.3.1 Foreword

In addition to fluorescence, UV spectroscopy was used to simulate the aromatic groups of NapD under different conditions, that is with or without NapA114/35. Fourth derivative UV spectra were calculated (Lange et al., 1996) and analyzed, to obtain the structural environment of the aromatic groups of NapD, and the complex.

### 3.3.2 Materials and Methods

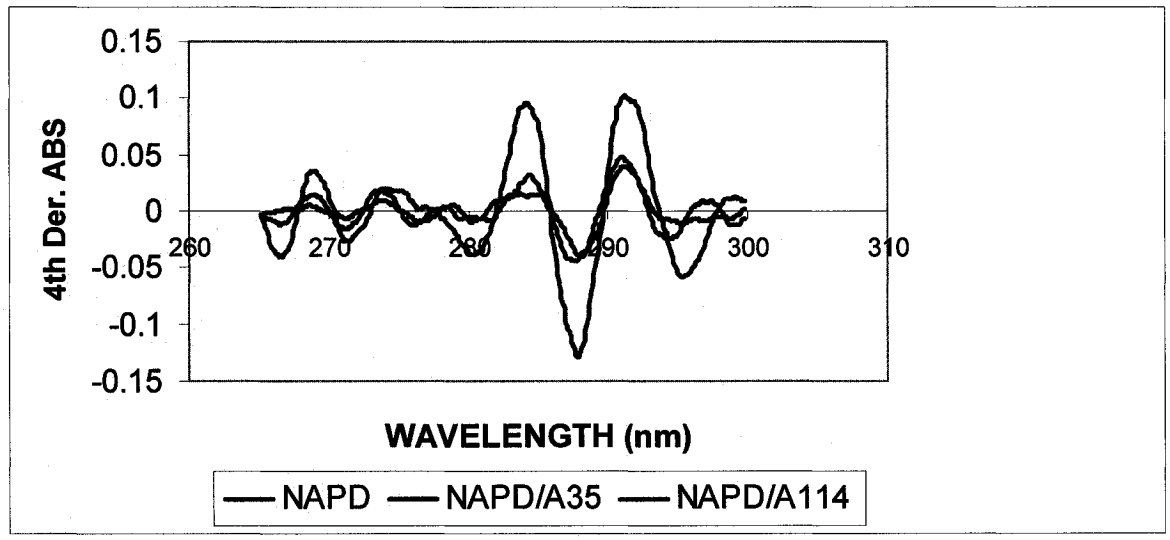
Ultraviolet spectra were recorded on a Varian Cary 1 spectrophotometer. Samples were scanned from 300 to 265 nm. The spectral bandwidth was 1.0nm; the step size was 0.1nm, averaging time 1.0s and scan rate 6.0nm/min. Experiments were performed at 25 $^{o}$C with the same protein concentrations and buffer as fluorescence. The spectrum of the buffer was recorded using the same settings and subtracted from all other spectra. The number of spectra taken was dependent on the $OD_{280nm}$, a minimum of 10 scans were taken. Spectra were averaged and the fourth derivatives calculated (Lange *et al.*, 1996). Calculating the 4th derivative of the UV spectrum makes the subtle differences in the environments of the aromatic amino acids more evident. The results from fourth derivatives were compared to spectra with different polarities, including water, 50% ethanol, 100% ethanol, 1-Butanol and Hexane to simulate the environment of the aromatic groups.

### 3.3.3 Results and Discussion

Fourth derivative spectra have been used in an effort to determine as much information as possible concerning the local environment of the aromatic groups. As can

be seen in Figure 3.3, the 4th derivative spectra for NapD and NapD/NapA35 are approximately the same. This suggests that the environment of the aromatic groups of NapD are not altered in the presence of NapA35. Considering NapD/NapA114 we can see that the intensities have increased, which will later be shown to be due to the presence of the three Tyrosines and one Tryptophan of NapA114.

Figure 3.3 4th derivative ultraviolet spectra for NapD, NapD/NapA35 and NapD/NapA114

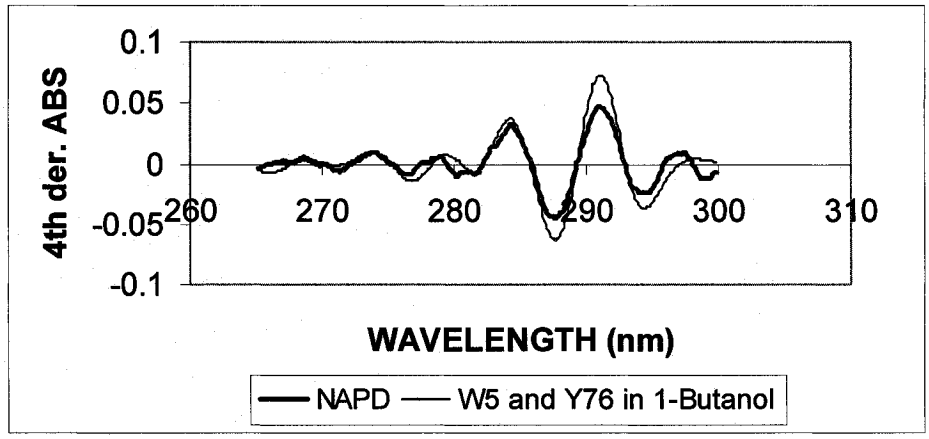



The common measure of the polarity of solvents is the dipole moment, dielectric constant and their miscibility in water (Streitwieser *et al.,* 1992). The most widely used polar solvent is water, with a dielectric constant of 78.5. 50% ethanol has a dielectric constant of 54 (Besteman *et al.,* 2005) and ethanol, with a dielectric constant of 24.3, has intermediate polarity. 1-butanol has a dielectric constant of 17.8, whereas, hexane has a dielectic constant of 1.9 and is known as a non-polar solvent. Generally polar or ionic compounds will only dissolve in polar solvents. Simulation of the obtained 4th derivative spectra with known standards in different solvents indicated that both aromatic residues of NapD are found in an environment that is hydrophobic, similar to conditions in 1-

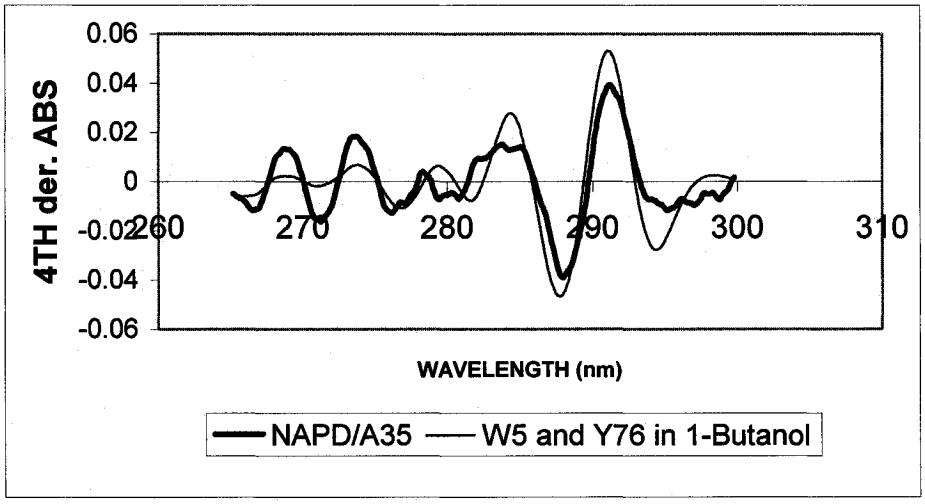
butanol (Figure 3.4). Furthermore the presence of NapA35 or NapA114 did not change the environment of Trp5 or Tyr76 of NapD (Figure 3.4B,C), which is coherent with fluorescence results. In other words the aromatic residues of NapD were simulated in an environment similar to that of 1-butanol for all conditions. This is consistent with Trp5 and Tyr76 being partially protected in the NMR structure of NapD. UV 4th derivative spectra suggests that Trp5 in NapD and Trp41 in NapA are found in a hydrophobic environment. However the fact that fluorescence quenching was not observed (Figure 3.2B) suggests that they are not found in close proximity. UV also showed that Tyr85, 92 and 109 of NapA114 are in a relatively polar environment when bound to NapD. This is in agreement with results obtained from CD, which suggest that NapA114 remains quite unstructured in the presence of NapD.

Figure 3.4

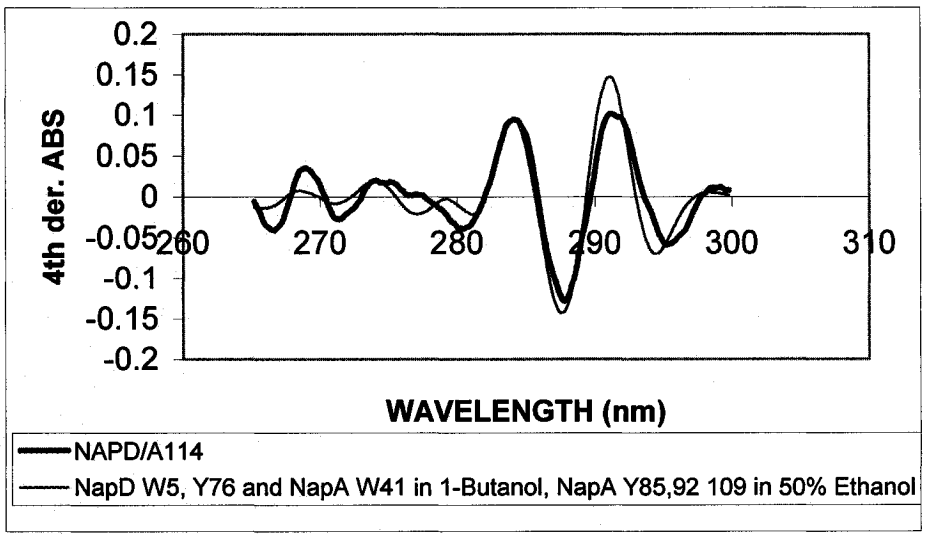
A. 4th derivative simulation of NapD.




B. 4th derivative simulation of NapD/NapA35




C. 4th derivative simulation of NapD/NapA114



### 3.4 Conclusions from Biophysical Techniques

Ultimately, the results indicate that the secondary structure of NapD is one that is composed of a mixture of alpha-helices and beta-sheets. The results are in agreement with the NMR structure of NapD. As previously stated the limitations of CD cannot allow us to conclude the exact percentages of secondary structure without some degree of uncertainty. The small changes observed in the CD spectrum for NapD/NapA35 will be examined in chapter five. Tertiary structural analysis revealed that the aromatic groups of NapD were found in a hydrophobic environment, which did not change significantly in the presence of NapA fragments, indicating that the aromatic residues of NapD are not at the binding interface. The results do suggest that that structure of NapD does not undergo a major conformational change during complex formation, which is consistent with the NMR structure of NapD. In the end, this chapter provides a foundation for in depth structural analysis by NMR or X-Ray crystallography studies. More studies need to be done in order to understand the mechanism of interaction and the role of the N-terminal consensus sequence of NapA.

# Chapter IV Probing Binding between NapD and NapA35 By Nuclear Magnetic Resonance

## 4.1 Foreword

For a more detailed study of the interactions between NapD and NapA35 Nuclear Magnetic Resonance (NMR) was selected. This technique allows us to probe the residues of both protein and peptide that are involved in binding. This can be accomplished by examining changes in chemical shift upon complex formation for individual amino acids. To realize this the signals obtained from NMR must be assigned (Kay *et al.*, 1990).

This chapter deals with the assignment of the backbone signals of NapA35 free and in complex with NapD. Additionally chemical shifts will be used to determine which part of the NapA leader peptide participates in chaperone binding.

## 4.2 Materials and Methods

To obtain information on the backbone of NapA35 $^{15}$N-labeled sample is needed. To prepare isotopically labeled protein, cultures were grown in minimal medium (M9). One liter of M9 is composed of 6g $Na_2HPO_4$, 3g $KH_2PO_4$ and 0.5g NaCl. 2mM MgSO4, 0.1mM $CaCl_2$, 5mg/L FeSO4, 0.02mg/L Vitamin B1, 1g/L $NH_4SO_4$ and 2g/L glucose is added to M9. For $^{15}$N-labeled samples 1g of $^{15}NH_4SO_4$ is added as the nitrogen source instead of $NH_4SO_4$. The remaining protein expression and purification steps are the same as described in Chapter 2. NMR Samples were 300µl and had protein concentrations of 1.5mM. $D_2O$ (30µl) was added for locking purposes. All NMR experiments unless stated otherwise were carried out at 312.4 K and were recorded with a Bruker-DRX 500 NMR

spectrometer equipped with pulse field gradients and a cryo probe. The NMR spectra were processed with XWINNMR software version 3.5 (Bruker Biospin) and analyzed with XEASY program (Bartels *et al.*, 1995).

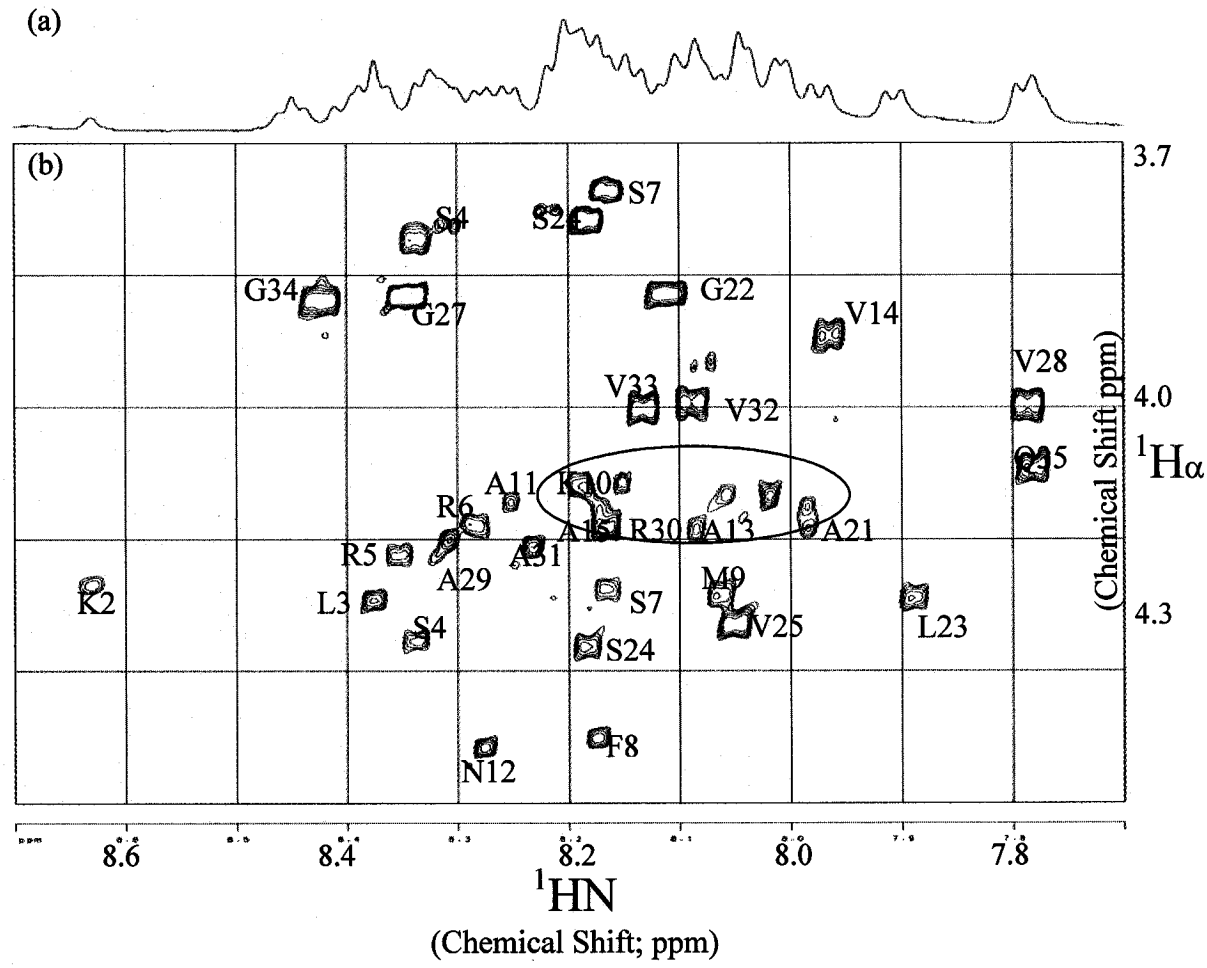
## 4.3 NMR Spectroscopy

### 4.3.1 Homonuclear NMR for NapA35

To assess the feasibility of NapA35 for NMR, a one-dimensional experiment with unlabeled peptide was carried out (Wuthrich, 1986). As can be seen in Figure 4.1a, the sample gave spectra with dispersed and sharp signals. Furthermore, NapA35 was soluble and present in reasonable amounts making it suitable for the planned NMR.

Assignment is needed in order to associate each NMR resonance with a specific nucleus. Homonuclear 2D experiments were used for the assignment of the unlabeled peptide. In the 2D TOCSY (Total Correlated Spectroscopy) experiment magnetization is transferred over a complete spin system of an amino acid by successive scalar coupling through bonds (Figure 4.2). Therefore, a characteristic pattern of signals results from which the type of amino acid can be identified. The 2D NOESY (Nuclear Overhause Effect Spectroscopy) experiment however, uses dipolar interaction of spins for correlation of protons that have a distance of less than 5 Å between them. In other words the nuclear Overhauser effect correlates protons of neighbouring amino acids in the sequence, as well as those close in space due to tertiary structure.

Figure 4.1 (a) 1D NMR spectra of NapA35. As can the seen the signals are sharp. The presence of doublets shows that the signals are well resolved and dispersed. (b) The NH-Hα region of 2D Homonuclear TOCSY for NapA35. Note that most of the signals are dispersed expect for some of the alanines in the poly-alanine stretch (circled). (c) Sequence of NapA35, where the residues in red are the ones that have been assigned.

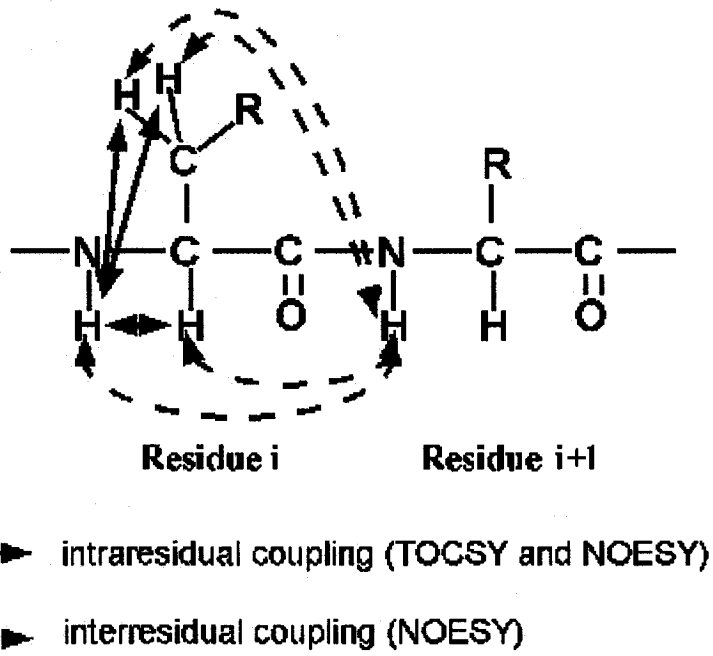



(c)

N-term-**MKLSRRSFMK ANAVAAAAAA AGLSVPGVARAVVGQ**-C-term

Homonuclear 2D experiments were used to assign the amino acid signals of the free peptide in solution. The 2D TOCSY experiment was used to assign the residues of unlabeled NapA35 based on their identity, whereas the 2D NOESY experiment was used to sequentially assign the residues in the sequence.

Figure 4.2 Schematic diagram of the information obtained from TOCSY and NOESY experiments. As can be seen from the blue arrows, TOCSY gives information through bonds and allows for the identification of amino acids types. NOESY on the other hand gives information through space allowing for sequence specific assignments.



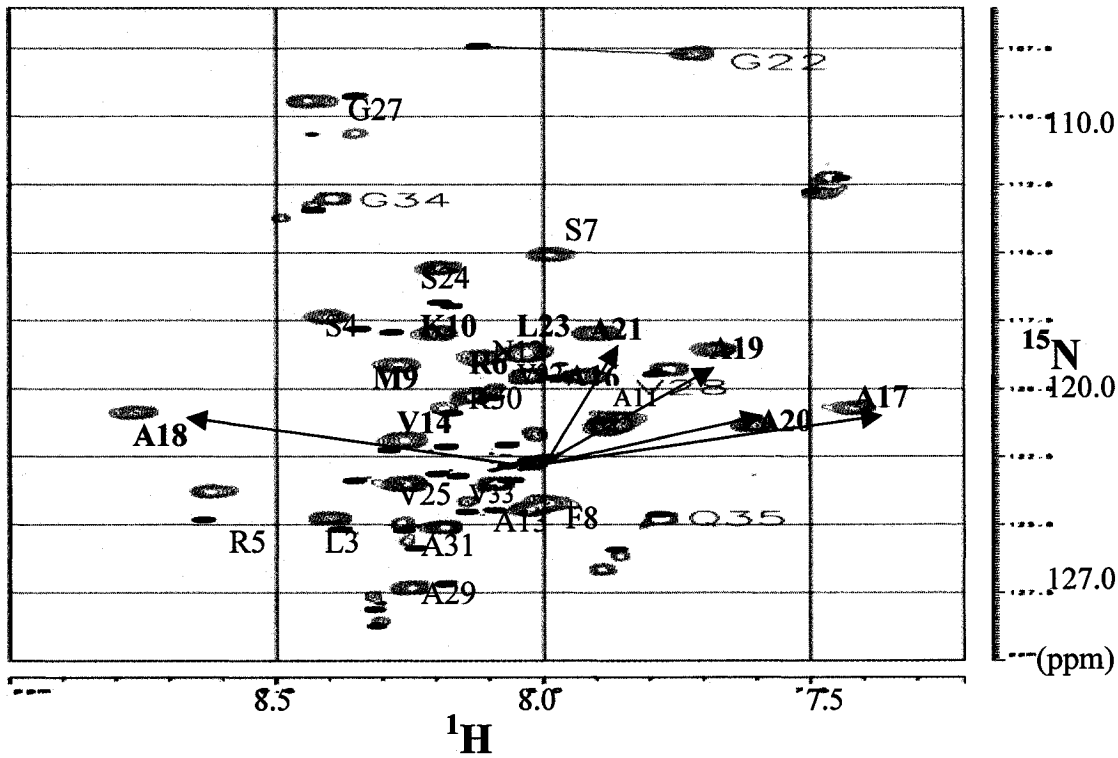

(http://www.cryst.bbk.ac.uk/PPS2/projects/schirra/html/assign.htm)

As can be seen in Figure 4.1b, which shows the NH-Hα region of 2D TOCSY, that all the peaks except for Met1, Ala16-20, and Pro26 have been assigned. The amino groups of the first residue is normally not present in NMR spectra. Also prolines are not seen because they do not have an amide proton. Furthermore the cluster of alanines that are unassigned is due to the overlap of the signals as can be seen in Figure 4.1b. NMR data suggest that the free peptide is unstructured in this region since the chemical shifts are similar to those of alanine when free in solution. To complete the sequential assignments 3D experiments are needed (Dubs *et al.*, 1979; Billeter *et al.*, 1982; Wagner and Wuthrich, 1982).

### 4.3.2 Heteronuclear NMR spectra and Backbone Signal Assignments

Heteronuclear NMR spectroscopy facilitates structure determination. An important NMR experiment is the $^{15}$N-HSQC (heteronuclear single quantum correlation), which correlates the nitrogen atom from the backbone with the proton attached to it. It is very useful for studying complexes. To be able to probe the amino acids of NapA35 involved in complex formation, $^{15}$N-HSQC experiment was run where NapA35 was $^{15}$N labeled and NapD was unlabeled (Figure 4.3). $^{15}$N-HSQC of free NapA35 was also acquired.

Figure 4.3 $^{15}$N-HSQC of NapA35 free (black crosspeaks) and in complex with NapD (green crosspeaks). Residues that underwent the largest changes in chemical shifts are labeled in red.



Each peak on $^{15}$N-HSQC, corresponds to a single amino acid (Figure 4.3). Here we can see only the amino acid residues of NapA35 because NapD is not $^{15}$N labeled. Amino acids of NapA35 in complex with NapD are labeled. As can be seen, amino acids towards the C-terminal do not undergo much changes in chemical shift. However, residues from the N-terminal all the way to approximately Leu23 do undergo significant changes in chemical shift, suggesting that these residues are important for formation of a complex with NapD. Those that underwent the greatest changes in chemical shift are labeled in red. The cluster of alanines in the poly-alanine stretch have been shifted as can be seen from the arrows.

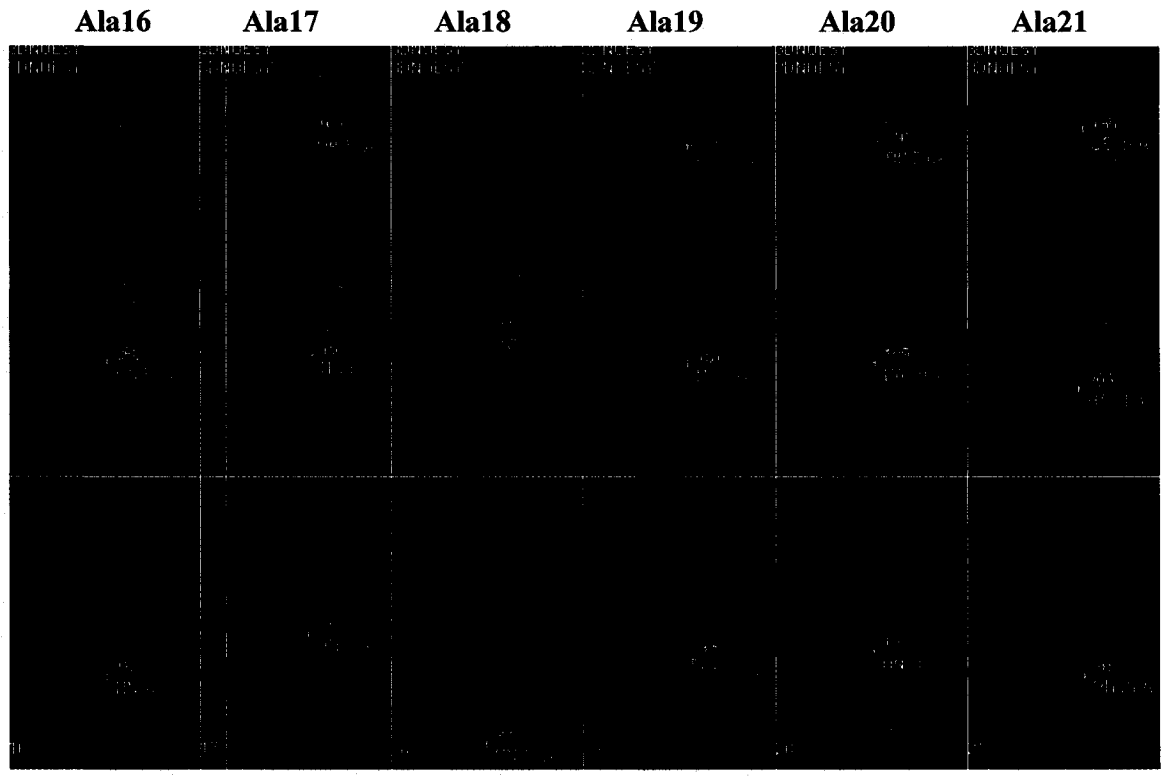
Ultimately, heteronuclear NMR was used to complete the assignment of free NapA35, however for the complex, heteronuclear experiments were necessary in order to distinguish between NapA35 and NapD.

The process of resonance assignment for $^{15}$N NapA35 is similar to the homonuclear case. The difference is that we can record three-dimensional experiments. Resonance assignments for the amide crosspeaks of $^{15}$N-HSQC were obtained using 3D $^{15}$N-TOCSY-HSQC and 3D $^{15}$N-NOESY-HSQC (Kay *et al.*, 1990).

Similarly to homonuclear NMR, TOCSY allowed for the identification of the type amino acid, whereas NOESY allowed for the assignments of individual amino acids in the sequence. These experiments build on the $^{15}$N-HSQC experiment, with the additional dimension corresponding to protons. The peaks that were shown to overlap in the proton dimension, can therefore be resolved if the corresponding nitrogens have different chemical shifts.

To be able to analyze these 3D experiments the cube is divided into strips where each strip corresponds to a single amino acid from $^{15}$N-HSQC. Figure 4.4, shows a segment of the 3D $^{15}$N-NOESY-HSQC strips, particularly in the region from Ala16 to Ala21. Recall in the free peptide, residues Ala16-20 had similar chemical shifts, however, in the complex they are greatly dispersed as seen in Figure 4.3, allowing for the complete assignment of NapA35.

Figure 4.4 3D $^{15}$N-NOESY-HSQC strips of NapA35 in complex with NapD. Ala16 to Ala21 is indicated and the sequential connectivities in the NH region are shown in purple. Also note in yellow the NH, Hα, and Hβ of the poly-alanine stretch, which were obtained from TOCSY.

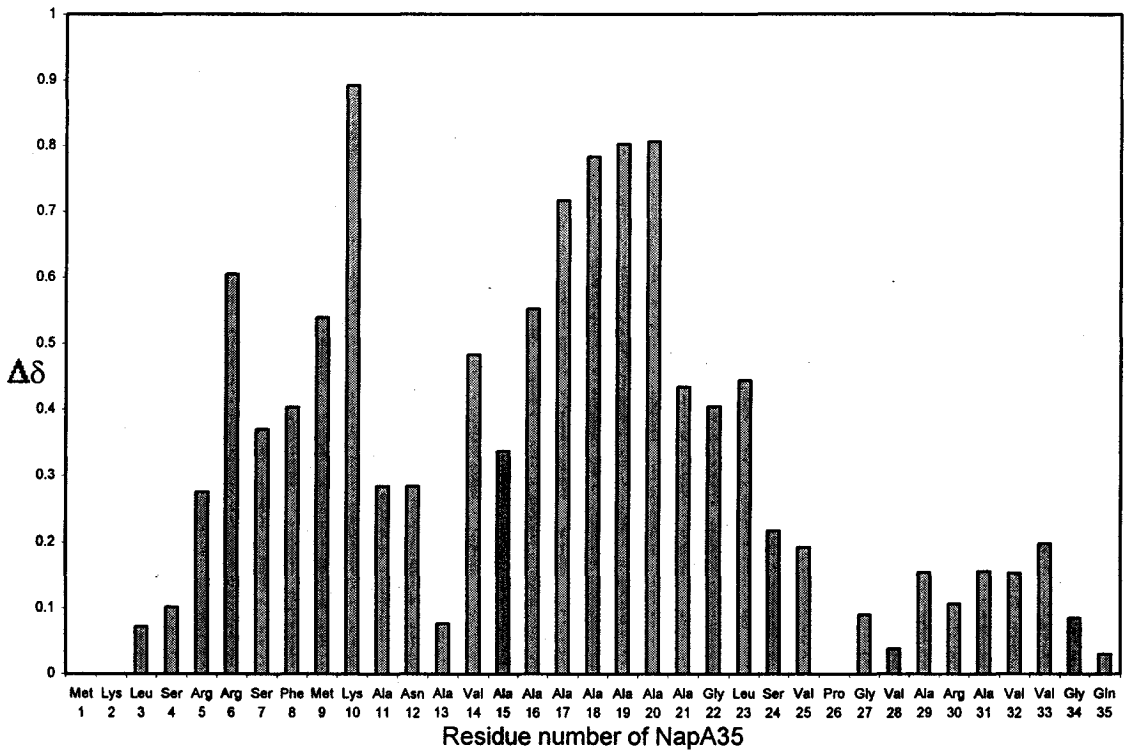



### 4.3.3 Mapping of the Binding between NapA35 and NapD

Once all the residues of the NapA35 have been assigned (Figure 4.3) we can use the assignments to probe the binding site. From co-purification in Chapter 2 we showed

that from the 828 amino acid residues of NapA, the first 35 amino acids are sufficient to bind NapD. Now we wish to determine exactly which of the NapA amino acids are involved in binding or in the formation of a potential structure upon NapD binding.

Figure 4.5 Histogram showing amide chemical shift differences between free NapA35 and its complex with NapD, calculated according to the formula $[(0.17N)^2 + (H)^2]^{1/2}$ (Farmer *et al.,* 1996).




As can be seen from the histogram (Figure 4.5) practically all 35 amino acids do undergo some changes in chemical shift, however the largest changes occur in the region from the N-terminal until Pro26. Specifically residues from Arg5 until Leu23 are involved in binding either directly or by aiding in the formation of a potential structure in complex. Residues that undergo the greatest changes are Arg6, Phe8, Met9, Lys10, Val14, Ala16-21, Gly22 and Leu23.

To further understand and discriminate between those amino acids that directly participate in NapD binding versus those that undergo a change in chemical shift due to the formation of a structure, a series of edited/filtered NMR experiments will be discussed in the upcoming chapter.

Chemical shift results are quite consistent with those obtained from co-purification experiments in chapter two where it was shown that the NapA27 peptide was enough to bind NapD whereas the NapA18 peptide was not. To reiterate, residues between Ala18 and Gly27 of NapA are important in binding as can be seen from the histogram of chemical shifts (Figure 4.5) and pull-down results. However, residues before Ala18 are equally important as can be seen in the aforementioned histogram.

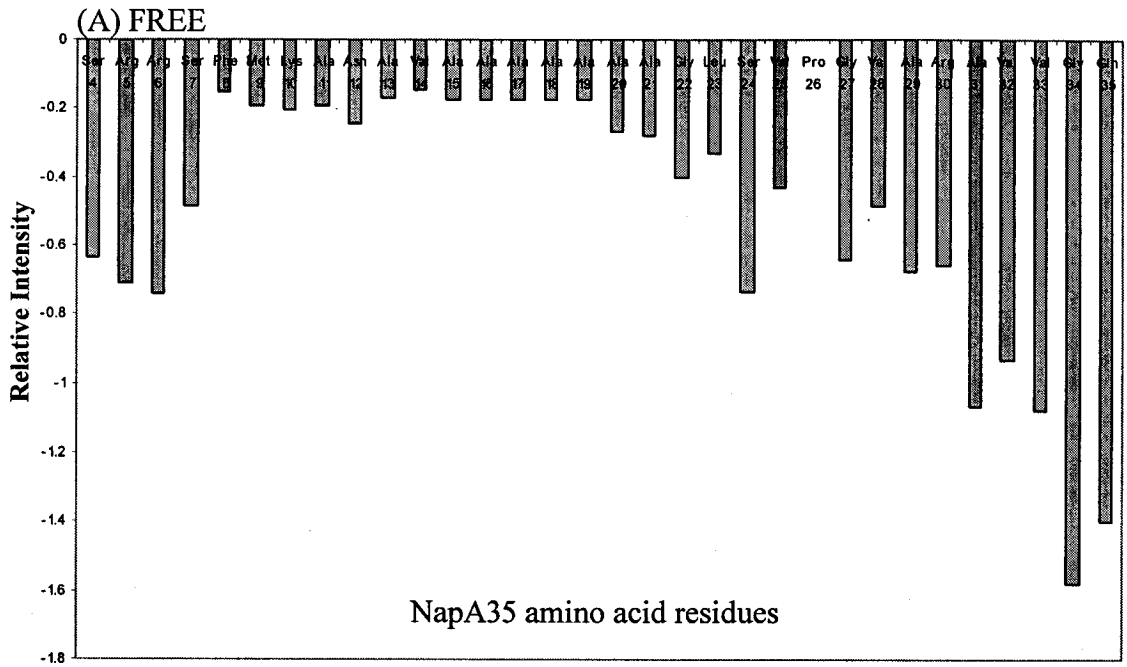
### 4.3.4 Characterizing the Dynamics of NapA35

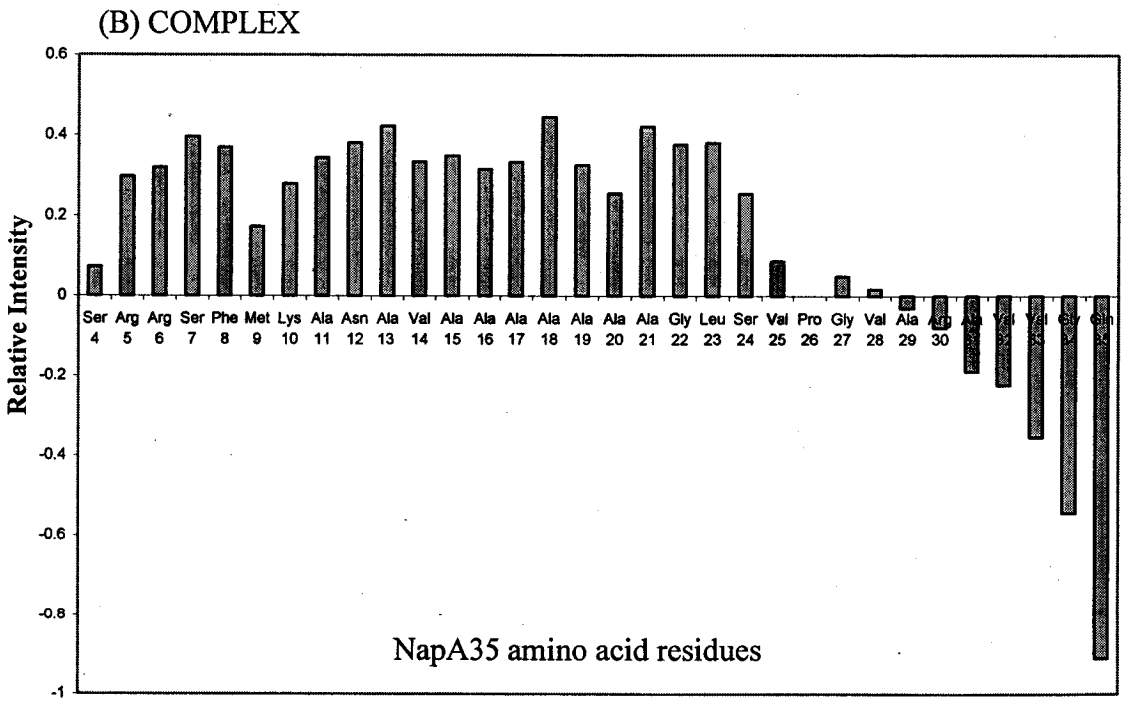
To characterize the molecular mobility of NapA35 both when free in solution and when in complex with NapD, a Heteronuclear NOE experiment was run (Peng *et al.*, 1994). This experiment measures the dynamics of the backbone NH. Here we basically run two HSQC's, a standard one ($^{15}$N-HSQC ) and another that has both positive and negative cross peaks. These positive and negative cross-peaks (hNOEs) are then integrated and divided by the values obtained from $^{15}$N-HSQC. Positive values indicate that the particular amino acid is not mobile while negative hNOEs are indicative of mobility (Figure 4.6).

When the NapA35 peptide was free in solution the results throughout the entire sequence were negative suggesting that the peptide is quite mobile when unbound (Figure 4.6a). However, when NapA35 is complexed with NapD (Figure 4.6b) a dramatic change in mobility is observed. Residues between Ser4 and Val28 have positive heteronuclear

NOEs, indicating that the peptide is immobile in that region. After that the peptide is mobile again. This experiment suggests that NapA35 is immobilized only upon binding to NapD, since it is mobile when free in solution.

Figure 4.6 Histogram showing the results from the Heteronuclear NOE experiment. (A) NapA35 free in solution. (B) NapA35 in complex with NapD.



### 4.3.5 $^{15}$N-HSQC of NapD free and in complex with NapA35

To analyze the interactions between NapD and NapA35 from the perspective of NapD, $^{15}$N-HSQC was also performed, where NapD was $^{15}$N labeled and the peptide was not (data not shown). Similarly to NapA35 there are changes in chemical shift, however they are not as profound as in the peptide. Overall, the structure of NapD when complexed does not undergo significant conformational changes (Ovidiu Minailiuc, unpublished results).

## 4.4 Conclusions from Binding studies by NMR

NMR is useful in that it allowed us to probe the interaction between NapD and NapA35 at the amino acid level. This chapter provides us with a lot of complementary results to chapters 2 and 3. The assignment of the backbone of NapA35 was accomplished for both free form and in complex with NapD. We noted from the $^{15}$N-HSQC of NapA35 complex versus free form, that many of the residues of the peptide were involved in binding, in particular amino acids before Pro26. Also from NMR we note that the poly alanine stretch undergoes significant changes in chemical shift upon complex formation. Furthermore the consensus sequence of the Tat leader peptide; KL**SRRSFMK** is involved in complex formation with NapD. As was seen both arginines of the twin arginine sequence did undergo chemical shift changes when bound to NapD. Residues of NapA that underwent the greatest change were Arg6, Phe8, Met9, Lys10, Val14, Ala16 to Ala21, Gly22 and Leu23. Since so many amino acids exhibit large changes in chemical shifts there are two possibilities; either the binding site between NapA35 and NapD is very extensive or NapA35 becomes structured upon binding.

Further structural analysis is needed to distinguish between the two explanations. This will be the focus of chapter five.

The analysis of the mobility of NapA35 in complex showed that the peptide is immobile when bound to NapD. However, further analysis of the Heteronuclear NOE data indicates that complexed NapA35 seems to have a higher mobility after Pro26. Perhaps Pro26 serves to break the structure. A similar experiment was run with free peptide as a control and it indicated that the peptide is mobile when free in solution implying that NapA35 becomes immobilized only during complex formation. This forms the basis for the next chapter where the goal is to determine whether in fact NapA35 becomes structured upon binding NapD.

# Chapter V Structural Characterization of NapA35 in complex

## 5.1 Foreword

The primary goal of this chapter is to characterize the tertiary structure of NapA35 in complex with NapD. To accomplish this we need to obtain the complete NMR signal assignments to analyze restraints for structure calculations. (Dubs *et al.*, 1979; Billeter *et al.*, 1982; Wagner and Wuthrich, 1982) Furthermore, interatomic distances and torsion angles are needed.

In the previous chapter the sequential backbone assignment of the amino acids of NapA35 was accomplished, whereas in this chapter the assignments of the amino acid side chains will be performed. The restraints obtained from NMR will be used to solve the structure of NapA35 in complex using the CYANA program. Also in this chapter, signals corresponding to interactions with NapD will be assessed to determine the residues of NapA35 at the binding interface.

## 5.2 Materials and Methods

To assign the side chains of NapA35, $^{15}N/^{13}C$-double labeled sample is needed. Isotopically labeled cultures were prepared in minimal media as previously mentioned except that 1g/L $^{15}NH_4SO_4$ and 2g/L [$^{13}C_6$]-glucose were used as nitrogen and carbon sources, respectively. $^{15}N/^{13}C$ experiments were recorded with the same Bruker-DRX 500 NMR spectrometer as in Chapter 4 except for $^{13}C$-HSQC and 3D $^{15}N/^{13}C$-NOESY-HSQC, which were recorded with a Varian 800MHz NMR spectrometer equipped with pulse field gradient probes (McGill University). As previously stated, NMR spectra were processed with XWINNMR software version 3.5 (Bruker Biospin) and analyzed with

XEASY program (Bartels *et al.,* 1995). Preparation and purification steps of NapA35 and NapD were the same as previously indicated.

**5.3 NMR Spectroscopy**

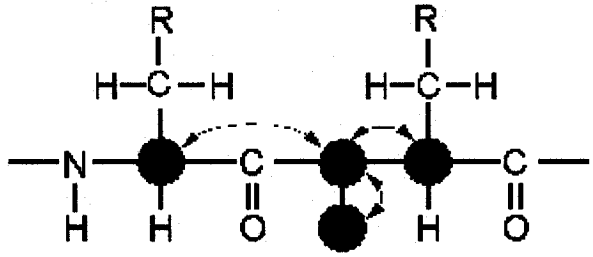
When a protein is $^{15}N/^{13}C$ labeled it is possible to run experiments that transfer magnetization over the peptide bond. The first experiment performed with $^{15}N/^{13}C$-labeled NapA35 and NapD complex was $^{13}C$-HSQC (data not shown). This 2D spectrum contains information about all carbon-attached hydrogens in NapA35.

A set of 3D experiments were performed where the base of the cube once again corresponds to $^{15}N$-HSQC, and the third dimension corresponds to carbon. The HNCA (Figure 5.1a) was carried out to verify sequential assignments and obtain carbon alpha chemical shifts (Grzesiek *et al.,* 1992; Constantine *et al.*, 1993).

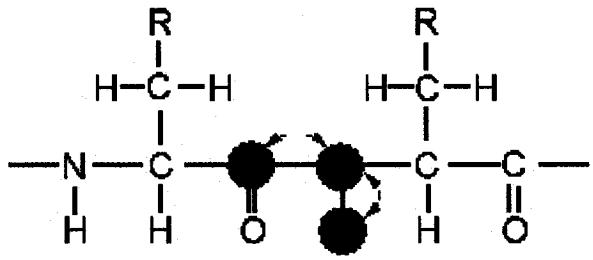
The HNCA experiment is the prototype of all 3D experiments. Magnetization begins at the amide proton, and is transferred to the nitrogen atom attached to it. Next magnetization is transferred to the C alpha nucleus. Finally, magnetization is transferred back to the amide proton. HNCA correlates an amide proton with both self and previous carbon alpha's. Figure 5.2 shows the strips of HNCA and the connectivities for the Cα's.

Figure 5.1 Schematic diagram of the amide coupling for 3D experiments. Correlations in the peptide bond for the (a) HNCA, (b) HNCO and (C) CBCA(CO)NH experiments. (a) HNCA correlates NH of one amino acid with the Cα of the self and previous amino acid. (b) HNCO correlates NH with the carbonyl group of the previous amino acid. (c) CBCA(CO)NH correlates the amide proton with the α and β carbon's of the previous amino acid.

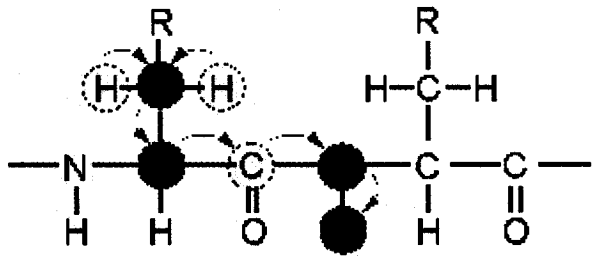
(a) **HNCA**



(b) **HNCO**

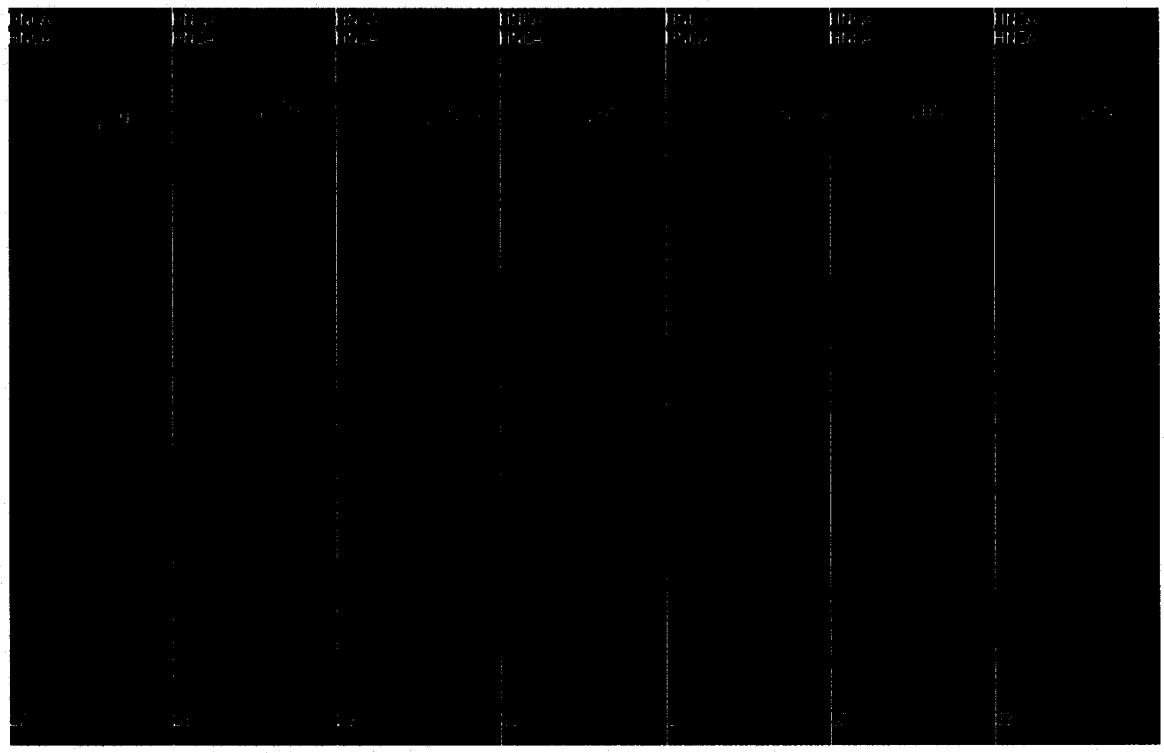


(c) **CBCA(CO)NH**



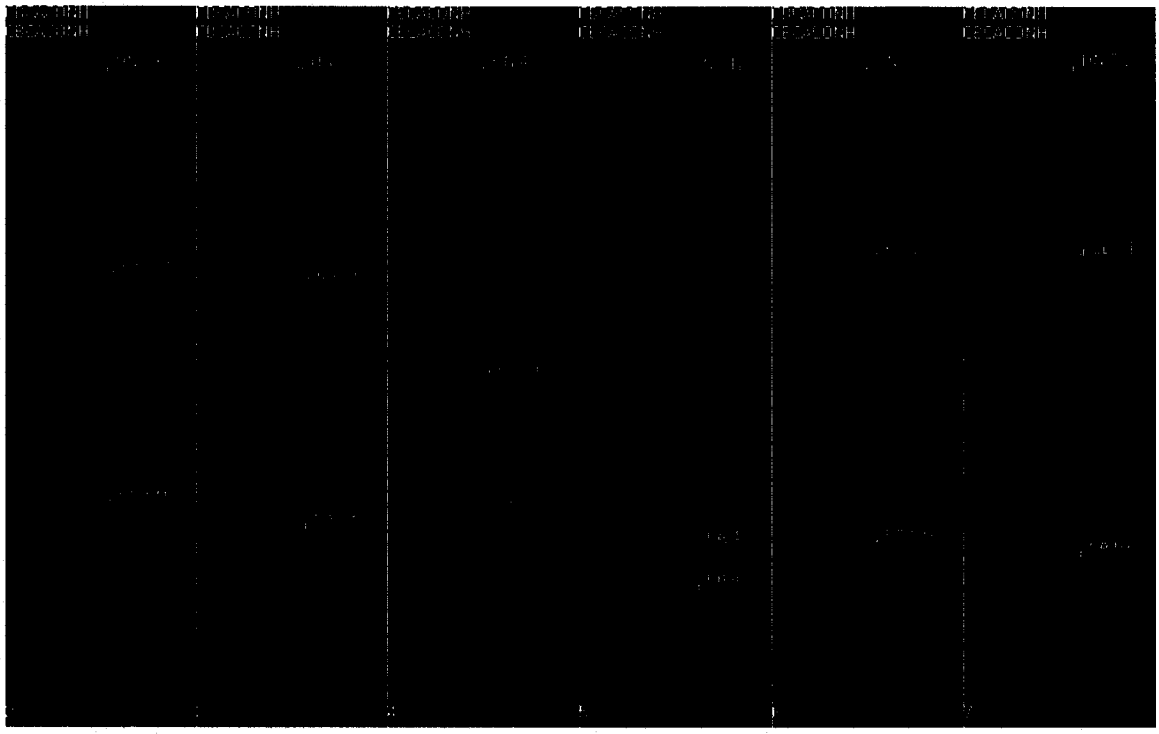
(http://www.cryst.bbk.ac.uk/PPS2/projects/schirra/html/3dnmr.htm)

Figure 5.2 HNCA strips for NapA35 in complex. Region from Gly27 to Val33 is indicated and connectivities are shown in purple. Each strip contains signals pertaining to self Cα and that to the previous amino acid.



The HNCO 3D experiment (Grzesiek *et al.,* 1992b) (Figure 5.1b) was also performed to identify the carbonyl group of each amino acid as well as resolve any overlap of signals that may have occurred in $^{15}$N-HSQC, since only one signal is observed per amino acid on each strip. CBCA(CO)NH 3D experiment (Figure 5.1c) was completed to obtain information on the carbon alpha and carbon beta of the previous amino acid. Figure 5.3 shows the strips for the CBCA(CO)NH experiment.

Figure 5.3 CBCA(CO)NH strips for NapA35 in complex. Region from Lys2 to Ser7 is indicated. The CA and CB chemical shifts we get from each strip correspond to the Cα and β of the previous amino acid. Although we normally cannot see the first residue (Met1) on $^{15}$N-HSQC, this experiment allows us to obtain the Cα and β co-ordinates of Met1 from Lys2.



The $^{1}$H/$^{15}$N projection of all these experiments looks like $^{15}$N-HSQC, which was assigned in the previous chapter. These experiments were done to assign the side chain protons of NapA35, and to obtain heteronuclear ($^{15}$N and $^{13}$C) chemical shifts in order to acquire constraints for secondary structure predication used in tertiary structure calculation. The $^{1}$H chemical shifts are tabulated in Table 5.1. Note that the positions of $^{1}$H for all the amino acids were determined except for the Phe8 ring and the gamma hydrogen of Val25 and Arg30. The complete list including $^{15}$N and $^{13}$C chemical shifts can be found in the Appendix (Table A).

Table 5.1 $^{1}H$ Chemical Shifts obtained from 3D NMR experiments of NapA35 in complex. (Note that "-" corresponds to protons that were unassigned).

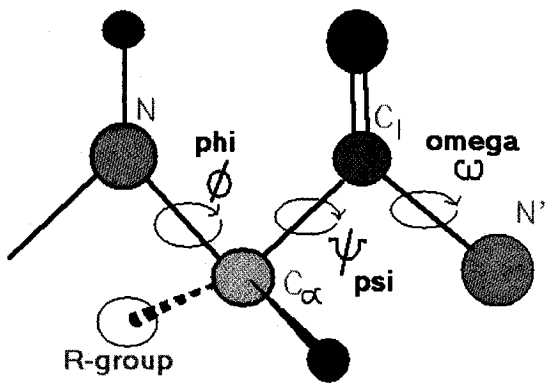
| Number in sequence | Amino Acid | Chemical Shift of Hydrogen (ppm) | | | | |
|---|---|---|---|---|---|---|
| | | alpha | beta | gamma | delta | epsilon |
| 1 | M | 4.143 | 1.963 | 3.122/2.875 | | |
| 2 | K | 4.183 | 1.639 | 1.404/1.287 | 1.742 | 2.944 |
| 3 | L | 4.18 | 1.661/1.557 | 1.64 | 0.829/0.896 | |
| 4 | S | 4.339 | 4.024/3.886 | | | |
| 5 | R | 4.245 | 2.002/1.949 | 1.795/1.731 | 3.273 | |
| 6 | R | 3.919 | 1.827/1.766 | 1.67/1.608 | 3.209 | |
| 7 | S | 4.053 | 3.733 | | | |
| 8 | F | 4.319 | 3.317/3.234 | | | |
| 9 | M | 4.085 | 2.395 | 3.025/2.701 | | |
| 10 | K | 3.789 | 1.963 | 1.803/1.338 | 1.948 | 2.872 |
| 11 | A | 3.913 | 1.419 | | | |
| 12 | N | 4.173 | 2.704/2.451 | | | |
| 13 | A | 3.079 | 1.3 | | | |
| 14 | V | 3.226 | 2.183 | 0.892/0.799 | | |
| 15 | A | 4.138 | 1.375 | | | |
| 16 | A | 4.068 | 1.486 | | | |
| 17 | A | 4.164 | 1.433 | | | |
| 18 | A | 3.77 | 1.334 | | | |
| 19 | A | 4.192 | 1.491 | | | |
| 20 | A | 4.135 | 1.538 | | | |
| 21 | A | 4.439 | 1.356 | | | |
| 22 | G | 3.967 | | | | |
| 23 | L | 4.613 | 1.521/1.403 | 1.482 | 0.759/0.658 | |
| 24 | S | 4.49 | 3.863/3.756 | | | |
| 25 | V | 4.278 | 2.025 | - | | |
| 26 | P | 4.407 | 2.311 | 2.091/1.996 | 3.851/3.64 | |
| 27 | G | 3.967 | | | | |
| 28 | V | 4.118 | 2.081 | 0.914 | | |
| 29 | A | 4.338 | 1.367 | | | |
| 30 | R | 4.285 | 1.86/1.72 | - | 3.209 | |
| 31 | A | 4.38 | 1.523 | | | |
| 32 | V | 4.142 | 2.081 | 0.938 | | |
| 33 | V | 4.159 | 2.106 | 0.963 | | |
| 34 | G | 3.967 | | | | |
| 35 | Q | 4.215 | 2.147/1.924 | 2.298 | | |

Chemical shifts were measured relative to internal 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) for $^{1}H$ and calculated for $^{13}C$ and $^{15}N$ with the assumptions $\gamma^{15}N/\gamma^{1}H = 0.101329118$ and $\gamma^{13}C/\gamma^{1}H = 0.251449530$ (Wishart *et al.*, 1995; Volpon *et al.*, 2003).

### 5.4 Secondary Structure Prediction

The secondary structure prediction of NapA35 in complex was based on a TALOS analysis of chemical shifts from HNCA, HNCACB, CACB(CO)NH, HNCO and 3D $^{15}N/^{13}C$-TOCSY-HSQC experiments. TALOS (Torsion Angle Likelihood Obtained from (chemical) Shift and Sequence Similarity) (Cornilescu *et al.*, 1999) is a program which compares experimental chemical sifts to those contained in a database of 78 proteins and yields quantitative secondary structure predictions based on phi and psi backbone dihedral angles (Figure 5.4).

Figure 5.4 Schematic showing the backbone dihedral angles of amino acids. Phi describes rotation between the Cα and N bond whereas psi describes rotation between the Cα and the carbonyl carbon.




(http://bmbiris.bmb.uga.edu/wampler/tutorial/prot2.html#torsions)

TALOS uses the sequence of NapA35 as well as the chemical shifts of Hα, Cα, Cβ, CO, and N (Table 5.1 and Appendix Table A) to predict the phi and psi backbone angles. As can be seen in table 5.2 TALOS predicted the range of phi and psi angles for Arg6 to Ala19, Ser24 and Val25. The dihedral backbone angles for other amino acids could not be unambiguously predicted by TALOS.

Table 5.2 Dihedral angles predicted from TALOS for NapA35 in complex. The Phi and Psi angles for these 16 amino acids will be used for structure calculations.

| Number in sequence | Amino Acid | Dihedral Angles | |
|---|---|---|---|
| | | PHI | PSI |
| 6 | R | -72.7 to -52.7 | -51.2 to -31.2 |
| 7 | S | -85.5 to -50.5 | -47.8 to -27.0 |
| 8 | F | -75.4 to -55.4 | -58.0 to -26.7 |
| 9 | M | -75.1 to -55.1 | -51.2 to -31.2 |
| 10 | K | -72.4 to -52.4 | -53.1 to -27.0 |
| 11 | A | -72.5 to -52.5 | -50.8 to -30.8 |
| 12 | N | -75.4 to -55.4 | -57.7 to -27.6 |
| 13 | A | -76.7 to -49.5 | -51.6 to -29.6 |
| 14 | V | -91.3 to -50.2 | -59.4 to -29.3 |
| 15 | A | -79.3 to -54.3 | -62.7 to -1.3 |
| 16 | A | -95.8 to -54.2 | -57.2 to 3.0 |
| 17 | A | -106.4 to -49.9 | -52.9 to 30.3 |
| 18 | A | -86.1 to -44.8 | -67.6 to 1.6 |
| 19 | A | -75.4 to -54.5 | -63.3 to -3.8 |
| 24 | S | -159.7 to -49.9 | 83.4 to 153.9 |
| 25 | V | -151.4 to -62.5 | 86.5 to 159.6 |

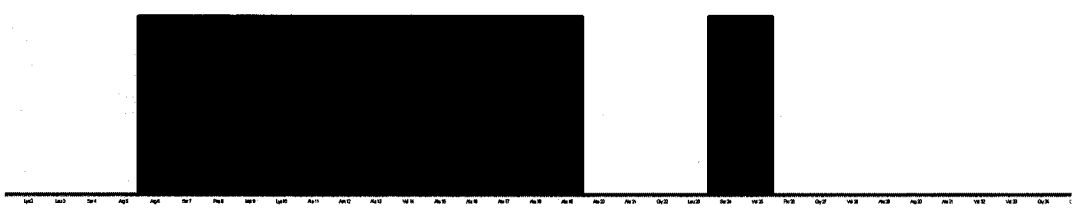
The elements of secondary structure have characteristic values of phi and psi angles. For example values of phi = -45 to - 130 and psi = -55 to -85 are representative for alpha-helical structures, whereas repeating values of phi = -110 to -140 and psi = +110 to +135 are representative of extended conformations; those found in beta sheet

structures for example. The phi and psi angles predicted for Arg6 to Ala19, corresponded to those that are found in alpha-helical structures. Whereas those predicted for Ser24 and Val25 corresponded to extended conformations.

From TALOS we were able to conclude that NapA35 is likely alpha-helical from Arg6 to Ala19 (Figure 5.5). However the two residues in extended conformations, Ser24 and Val25, cannot be used directly because they are not found within a stretch of beta-sheet residues. The TALOS results for the structured portion of NapA35 were verified manually, where the NH-Hα cross-peaks observed in 3D $^{15}$N-NOESY-HSQC experiment were consistent with alpha-helical secondary structures; in other words NOEs were found between i, i + 3 and i + 4 residues.

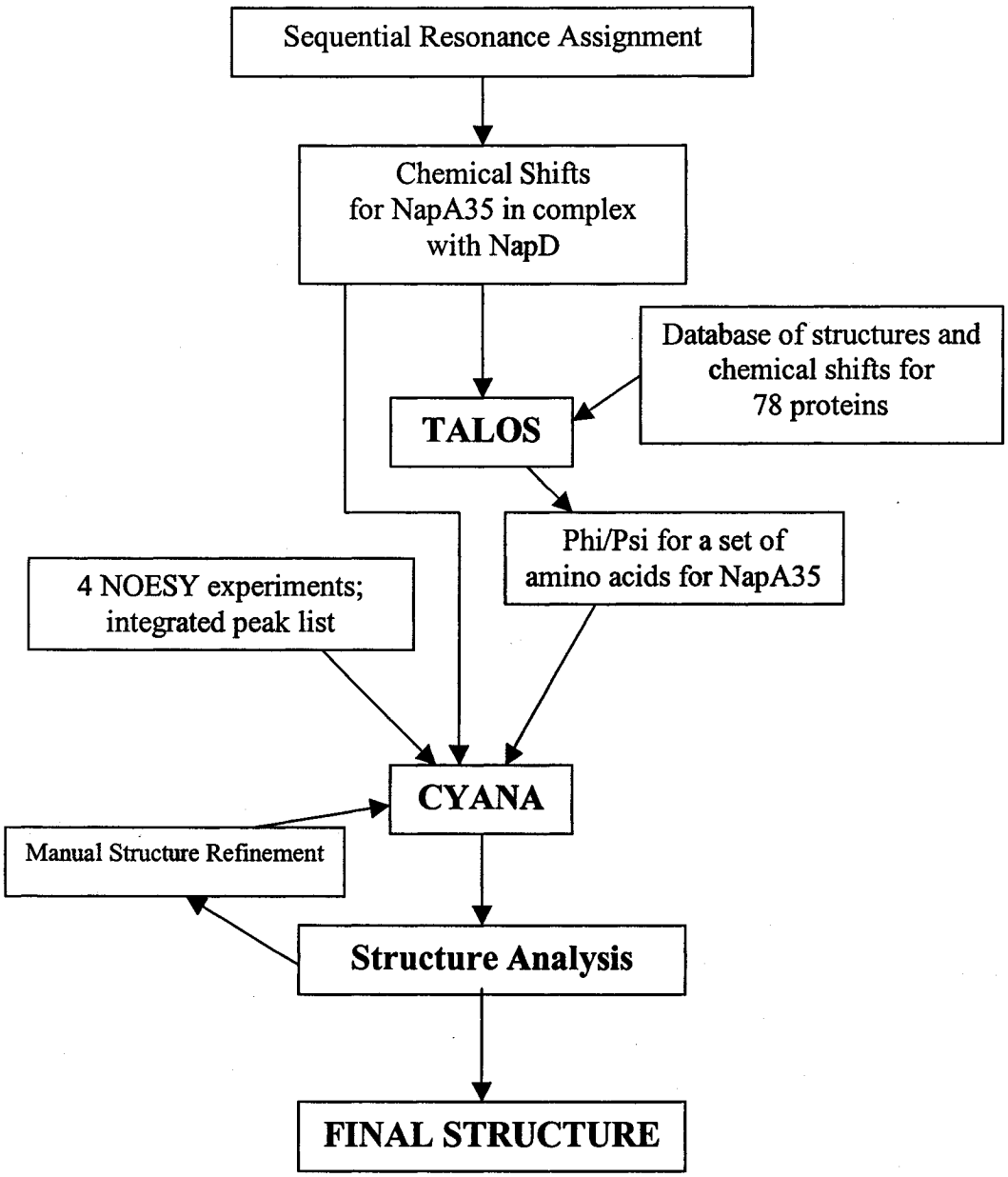
Figure 5.5 TALOS prediction for NapA35 in complex. As can be seen the N-terminal portion is alpha-helical until Ala19. Red corresponds to alpha-helical structure, and blue corresponds to extended structure.

***M*KLSR $^{6}$RSFMKANAVAAAA$^{19}$AAAGLSVPGVARAVVGQ**



The restraints obtained from TALOS in Table 5.2 will be used as input for structure calculations in the next section.

Figure 5.6 Summary of NMR structure determination.



## 5.5 3D Structure Calculations

For structure calculations by NMR both distance restraints and angle restraints are needed. A cross peak in a NOESY experiment means that the two hydrogens are close in space. Cross-peaks of protons for self or neighboring residues are identified as short range NOEs. Cross-peaks of protons that are between two and five amino acids apart in the protein sequence, known as medium range NOE's are mainly indicative of the protein backbone conformation and are important for secondary structure determination. Furthermore, cross-peaks of protons that are five or more residues apart also known as long-range NOE's contain the most important information for tertiary structure calculations. Since the structure of NapA35 in complex, as predicted by TALOS, is mostly α-helical, medium range NOE'S will be the most useful. Another variable for structure calculations is the intensity of a NOESY peak, which is proportional to $1/(\text{distance}^6)$. Chemical shifts on the other hand are used to generate angle restraints (phi and psi angles) as seen in the previous section.

A summary of structure determination by NMR is shown in Figure 5.6. CYANA (Combined assignment and dynamics algorithm for NMR applications) (Herrmann *et al.*, 2002) will be used to calculate the structure of NapA35 in complex. It is a program for automated structure calculation of proteins on the basis of conformational constraints set from NMR. CYANA uses the integrated NOEs obtained from NMR experiments, the complete list of chemical shifts, as well as the angle restraints generated from TALOS (Figure 5.6). From the data that is inputted and the amino acid sequence CYANA automatically assigns the NOESY cross-peaks using a fast torsion angle dynamics

iterative program (Güntert *et al.,* 1997). The program attempts to fold the starting structure into one with as small an energy potential as possible.

The angle constraints obtained from TALOS as well as cross-peaks from 3D $^{15}$N-NOESY-HSQC, 3D $^{15}$N/$^{13}$C-NOESY-HSQC, 2D Edited/Filtered NOESY and 2D $^{1}$H$^{1}$H NOESY were used for structure calculation using CYANA. The cross-peaks from these four experiments were selected manually and integrated in XEASY using the ellipticity footprint to determine their intensity. Also inputted into structure calculation was the complete list of chemical shifts (Appendix Table A) for $^{1}$H, $^{13}$C and $^{15}$N. The input for CYANA is summarized in Figure 5.6.

The cross-peaks of 3D $^{15}$N-NOESY-HSQC and 3D $^{15}$N/$^{13}$C-NOESY-HSQC experiments were particularly useful because they contained the majority of information regarding the backbone and side chains. Also inputted into CYANA was a set of manually assigned NOEs obtained from 2D $^{1}$H$^{1}$H NOESY. Furthermore, the 2D Edited/Filtered NOESY was used in order to discriminate between inter and intra-molecular signals. Since we were dealing with NMR spectra of NapA35 in complex with NapD, signals from NapA35 to NapD were present. These cross-peaks had to be removed in order to determine the intra-molecular interactions within NapA35 and calculate the structure. A series of edited/filtered NMR experiments will be discussed in section 5.6, but briefly, two NMR experiments were performed with different samples. The first was a 2D $^{15}$N edited/edited experiment with $^{15}$N NapA35/NapD unlabeled complex, which contained cross-peaks corresponding to both NapA35 and NapD. The second experiment was a 2D filtered/filtered experiment with $^{15}$N NapD/NapA35 unlabeled complex, which contained cross-peaks corresponding only to NapA35.

Initially, a total of 841 NOE distance constraints were derived from these four experiments. Thirty-two phi and psi backbone dihedral angles restraints were obtained from TALOS, from which 14 pairs of angles corresponded to the alpha-helical portion of the peptide. CYANA started with the calculation of 100 structures, where the 40 lowest energy structures were retrieved. The structural statistics for NapA35 in complex obtained from CYANA is shown in Table 5.3. Run 1 indicates the results for the first CYANA structure calculations, whereas Run 2 indicates the results obtained after manual structure analysis. Sybyl 6.4 software (Tripos, St. Louis, MO) was used to display the structure for analysis.

Of the 841 NOEs selected from run one, 566 were assigned whereas 275 were not. Of the 566 peaks assigned 353 were short range NOEs; corresponding to the self or neighboring residue, 62 were medium range NOEs corresponding to amino acids within four residues, and 24 were long-range NOEs corresponding to amino acids five or more residues apart. Although the cross-peaks obtained from the experiments are unique, the distances or restraints they represent may be redundant and must be sorted out.

Table 5.3 Structural Statistics for NapA35 in complex. *Note that the results for *RMSDs* and *Ramachandran plot* are those obtained from the final structure.

*Restraints for structure calculations*

| | | Run 1 | Run 2 |
|---|---|---|---|
| NOE Peaks: | | | |
| Used in Calculations | | 841 | 819 |
| in 3D $^{15}N/^{13}C$-NOESY-HSQC | | 479 | 462 |
| in 3D $^{15}N$-NOESY-HSQC | | 221 | 219 |
| in 2D Edited/Filtered NOESY | | 127 | 124 |
| in 2D $^{1}H^{1}H$ NOESY | | 14 | 14 |
| assigned | | 566 | 549 |
| unassigned | | 275 | 270 |
| Cross peaks: | | | |
| with diagonal assignment | | 127 | 127 |
| with off-diagonal assignment | | 439 | 422 |
| with unique assignment | | 388 | 366 |
| with short-range assignment | \|i-j\|<=1 | 353 | 355 |
| with medium-range assignment | 1<\|i-j\|<5 | 62 | 65 |
| with long-range assignment | \|i-j\|>=5 | 24 | 2 |

*Dihedral angles restraints from TALOS*

| | |
|---|---|
| Phi dihedral angles | 16 |
| Psi dihedral angles | 16 |

**RMSDs for residues Arg5 to Ala21 for 40 structures* (Å):

| | |
|---|---|
| Average backbone RMSD to mean | 0.45+/-0.16Å |
| Average heavy atom RMSD to mean | 1.09+/-0.23Å |

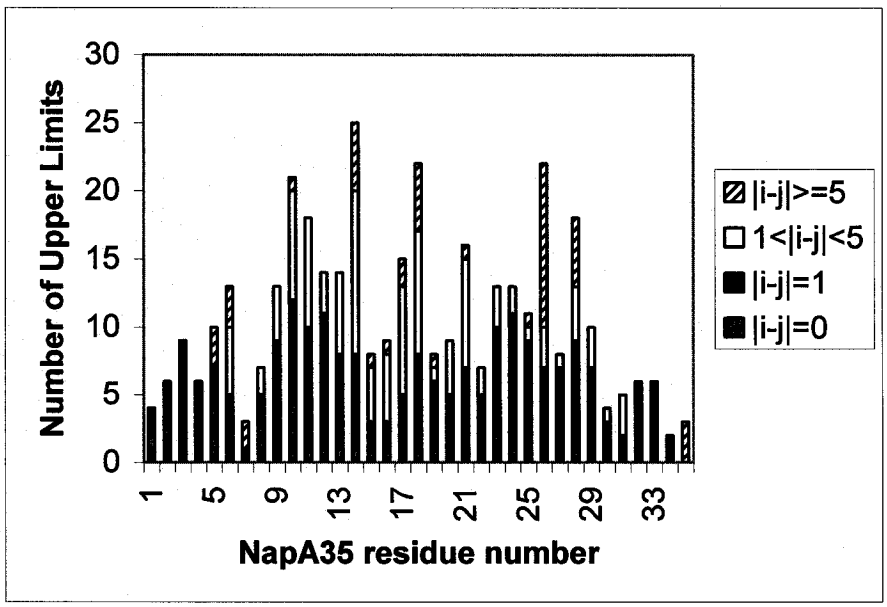
**Ramachandran plot statistics*

| | |
|---|---|
| Residues in most favoured regions | 64.9% |
| Residues in additional regions | 23.9% |
| Residues in generously allowed regions | 7.4% |
| Residues in disallowed regions | 3.8% |

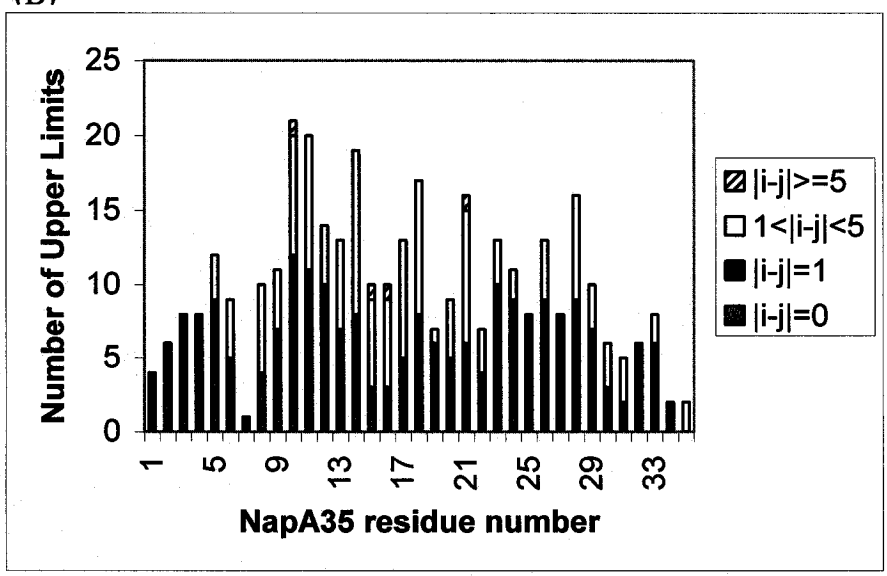
Figure 5.7 shows the distribution for non-redundant distances of short, medium and long-range NOEs as a function of the amino acid sequence. Therefore, if these distances were totaled the results would actually be smaller than in Table 5.3, because they correspond to actual distances and not cross peaks.

Figure 5.7 Histogram showing the distribution of non-redundant short ($|i-j| \leq 1$), medium ($1 < |i-j| < 5$) and long-range NOEs ($|i-j| \geq 5$) obtained from CYANA. (A) Results obtained from first round of calculations. (B) Results after manual refinement.

(A)




(B)

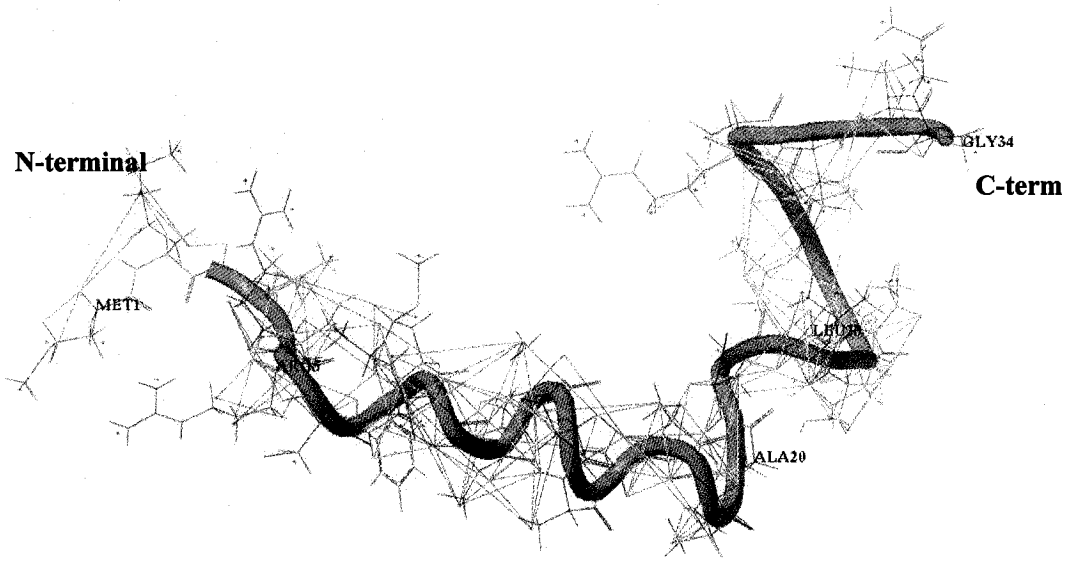


Looking at Figure 5.7A shows us the distribution of the distances of NOEs within the peptide obtained from the first run of CYANA. For tertiary structure determination medium and long-range NOEs are the most important. Once the initial structure was calculated by CYANA, the long-range NOEs corresponding to Arg5, Arg6, Ser7, Val14, Ala15-19, Ala21, Val25, Pro26, Val28 and Qln35 were examined. The first concern was the presence of long-range NOEs for the portion that seemed to be unstructured. As previously demonstrated from chemical shifts and Heteronuclear NOE experiments the peptide is mobile after Pro26. However as can be seen from the histogram Pro26, Val28 and Qln35 were calculated by CYANA to be structured due to the presence of long-range NOEs. Therefore, the NOEs were verified manually. Several changes were completed, which were for the most part corresponding to intermolecular signals to NapD that had not been deleted, initially. The structure was recalculated using CYANA, where the results are tabulated in Table 5.3, under Run 2. Although the statistics did not change significantly, the total NOEs used for calculations were 819. More importantly, is the fact that the number of long-range NOEs are much lower after manual refinement. Cross-peaks with long-range assignments from CYANA were reduced from 24 in Run 1 to 2 in Run 2 (Figure 5.7B). The two long-range NOEs obtained in Run 2 were verified. They correspond to interactions between Lys10 and Ala15, as well as between Ala16 and Ala21.

Figure 5.8 shows the network of constraints obtained from CYANA for run 2. These distance limits dictate the tertiary structure of NapA35 in complex. As can be seen, the majority of distances correspond to those that are short and medium range, since the structured portion is alpha-helical. This is also shown in Figure 5.7B. For the region from

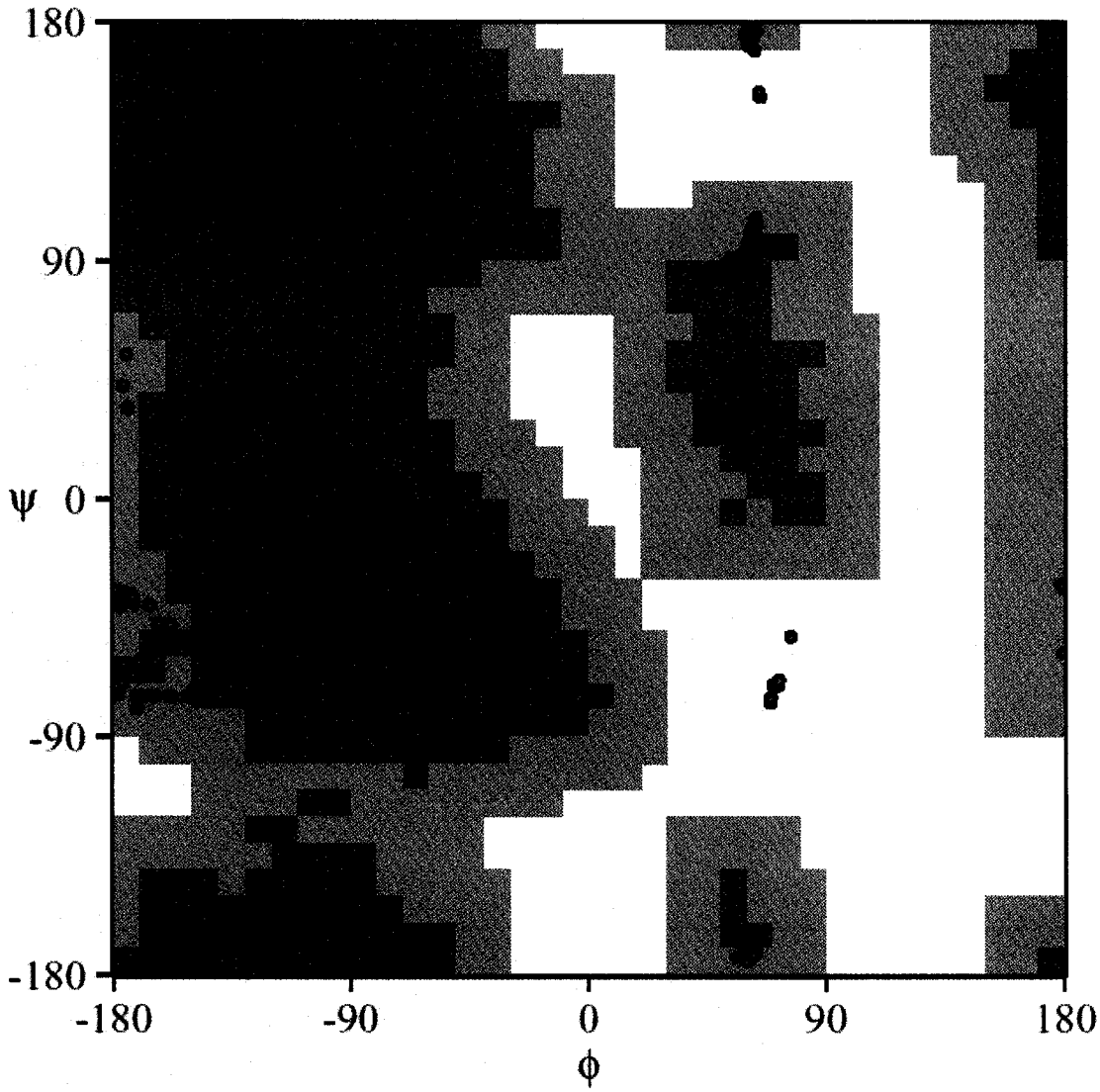
Arg5 to Ala20, 53% of distances correspond to short range NOEs, 45% correspond to medium range NOEs and 2% correspond to long range NOEs. However, looking at the C-terminal we can see that after Leu23 the majority of distances correspond to short range NOEs, since that region is unstructured. Figure 5.8B indicates that after Pro26, 68% are short range and 32% are medium range.

Figure 5.8 Schematic showing the disposition of upper distance limits obtained from CYANA (Run 2) for NapA35 in complex. 3D structure of NapA35 is shown in cyan, the amino acid residues are shown in blue and the network of distances are in orange.




The *Ramachandran* plot is a plot of phi vs psi angles. It is often used during the last stages of structure determination because it points out the residues that need more work (Figure 5.9). Each peak on the *Ramachandran* plot corresponds to a single residue, however in this plot we see the peaks of 40 conformers of NapA35.

Figure 5.9 *Ramachandram* Plot for CYANA output (Run 2) of NapA35 in complex.



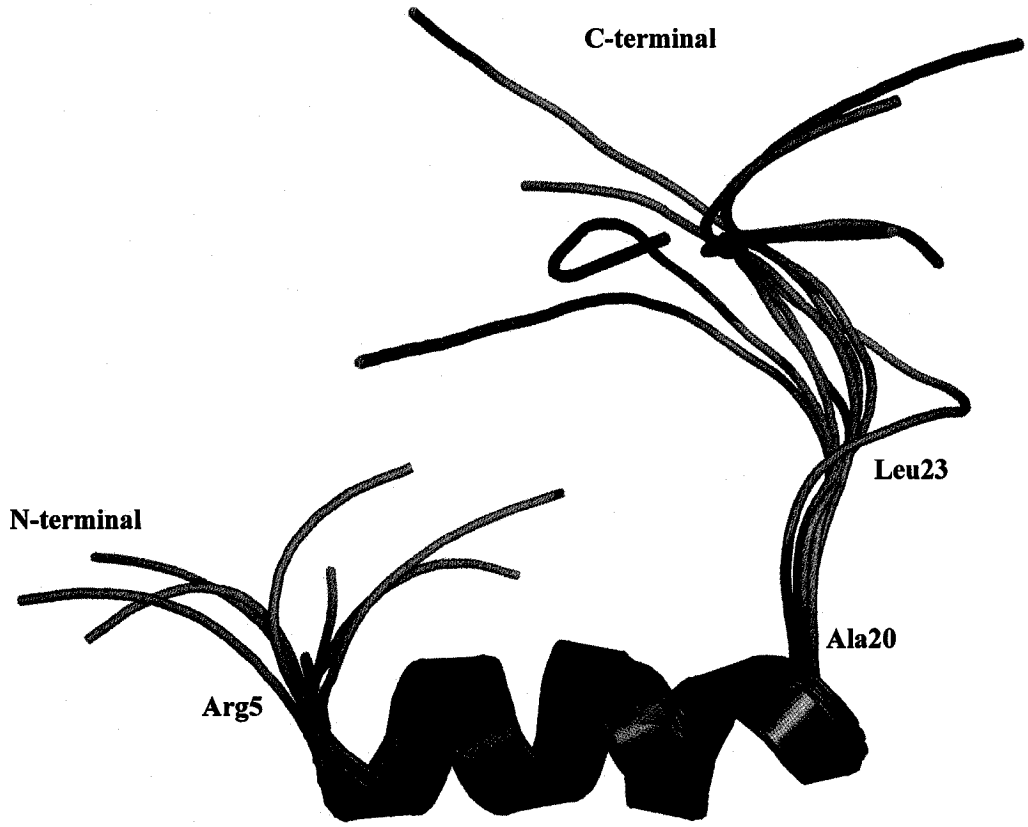

As can be seen on figure 5.10 the majority of signals are in the alpha-helical region (-45 to – 130, -55 to -85) and in the portion corresponding to extended conformations (-110 to -140, 110 to 135). This is as expected because NapA35 is alpha-helical at the N-terminal and mostly in the extended conformation at the C-terminal. The statistics for Ramachandran are as follows; 64.9% of residues are in most favored

regions, 23.9% are in additional regions, 7.4% are in generously allowed regions and 3.8% of residues are in disallowed regions.

The program Pymol (DeLano) was used to analyze the structure and generate the figures. The 3D structures of the eight lowest energy conformers of NapA35 in complex are shown in Figure 5.10. Upon binding to NapD, NapA35 forms an alpha-helix, which extends from the Arg5 to Ala20. It is quite remarkable that the N-terminal consensus sequence; KL**SRRSFMK** of the Tat signal peptide is actually structured starting from the twin arginines, which are the most conserved residues. Furthermore the hydrophobic poly-alanine region is also structured in the complex. The results predicted by TALOS as well as those obtained from chemical shift changes and Heteronuclear NOE experiments are consistent with the structure of NapA35. In Chapter 4, we showed that the residues from Arg5 to Leu23 undergo the largest changes in chemical shifts upon complex formation (Figure 4.5). This is quite interesting since the residues that are structured correspond to those in that region. In other words the alpha-helix explains the dispersion in chemical shifts obtained. Further analysis also indicated that the amino acids at the C-terminal since this fragment does not interact with NapD and remains unstructured. As can be seen in Figure 5.10, the structure of the peptide in complex is mobile at the C-terminal after Leu23. However, analyzing the dynamics of the peptide we can see that Heteronuclear NOE's are drastically changed, comparing to free peptide, suggesting that the peptide is immobilized from the N-terminal to Pro26 (Figure 4.6). This is in accordance with the structure since the lack of motion in the peptide when in complex is due to the structuring of NapA35 and not due to an extensive binding interface. Interestingly the results demonstrate that NapA35 only becomes structured when it is

bound to NapD. Examination of the final structure of NapA35, shows that the alpha-helical portion is very well defined, whereas the C-terminal portion is mobile and present in many conformations.

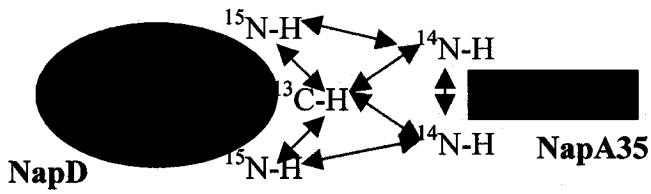
Figure 5.10 The eight lowest energy structures of NapA35 in complex with NapD. The alpha-helix at the N-terminal is red, whereas the mobile portion at the C-terminal is cyan.



## 5.6 Binding Interface of NapA35

The NMR structure of NapD was recently determined to consist of an α/β sandwich composed of two alpha-helices and four anti-parallel beta-strands (Ovidiu Minailiuc, unpublished results). The binding surface of NapD is found almost exclusively on the exposed surfaces of the beta-strands, implying that the alpha-helical N-terminus of NapA35 interacts with the beta-sheet of NapD.

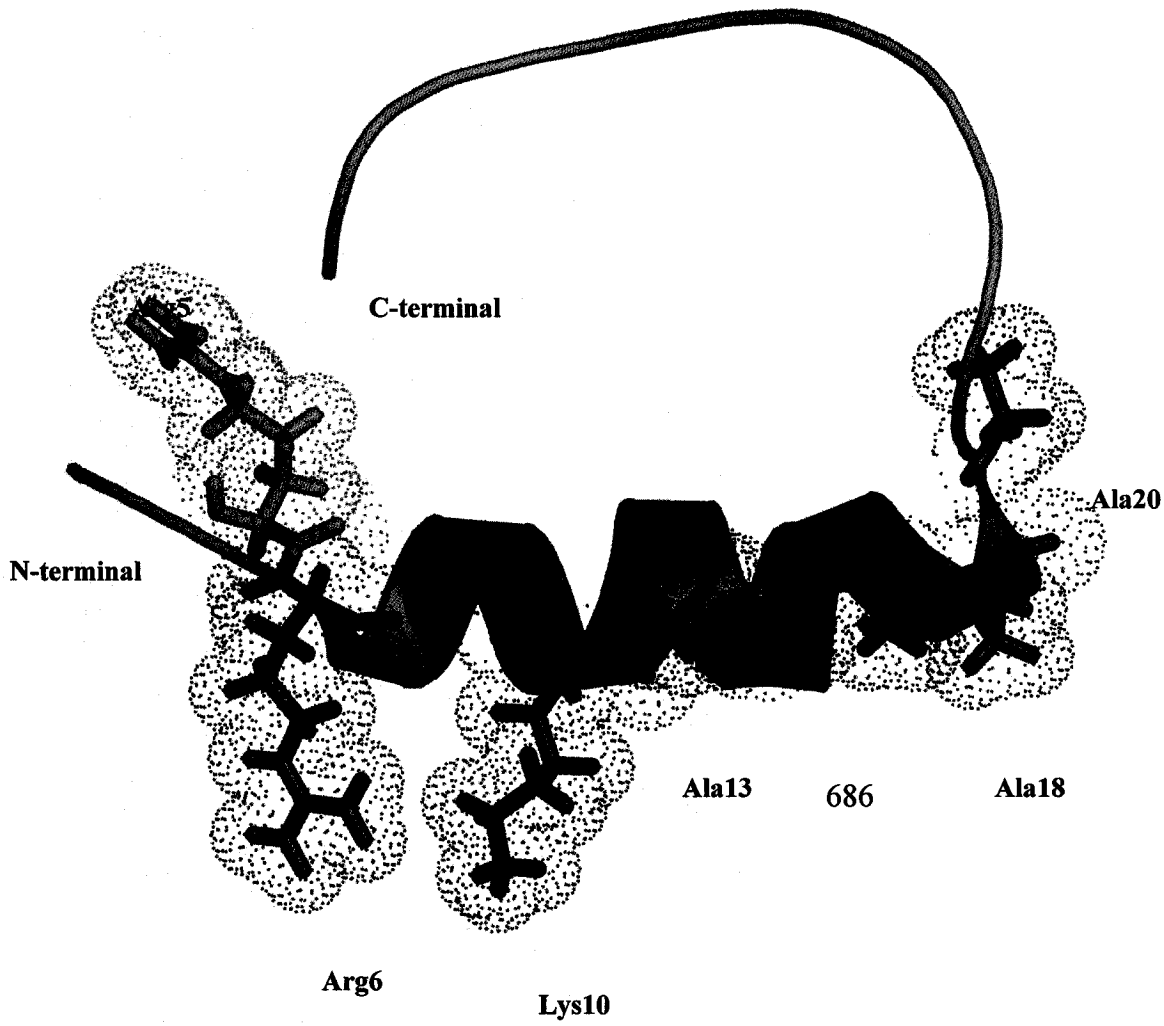
Figure 5.11 Schematic representation of the NapD/NapA complex for edited/filtered experiments. In this figure NapD is $^{15}N/^{13}C$ labeled whereas NapA35 is unlabeled. Correlations within NapD are observed by isotope-*edited* 2D experiments (red arrows). Whereas, correlations within NapA35 (blue arrow) are selected by isotope-*filtered* experiments. Intermolecular interactions between NapD and NapA35, are shown in green.




To be able to identify the amino acids of NapA35 at the binding interface with NapD, a series of *edited/filtered* NMR experiments was carried out. The term *edited* means the isotope is selected, whereas, *filtered* is used to denote rejection. The complex was prepared under different labeling conditions so as to be able to obtain cross-peaks that are pertinent to obtaining intra-molecular interactions within NapA35. Two samples were prepared, one in which NapA35 is labeled and the other where NapD is labeled. Figure 5.11 shows a representation of the isotope-filtered and edited $^{1}H$ NMR for NapD doubly labeled and NapA35 complex.

By filtering out the labeled isotopes (in this case found on NapD) the intramolecular signals within NapA35 (blue arrow) can be accessed. In other words, intermolecular signals between NapA35 and NapD (green arrows) as well as intramolecular signals within NapD (red arrows) are filtered out or eliminated. To see all the signals from NapA35, the opposite experiment was done. Here we prepared a complex with labeled NapA35 and unlabeled NapD and ran a series of edited/filtered experiments. The experiment of interest was the edited/edited 2D, which we observed correlations between intramolecular NapA35 signals and intermolecular signals between NapA35 and NapD. From there we were able to identify those amino acids that were likely at the binding interface (Figure 5.12).

Figure 5.12 Binding interface of NapA35. Residues of NapA35 that are directly interacting with NapD are displayed. Those in red are found on the alpha-helix, whereas those in cyan correspond to residues on the unstructured portion.




Residues that had signals representing NOEs to NapD included Arg5, Arg6, Lys10, Ala13, Ala15, Ala17, Ala18 and Ala20 (Figure 5.12). As expected the mobile portion of NapA35 seems to be relatively far from NapD, which is consistent with the small changes in chemical shifts observed upon NapD binding.

## 5.7 Conclusions from Structure Calculations

In this chapter the complete assignment of NapA35 in complex was accomplished. This allowed us to predict the secondary structure by TALOS. The predicted phi and psi angles from TALOS as well as the NOE restraints set from NMR were used for tertiary structure calculations using the CYANA program. The structure of NapA35 in complex with NapD is alpha-helical from Arg5 to Ala20. NapA35 forms this complex tertiary structure only when it is bound to NapD.

Overall determination of the structure of NapA35 helped answer a lot of questions and confirm many of the results from previous chapters. Two main hypotheses were postulated in Chapter 4; perhaps the binding site of NapA35 is extensive, in that many residues interact with NapD, or NapA35 forms a tertiary structure upon binding. This chapter allowed us to conclude that NapA35 does in fact form a helical structure upon binding to NapD. Therefore the changes in mobility and chemical shifts can be attributed to the structuring of NapA35.

To determine the binding interface of NapA35 with NapD a series of edited/filtered experiments were carried out. The results indicate that Arg5, Arg6, Lys10, Ala13, Ala15, Ala17, Ala18 and Ala20 are at the binding interface with NapD. Furthermore, the unstructured mobile C-terminal fragment does not participate in NapD binding.

In the end we have shown for the first time that the Tat consensus sequence; KLSRRSFMK as well as the hydrophobic poly-alanine core of NapA35 is alpha-helical and at the binding interface with the beta-sheet of NapD.

# Chapter VI Overview of NapD and NapA35 complex

## 6.1 Final Conclusions

The detailed conclusions deduced from the individual experiments were stated at the end of each chapter however this section contains conclusions that were mainly done by comparing two or more methods. Overall the results obtained from the structural investigation of the periplasmic nitrate reductase binding its maturation chaperone are consistent.

It was previously demonstrated that NapD is essential for periplasmic nitrate reductase activity in *E. coli* (Potter and Cole 1999). However the role of NapD in nitrate reduction was uncharacterized. To fill this gap, we have undertaken binding studies between NapA and its potential chaperone NapD. Although numerous structures are available for proteins exported by the Tat system, most of them exclude the signal peptide. The only study in which the Tat leader peptide was incorporated was for a protein that was embedded in the membrane (Kurisu *et al.*, 2003). Therefore no structural characterization is available for the Tat leader peptide except for the fact that it remains unfolded in the cytoplasm. As for the chaperones for these Tat containing proteins only one has been characterized, TorD. Clearly more structural data would help in understanding their function.

In this study binding was shown to occur between NapD and NapA. Different constructs of NapA proved to be important not only as a means of exploring the binding taking place but also for biophysical and NMR experiments. For the first time NapD was shown to interact with the leader peptide of NapA. This is similar to the interaction

reported between the signal peptide of TorA and its chaperone TorD (Hatzixanthis *et al.*, 2005). The 3D structure of TorD in Schewanella massila was shown to be composed of two connected alpha-helical domains (Tranier et al., 2003) (PDB accession code 1N1C). Recently Ovidiu Minailiuc, solved the NMR structure of NapD in *E. coli* (unpublished results). The structure was shown to consist of four beta-strands and two alpha-helices. Although the structure of NapD is different than that for TorD, the results suggest some commonality in the function of these chaperones. The interactions on the interface of TorD must involve the alpha-helical domains, whereas the binding interface of NapD is exclusively through the beta-sheets. This is likely due to evolutional convergence.

Results from co-purification showed that NapA27 was the shortest peptide strongly binding NapD. A slightly longer peptide, NapA35 was chosen for further characterization studies of the NapD/NapA. NMR chemical shift analysis indicates that the majority of the residues in this peptide do undergo changes when in complex. The largest changes in chemical shift from NMR were found in amino acids Arg5 to Leu23, which is consistent with the results obtained from co-purification. Never the less co-purification was a quick method in that it allowed us to probe the binding between NapD and NapA. Conversely, the use of NMR did allow for a more detailed picture of what is occurring.

The biophysical techniques applied to this system provided additional information that was coherent with the rest of the findings. CD indicated that NapD is composed of both alpha-helical and beta-sheet secondary structures, which was consistent with the recent results obtained from the tertiary structure of NapD obtained in our lab by NMR. Results from spectroscopic methods, including fluorescence and UV, showed that the

conformation of NapD is relatively unchanged during complex formation and that the aromatic residues of NapD are not at the binding interface with NapA. The results obtained from the CD spectra of NapD/NapA35 complex showed a slight change in the secondary structure. This is consistent with NMR, which showed that NapD has a slightly higher Beta-sheet content when in complex. Furthermore NapA35 was shown to be alpha-helical when in complex by NMR.

NMR not only showed that the leader peptide is alpha-helical in complex but also allowed us to resolve the residues that are the most important in binding or formation of the complex. Similar to the results obtained from co-purification the first 27 amino acids seem to be sufficient for binding. It is quite remarkable that out of the 828 amino acids of NapA, that the interaction is localized within the signal peptide and not the catalytic domains. These results indicate that the consensus sequence of the Tat translocon system responsible for exporting NapA out into the periplasm is actually also involved in binding the NapD chaperone. This confirms that the leader sequence has a dual role; in that it serves to target proteins to the membrane and bind to the maturation chaperone NapD.

Previous studies on the structure of Tat dependent proteins have suggested that the leader peptide is unstructured in the cytoplasm. The NMR structure of NapA35 in complex with its maturation chaperone NapD suggests that the leader peptide only becomes structured when in complex. Results from structural studies on the Tat leader peptide of the *E. coli* SufI protein have indicated that the signal peptide switches between helical and unstructured conformations depending on the environment (San Miguel et al., 2003), where it was determined to be unstructured when in an aqueous solution and alpha-helical when placed in a hydrophobic environment. This is similar to the

structuring of the N-terminal leader sequence reported in the mitochondrial protein import receptor Tom20 (Abe et al., 2000). In this system the targeting signal is unstructured in aqueous solution, however it forms an alpha-helix when bound to the receptor. Another example of the structuring of the leader peptide was demonstrated in the cytochrome $b_6f$ complex from the cyanobacterium *Mastigocladus laminosus* which showed that the Tat signal peptides is alpha-helical when embedded in the membrane (Kurisu et al., 2003). The results for the structuring of the leader peptide of periplasmic nitrate reductase are consistent with the notion that the signal peptides undergoes conformational changes depending on their environment. For the first time we have shown that the structuring of the leader peptide occurs upon complex formation with its maturation chaperone.

By obtaining the structure of NapA35 we were able to strengthen all the results obtained from the other methods. In fact understanding the mechanism is impossible without proper structural characterization of the system in question. By knowing the structure of NapA35 and NapD we are able to conclude that the alpha-helical structure of NapA interacts with the Beta-sheets of NapD. This is fundamental in understanding the binding taking place between the Tat signal peptide and its maturation chaperone.

## 6.2 Future Directions

The mechanism of Tat transport is quite striking since it is the only protein export system known to transport folded proteins (Berks *et al.*, 2005). How this is accomplished without the membrane becoming fully permeable is definitely a challenge for the future. The Tat pathway deserves more studying since it exports proteins involved in cell division, mobility, symbiosis and pathogenicity (Turner *et al.*, 2004).

Once NapA is properly folded and loaded with cofactors, perhaps NapB results in the dissociation of NapD, thereby exposing the twin-arginine signal peptide allowing for export. Future challenges include determining a mechanism by which NapD releases the signal peptide so that it can be targeted to the membrane to be exported. Both detailed structural and biochemical experiments are needed to determine the mechanism of release so that the signal peptide can interact with the Tat translocase to be exported. Our results on the structuring of the Tat leader sequence when in complex with the chaperone NapD, constitutes a first step in this direction.

# Appendix

Table A. Complete table of chemical shifts for $^{1}H$, $^{15}H$ and $^{13}C$ (in ppm) of all the amino acids for NapA35 in complex. This was used for structure calculations.

| Peak # | Chemical Shift | Atom Type | Residue # | Peak # | Chemical Shift | Atom Type | Residue # | Peak # | Chemical Shift | Atom Type | Residue # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | - | N | 1 | 62 | - | HD22 | 3 | 123 | - | NH2 | 6 |
| 2 | - | H | 1 | 63 | - | HD23 | 3 | 124 | - | HH21 | 6 |
| 3 | 54.994 | CA | 1 | 64 | - | QQD | 3 | 125 | - | HH22 | 6 |
| 4 | 4.143 | HA | 1 | 65 | 175.456 | C | 3 | 126 | - | QH2 | 6 |
| 5 | 32.215 | CB | 1 | 66 | 116.493 | N | 4 | 127 | 175.888 | C | 6 |
| 6 | - | HB2 | 1 | 67 | 8.232 | H | 4 | 128 | 114.960 | N | 7 |
| 7 | - | HB3 | 1 | 68 | 58.837 | CA | 4 | 129 | 8.008 | H | 7 |
| 8 | 1.963 | QB | 1 | 69 | 4.339 | HA | 4 | 130 | 63.856 | CA | 7 |
| 9 | 31.320 | CG | 1 | 70 | 63.229 | CB | 4 | 131 | 4.053 | HA | 7 |
| 10 | 3.122 | HG2 | 1 | 71 | 4.024 | HB2 | 4 | 132 | 64.770 | CB | 7 |
| 11 | 2.875 | HG3 | 1 | 72 | 3.886 | HB3 | 4 | 133 | - | HB2 | 7 |
| 12 | - | QG | 1 | 73 | - | QB | 4 | 134 | - | HB3 | 7 |
| 13 | - | QE | 1 | 74 | - | HG | 4 | 135 | 3.733 | QB | 7 |
| 14 | - | CE | 1 | 75 | 173.165 | C | 4 | 136 | - | HG | 7 |
| 15 | - | HE1 | 1 | 76 | 123.090 | N | 5 | 137 | 174.376 | C | 7 |
| 16 | - | HE2 | 1 | 77 | 8.503 | H | 5 | 138 | 123.742 | N | 8 |
| 17 | - | HE3 | 1 | 78 | 58.503 | CA | 5 | 139 | 8.019 | H | 8 |
| 18 | 174.808 | C | 1 | 79 | 4.245 | HA | 5 | 140 | 60.481 | CA | 8 |
| 19 | 120.421 | N | 2 | 80 | 29.464 | CB | 5 | 141 | 4.319 | HA | 8 |
| 20 | 8.333 | H | 2 | 81 | 2.002 | HB2 | 5 | 142 | 38.576 | CB | 8 |
| 21 | 57.054 | CA | 2 | 82 | 1.949 | HB3 | 5 | 143 | 3.317 | HB2 | 8 |
| 22 | 4.183 | HA | 2 | 83 | - | QB | 5 | 144 | 3.234 | HB3 | 8 |
| 23 | 28.832 | CB | 2 | 84 | 27.033 | CG | 5 | 145 | - | QB | 8 |
| 24 | - | HB2 | 2 | 85 | 1.795 | HG2 | 5 | 146 | - | QD | 8 |
| 25 | - | HB3 | 2 | 86 | 1.731 | HG3 | 5 | 147 | - | QE | 8 |
| 26 | 1.639 | QB | 2 | 87 | - | QG | 5 | 148 | - | QR | 8 |
| 27 | 24.345 | CG | 2 | 88 | 43.060 | CD | 5 | 149 | - | CG | 8 |
| 28 | 1.404 | HG2 | 2 | 89 | - | HD2 | 5 | 150 | - | CD1 | 8 |
| 29 | 1.287 | HG3 | 2 | 90 | - | HD3 | 5 | 151 | - | HD1 | 8 |
| 30 | - | QG | 2 | 91 | 3.273 | QD | 5 | 152 | - | CE1 | 8 |
| 31 | 32.265 | CD | 2 | 92 | - | NE | 5 | 153 | - | HE1 | 8 |
| 32 | - | HD2 | 2 | 93 | - | HE | 5 | 154 | - | CZ | 8 |
| 33 | - | HD3 | 2 | 94 | - | CZ | 5 | 155 | - | HZ | 8 |
| 34 | 1.742 | QD | 2 | 95 | - | NH1 | 5 | 156 | - | CE2 | 8 |
| 35 | 41.682 | CE | 2 | 96 | - | HH1 | 5 | 157 | - | HE2 | 8 |
| 36 | - | HE2 | 2 | 97 | - | NH2 | 5 | 158 | - | CD2 | 8 |
| 37 | - | HE3 | 2 | 98 | - | HH21 | 5 | 159 | - | HD2 | 8 |
| 38 | 2.944 | QE | 2 | 99 | - | HH22 | 5 | 160 | 174.808 | C | 8 |
| 39 | - | NZ | 2 | 100 | - | QH2 | 5 | 161 | 119.132 | N | 9 |
| 40 | - | HZ1 | 2 | 101 | 175.932 | C | 5 | 162 | 8.339 | H | 9 |
| 41 | - | HZ2 | 2 | 102 | 118.542 | N | 6 | 163 | 59.624 | CA | 9 |
| 42 | - | QZ | 2 | 103 | 8.111 | H | 6 | 164 | 4.085 | HA | 9 |
| 43 | 174.376 | C | 2 | 104 | 59.934 | CA | 6 | 165 | 33.421 | CB | 9 |
| 44 | 121.585 | N | 3 | 105 | 3.919 | HA | 6 | 166 | - | HB2 | 9 |
| 45 | 8.171 | H | 3 | 106 | 29.507 | CB | 6 | 167 | - | HB3 | 9 |
| 46 | 55.665 | CA | 3 | 107 | 1.827 | HB2 | 6 | 168 | 2.395 | QB | 9 |
| 47 | 4.180 | HA | 3 | 108 | 1.766 | HB3 | 6 | 169 | 32.450 | CG | 9 |
| 48 | 41.629 | CB | 3 | 109 | - | QB | 6 | 170 | 3.025 | HG2 | 9 |
| 49 | 1.661 | HB2 | 3 | 110 | 26.744 | CG | 6 | 171 | 2.701 | HG3 | 9 |
| 50 | 1.557 | HB3 | 3 | 111 | 1.670 | HG2 | 6 | 172 | - | QG | 9 |
| 51 | - | QB | 3 | 112 | 1.608 | HG3 | 6 | 173 | - | QE | 9 |
| 52 | 26.325 | CG | 3 | 113 | - | QG | 6 | 174 | - | CE | 9 |
| 53 | 1.640 | HG | 3 | 114 | 43.192 | CD | 6 | 175 | - | HE1 | 9 |
| 54 | 0.829 | QD1 | 3 | 115 | - | HD2 | 6 | 176 | - | HE2 | 9 |
| 55 | 0.896 | QD2 | 3 | 116 | - | HD3 | 6 | 177 | - | HE3 | 9 |
| 56 | 23.287 | CD1 | 3 | 117 | 3.209 | QD | 6 | 178 | - | C | 9 |
| 57 | - | HD11 | 3 | 118 | - | NE | 6 | 179 | 117.954 | N | 10 |
| 58 | - | HD12 | 3 | 119 | - | HE | 6 | 180 | 8.273 | H | 10 |
| 59 | - | HD13 | 3 | 120 | - | CZ | 6 | 181 | 60.487 | CA | 10 |
| 60 | 24.605 | CD2 | 3 | 121 | - | NH1 | 6 | 182 | 3.789 | HA | 10 |
| 61 | - | HD21 | 3 | 122 | - | HH1 | 6 | 183 | 30.108 | CB | 10 |

| Peak # | Chemical Shift | Atom Type | Residue # |
|---|---|---|---|
| 184 | - | HB2 | 10 |
| 185 | - | HB3 | 10 |
| 186 | 1.963 | QB | 10 |
| 187 | 26.406 | CG | 10 |
| 188 | 1.803 | HG2 | 10 |
| 189 | 1.338 | HG3 | 10 |
| 190 | - | QG | 10 |
| 191 | 31.741 | CD | 10 |
| 192 | - | HD2 | 10 |
| 193 | - | HD3 | 10 |
| 194 | 1.948 | QD | 10 |
| 195 | 41.763 | CE | 10 |
| 196 | - | HE2 | 10 |
| 197 | - | HE3 | 10 |
| 198 | 2.872 | QE | 10 |
| 199 | - | NZ | 10 |
| 200 | - | HZ1 | 10 |
| 201 | - | HZ2 | 10 |
| 202 | - | QZ | 10 |
| 203 | 175.305 | C | 10 |
| 204 | 119.898 | N | 11 |
| 205 | 8.043 | H | 11 |
| 206 | 54.886 | CA | 11 |
| 207 | 3.913 | HA | 11 |
| 208 | 1.419 | QB | 11 |
| 209 | 17.590 | CB | 11 |
| 210 | - | HB1 | 11 |
| 211 | - | HB2 | 11 |
| 212 | - | HB3 | 11 |
| 213 | 178.309 | C | 11 |
| 214 | 118.307 | N | 12 |
| 215 | 8.137 | H | 12 |
| 216 | 55.256 | CA | 12 |
| 217 | 4.173 | HA | 12 |
| 218 | 37.576 | CB | 12 |
| 219 | 2.704 | HB2 | 12 |
| 220 | 2.451 | HB3 | 12 |
| 221 | - | QB | 12 |
| 222 | - | CG | 12 |
| 223 | - | ND2 | 12 |
| 224 | - | HD21 | 12 |
| 225 | - | HD22 | 12 |
| 226 | - | QD2 | 12 |
| 227 | 174.678 | C | 12 |
| 228 | 123.947 | N | 13 |
| 229 | 8.055 | H | 13 |
| 230 | 54.237 | CA | 13 |
| 231 | 3.079 | HA | 13 |
| 232 | 1.300 | QB | 13 |
| 233 | 17.594 | CB | 13 |
| 234 | - | HB1 | 13 |
| 235 | - | HB2 | 13 |
| 236 | - | HB3 | 13 |
| 237 | 179.000 | C | 13 |
| 238 | 121.401 | N | 14 |
| 239 | 8.290 | H | 14 |
| 240 | 67.143 | CA | 14 |
| 241 | 3.226 | HA | 14 |
| 242 | 31.553 | CB | 14 |
| 243 | 2.183 | HB | 14 |
| 244 | 0.892 | QG1 | 14 |
| 245 | 0.799 | QG2 | 14 |
| 246 | 23.380 | CG1 | 14 |
| 247 | - | HG11 | 14 |
| 248 | - | HG12 | 14 |
| 249 | - | HG13 | 14 |
| 250 | 20.907 | CG2 | 14 |
| 251 | - | HG21 | 14 |
| 252 | - | HG22 | 14 |
| 253 | - | HG23 | 14 |
| 254 | - | QQG | 14 |
| 255 | 174.030 | C | 14 |
| 256 | 121.177 | N | 15 |
| 257 | 7.920 | H | 15 |
| 258 | 54.371 | CA | 15 |
| 259 | 4.138 | HA | 15 |
| 260 | 1.375 | QB | 15 |
| 261 | 18.255 | CB | 15 |
| 262 | - | HB1 | 15 |
| 263 | - | HB2 | 15 |
| 264 | - | HB3 | 15 |
| 265 | 177.660 | C | 15 |
| 266 | 119.384 | N | 16 |
| 267 | 8.026 | H | 16 |
| 268 | 54.477 | CA | 16 |
| 269 | 4.068 | HA | 16 |
| 270 | 1.486 | QB | 16 |
| 271 | 18.509 | CB | 16 |
| 272 | - | HB1 | 16 |
| 273 | - | HB2 | 16 |
| 274 | - | HB3 | 16 |
| 275 | 177.790 | C | 16 |
| 276 | 120.253 | N | 17 |
| 277 | 7.452 | H | 17 |
| 278 | 53.468 | CA | 17 |
| 279 | 4.164 | HA | 17 |
| 280 | 1.433 | QB | 17 |
| 281 | 18.191 | CB | 17 |
| 282 | - | HB1 | 17 |
| 283 | - | HB2 | 17 |
| 284 | - | HB3 | 17 |
| 285 | 175.154 | C | 17 |
| 286 | 120.680 | N | 18 |
| 287 | 8.799 | H | 18 |
| 288 | 55.583 | CA | 18 |
| 289 | 3.770 | HA | 18 |
| 290 | 1.334 | QB | 18 |
| 291 | 16.975 | CB | 18 |
| 292 | - | HB1 | 18 |
| 293 | - | HB2 | 18 |
| 294 | - | HB3 | 18 |
| 295 | 176.623 | C | 18 |
| 296 | 118.518 | N | 19 |
| 297 | 7.794 | H | 19 |
| 298 | 54.403 | CA | 19 |
| 299 | 4.192 | HA | 19 |
| 300 | 1.491 | QB | 19 |
| 301 | 17.547 | CB | 19 |
| 302 | - | HB1 | 19 |
| 303 | - | HB2 | 19 |
| 304 | - | HB3 | 19 |
| 305 | 178.352 | C | 19 |
| 306 | 121.126 | N | 20 |
| 307 | 7.690 | H | 20 |
| 308 | 54.745 | CA | 20 |
| 309 | 4.135 | HA | 20 |
| 310 | 1.538 | QB | 20 |
| 311 | 17.919 | CB | 20 |
| 312 | - | HB1 | 20 |
| 313 | - | HB2 | 20 |
| 314 | - | HB3 | 20 |
| 315 | 176.277 | C | 20 |
| 316 | 117.805 | N | 21 |
| 317 | 7.970 | H | 21 |
| 318 | 51.434 | CA | 21 |
| 319 | 4.439 | HA | 21 |
| 320 | 1.356 | QB | 21 |
| 321 | 19.817 | CB | 21 |
| 322 | - | HB1 | 21 |
| 323 | - | HB2 | 21 |
| 324 | - | HB3 | 21 |
| 325 | 174.548 | C | 21 |
| 326 | 107.514 | N | 22 |
| 327 | 7.797 | H | 22 |
| 328 | 45.416 | CA | 22 |
| 329 | 3.967 | HA2 | 22 |
| 330 | - | HA3 | 22 |
| 331 | - | QA | 22 |
| 332 | 172.085 | C | 22 |
| 333 | 118.702 | N | 23 |
| 334 | 8.054 | H | 23 |
| 335 | 52.958 | CA | 23 |
| 336 | 4.613 | HA | 23 |
| 337 | 43.530 | CB | 23 |
| 338 | 1.521 | HB2 | 23 |
| 339 | 1.403 | HB3 | 23 |
| 340 | - | QB | 23 |
| 341 | 26.084 | CG | 23 |
| 342 | 1.482 | HG | 23 |
| 343 | 0.759 | QD1 | 23 |
| 344 | 0.658 | QD2 | 23 |
| 345 | 22.149 | CD1 | 23 |
| 346 | - | HD11 | 23 |
| 347 | - | HD12 | 23 |
| 348 | - | HD13 | 23 |
| 349 | 24.941 | CD2 | 23 |
| 350 | - | HD21 | 23 |
| 351 | - | HD22 | 23 |
| 352 | - | HD23 | 23 |
| 353 | - | QQD | 23 |
| 354 | 173.295 | C | 23 |
| 355 | 115.206 | N | 24 |
| 356 | 8.206 | H | 24 |
| 357 | 57.983 | CA | 24 |
| 358 | 4.490 | HA | 24 |
| 359 | 63.551 | CB | 24 |
| 360 | 3.863 | HB2 | 24 |
| 361 | 3.756 | HB3 | 24 |
| 362 | - | QB | 24 |
| 363 | - | HG | 24 |
| 364 | 171.177 | C | 24 |
| 365 | 122.925 | N | 25 |
| 366 | 8.239 | H | 25 |
| 367 | 59.521 | CA | 25 |
| 368 | 4.278 | HA | 25 |
| 369 | 32.219 | CB | 25 |
| 370 | 2.025 | HB | 25 |
| 371 | - | QG1 | 25 |
| 372 | - | QG2 | 25 |
| 373 | - | CG1 | 25 |
| 374 | - | HG11 | 25 |
| 375 | - | HG12 | 25 |
| 376 | - | HG13 | 25 |
| 377 | - | CG2 | 25 |
| 378 | - | HG21 | 25 |
| 379 | - | HG22 | 25 |
| 380 | - | HG23 | 25 |
| 381 | - | QQG | 25 |
| 382 | - | C | 25 |
| 383 | - | N | 26 |
| 384 | 50.454 | CD | 26 |
| 385 | 63.294 | CA | 26 |
| 386 | 4.407 | HA | 26 |
| 387 | 31.767 | CB | 26 |
| 388 | - | HB2 | 26 |
| 389 | - | HB3 | 26 |
| 390 | 2.311 | QB | 26 |

| Peak # | Chemical Shift | Atom Type | Residue # |
|---|---|---|---|
| 391 | 27.066 | CG | 26 |
| 392 | 2.091 | HG2 | 26 |
| 393 | 1.996 | HG3 | 26 |
| 394 | - | QG | 26 |
| 395 | 3.851 | HD2 | 26 |
| 396 | 3.640 | HD3 | 26 |
| 397 | - | QD | 26 |
| 398 | 174.721 | C | 26 |
| 399 | 108.911 | N | 27 |
| 400 | 8.433 | H | 27 |
| 401 | 45.074 | CA | 27 |
| 402 | 3.967 | HA2 | 27 |
| 403 | - | HA3 | 27 |
| 404 | - | QA | 27 |
| 405 | 171.609 | C | 27 |
| 406 | 118.956 | N | 28 |
| 407 | 7.820 | H | 28 |
| 408 | 62.064 | CA | 28 |
| 409 | 4.118 | HA | 28 |
| 410 | 32.680 | CB | 28 |
| 411 | 2.081 | HB | 28 |
| 412 | - | QG1 | 28 |
| 413 | - | QG2 | 28 |
| 414 | 21.319 | CG1 | 28 |
| 415 | - | HG11 | 28 |
| 416 | - | HG12 | 28 |
| 417 | - | HG13 | 28 |
| 418 | 20.249 | CG2 | 28 |
| 419 | - | HG21 | 28 |
| 420 | - | HG22 | 28 |
| 421 | - | HG23 | 28 |
| 422 | 0.914 | QQG | 28 |
| 423 | 173.209 | C | 28 |
| 424 | 126.648 | N | 29 |
| 425 | 8.272 | H | 29 |
| 426 | 52.303 | CA | 29 |
| 427 | 4.338 | HA | 29 |
| 428 | 1.367 | QB | 29 |
| 429 | 18.876 | CB | 29 |
| 430 | - | HB1 | 29 |
| 431 | - | HB2 | 29 |
| 432 | - | HB3 | 29 |
| 433 | 174.851 | C | 29 |
| 434 | 119.851 | N | 30 |
| 435 | 8.153 | H | 30 |
| 436 | 55.727 | CA | 30 |
| 437 | 4.285 | HA | 30 |
| 438 | 31.332 | CB | 30 |
| 439 | 1.860 | HB2 | 30 |
| 440 | 1.720 | HB3 | 30 |
| 441 | - | QB | 30 |
| 442 | - | CG | 30 |
| 443 | - | HG2 | 30 |
| 444 | - | HG3 | 30 |
| 445 | - | QG | 30 |
| 446 | 42.939 | CD | 30 |
| 447 | - | HD2 | 30 |
| 448 | - | HD3 | 30 |
| 449 | 3.209 | QD | 30 |
| 450 | - | NE | 30 |
| 451 | - | HE | 30 |
| 452 | - | CZ | 30 |
| 453 | - | NH1 | 30 |
| 454 | - | HH1 | 30 |
| 455 | - | NH2 | 30 |
| 456 | - | HH21 | 30 |
| 457 | - | HH22 | 30 |
| 458 | - | QH2 | 30 |
| 459 | 173.338 | C | 30 |
| 460 | 124.628 | N | 31 |
| 461 | 8.203 | H | 31 |
| 462 | 52.024 | CA | 31 |
| 463 | 4.380 | HA | 31 |
| 464 | 1.523 | QB | 31 |
| 465 | 19.403 | CB | 31 |
| 466 | - | HB1 | 31 |
| 467 | - | HB2 | 31 |
| 468 | - | HB3 | 31 |
| 469 | 174.808 | C | 31 |
| 470 | 118.945 | N | 32 |
| 471 | 8.038 | H | 32 |
| 472 | 62.187 | CA | 32 |
| 473 | 4.142 | HA | 32 |
| 474 | 32.533 | CB | 32 |
| 475 | 2.081 | HB | 32 |
| 476 | - | QG1 | 32 |
| 477 | - | QG2 | 32 |
| 478 | 21.471 | CG1 | 32 |
| 479 | - | HG11 | 32 |
| 480 | - | HG12 | 32 |
| 481 | - | HG13 | 32 |
| 482 | 19.791 | CG2 | 32 |
| 483 | - | HG21 | 32 |
| 484 | - | HG22 | 32 |
| 485 | - | HG23 | 32 |
| 486 | 0.938 | QQG | 32 |
| 487 | 173.468 | C | 32 |
| 488 | 122.708 | N | 33 |
| 489 | 8.084 | H | 33 |
| 490 | 62.102 | CA | 33 |
| 491 | 4.159 | HA | 33 |
| 492 | 32.680 | CB | 33 |
| 493 | 2.106 | HB | 33 |
| 494 | - | QG1 | 33 |
| 495 | - | QG2 | 33 |
| 496 | 21.471 | CG1 | 33 |
| 497 | - | HG11 | 33 |
| 498 | - | HG12 | 33 |
| 499 | - | HG13 | 33 |
| 500 | 19.944 | CG2 | 33 |
| 501 | - | HG21 | 33 |
| 502 | - | HG22 | 33 |
| 503 | - | HG23 | 33 |
| 504 | 0.963 | QQG | 33 |
| 505 | 173.770 | C | 33 |
| 506 | 112.606 | N | 34 |
| 507 | 8.406 | H | 34 |
| 508 | 44.727 | CA | 34 |
| 509 | 3.967 | HA2 | 34 |
| 510 | - | HA3 | 34 |
| 511 | - | QA | 34 |
| 512 | 170.529 | C | 34 |
| 513 | 124.636 | N | 35 |
| 514 | 7.805 | H | 35 |
| 515 | 56.853 | CA | 35 |
| 516 | 4.215 | HA | 35 |
| 517 | 30.217 | CB | 35 |
| 518 | 2.147 | HB2 | 35 |
| 519 | 1.924 | HB3 | 35 |
| 520 | - | QB | 35 |
| 521 | 34.118 | CG | 35 |
| 522 | - | HG2 | 35 |
| 523 | - | HG3 | 35 |
| 524 | 2.298 | QG | 35 |
| 525 | - | CD | 35 |
| 526 | - | NE2 | 35 |
| 527 | - | HE21 | 35 |
| 528 | - | HE22 | 35 |
| 529 | - | QE2 | 35 |
| 530 | - | C | 35 |

# REFERENCES


Abe Y., Shodai T., Muto T., Mihara K., Torii H., Nishikawa S., Endo T. and Kohda D. (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell.* **3,** 551-60.

Alami, M., Luke, I., Deitermann, Eisner, G., Koch,H.G., Brunner, J. and Muller, M., (2003) Differential Interactions between a Twin-Arginine Signal Peptide and its Translocase in *Escherichia coli. Mol. Cell*, **12,** 937-946.

Alami, M., Trescher, D., Wu, L.F., and Muller, M. (2002) Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli. J. Biol. Chem.* **277,** 20499-20503.

Anderson, D., Carisson, U., and Freskgard, P. (2001) Contribution of Tryptophan residues to the CD spectrum of the extracellular domain of human tissue factor. *European J. Biochem* **268,** 1118.

Andrade, M.A., Chacon, P., Merelo J.J., and F. Moran. (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Engineering.* **6,** 383-390.

Anfinsen C.B., (1972) The formation and stabilization of protein structure. *Biochem J.* **128,** 737-49.

Applequist, J., Bode, K., Appella, D.H., Christianson, L.A., and Gellman S.H., J. (1998) Theoretical and Experimental Circular Dichroic Spectra of the Novel Helical Foldamer Poly[(1R,2R)-trans-2-aminocyclopentanecarboxylic acid]. *Am. Chem. Soc.,* **120,** 4891-4892.

Arnoux, P., Sabaty, M., Alric, J., Frangioni, B., Guigliarelli, B., Adriano, J.M. and Pignol, D. (2003) Structural and redox plasticity in the heterodimeric periplasmic nitrate reductase. *Nat. Struct. Biol.* **10,** 928–934.

Bairoch, A., Bucher, P., and Hofman, K., (1997) The Prosite database. *Nucl. Acids Res.,* **25,** 217-221.

Boyington, J.C., Gladyshev, V.N., Khangulov, S.V., Stadtman, T.C., Sun P.D. (1997) Crystal Structure of Formate Dehydrogenase H: Catalysis Involving Mo, Molybdopterin, Selenocysteine, and an Fe4S4 Cluster. *Science*, **275,** 1305-08.

Bartels, C., Xia, T.-H., Billeter, M., Guntert, P. and Wuthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **5,** 1-10.

Besteman, K., Zevenbergen M.AG., and Lemay S.G. (2005) Charge inversion by multivalent ions: Dependence on dielectric constant and surface-charge density. *Phys Review* **72,** 061501-1-9.

Berks, B.C., Ferguson, S. J., Moir, J. W. and Richardson, D. J. (1995) Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions, *Biochim. Biophys. Acta* **1232,** 97–173.

Berks, B.C., Palmer, T, and Sargent, F. (2005) Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr Opin Microbiol* **8,** 174-81.

Berks, B. C., Palmer, T., and Sargent, F. (2003) The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol.* **47,** 187-254.

Berks, B.C., Richardson, D. J., Reilly, A., Willis, A.C. and Ferguson, S. J. (1995) The napEDABC gene cluster encoding the periplasmic nitrate reductase system Thiosphaera pantotropha. *Biochem J.* **309**, 983-992.

Berks, B. C., Sargent, F., and Palmer, T. (2000) The Tat protein export pathway. *Mol Microbiol.* **35,** 260-74.

Billeter, M., Braun, W. and Wuthrich, K., (1982) Sequential Resonance Assignments in Protein 1H Nuclear Magnetic Resonance Spectra. Computation of Sterically Allowed Proton-Proton Distances and Statistical Analysis of Proton-Proton Distances in Single Crystal Protein Conformations. *J. Mol. Biol.* **155,** 321-346.

Brinkmann, U., Mattes, R.E. and Buckel, P. (1989) High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. *Gene* **85**, 109-14.

Brondijk, T.H.C., Fiegen, D., Richardson, D.J., and Cole, J.A., (2002) Roles of NapF, NapG and NapH subunits of the *Escherichia coli* periplasmic nitrate reductase, in ubiquinol oxidation. *Mole. Microbio.* **44,** 245-55.

Brondijk, T.H.C., Nilavongse, A., Filenko, N., Richardson, D.J., and Cole, J.A. (2004) NapGH components of the periplasmic nitrate reductase of *Escherichia coli* K-12: location, topology and physiological roles in quinol oxidation and redox balancing. *Biochem. J.* **379**, 47-55.

Buchanan, G. et al. (2001) A genetic screen for suppressors Escherichia coli Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. *Arch. Microbiol.* **177**, 107–112.

Byerly, D.W., McElroy, C.A. and Foster, M.P. (2002) Mapping the surface of *Escherichia coli* peptide deformylase by NMR with organic solvents. *Protein Science* **11,**1850-1853.

Cole, J. (1996) Nitrate Reduction to ammonia by enteric bacteria: redundancy or a strategy for survival during oxygen starvation? *FEMS Microbiol Lett* **136**, 1-11.

Constantine, K.L., Goldfarb, V., Wittekind, M., Friedrichs, MS, Anthony, J., Ng, SC, and Mueller, L. (1993) Aliphatic 1H and 13C resonance assignments for the 26-10 antibody VL domain derived from heteronuclear multidimensional NMR spectroscopy. *J Biomol NMR* **3,** 41-54.

Cornilescu, G., Delaglio F., and Bax, A., (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR* **13,** 289–302.

Creighton, T.E. (1993) Proteins: Structures and Molecular Properties. 2nd edition. WH Freeman and Company, New York.

de Keyzer, J., C. van der Does, and A. J. Driessen. (2003) The bacterial translocase: a dynamic protein channel complex. *Cell. Mol. Life Sci.* **60**, 2034-2052.

DeLano WL. The PyMOL Molecular Graphics System. DeLano Scientific LLC, CA, http://www.pymol.org.

Delgado, M.J., Bonnard, N., Ayala, A.T., Bedmar, E.J., and Muller, P. (2003) The Bradyrhizobium japonicum napEDABC genes encoding the periplasmic nitrate reductase are essential for nitrate respiration. *Microbiology, 149, 3395-3403.*

Dias, J.M. Than, M.E., Humm, A., Huber, R., Bourenkov, G., Bartunik, H., Bursakov, S., Calvete, J. Caldeira, J. Carneiro, C., Moura, J.J.G., Moura I., and Romão M.J. (1999)

Crystal structure of the first dissimilatory nitrate reductase at 1.9 Å solved by MAD methods", *Structure Fold Des*, **7**, 65-79.

Dill, K.A. (1990) Dominant forces in protein folding. Biochemistry 29, 7133-55.

Ding Z., and Christie P. J. (2003) Agrobacterium tumefaciens twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. J Bacteriol. **185**, 760-71.

Draper D.E., (1996) Escherichia coli and Salmonella Cellular and Molecular Biology, Translational Initiation, 2nd ed. Vol.1, 905-6.

Drenth, J. (1994) Principles of protein X-ray crystallography. New York: Springer Verlag. 1-304.

Dubs, A., Wagner, G., and Wuthrich, K., (1979) Individual assignments of amide proton resonances in the proton NMR spectrum of the basic pancreatic trypsin inhibitor. *Biochim. Biophys. Acta* **577,** 177-194.

Farmer, B., II, Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J. J., Robertson, J. G., and Mueller, L. (1996) Localizing the NADP+ binding site on the MurB enzyme by NMR. *Nat. Struct. Biol.* **3,** 995-997.

Forsberg, M., Blomgran, R., Lerm, M., Sarndahl, E., Sebti, S. M., Hamilton, A., Stendahl, O., and Zheng, L. J. (2003) Differential effects of invasion by and phagocytosis of Salmonella typhimurium on apoptosis in human macrophages: potential role of Rho-GTPases and Akt. Leukoc *Biol.* **74,** 620-9.

Geddes, C.D. (2001) Optical halide sensing using fluorescence quenching: theory, simulations and applications—a review. *Meas. Sci. Technol.* **12,** R53–R88.

Greenfield, N.J. (2004) Circular dichroism analysis for protein-protein interactions. *Methods Mol Biol.* **261,** 55-78.

Greenfield, N.J. (2004) Analysis of circular dichroism data. *Methods Enzymol.* **383**:282-317.

Greenfield, N.J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, *Anal Biochem* **235**:1-10

Greenfield, N.J. and Fasman, G.D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry,* **8,** 4108-16.

Grimsley, J.K., Scholtz, J.M., Pace, C.N., and Wild, J.R., (1997) Organophosphorus Hydrolase Is a Remarkably Stable Enzyme That Unfolds through a Homodimeric Intermediate. *Biochemistry,* **36,** 14366-14374.

Grzesiek, S., Dobeli H., Gentz, R., Garotta G., Labhardt A.M., and Bax A., (1992) 1H, 13C, and 15N NMR backbone assignments and secondary structure of human interferon-gamma. *Biochemistry,* **31,** 8180-90.

Grzesiek, S., and Bax, A. (1992a) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114,** 6291-6293.

Grzesiek, S., and Bax, A. (1992b) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96,** 432-440.

Guntert, P., Mumenthaler C., and Wuthrich, K., (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol.* **273,** 283–298.

Halbig, D. et al. (1999) The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact

twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Eur. J. Biochem.* **263**, 543–551.

Hatzixanthis, K., Clarke, T. A., Oubrie, A, Richardson, D.J., Turner, R.J., and Sargent, F. (2005) Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. *Proc Natl Acad Sci. U S A.* **102,** 8460-5.

Herrmann, T., Güntert, P. and Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209-227.

Ilbert, M., Mejean, V., Giudici-Orticoni, M. T., Samama, J. P. and Iobbi-Nivol, C. (2003) Involvement of a Mate Chaperone (TorD) in the Maturation Pathway of Molybdoenzyme TorA. *J. Biol. Chem.* **278**, 28787-28792

Ilbert, M., Mejean, V. and Iobbi-Nivol, C. (2004) Functional and structural analysis of members of the TorD family, a large chaperone family dedicated to molybdoproteins. *Microbiology,* **150**, 935-943

Ize, B., Porcelli, I., Lucchini, S., Hinton, J.C., Berks, B.C., and Palmer, T. (2004). Novel phenotypes of Escherichia coli tat mutants revealed by global gene expression and phenotypic analysis. *J Biol Chem.* **279,** 47543-54.

Jack, R.L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins. *EMBO J.* **23,** 3962-72.

Jack, R.L., Dubini, A., Palmer, T., and Sargent, F. (2005) Common principles in the biosynthesis of diverse enzymes. *Biochem Soc Trans.* **33,** 105-7.

Kay, L.E., Gardner, K.H. (1997) Solution NMR spectroscopy beyond 25 kDa. *Curr Opin Struct Biol,* **7,** 722-731.

Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) Three-dimensional triple resonance NMR-spectroscopy of isotopically enriched proteins. *J. Magn. Res.* **89,** 496-512.

Kipping M., Lilie H., Lindenstrauss U., Andreesen J.R., Griesinger C., Carlomagno T., and Bruser T. (2003) Structural studies on a twin-arginine signal sequence. *FEBS Lett.,* **550,** 18-22.

Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55.

Kozlov, G., Perreault, A., Schrag, J.D., Park, M., Cygler, M., Gehring, K., and Ekiel, I., (2004) Insights into the function of PSI domains from structure of the Met receptor PSI domain. *Biochem. And Biophys. Research Comm.* **321,** 234-40.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.

Kurisu G, Zhang H, Smith JL & Cramer WA (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* **302,** 1009–1014.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680–685.

Lange, R., Frank, H., Saldana, J.L. and Balny, C. (1996) Fourth derivative UV-spectroscopy of proteins under high pressure: 1. Factors affecting the fourth derivative spectrum of the aromatic amino acids. *Eur. J. Biophys.* **24**, 277–283.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486.

Loo, J.A. (1995) Observation of Large Subunit Protein Complexes by Electrospray Ionization Mass Spectrometry, *Journal of Mass Spectrometry* **30,** 180-83.

Liu, H.P., Takio, S., Satoh, T., and Yamamoto, I. (1999) Involvement in Denitrification of the napKEFDABC Genes Encoding the Periplasmic Nitrate Reductase System in the Denitrfying Phototrophic Bacterium *Rhodobacter sphaeroides* f. sp. *Denitrificans*. *Biosci. Biotechnol. Biochem.*, **63,** 530-6.

Marietou, A., Richardson, D., Cole, J., and Mohan, S., (2005) Nitrate reduction by *Desulfovibrio desulfuricans*: A periplasmic nitrate reductase system that lacks NapB, but includes a unique tetraheme c-type cytochrome, NapM. *FEMS Microbio. Lett.* **248,** 217-25.

Marti-Renom, M.A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29,** 291-325.

Matsudaira, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes *J. Biol. Chem.* **262**, 10035-10038.

McAlpine, A.S., McEwan A.G., and Bailey S., (1998) The High Resolution Crystal Structure of DMSO Reductase in Complex with DMSO. *J.Mol.Biol.*, **275**, 613-23.

Nurizzo, D., Halbig, D., Sprenger, G.A. and Baker, E.N. (2001) Crystal structures of the precursor form of glucose-fructose oxidoreductase from Zymomonas mobilis and its complexes with bound ligands. *Biochemistry,* **40,** 13857-67.

Ochsner U. A., Snyder A., Vasil A. I., and Vasil M. L. (2002) Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc Natl Acad Sci U S A.* **99,** 8312-7.

Oresnik I.J., Ladner C.L., and Turner R.J. (2001) Identification of a twin-arginine leader-binding protein. *Mol Microbiol.,* **40,** 323-31.

Palmer III A.G. (1997) Probing molecular motion by NMR. *Curr Opin Struct Biol.* **7,** 732-737.

Palmer T, Sargent F, and Berks BC. (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol.* **13,** 175-80.

Papish A.L., Ladner C.L., and Turner R.J. (2003) The twin-arginine leader-binding protein, DmsD, interacts with the TatB and TatC subunits of the Escherichia coli twin-arginine translocase. *J Biol Chem.* **29,** 2501-6.

Peng, J.L., and Wagner G. (1994) Investigation of protein motions via relaxation measurements. *Methods Enzymol.* **239,** 563-596.

Pommier, J., Mejean, V., Giordano, G. and Iobbi-Nivol, C. (1998) TorD, A Cytoplasmic Chaperone That Interacts with the Unfolded Trimethylamine *N*-Oxide Reductase Enzyme (TorA) in *Escherichia coli J. Biol. Chem.* **273,** 16615-16620.

Potter, L.C, Angove, H., Richardson, D. and Cole, J.A. (2001), Nitrate reduction in the periplasm of Gram-negative bacteria, *Adv. Microbial Physiol.* **45,** 51–112.

Potter, L. C. and Cole, J. A. (1999). Essential roles for the products of the *napABCD* genes, but not *napFGH*, in periplasmic nitrate reduction by *Escherichia coli* K-12. *Biochem. J.* **344**, 69–76.

Pradel, N., Ye, C., Livrelli, V., Xu, J. and Wu, L F. (2003) Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic Escherichia coli O157:H7. *Infect Immun.* **71,** 4908-16.

Richardson, D.J., Berks, B.C., Russel, D.A., Spiro, S. and Taylor, C.J. (2001). Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. *Cell mol. Life Sci.* **85**, 165-178.

Richardson, D. J., and Watmough, N.J. (1999). Inorganic nitrogen metabolism in bacteria. *Curr Opin Chem Biol* **3,** 207-219.

Rossier O., and Cianciotto N. P. (2005) The Legionella pneumophila tatB gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infect Immun.* **73,** 2020-32.

San Miguel M., Marrington R., Rodger P.M., Rodger A., and Robinson C. (2003) An Escherichia coli twin-arginine signal peptide switches between helical and unstructured conformations depending on the hydrophobicity of the environment. *Eur J Biochem.* **270,** 3345-52.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, **2nd edn**. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Sarver, R.W. and Krueger, W.C. (1991) An infrared and circular dichroism combined approach to the analysis of protein secondary structure. *Anal Biochem* **199**: 61-7.

Schöneich, C., Hühmer, A., Rabel, S., Stobaugh, J., Jois, S., Larive, C., Siahaan, T., Squier, T., Bigelow, D., and Williams, T. (1995) *Analytical Chemistry* **67**, 155R-181R.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and Clore, G.M. (2003) The Xplor-NIH NMR Molecular Structure Determination Package. *J. Magn. Res.* **160**, 66-74.

Staunton, D., Owen, J. and Campbell, I.D. (2003) NMR and structural genomics. *Acc. Chem. Res.* **36**, 207-214.

Sternberg, MJE, Grace DEP, Phillips DC (1979) Separation and Analysis of Peptides and Proteins. Dynamic information from protein crystallography. *J Mol Biol,* **130,** 231-253.

Stewart, V., Lu, Y., and Darwin, A.J., (2002) Periplasmic Nitrate Reductase (NapABC) Supports Anaerobic Respiration by *Escherichia coli* K-12, *J. Bacteriology* **184,** 1314-23.

Terwilliger, T. C. (2003) Automated structure solution, density modification and model building. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 1937-40.

Thomasson, W.A. Unraveling the Mystery of Protein Folding. Breakthroughs in Bioscience (http://opa.faseb.org/pdf/protfold.pdf)

Tranier, S., Iobbi-Nivol, C., Birck, C., Ilbert, M., Mortier-Barriere, I., Mejean, V. & Samama, J. P. (2003) A Novel Protein Fold and Extreme Domain Swapping in the Dimeric TorD Chaperone from Shewanella massilia. *Structure,* **11,** 165-174.

Turner R.J., Papish A.L., and Sargent F. (2004) Sequence analysis of bacterial redox enzyme maturation proteins (REMPs). *Can J Microbiol.* **50**:225-38.

Volpon, L., Lievre, C., Osborne, M.J., Gandhi, S., Iannuzzi, P., Larocque, R., Cygler, M., Gehring, K., and Ekiel, I., (2003) The Solution Structure of YbcJ from Escherichia coli Reveals a Recently Discovered αL Motif Involved in RNA Binding. *Journal of Bacteriology*, **185,** 4204-10.

Voulhoux, R., Ball, G., Ize, B., Vasil, M.L., Lazdunski, A., Wu, L. F., and Filloux, A. (2001) Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J.* **20,** 6735-41.

Waddell W.J. (1956) A simple ultraviolet spectrophotometric method for the determination of protein. *J Lab Clin Med.* **48,** 311-4.

Wagner, G. (1997) An account of NMR in structural biology. *Nat. Struct. Biol.*; Oct. Suppl, 841-844.

Wagner, G. and Wuthrich, K., (1982) Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. *J.Mol.Biol.* **155,** 347-366.

Walter, S. and Buchner, J. (2002) Molecular chaperones-cellular machines for protein folding. *Andrew Chem Int Ed. Engl* **41,** 1098-113.

Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H. Cole, J.A., and Turner, R.J., (1998) A Novel and Ubiquitous System for Membrane Targeting and Secretion of Cofactor-Containing Proteins. *Cell* **93,** 93-101.

Wishart, D.S., C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. (1995) $^{1}H$, $^{13}C$ and $^{15}N$ chemical shift references in biomolecular NMR. *J. Biomol. NMR* **6,** 135-140.

Woody, R.W. and Dunker, A.K. (1996) Aromatic and cystine side-chain circular dichroism in proteins. *Circular Dichroism and the Conformational Analysis of Biomolecules.* **109,** 157.

Wuthrich, K., (1986) NMR of Proteins and Nucleic Acids. *Wiley, New York.*

Zhuang, X., Ha, T., Kim, H.D., Centner, T., Labeit, S., and Chu, S. (2000) Fluorescence quenching: A tool for single-molecule protein-folding study. *PNAS* **97,** 14241–14244

WEBSITES

http://dicroprot-pbil.ibcp.fr/Documentation_dicroprot.html
http://maple.bioc.columbia.edu/pp/
http://s-provencher.com/pages/contin-cd.shtml
http://www.ahpcc.unm.edu/~aroberts/main/pc_software.htm
http://www.cryst.bbk.ac.uk/PPS2/projects/schirra/html/3dnmr.htm
http://www.cryst.bbk.ac.uk/PPS2/projects/schirra/html/assign.htm
http://www.embl-heidelberg.de/~andrade/k2d.html
http://www.med.unc.edu/wrkunits/2depts/biochem/MACINFAC/cdprep.html
http://www.ncbi.nlm.nih.gov/
http://www.newark.rutgers.edu/chemistry/grad/chem585/lecture1.html#1
http://www.rcsb.org/pdb/
http://www2.umdnj.edu/cdrwjweb/