Initiation of DNA Replication in *Bacillus subtilis*: Structural Studies of the DnaA-DnaD Interaction

Eleyna Martin, Msci. (Hons.)

Thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

June 2018

Abstract

Replication of genetic information is a vital process across all domains of life. *Bacillus subtilis* is considered the gram-positive model bacterium for studying DNA replication (*Escherichia coli* has been studied extensively as the gram-negative model) and is most representative of the ancestral phylum of prokaryotes. DNA replication has three distinct stages; initiation, elongation and termination. Replication initiation is the focus of this research and this process occurs at a single origin conserved throughout bacteria, termed *oriC*. *B. subtilis* primosomal machinery is formed of replication initiator proteins DnaA, DnaD and DnaB, the helicase loader DnaI, replicative helicase DnaC and primase DnaG. The role of the initiator proteins is to melt the DNA double helix and enable loading of the hexameric ring helicase onto each strand of DNA for bidirectional replication. Initiation is the first stage in DNA replication and despite its importance the molecular mechanisms of replication initiation remain largely unclear.

The work presented in this thesis has focussed on the essential interaction between replication initiator proteins DnaA and DnaD, with an aim to characterise their binding interface and reveal molecular details of their mechanisms of interaction during DNA replication initiation. The direct interaction between isolated DnaA domain I and DnaD DDBH2 domain was detected by NMR spectroscopy which was subsequently used to identify the specific residues involved and characterise the nature of the binding interface. The kinetics of the interaction were investigated by SPR and computational techniques were used to model the DnaA-DnaD complex. This structural characterisation of the DnaA-DnaD interaction provides greater understanding of the molecular mechanisms of DnaA and DnaD during DNA replication initiation.

Acknowledgements

I would first and foremost like to thank my supervisors Professor Panos Soultanas and Professor Mark Searle for giving me the opportunity to work on this project and for their support and guidance throughout my research. A great appreciation to Dr Huw Williams for his extensive NMR support that has been vital throughout this work. I am very grateful to Dr Alex Slater for his assistance within the lab and especially with circular dichroism, ITC and SPR experiments. Thank you to Dr Vaselious Paschalis for his guidance on protein expression and purification and Dr Juliet Morgan for her expertise on NMR assignment theory. I am grateful to Dr Dan Scott and Jedd Bellamy-Carter for their assistance with mass spectrometry. An additional thank you to all members of the Searle, Soultanas and Thomas research groups for their friendship and creating such an enjoyable work environment throughout my time here. Finally, I must thank my family for their constant support and encouragement.

Abbreviations

A_{280}	Absorbance at 280 nm
AAA+	ATPases associated with diverse cellular activities
ADP	Adenosine diphosphate
AMU	Atomic mass unit
APS	Ammonium persulphate
AT	Adenine and Thymine
ATP	Adenosine triphosphate
AU	Absorbance unit
Bis-Tris	2,2-Bis(hydroxymethyl)-2,2',2"-nitrotriethanol
BMOE	Bis(maleimido)ethane
B. subtilis	*Bacillus subtilis*
bp	Base pair
CD	Circular Dichroism
CSP	Chemical shift perturbation
CTD	C-terminal domain
DDBH1	DnaD DnaB homology 1
DDHB2	DnaD DnaB homology 2

DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
DUE	DNA unwinding element
ds	Double stranded
D_2O	Deuterium oxide
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
FPLC	Fast protein liquid chromatography
FRET	Fluorescent resonant energy transfer
g	Gravitational acceleration
GC	Guanine and Cytosine
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulphonic acid
HSQC	Heteronuclear single quantum correlation
IPTG	Isopropyl β-D-1thiogalactopyranoside
ITC	Isothermal titration calorimetry
K	Kelvin
kb	Kilobase
kDa	Kilodalton
KLD	Kinase, ligase, DpnI
LB	Luria Broth
MHz	Megahertz
MWCO	Molecular weight cut off

mS	millisiemens
NMR	Nuclear magnetic resonance
NTD	N-terminal domain
OD_{595}	Optical density at 595 nm
oriC	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PNK	Polynucleotide kinase
ppm	Parts per million
RCF	Relative centrifugal force
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
ss	Single stranded
Ta	Association temperature
TAE	Tris-acetate-EDTA
TEMED	Tetramethylenediamine
Tris	Tris(hydroxymethyl)aminoethane
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume
1D	One dimensional
2D	Two dimensional
3D	Three dimensional

Amino Acid	Three Letter Code	One Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table 1 *Abbreviations and one letter notations for the 20 naturally occurring amino acids.*

Contents

Chapter 1

Introduction

1.1 DNA Replication

All living organisms require the ability to replicate their genetic information. This process must occur so that a cell can produce genetically identical daughter cells. With the exception of a few viruses, the DNA double helix carries the genetic information. Vast quantities of DNA must be duplicated with extreme accuracy making replication a complex task for all organisms, hence, DNA replication requires the cooperation of many proteins. DNA duplication rates of 1000 nucleotides per second can be achieved, with high accuracy, by elaborate replication machinery[1]. DNA replication must be coordinated to ensure it only occurs once per cell cycle and this regulation adds an additional level of complexity to the process. Numerous mechanisms ensure the DNA sequence is not modified during replication but occasional changes do occur and over time provide genetic variation. The mutation rate is roughly the same across species at 1 change per 10^9 nucleotides[1]. Features of DNA replication have been conserved between prokaryotes and eukaryotes and DNA replication is fundamentally similar across all domains of life. However, the specific proteins involved and regulatory mechanisms differ greatly between the domains of life and so, the main focus here will be on DNA replication in bacteria.

1.1.1 Semi-conservative, Semi-discontinuous Replication

DNA replication across all domains of life occurs through a semi-conservative mechanism. At the origin of replication (a single origin in bacteria or multiple for archaea and eukaryotes) parental duplex DNA is separated progressively and used as a template for synthesis of new complementary DNA strands. This results in two double-stranded daughter chromosomes containing one conserved parental strand and one newly synthesised DNA strand[2] (Figure 1.1). An asymmetric replication fork is formed at the origin of replication and moves progressively along the parental DNA double helix. Eukaryotic DNA replication is unidirectional, however, during bacterial DNA replication two replication forks are formed and replication proceeds bidirectionally.

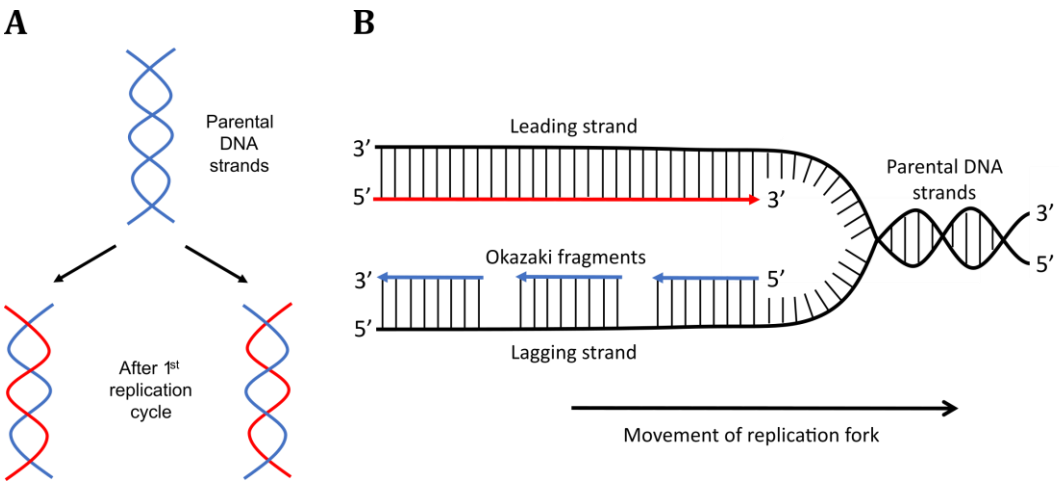


Figure 1.1 A. *Semi-conservative DNA replication. After each round of replication, the daughter chromosome contains one conserved parental stand and one newly synthesised DNA strand.* ***B.*** *Semi-discontinuous DNA replication. DNA polymerases synthesise DNA in the 5' - 3' direction. Consequently, the leading strand synthesis occurs continuously while the lagging strand is synthesised as a series of Okazaki fragments.*

DNA strands are synthesised by the action of DNA polymerases that extend DNA in a 5' – 3' direction. Consequently, each strand of duplex DNA requires a different mechanism of synthesis to allow simultaneous replication of the antiparallel template strands while the DNA polymerase travels only in one direction. The nascent DNA strand extending in the 5' – 3' direction is deemed the leading strand and is continuously synthesised with the replication fork movement. In contrast, the nascent

DNA strand extended in the 3' – 5' direction is synthesised as a series of 5' – 3' segments known as Okazaki fragments[3] (1000 – 2000 nucleotides in bacteria and 100 – 200 nucleotides in eukaryotes) and this is deemed the lagging strand (Figure 1.1). DNA ligase then catalyses the formation of phosphodiester bonds between the Okazaki fragments. This process is termed 'semi-discontinuous replication' as only the leading strand is replicated simultaneously.

1.1.2 *Bacillus subtilis* as a Model for Bacterial DNA Replication

The gram-negative model organism *Escherichia coli (E. coli)* was previously deemed the universal model for the investigation of bacterial DNA replication. However, recent studies suggest that the phylum of *Firmicutes* provide a better model[4]. The *Firmicutes* are gram-positive bacteria of three classes; *Bacilli*, *Clostridia* and *Mollicutes*. Their genomes have low GC content and are most representative of the ancestral phylum of prokaryotes. Although the gram-negative *E. coli* is the most studied model organism for DNA replication, the replication machinery of *Bacillus subtilis (B. subtilis)* has also been studied extensively and can be considered the gram-positive model bacterium. There is a divergence time of approximately 1.2 billion years between *B. subtilis* and *E. coli*[5] and in terms of cellular functions and mechanisms the majority of bacteria show greater resemblance to *B. subtilis*[6].

1.1.3 DNA Replication as an Antimicrobial Target

The bacterial DNA replication machinery, along with mechanisms of regulation for DNA replication, varies between species within the prokaryotic domain of life. Most of our current understanding has been derived from the gram-negative model organism *E. coli*. However, widening our understanding to include gram-positive organisms, using the *B. subtilis* model, is desirable. DNA replication is an underexploited antimicrobial target given that the replicative proteins are highly conserved and the mechanisms of DNA replication are distinct between prokaryotes and eukaryotes[7, 8]. Gram-positive pathogens such as *C. difficile* and *S. aureus* are important targets for new antibiotic drugs and antimicrobial therapies. Current therapies are limited to inhibitors of topoisomerase II and with the challenge of antimicrobial resistance, exploration of new

targets is essential for development of novel DNA replication-targeting antimicrobials[9]. Nevertheless, full understanding of the mechanisms of DNA replication is required before effective therapies can be produced.

1.2 The Bacterial Chromosomal Origin of Replication

Typically, bacteria have circular chromosomes and DNA replication initiates at a single origin, *oriC*. This unique site consists of an AT rich region known as the DNA unwinding element (DUE) where initial melting of the DNA double helix occurs[10]. Subsequently, the replication forks are formed, on each exposed DNA strand, and progress in opposite directions along the chromosome for bidirectional replication. Replicative origins can vary significantly in length, sequence and organisation. Nevertheless, conserved features include the DUE, presence of DnaA-box clusters and the locations for binding of regulatory proteins[11]. The position of the *oriC* on the chromosome is typically found in close proximity to the *dnaA* gene, encoding master replication initiator protein DnaA. The genes around *oriC* and *dnaA* are also commonly conserved and are typically flanked by *dnaN* [11, 12]. Chromosomal origins of replication can be continuous, such as with *E. coli*, where all functional elements are within a single intergenic region. Alternatively, they can be bipartite, such as with *B. subtilis*, where two regions of functional elements are separated by the *dnaA* gene (as displayed in Figure 1.2). The bipartite origin is important for *B. subtilis,* where looped structures identified by electron microscopy are thought to be a consequence of the nature of the two sub-regions of the origin[13].

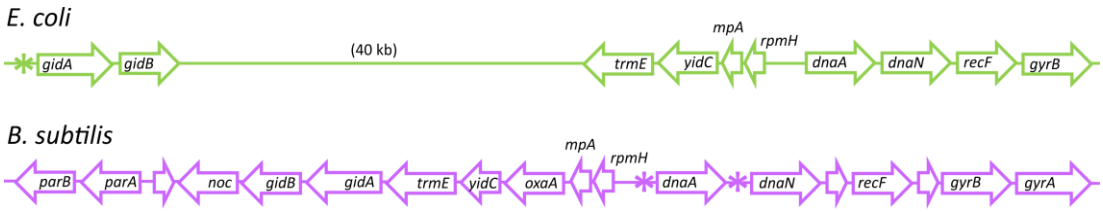


Figure 1.2 *Gene organisation around the oriC for B. subtilis and E. coli[4]. DnaA-box clusters forming the origin of replication are indicated by asterisks.*

As mentioned previously, a conserved feature within bacterial origins of replication are DnaA-boxes, recognition sequences for binding of the initiator protein DnaA. DnaA binds double stranded DNA, at specific 9 base pair sequences, note *Thermotoga maritima* is an exception where a 12 base pair sequence was identified[11]. The number and spacing of DnaA-boxes varies between organisms, they share a conserved core sequence and deviate by only one or two bases from the *E. coli* and *B. subtilis* shared consensus sequence 5' – TTATCCACA – 3'[14]. The affinity of DnaA for the DnaA-boxes within the *oriC* can also differ, due to sequence variations, with high and low affinity boxes as well as DnaA boxes that are dependent on the presence of ATP-bound DnaA[15]. DnaA-box orientation and locations are imperative for replication initiation. DnaA-box spacing altered by part of a helical turn is detrimental to initiation in *E. coli*[16]. Spatial preferences have been observed for certain bacteria, such as two closely spaced DnaA-boxes over a single DnaA-box[17, 18]. Within *B. subtilis*, DnaA-box clusters are located in the intergenic regions both upstream and downstream of the *dnaA* gene, both of which are essential to the origin function[19].

DNA unwinding elements are another conserved feature of replication origins. These are AT rich repeat sequences consisting of typically 13 nucleotides which provide reduced thermo-stability compared to neighbouring sequences. DUE's are located upstream or downstream of a DnaA-box cluster where the location of the DUE and its neighbouring DnaA-box cluster is essential for DNA unwinding[20]. Despite the difference in origin composition between bacterial species, all function to; allow DNA unwinding, direct formation of replication complexes and contain sequences for regulation of DNA replication. The location of *oriC* within the genome is conserved in close proximity to the *dnaA* gene across many bacteria. However, within *E. coli* and related gram-negative organisms, major genomic rearrangement has removed the proximity of *oriC* to *dnaA* (Figure 1.2). Replication origins likely co-evolved along with the cognate DnaA protein to optimise their interactions, accounting for the species specific mechanisms of DNA replication initiation[11].

1.3 Stages of DNA Replication

DNA replication has three distinct stages; initiation, elongation and termination. Initiation involves assembly of the replisome at the chromosomal origin of replication. Elongation is the progression of the replication forks along the chromosome, during which complementary DNA is synthesised. Finally, termination is the physiological arrest of replication when chromosome duplication is completed. The components promoting initiation, elongation and termination are well conserved throughout bacteria, though differences can be found in the mechanisms of genome duplication across species. The majority of knowledge surrounding DNA replication in bacteria has been derived from the gram-negative model organism *E. coli*, here differences observed in the gram-positive model organism *B. subtilis* will also be discussed.

1.3.1 Initiation of DNA Replication

The first stage of DNA replication is initiation. This involves assembly of the orisome (pre-replicative complex) at the *oriC*, which is highly regulated and involves a number of proteins known as replication initiators[4]. During genome replication, duplex DNA must be separated for DNA polymerase to copy the genetic information. The purpose of orisome assembly is to locally melt the DNA double helix and allow loading of the replicative helicase onto each template strand of DNA for bidirectional replication using two distinct replication forks.

1.3.1.1 Initiation in *Escherichia coli*

The initiator protein DnaA binds to the *oriC* to initiate DNA replication. DnaA binds at multiple recognition sites within the origin termed DnaA-boxes[21] (as described in section 1.2) where it forms a right handed helical oligomer[22, 23]. The DNA of the *oriC* is bound around the outside of the helical filament which induces localised unwinding of the duplex DNA at the AT rich DUE[24, 25]. DnaA then recruits the DNA helicase, DnaB, to be loaded onto the exposed ssDNA by the helicase loader DnaC[26, 27]. The helicase DnaB then recruits the primase DnaG and the β-clamp DnaN which in turn recruits the replicative machinery as discussed in section 1.3.2.

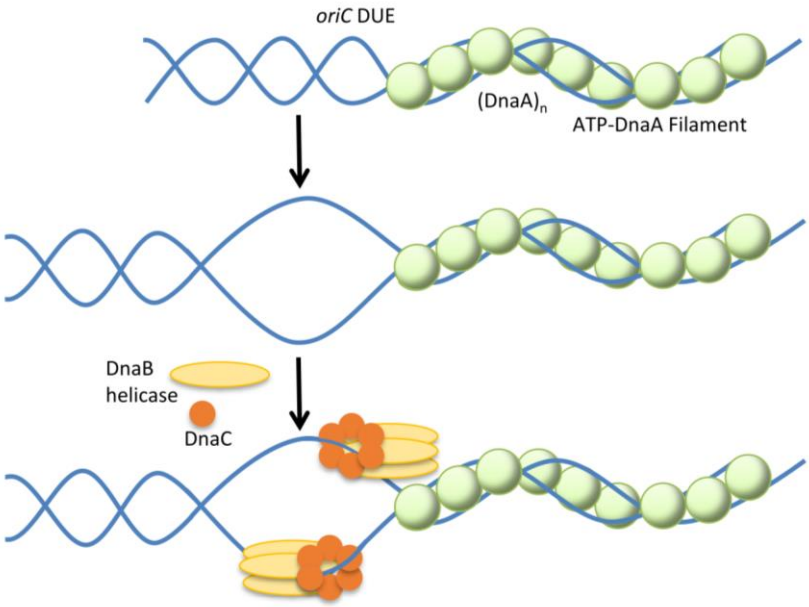


Figure 1.3 *Initiation of DNA replication in E. coli. Initiator protein DnaA specifically binds DnaA-box sites within oriC. A right handed oligomeric filament structure is formed by ATP-DnaA which induces local melting of the AT rich DUE. DnaC mediates loading of the replicative helicase DnaB onto ssDNA at the origin via a 'ring-breaking' mechanism.*

DnaA is a member of the AAA+ family of proteins (ATPases associated with various cellular activities) due to its ability to bind and hydrolyse ATP[14]. The active ATP-bound form of DnaA is required for oligomerisation at the *oriC* which subsequently induces DNA unwinding and replisome assembly[28, 29]. As mentioned in section 1.2, the affinity of DnaA for DnaA-boxes within the origin of replication can vary depending on sequence as well as the ATP/ADP bound state of DnaA[30, 31]. DnaA-boxes termed 'strong' can bind ATP and ADP bound DnaA whereas 'weak' DnaA-boxes have a much greater affinity for ATP-bound DnaA[32]. DnaA recruitment to the origin is thought to be cooperative whereby sites with 'strong' DnaA-boxes act as anchor points to recruit ATP-DnaA to the 'weak' sites for oligomerisation into the helical filament[15, 33]. Local melting of the DUE is thought to be directly mediated by the DnaA oligomer. Positive writhe is introduced into the dsDNA wrapped around the outside of the DnaA filament which is compensated by negative writhe just ahead of the filament at the DUE[22, 23, 34]. The ATPase domain of DnaA can bind unwound ssDNA to stabilise strand separation.

The NTD of DnaA interacts with the helicase DnaB for recruitment to the *oriC*[26, 35, 36]. DnaB is active as a homo-hexamer which is loaded onto the exposed ssDNA by the helicase loader DnaC. Loading of the helicase is proposed to occur by a 'ring-breaking' mechanism where DnaC forms an oligomer that remodels the hexameric ring of DnaB to open a gap for loading onto the ssDNA[37, 38]. Once loaded, DnaB recruits DnaG and DnaN which then recruits the DNA polymerase to begin the elongation phase of DNA replication.

1.3.1.2 Initiation in *Bacillus subtilis*

Initiation of DNA replication in *B. subtilis* also utilised the initiator protein DnaA, the helicase DnaC akin to *E. coli* DnaB, a helicase loader DnaI and the primase DnaG. *B. subtilis* DnaA forms right handed helical oligomers on dsDNA, can bind ssDNA within the filament and requires ATP-bound DnaA for cooperative binding to the origin of replication. However, replication initiation in *B. subtilis* requires two additional initiator proteins DnaD and DnaB (note *B. subtilis* DnaB is not related to the *E. coli* DnaB helicase, see Table 1.1). These are homologous proteins which exhibit DNA remodelling activity (further details on *B. subtilis* DnaA, DnaD and DnaB proteins are discussed in section 1.5). Table 1.1 displays the equivalent components essential to *E. coli* and *B. subtilis* DNA replication initiation. This mechanism of replication initiation is conserved across low GC content gram-positive organisms where initiator proteins DnaA, DnaD, DnaB and DnaI have been found[39]. Figure 1.4 illustrates the hierarchical assembly of the primosomal machinery at the *B. subtilis* origin of replication.

Role	***E. coli***	***B. subtilis***
Initiator	DnaA	DnaA
DNA remodelling	-	DnaD, DnaB
Helicase loader	DnaC	DnaI
Helicase	DnaB	DnaC
Primase	DnaG	DnaG

Table 1.1 *The essential DNA replication initiation machinery of E. coli and B. subtilis. Note the somewhat confusing nomenclature differences between the two bacterial replication systems.*

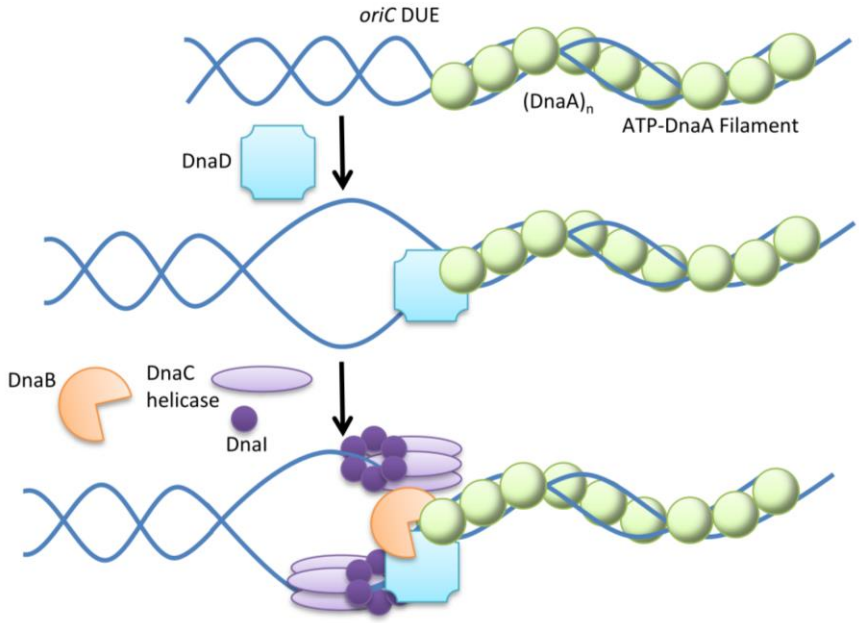


Figure 1.4 *Model for B. subtilis DNA replication initiation at the oriC. Sequential recruitment of initiator proteins occurs beginning with localisation of master initiator protein DnaA at the oriC. DnaA specifically binds DnaA-box sites forming a high order oligomeric filament structure. This locally melts the AT rich DUE with the help of initiator protein DnaD. DnaB is recruited and assists DnaI mediated loading of the replicative helicase DnaC onto ssDNA at the origin. The 'ring-assembly' loading mechanism of the hexameric helicase occurs in an ATP dependent manner. Recruitment of DnaG completes the active primosome.*

Loading of the hexameric helicase DnaC onto DNA in *B. subtilis* is thought to occur via a 'ring-assembly' mechanism[38]. This mechanism was proposed in contrast to the 'ring-breaking' mechanism for *E. coli* as the pre-formed hexameric DnaC is inactive *in vitro*, even in the presence of loaders DnaB and DnaI, yet monomeric DnaC displays helicase and translocase activity[40]. Helicase loader DnaI and co-loader DnaB cooperate to assemble the helicase monomers into functional hexameric rings directly onto the ssDNA. DnaI contains an N-terminal helicase interaction domain and a C-terminal AAA+ domain, the ATPase activity of DnaI is stimulated by the presence of ssDNA but only in the absence of the N-terminal domain. It is thought that binding of the N-terminal domain to the DnaC helicase alters the conformation to reveal the ssDNA binding site within the C-terminal domain which subsequently stimulates loading of the helicase onto ssDNA. The ATPase activity of DnaI may stimulate its release from DnaC once loading has occurred. Figure 1.5 illustrates the different mechanism for helicase loading in *E. coli* and *B. subtilis*.

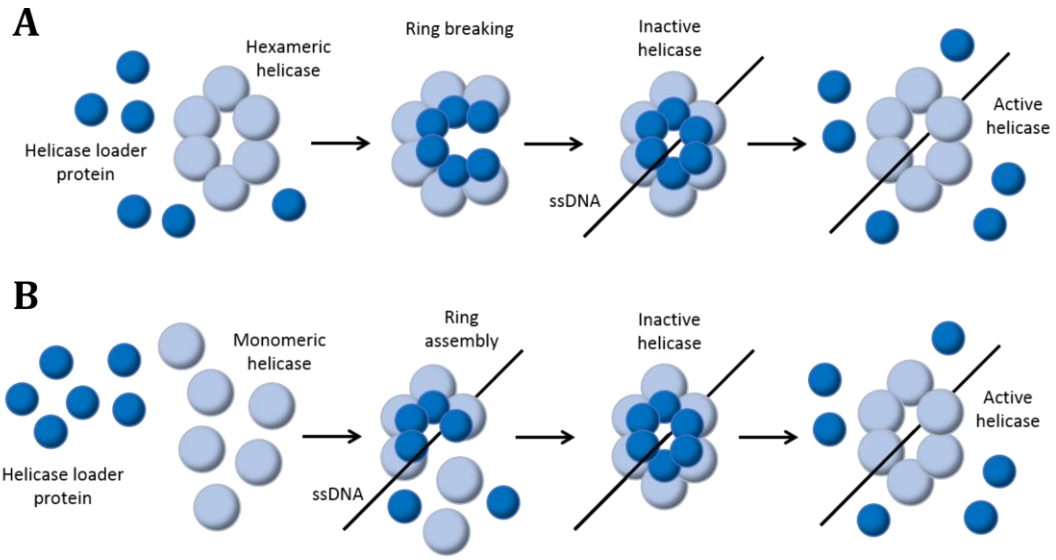


Figure 1.5 *Schematic representation of helicase loading mechanisms.* ***A.*** *The 'ring-breaking' mechanism. The helicase loader protein binds the hexameric helicase and alters the conformation, forming a 'break' in the ring, to allow loading onto ssDNA.* ***B.*** *The 'ring-assembly' mechanism. The helicase loader protein binds helicase monomers and assembles the hexameric ring around ssDNA. In both cases the helicase loader protein dissociates from the helicase once loaded onto the DNA.*

1.3.2 The Elongation Phase of DNA Replication

During the elongation phase of DNA replication, the replisome (a large multi subunit complex) processively synthesises DNA. The replisome consists of DNA polymerases, the helicase, primase, processivity clamps and the clamp loader complex (Figure 1.6). The helicase (*E. coli* DnaB, *B. subtilis* DnaC) is located at the head of the replication fork, tethering through the hexameric ring the lagging strand of the template DNA. ATP-hydrolysis drives the separation of the dsDNA by translocation of the helicase along the template strand. The exposed ssDNA is then coated with single-stranded binding protein (SSB) to prevent re-annealing of the template DNA while also providing protection from nucleases[7, 41].

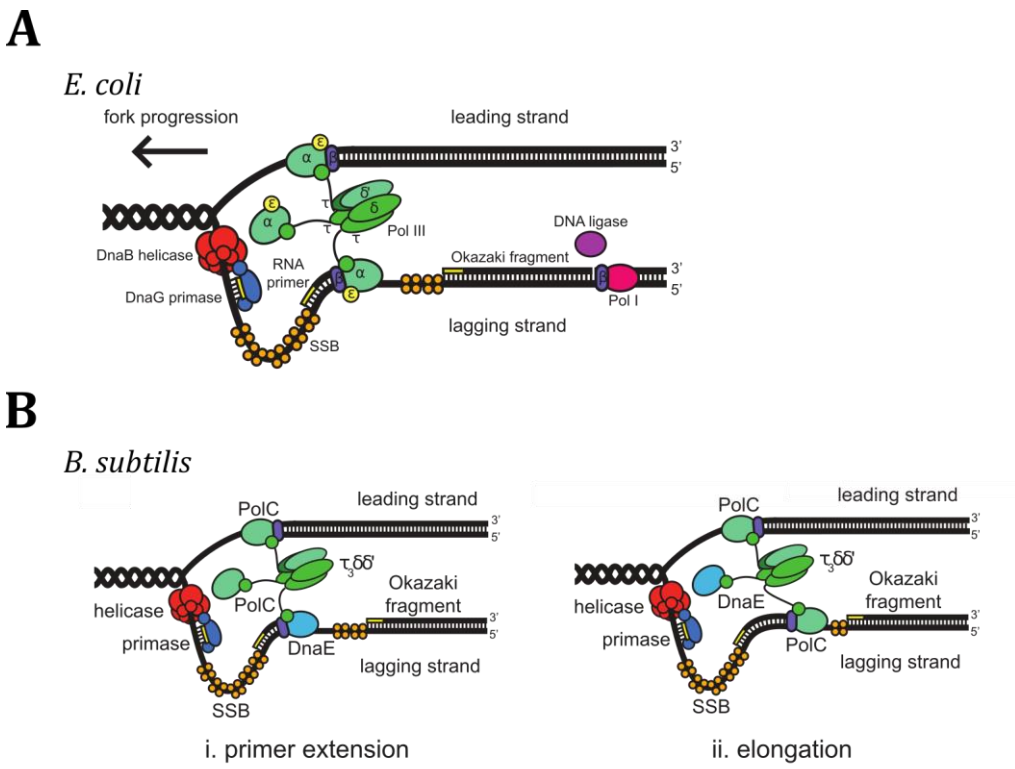


Figure 1.6 *The replisomal machinery of E. coli and B. subtilis.* ***A.*** *Schematic representation of the E. coli replisome showing lagging and leading strand synthesis at the replication fork.* ***B.*** *Schematic representation of the B. subtilis replisome showing polymerase handover during lagging strand synthesis. (i) DnaE extends RNA primers synthesised by DnaG (ii) PolC continues to extend the DNA primers synthesised by DnaE. Figure adapted from Robinson et al. 2012[7].*

The *E. coli* DNA polymerase (DNA pol III) extends the short RNA fragments, synthesised by the primase DnaG, with the assistance of the processivity clamp (DnaN/the β-clamp). DNA polymerase III consists of three subunits; the catalytic subunit α, ε providing the proofreading function, and θ thought to stimulate the activity of α. DNA polymerase III extends RNA fragments which are later replaced with DNA by DNA polymerase I. *B. subtilis* uses two distinct DNA polymerases, DnaE and PolC, these are homologous where PolC has undergone domain rearrangement and has Mg^{2+} dependent exonuclease activity. PolC is considered the main replicative polymerase for both the leading and lagging strand but is only able to extend DNA primers. While DnaE extends RNA primers, synthesised by DnaG, before handoff to PolC[42]. *E. coli* DNA pol III is related more closely to DnaE than to PolC[43]. The *B. subtilis* mechanism for lagging strand synthesis, whereby DnaE extends RNA primers, is analogous to the eukaryotic system during which DNA polymerase α extends RNA primers with DNA and subsequent DNA synthesis is continued by polymerase δ[7].

DnaN forms a closed ring structure (via a homodimer of C-shaped monomers) on the DNA directly behind the DNA polymerase. The interaction between DnaN and the dsDNA is minimised by binding across the major and minor grooves to permit 'sliding' along the DNA and subsequently allowing the polymerase to synthesise up to 1000 bases per second. The clamp loader is a pentameric complex ($\tau_3\delta\delta'$) essential as each new lagging strand requires loading of a β-clamp. The τ subunit is a product of the *DnaX* gene and forms an interaction with both the helicase and DNA polymerase and is thought to play an architectural role to couple DNA unwinding and DNA extension[7, 41].

1.3.3 Termination of DNA Replication

1.3.3.1 Termination in *Escherichia coli*

The termination of DNA replication occurs at a locus positioned directly opposite the *oriC*. *Ter* sites flank either side of the replication terminus, these are non-palindromic 23 bp sequences, termed *Ter*A-J, which are bound by the Tus (terminator utilisation substance) protein. The moving replication fork can pass these *Ter* sites from one direction only as determined by the orientation of the site. Therefore, DNA replication is halted when these sites are encountered in the non-permissive direction but the replication fork will bypass a *Ter* site when travelling in the permissive direction[44-46] as illustrated in Figure 1.7. The Tus protein is a 36 kDa monomeric protein that binds asymmetrically to the *Ter* sites[47]. In the permissive direction contact between *E. coli* DnaB helicase and Tus causes dissociation of the Tus protein from the *Ter* site. In contrast, contact between DnaB and Tus in the non-permissive direction prevents translocation of DnaB therefore halting the replication fork. This is a consequence of the asymmetric binding of Tus at the dsDNA in the non-permissive orientation where unwinding of the dsDNA by DnaB helicase causes a specific cytosine base within the Ter site to move into a binding site within Tus. This results in a locked Tus-*Ter* complex which prevents progression of the replication fork[48, 49].

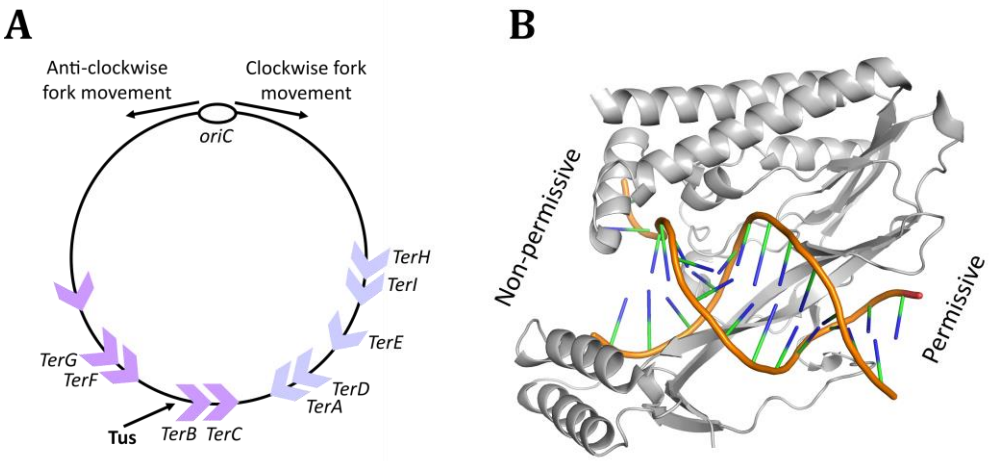


Figure 1.7 *Termination of DNA replication in E. coli.* ***A.*** *Location and orientation of the Ter sites within the E. coli chromosome, permissive direction indicated by arrow shape.* ***B.*** *Structure of the E. coli Tus-Ter complex showing the non-permissive face (left) and permissive face (right) [PDB - 2EWJ].*

1.3.3.2 Termination in *Bacillus subtilis*

For *B. subtilis*, replication termination occurs via a similar mechanism to *E. coli* whereby *Ter* sites are orientated to create a permissive and non-permissive direction. However, rather than the Tus protein binding at *Ter* sites, two RTP (replication termination proteins) bind within the *Ter* region at specific sites A and B[50, 51]. The moving replication fork is halted when approaching from the B site or non-permissive direction, but can continue passed the site when approaching from the A site or permissive direction (Figure 1.8). The RTP protein is a homodimer that displays asymmetry in the winged-helix domain, and thereby contacts with the dsDNA, where a 'wing up' and 'wing down' conformation is taken by the A and B sites respectively[52, 53]. The binding of the A site is cooperative following binding of the B site and is essential to halt the replication fork, yet, the molecular mechanism used by RTP to terminate DNA replication is unknown.

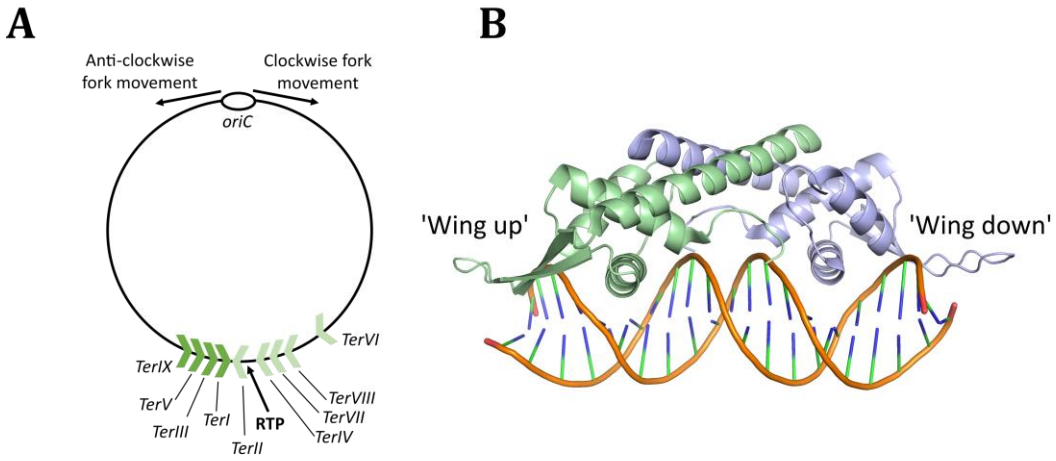


Figure 1.8 *Termination of DNA replication in B. subtilis.* ***A.*** *Location and orientation of the Ter sites within the B. subtilis chromosome, permissive direction indicated by arrow shape.* ***B.*** *Structure of the B. subtilis RTP dimer bound to a B-site region of dsDNA, The 'wing up' (left) and 'wing down' (right) conformations are displayed by individual monomers (highlighted green and blue) [PDB - 1F4K].*

1.4 Regulation of DNA Replication

DNA replication is tightly regulated to ensure that initiation occurs only once per cell cycle. DNA replication initiation is coupled to the cell cycle to ensure DNA replication coordinates with cellular growth[54]. Regulation of replication initiation involves a balance between activating initiation at a specific time in the cell cycle and repressing additional initiation events at all other times. Regulation of DNA replication occurs mainly during the initiation phase, prior to the recruitment of the replication machinery, with regulatory mechanism typically affecting the initiator protein DnaA and its ability to form a nucleoprotein complex at the origin of replication. It was previously thought that the cellular concentration of DnaA was a limiting factor for the rate of DNA replication[55]. However, studies in both *E. coli* and *B. subtilis* suggest that increasing the concentration alone cannot trigger DNA replication initiation[56]. Nutrient mediated growth rate regulation is dependent on both DnaA and *oriC* in *B. subtilis* with multiple regulatory mechanisms reported involving various systems for coupling growth rate with replication[57].

DnaA acts as a transcription regulator in *E. coli* and *B. subtilis* through the auto-regulation of the *dnaA* gene promoter whereby expression of DnaA is reduced when its

concentration rises above a threshold[58, 59]. DnaA has also been shown to regulate the expression of a number of genes within *E. coli* and *B. subtilis*. It is proposed that many DnaA-binding sites utilised for transcription regulation within *B. subtilis* are conserved in other organisms[58]. DnaA transcription mediated coupling of replication with the cell cycle has also been observed in *C. cresentus*[60, 61].

Though a variety of protein regulators and cis-acting DNA elements have been identified across various organisms, they are highly species specific where the same regulators are observed only in very closely related bacteria[62, 63]. Gram-negative and gram-positive bacteria employ separate mechanisms for regulation of replication initiation and the model organisms *E. coli* and *B. subtilis* have provided most of our current understanding[64].

1.5 Understanding the Mechanism of Replication Initiation in *Bacillus subtilis*

As described in section 1.3.1, DNA replication initiation in *B. subtilis* utilises the conserved initiator protein DnaA but also the essential initiator proteins DnaD and DnaB. The exact role of the proteins DnaD and DnaB during replication initiation remain unknown. DnaD was previously thought to inhibit the ATP-dependent binding of DnaA to *oriC* DNA[65, 66]. Suggesting that DnaD may negatively regulate *B. subtilis* DNA replication initiation by affecting the formation of the helical DnaA filament at the origin of replication. However, we hypothesise that DnaD plays a cooperative role during replication initiation where its DNA remodelling activity can aid DUE unwinding by DnaA. Knowledge of the molecular details of the interaction between DnaA and DnaD is important in understanding the mechanism of DNA replication initiation in *B. subtilis* and related gram-positive organisms.

1.5.1 Master Initiator Protein DnaA

DnaA is a replication initiator protein conserved across all bacteria. It belongs to the AAA+ family of proteins and has a multi-domain architecture composed of four distinct domains[67] (Figure 1.9). N-terminal domain I is a helicase interaction domain that

also promotes self-association of DnaA monomers. It has been deemed an 'interaction hub' as various regulatory proteins binding to domain I can stimulate or repress DnaA assembly at *oriC*[68]. Domain I has a KH-domain fold (K homology) typically found in ssDNA binding proteins[36, 69, 70] but no ssDNA binding role has been demonstrated for this domain *in vivo*. Domain II is a linker region, poorly conserved and of variable length across species. NMR studies, on *E. coli* DnaA, suggest that domain II is unstructured and acts as a flexible tether between domains I and III. It is likely that this 'spacer' domain is required for DnaA to adopt the correct conformation during initiation[71] and the size may be related to the spacing of DnaA-boxes within the origin[17].

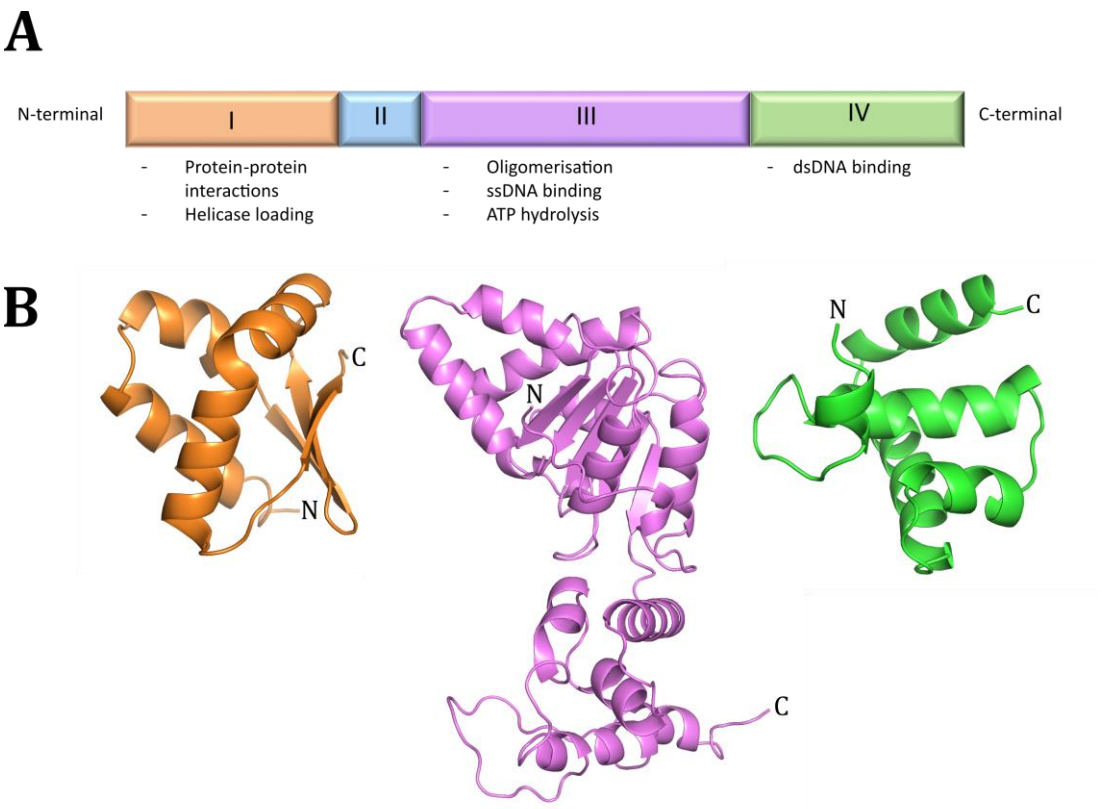


Figure 1.9 *Domain architecture of B. subtilis DnaA.* ***A.*** *Structure and function of DnaA domains I-IV.* ***B.*** *Left - The X-ray crystal structure of DnaA domain I has a KH-fold similar to that found in ssDNA binding proteins [PDB - 4TPS]. Centre – The AAA+ domain of DnaA contains Walker A and B type ATPase motifs [A. aeolicus DnaA domain III PDB - 1L8Q]. Right – Domain IV of DnaA confers its DNA binding activity through a basic loop that interacts with the minor groove and a helix-turn-helix that inserts into an adjacent major groove in the DNA [E. coli DnaA domain IV PDB – 1J1V].*

Domain III has Walker A and B type ATPase motifs, capable of ATP binding and hydrolysis activity, and confers the main oligomerisation activity for DnaA filament formation. It comprises two subdomains, structurally distinct, termed IIIa and IIIb where the ATP binding site is located at the interface. Domain IIIa has an αβ structure whereas domain IIIb has a helical structure and both contribute to nucleotide binding, IIIa via Walker A, Walker B and sensor I sequence motifs and IIIb by the sensor II motif[72, 73]. Domain III also plays a role in ssDNA binding[74, 75]. Domain IV is involved in specific double stranded DNA binding at the DnaA-boxes within *oriC*[76]. Essential for its activity are two motifs; a basic loop that interacts with the minor groove of DNA through an arginine residue, and a helix-turn-helix that inserts into an adjacent major groove in the DNA. Together these result in an approximately 40° bend in the DNA, as illustrated in figure 1.10.

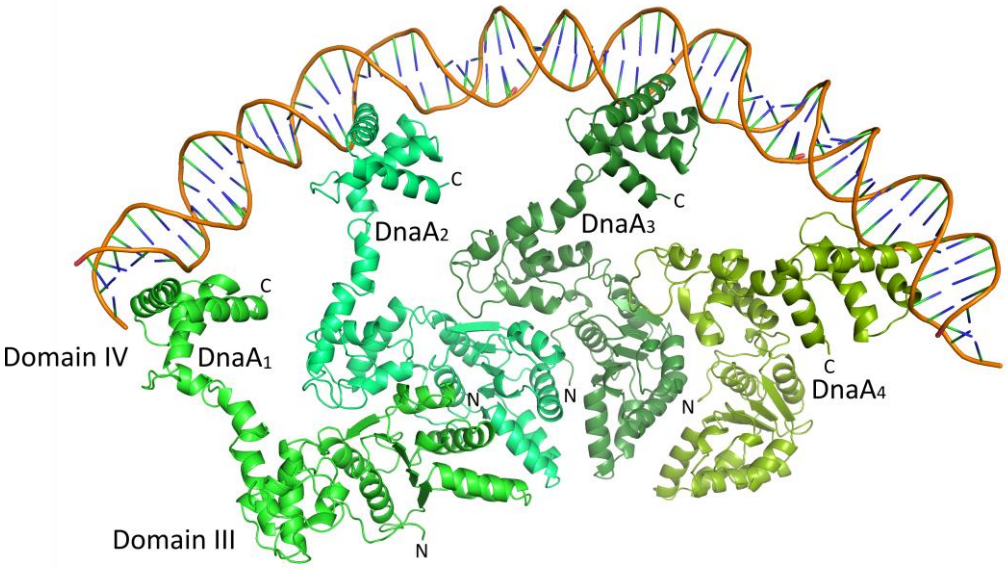


Figure 1.10 *Model for a B. subtilis DnaA tetramer bound to dsDNA. Domains III and IV of DnaA are shown where individual DnaA monomers are represented by different colours. Oligomerisation occurs via domain III interactions. DnaA-box sites within the oriC are bound specifically by domain IV. Model based on X-ray crystal structures of E. coli DnaA domain IV complexed to DnaA-box DNA [PDB - 1J1V] and A. aeolicus DnaA domain III [PDB - 1L8Q].*

Oligomerisation of DnaA molecules occurs through domain III in an ATP dependent manner. Binding and hydrolysis of ATP mediates a molecular switch that regulates DnaA filament formation. When ATP is bound to DnaA, an arginine of the

sensor II helix from subdomain IIIb is engaged by the γ-phosphate of ATP. This gives rise to an active site conformation that allows a second arginine of the box VII helix from subdomain IIIa of an adjacent DnaA molecule to interact in *trans* with the same γ-phosphate in the active site of the first DnaA molecule[23] (as illustrated in Figure 1.11). Dimer formation is facilitated and using the same conformation at the active site multiple DnaA molecules can oligomerise through sensor II box VII interactions. When ADP is bound at the active site of DnaA, the sensor II helix is no longer engaged with the γ-phosphate and this conformation prevents dimerisation due to steric clash between the sensor II arginine and box VII helix of the second DnaA molecule. ATP therefore acts as a molecular switch between the closed state and open conformation of the DnaA active site, permitting oligomerisation. DnaA is bound at the high affinity DnaA-boxes throughout the cell cycle. During replication initiation ATP-DnaA molecules bind to their respective DnaA-box clusters and oligomerise to form a nucleoprotein complex that facilitates melting of the DUE.

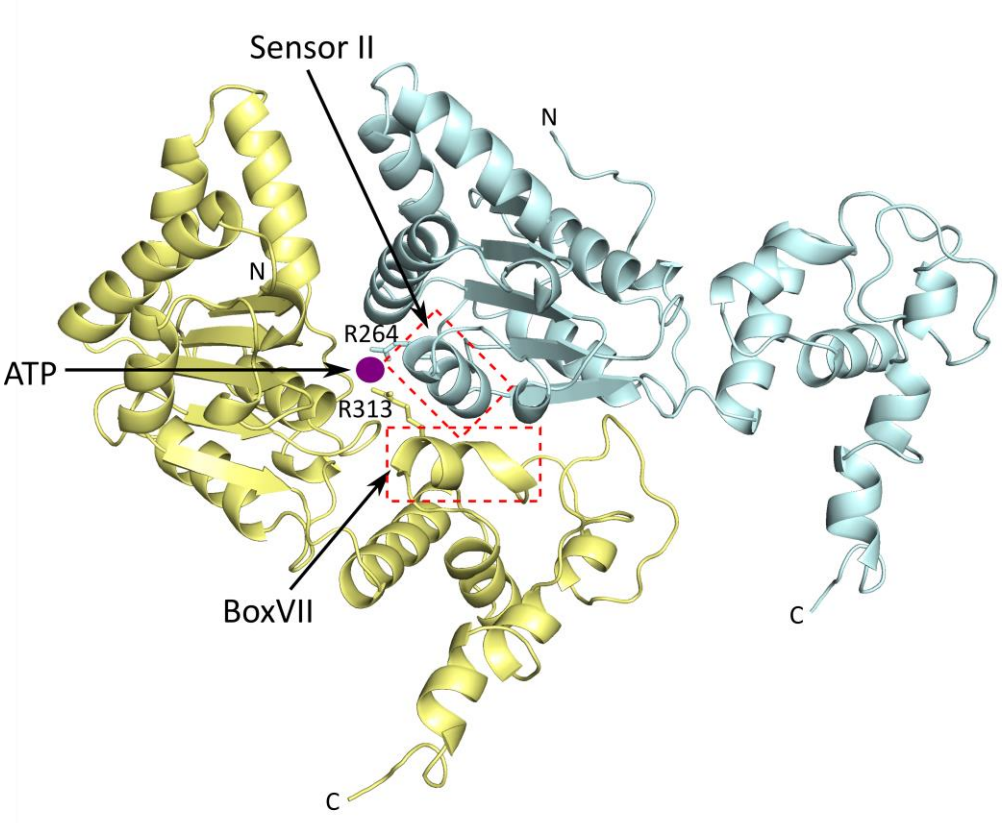


Figure 1.11 *Oligomerization of DnaA domain III through sensor II - box VII interaction. The cyan and yellow subunits represent single monomers of DnaA. Box VII and sensor II arginine residues are displayed as sticks, ATP is represented by a purple dot. A sensor II arginine of the first DnaA monomer interacts in cis with ATP bound in its active site. This conformation allows an arginine from box VII of the second DnaA monomer to interact in trans with the same ATP molecule. Thus, oligomerization is facilitated by ATP.*

The intrinsic ATPase activity of the AAA+ domain is low and following replisome formation the ATP hydrolysis activity of DnaA is stimulated, leading to its inactivation, and disassembly of the DnaA nucleoprotein complex is promoted by the ADP bound form of DnaA. ATP hydrolysis is promoted by Hda during regulatory inactivation of DnaA (RIDA) in *E. coli*[77-79]. Hda is homologous to DnaA domain III[78] and the interaction with DnaA occurs via the ATPase domain[80, 81] in a similar manner to the formation of DnaA oligomers. Hda also binds to the DNA-bound β-clamp, DnaN, and the DnaA-Hda-DnaN-DNA complex is required to mediate hydrolysis of DnaA-ATP to DnaA-ADP. Avidity for ssDNA has also been linked to the ATPase status of DnaA, ADP-DnaA has been reported to bind ssDNA with an approximately 10-fold weaker affinity than ATP-DnaA[82]. YabA is the principle regulator of replication initiation in *B. subtilis*[83]. YabA can interact with DnaA and the polymerase clamp DnaN[84, 85] reminiscent of *E. coli* Hda. However, YabA does not promote DnaA-ATP hydrolysis *in vitro*[78] and thus employs an alternative regulatory mechanism. YabA is proposed to prevent the co-operative binding of DnaA at *oriC* until a high concentration of free DnaN is reached in the cell and thus removes YabA from DnaA to enable DnaA binding at *oriC*. This model is consistent with the observation that YabA is able to disrupt DnaA-oligomer formation on *DNA in vitro*[66]. Inactivation of DnaA after replication initiation is essential to ensure the priming event occurs only once per cell cycle for accurate genome duplication.

DnaA filament formation has a critical role in DNA replication initiation. During filament assembly, insertion of the box VII helix into domain III of the adjacent DnaA molecule causes a right handed helical filament conformation to be adopted due to protrusion of the helical insertion away from the core[23, 86]. The double stranded DNA is located around the outside of the helical filament through DnaA domain IV interaction. This arrangement causes positive toroidal wrap of the DNA double helix that is compensated by inducing negative supercoiling ahead of the DNA-DnaA filament. The DNA double helix is then destabilised facilitating melting at the AT rich DUE within the *oriC*. DnaA may also enhance direct melting of the DUE by single stranded DNA binding. DnaA filaments are dynamic and on melting of the *oriC*, conformational change allows DnaA to invade ssDNA to further stabilise DUE melting. DNA binds non-specifically through domain III helices α3, α4, α5 and α6 along the inside of the DnaA filament (Figure 1.12). Each AAA+ domain binds a triplet of nucleotides separated from the next triplet by a 10 Å gap[75]. This extension relative to regular B-form DNA stabilises the unwound state of the origin, enhancing melting at

the DUE. DnaA-mediated unwinding of the origin is a crucial step in initiation that allows loading of the hexameric helicase DnaC onto ssDNA.

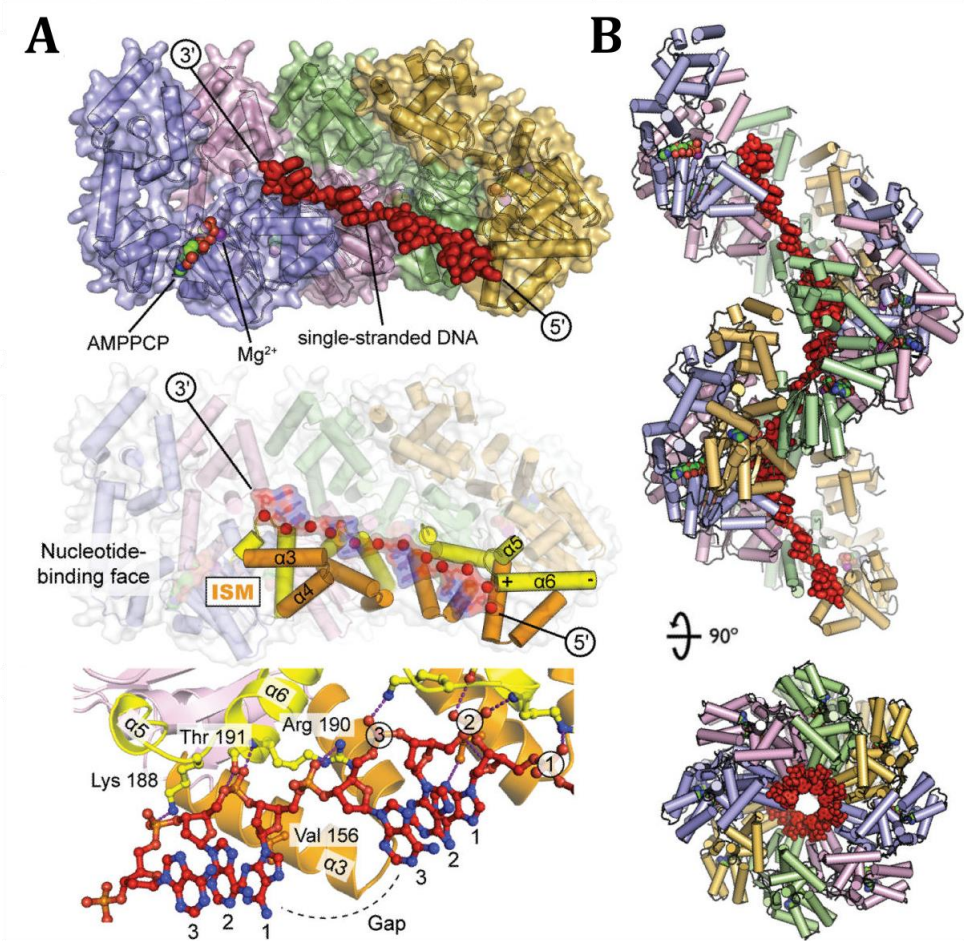


Figure 1.12 *Binding of ssDNA through domain III of DnaA. Model based on X-ray crystal structure of A. aeolicus DnaA bound to AMPPCP and ssDNA [PDB - 3R8F].* ***A.*** *Side view of the asymmetric unit. Top - DnaA subunits are coloured individually, ssDNA is displayed in red and AMPCC and Mg^{2+} are displayed as spheres coloured by element. Middle – Helices α3/α4 and α5/α6 are highlighted in orange and yellow respectively. ISM represents initiator specific motif. Bottom - Protein-DNA contacts are shown for DnaA subunits B and C.* ***B.*** *Side and top view of a 12 subunit DnaA oligomer bound to three strands of ssDNA, model reconstituted through crystal packing. Figure from Duderstadt et al. 2012[75].*

1.5.2 Initiator Proteins DnaD and DnaB

Along with the conserved protein DnaA, *B. subtilis* also has essential genes encoding initiator proteins DnaD and DnaB. Both proteins are required at the origin, recruited in a hierarchical manner, for loading of the replicative helicase DnaC onto ssDNA. DnaD is recruited to the *oriC* via a direct interaction with DnaA and is a key regulator of replication initiation[87]. DnaD is required for the recruitment of DnaB which functions as a co-loader, along with DnaI, of the replicative helicase DnaC. DnaD and DnaB share structural similarity and both exhibit DNA remodelling activity[40, 88] they are also components of the DnaA-independent replication restart machinery[89]. The exact roles of DnaD and DnaB during DNA replication initiation remain unclear. Although there is little sequence similarity between DnaD and DnaB, both have a modular architecture composed of two shared domains termed DDBH1 and DDBH2 (DnaD DnaB Homology 1 and 2). DnaD is composed of an N-terminal DDBH1 domain and a C-terminal DDBH2 domain, whilst DnaB is composed of a DDBH1 domain and two DDBH2 domains[90, 91], as illustrated in Figure 1.13. DNA binding activity occurs through the DDBH2 domain[92, 93] and the DDBH1 domain is responsible for DNA-independent oligomerisation. Both proteins are active as homo-tetramers and form high order oligomeric complexes that can remodel DNA.

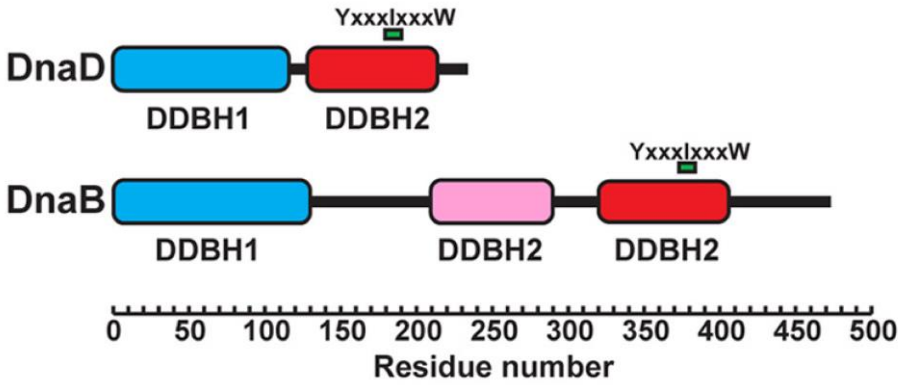


Figure 1.13 *Architecture of B. subtilis DnaD and DnaB. Structural homology is shown through the shared DDBH1 and DDBH2 domains, the conserved YxxxIxxxW motif is also highlighted. Figure adapted from Briggs et al. 2012[4].*

DnaB is thought to act together with DnaI to enable loading of DnaC helicase during replication initiation. DnaI alone has been shown to be sufficient to load DnaC onto DNA[94], however, DnaB is required for recruitment of the DnaC-DnaI complex to the *oriC*[88] and has been shown to stimulate the helicase and translocase activity of DnaC in the presence of DnaI[40]. DnaB may also direct association of the DNA replication machinery with the cell membrane[95, 96].

DnaD is a non-specific DNA binding protein that binds both doubled stranded and single stranded DNA, with a higher affinity to the latter[97]. AFM has shown that DnaD forms large scaffolds when bound to DNA[98, 99]. These are formed through DDBH1 domain interaction and reinforce duplex untwisting by the DDBH2 domain. DnaD remodels supercoiled DNA by eliminating writhe through negative twist. The DNA remodelling activity of DnaD occurs through both scaffold-forming activity in the N-terminal domain and DNA-binding and oligomerisation activity in the C-terminal domain[92, 100]. X-ray crystallography of the N-terminal domain of *B. subtilis* DnaD revealed a winged helix-turn-helix structure[84],[101] (Figure 1.14). This motif is common for proteins with DNA-binding activity, however, in the case of DnaD, the N-terminal domain facilitates the formation of extensive protein scaffolds and the C-terminal domain is involved in DNA interaction[102]. Also within the DDBH1 domain of DnaD is an N-terminal helix-strand-helix and single helix at the C-terminal[92]. Dimer and tetramer formation occurs through the helix-strand-helix motif and the winged-helix fold is involved in higher order oligomerisation forming large DnaD scaffolds in *B. subtilis*. DnaD and DnaB share a highly conserved motif (YxxxIxxxW), highlighted in Figure 1.14. The structure of the C-terminal domain of DnaD was modelled by Jeremy Craven (University of Sheffield) using coordinates from NMR spectroscopy and the crystal structure of the DnaD-like replication protein from *Streptococcus mutans*[90, 100] and revealed that the conserved YxxxIxxxW motif forms an exposed patch on the surface of the protein. This motif, the C-terminal helix and a region of unstructured C-terminal tail (residues 206-215) of DnaD contribute to DNA binding and are essential for *B. subtilis* viability[90].

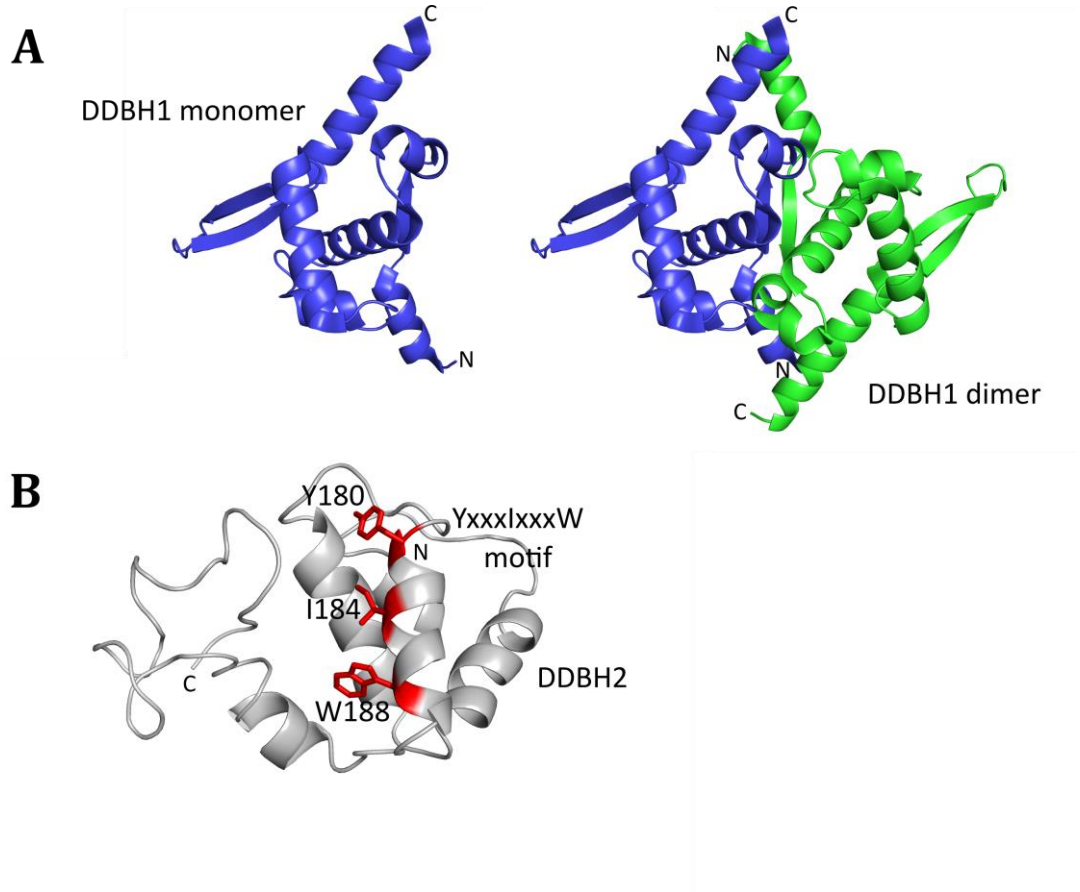


Figure 1.14 *Modular architecture of B. subtilis DnaD.* ***A.*** *The structure of the DDBH1 domain determined by X-ray crystallography [PDB - 2V79]. Left shows the monomer and right the dimeric conformation.* ***B.*** *The modelled structure of the DDBH2 domain with the conserved YxxxIxxxW motif residues highlighted in red (structure supplied by Dr Jeremy Craven, University of Sheffield, based on NMR coordinates and the crystal structure of the DnaD-like replication protein from Streptococcus mutans PDB – 2ZC2).*

The predicted model of a DnaD tetramer (Figure 1.15) resembles the structure of *Helicobacter pylori* (*H. pylori*) initiator protein HobA bound to DnaA[69]. Interaction between HobA and DnaA is mediated through domain I of DnaA. It has been suggested that the HobA-DnaA complex forms a super-structural scaffold at the *oriC* through DNA-DnaA-HobA interaction at multiple DnaA-box sites. We propose that DnaD interacts with DnaA at the *oriC* in a similar manner. DnaA forms a large nucleoprotein complex at the origin to initiate DNA replication. Electron microscopy has shown that when *B. subtilis* DnaA is bound to DnaA-box clusters both upstream and downstream of the *dnaA* gene, the two regions of the *oriC* interacted causing a loop in the DNA[13]. The looping of the DNA may be mediated by DnaD scaffold formation acting to bridge DnaA molecules at the DnaA-box clusters of the split origin (Figure 1.16). This looped structure may be regulatory to replication initiation or involved in loading of the two

replisomes[4]. The reason behind the split origin is yet to be understood, it may be regulatory or involved in loading of the two replisomes on the leading and lagging DNA stands.

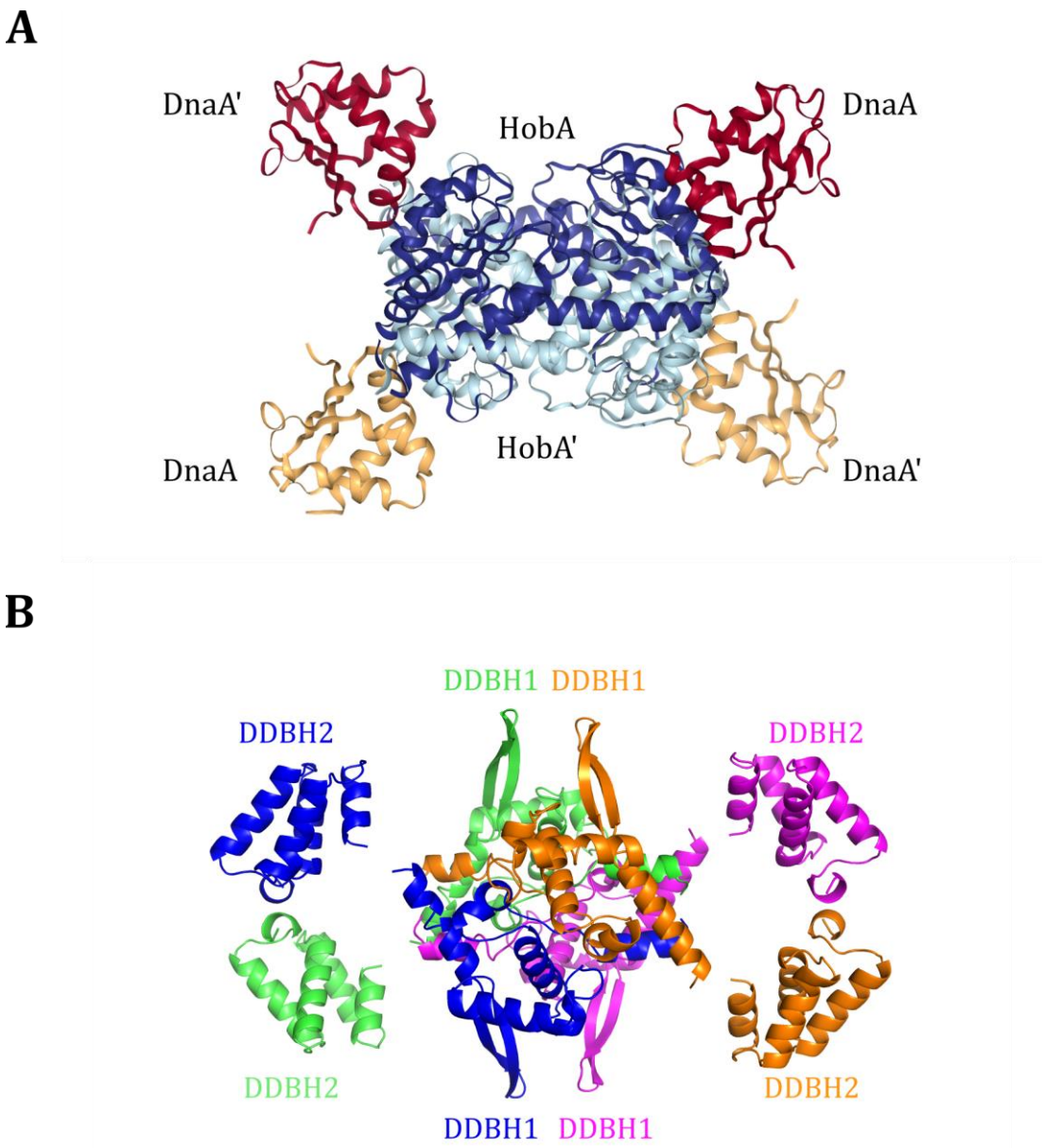


Figure 1.15 *Comparison of the H. pylori HobA-DnaA complex and the modelled B. subtilis DnaD tetramer.* ***A.*** *X-ray crystal structure of the H. pylori HobA bound to DnaA [PBD - 2WPO]. The HobA dimer is shown binding to four molecules of DnaA.* ***B.*** *Model of the B. subtilis DnaD tetramer predicted based on crystal lattice packing (Richard Rymer Soultanas group), individual DnaD monomers are identified by colour. Model based on the DDBH2 domain structure of E. faecalis [PDB - 2I5U].*

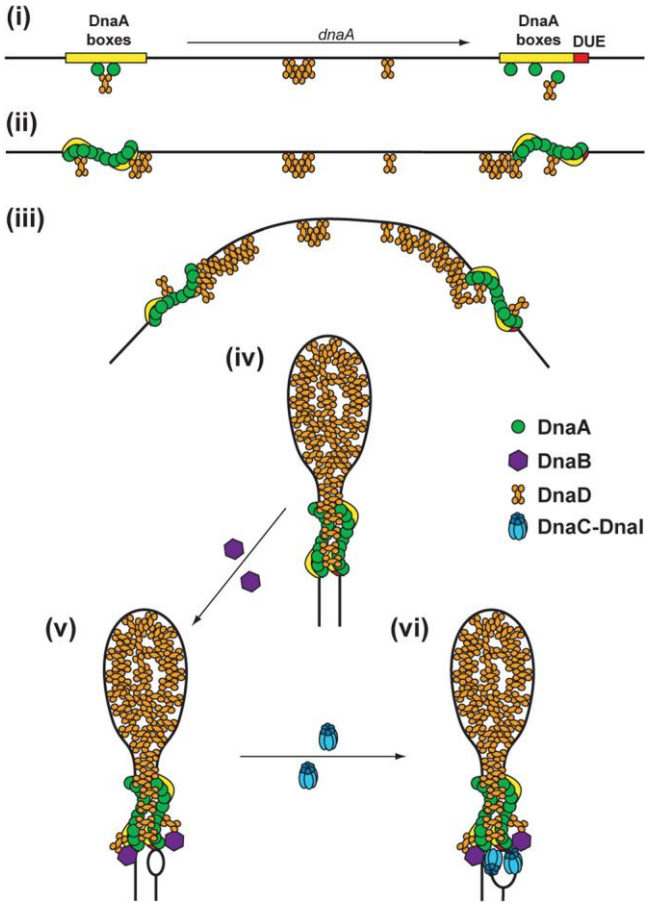


Figure 1.16 *Speculative model for the role of DnaD during replication initiation in B. subtilis. DnaA binds to the oriC at multiple DnaA-box sites forming nucleoprotein filaments. DnaD bound to the DnaA filament may recruit additional DnaD to the adjacent DNA region. DNA bending, induced by DnaD invasion, enables interaction between the DnaA filament complexes either side of the dnaA gene. DnaD scaffold formation stabilises looped DNA. The DnaA complex induces unwinding of the DUE where DnaD mediated DNA remodelling may facilitate DUE unwinding. DnaB then recruits the DnaC-DnaI complex to the ssDNA. Figure from Briggs et al. 2012[4].*

The requirement of the two additional proteins, DnaD and DnaB, for DNA replication initiation in *B. subtilis* provide additional opportunities for the regulation of replication initiation. The abundance of DnaD molecules *in vivo* (3000-5000 molecules per cell)[103] suggests it has additional roles outside of DNA replication, perhaps a global role in DNA remodelling. DNA remodelling by DnaD has been shown to stimulate DNA repair by Nth endonuclease in response to DNA damage by H_2O_2[104] and there is possibly a link between DnaD and DNA repair across all species[105]. DnaD and DnaB are also involved in re-initiation of replication at sites other than *oriC* as part of the PriA mediated primosome[89]. PriA promotes replication restart by the assembly of a single replisome on arrested and repaired chromosomal forks[97].

1.5.3 The DnaA-DnaD Interaction

The direct interaction between DnaA and DnaD was confirmed by yeast two-hybrid studies and residues were identified, in the C-terminal (DDBH2 domain) of DnaD, essential for the interaction with DnaA[106]. These included the region $E_{134}EEFARP_{140}$ and have been deemed the DnaA interaction patch (AIP). The residues form a distinct surface exposed loop on one face of DnaD, as highlighted in Figure 1.17. This interaction patch appears on the opposite face of DnaD to that involved in DNA binding suggesting DnaD interacts with both DNA and DnaA simultaneously through its DDBH2 domain. Domain I of DnaA is involved in helicase binding along with the binding of regulatory proteins, such as SirA[107], and it is likely that the interaction with DnaD also occurs through this domain. We propose the DNA unwinding activity of DnaD enhances DUE melting during replication initiation, by DnaA, through the *oriC*-DnaA-DnaD complex[4] and identification of the specific residues involved in the binding interface are key to understanding the molecular mechanism of DNA replication initiation.

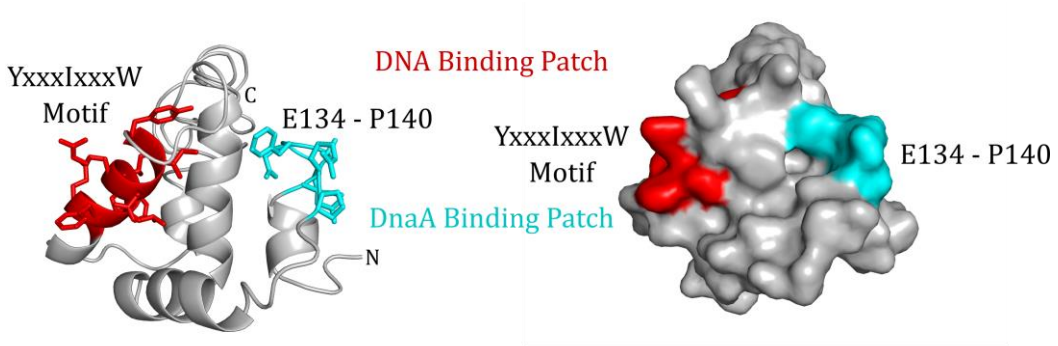


Figure 1.17 *The DnaA binding patch within B. subtilis DnaD DDBH2 domain. Cartoon (left) and surface (right) representations of the DnaD DDBH2 domain. The region E134 – P140 (highlighted cyan) within the DDBH2 domain identified as crucial for the interaction with DnaA by Yeast-two-hybrid studies[106]. The DNA binding patch and the conserved YxxxIxxxW motif for DNA binding (highlighted red). The structure of the B. subtilis DDBH2 domain was modelled by Jeremy Craven (University fo Sheffield) using NMR coordinates and the crystal structure of the DnaD-like replication protein from Streptococcus mutans PDB – 2ZC2).*

1.6 Research Aims

The aim of this research was to improve our understanding of the *Bacillus subtilis* mechanism of DNA replication initiation. Specifically, to characterise the interaction between essential initiator proteins DnaA and DnaD using structural and biophysical methods. Solution state NMR is a well-established technique used for structure determination of proteins and to study dynamics and binding properties of biological molecules. As mentioned in section 1.5.3, the interaction between DnaA and DnaD is proposed to occur through DnaA domain I and the DnaD DDBH2 domain, these single domains were identified as suitable candidates for NMR analysis. The mechanism of the interaction between DnaA and DnaD remains unknown, along with their combined role during DNA replication initiation to locally unwind duplex DNA at the origin of replication. The biophysical characterisation of the interaction aimed to initially detect the interaction between the isolated domains of DnaA and DnaD, followed by identification of the key residues involved in the interaction sites on both proteins. Greater understanding of the physiological interplay between DnaA and DnaD during DNA replication initiation could highlight these initiator proteins as targets for antimicrobial therapies.

Chapter 2

Materials and Methods

2.1 Chemicals

All chemicals were purchased from Sigma-Aldrich or Fisher Scientific unless stated otherwise.

2.2 Biological Materials

B. subtilis DnaA in pET-28b and DnaA-NTD 1-105 in pET-28b were supplied by Geoff Briggs (Soultanas Group, University of Nottingham). *B. subtilis* DnaD in pET-22b was supplied by William Grainger (Soultanas Group, University of Nottingham). DnaD-CTD 129-232 in pET-28b was supplied by Farhat Marston (Craven Group, Univeristy of Sheffield).

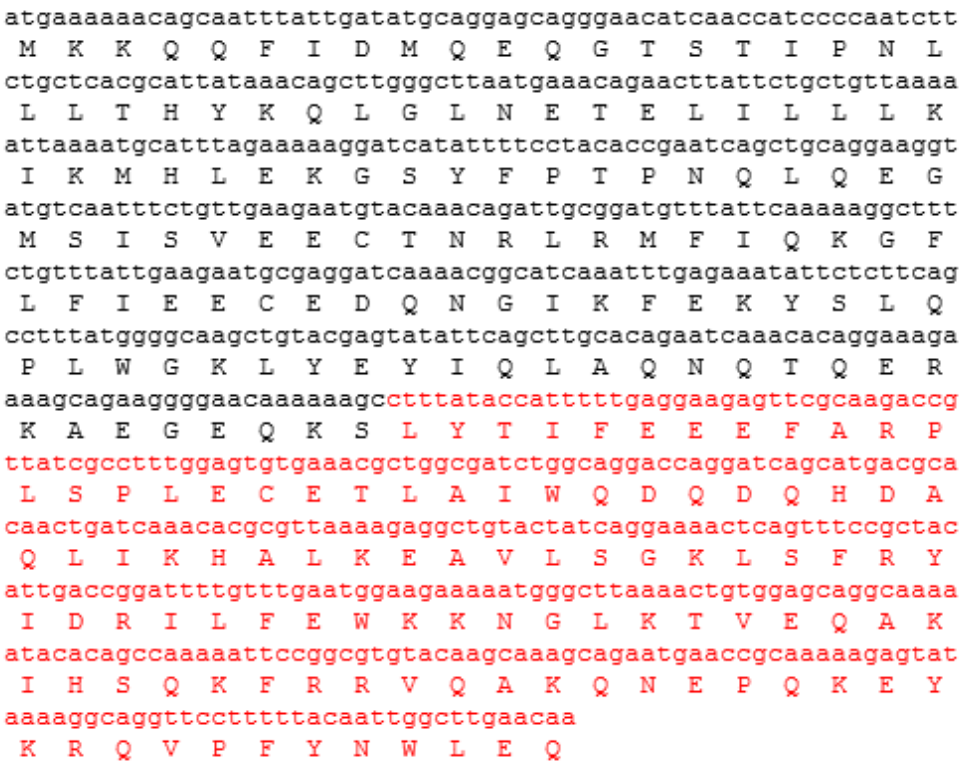


Figure 2.1 *DNA and protein sequence for B. subtilis DnaD. Residues 129-232, highlighted red, represent the C-terminal domain or DDBH2 domain.*

```
atggaaaatatattagacctgtggaaccaagcccttgctcaaatcgaaaaaaagttgagc
 M  E  N  I  L  D  L  W  N  Q  A  L  A  Q  I  E  K  K  L  S
aaaccgagttttgagacttggatgaagtcaaccaaagcccactcactgcaaggcgataca
 K  P  S  F  E  T  W  M  K  S  T  K  A  H  S  L  Q  G  D  T
ttaacaatcacggctcccaatgaatttgccagagactggctggagtccagatacttgcat
 L  T  I  T  A  P  N  E  F  A  R  D  W  L  E  S  R  Y  L  H
ctgattgcagatactatatatgaattaaccggggaagaattgagcattaagtttgtcatt
 L  I  A  D  T  I  Y  E  L  T  G  E  E  L  S  I  K  F  V  I
cctcaaaatcaagatgttgaggactttatgccgaaaccgcaagtcaaaaaagcggtcaaa
 P  Q  N  Q  D  V  E  D  F  M  P  K  P  Q  V  K  K  A  V  K
gaagatacatctgattttcctcaaaatatgctcaatccaaaatatacttttgatactttt
 E  D  T  S  D  F  P  Q  N  M  L  N  P  K  Y  T  F  D  T  F
gtcatcggatctggaaaccgatttgcacatgctgcttccctcgcagtagcggaagcgccc
 V  I  G  S  G  N  R  F  A  H  A  A  S  L  A  V  A  E  A  P
gcgaaagcttacaacccttttatttatctatgggggcgtcggcttagggaaaacacactta
 A  K  A  Y  N  P  L  F  I  Y  G  G  V  G  L  G  K  T  H  L
atgcatgcgatcggccattatgtaatagatcataatccttctgccaaagtggtttatctg
 M  H  A  I  G  H  Y  V  I  D  H  N  P  S  A  K  V  V  Y  L
tcttctgagaaatttacaaacgaattcatcaactctatccgagataataaagccgtcgac
 S  S  E  K  F  T  N  E  F  I  N  S  I  R  D  N  K  A  V  D
ttccgcaatcgctatcgaaatgttgatgtgcttttgatagatgatattcaatttttagcg
 F  R  N  R  Y  R  N  V  D  V  L  L  I  D  D  I  Q  F  L  A
gggaaagaacaaacccaggaagaatttttccatacatttaacacattacacgaagaaagc
 G  K  E  Q  T  Q  E  E  F  F  H  T  F  N  T  L  H  E  E  S
aaacaaatcgtcatttcaagtgaccggccgccaaaggaaattccgacacttgaagacaga
 K  Q  I  V  I  S  S  D  R  P  P  K  E  I  P  T  L  E  D  R
ttgcgctcacgttttgaatggggacttattacagatatcacaccgcctgatctagaaacg
 L  R  S  R  F  E  W  G  L  I  T  D  I  T  P  P  D  L  E  T
agaattgcaatttttaagaaaaaaggccaaagcagagggcctcgatattccgaacgaggtt
 R  I  A  I  L  R  K  K  A  K  A  E  G  L  D  I  P  N  E  V
atgctttacatcgcgaatcaaatcgacagcaatattcgggaactcgaaggagcattaatc
 M  L  Y  I  A  N  Q  I  D  S  N  I  R  E  L  E  G  A  L  I
agagttgtcgcttattcatctttaattaataaagatattaatgctgatctggccgctgag
 R  V  V  A  Y  S  S  L  I  N  K  D  I  N  A  D  L  A  A  E
gcgttgaaagatattattccttcctcaaaaccgaaagtcattacgataaaagaaattcag
 A  L  K  D  I  I  P  S  S  K  P  K  V  I  T  I  K  E  I  Q
agggtagtaggccagcaatttaatattaaactcgaggatttcaaagcaaaaaaacggaca
 R  V  V  G  Q  Q  F  N  I  K  L  E  D  F  K  A  K  K  R  T
aagtcagtagcttttccgcgtcaaatcgccatgtacttatcaagggaaatgactgattcc
 K  S  V  A  F  P  R  Q  I  A  M  Y  L  S  R  E  M  T  D  S
tctcttcctaaaatcggtgaagagtttggaggacgtgatcatacgaccgttattcatgcg
 S  L  P  K  I  G  E  E  F  G  G  R  D  H  T  T  V  I  H  A
catgaaaaaatttcaaaactgctggcagatgatgaacagcttcagcagcatgtaaaagaa
 H  E  K  I  S  K  L  L  A  D  D  E  Q  L  Q  Q  H  V  K  E
attaaagaacagcttaaa
 I  K  E  Q  L  K
```

Figure 2.2 *DNA and protein sequence for B. subtilis DnaA. Residues 1-81, highlighted blue, represent the N-terminal domain or domain I.*

2.2.1 *Escherichia coli* cells

E. coli Strain	Supplier
XL-1 Blue	Stratagene
BL21 (DE3)	Sigma-Aldrich
NEB 5-alpha	New England Biolabs

Table 2.1 *E. coli cell strains used and supplier of the original stock.*

2.2.2 Antibiotics

Antibiotic	Working Concentration	Supplier
Carbenicillin	50 µg/mL	Melford
Kanamycin	30 µg/mL	Life Technologies

Table 2.2 *Antibiotics used, their working concentration and supplier.*

2.2.3 Plasmid Vectors

Vector	Antibiotic Resistance	Supplier
pET-22b	Ampicillin	Novagen
pET-28b	Kanamycin	Novagen

Table 2.3 *Plasmid vectors used, the antibiotic resistance incorporated and vector supplier.*

2.2.4 Enzymes

All enzymes were purchased from NEB with the exceptions of Thrombin supplied by Enzyme Research Laboratories and Benzonase supplied by Novagen.

2.2.5 Oligonucleotides

Unmodified DNA oligonucleotides were supplied by either Eurofins or Sigma-Aldrich as high purity salt free lyophilised samples. Oligonucleotides, for use as PCR primers, were resuspended to 100 µM with sterile Milli-Q water and stock solutions of 10 µM were prepared for use in PCR reactions. Oligonucleotide solutions were stored at -20 °C.

2.3 Biochemical Methods

2.3.1 Molecular Cloning: Truncations using High Fidelity Q5 Polymerase

PCR mutagenesis was used to construct single domain constructs and truncated versions of the *B. subtilis* DnaA and DnaD proteins. PCR reactions were run in a thermocycler (Labnet MultiGene Mini) using the conditions described in tables 2.4 and 2.5. Template DNA and oligonucleotide primers used to incorporate mutations during PCR amplification are defined in tables 2.6 and 2.7.

Reaction Component	Final Concentration
Q5 reaction buffer	1 x
High GC enhancer	1 x
dNTP solution	200 µM
Forward primer	0.5 µM
Reverse primer	0.5 µM
Template DNA	< 1,000 ng
Q5 polymerase	0.02 U/µL

Table 2.4 *Reaction components used during PCR amplification of DNA. Q5 polymerase, Q5 reaction buffer, high GC enhancer and dNTP solutions were purchased from NEB. The conditions used for the PCR reactions were based on the NEB protocol for Q5 polymerase. DNA templates used are described for individual PCR reactions and oligonucleotide primers are defined in table 2.6.*

Step	Temperature (°C)	Time	Number of Cycles
Initial strand separation	98	2 min	1
Strand separation	98	15 sec	
Primer annealing	-	30 sec	30
Strand extension	72	30 sec / kb	
Final extension	72	5 min	1

Table 2.5 *Thermocyclic conditions used for PCR amplification of DNA. The primer annealing temperature and strand extension times were variable and are described for individual PCR reactions. The primer annealing temperature used was dependent on the specific oligonucleotide primers required for each reaction. The strand extension time was calculated by the length of DNA to be amplified and the activity of Q5 polymerase (1 unit per 50 µL reaction can process 1 kb per 20-30 seconds depending on complexity).*

Construct	Template DNA
DnaD 129-196	pET28a DnaD 129-232
DnaA 1-81	pET28b DnaA^{CC}-his
DnaD 1-196-his	pET22b DnaD
DnaA 1-105	pET28b DnaA^{CC}-his

Table 2.6 *Plasmid DNA templates used for PCR mutagenesis of B. subtilis DnaA and DnaD constructs.*

Construct	Primer	Sequence 5'–3'	Ta (°C)
DnaD 129-196	Forward	CGCGGCAGC**catatg**CTTTATACCATTTTTGAG	72
	Reverse	TGCCTGCTC**aagcttta**CACAGTTTTAAGCCC	
DnaA 1-81	Forward	GATATACC**catatg**GAAAATATATTAGACCTG	63
	Reverse	GGCCGC**aagctttta**AGGAATGACAAACTTAATGC	
DnaD 1-196-his	Forward	GCGGCCGCACTCGAGCAC	71
	Reverse	CACAGTTTTAAGCTTATTTTTCTTCCATTCAAACAAAATCC	
DnaA 1-105	Forward	GATATACC**catatg**GAAAATATATTACTG	59
	Reverse	GAGTGCGGCCGC**aagctttta**ATCAGATGTATC	

Table 2.7 *DNA oligonucleotide sequences used for PCR mutagenesis of B. subtilis DnaA and DnaD constructs. Insertion mutations are indicated in lowercase bold. Ta values shown correspond to the specific primer annealing temperature used during the PCR reactions.*

With the DnaD 1-196-his construct, the entire template plasmid was amplified by PCR with a deletion of the region immediately after V196 up to and including the HindIII site of the pET22b vector. The plasmid could then be re-ligated to incorporate the deletion. For all other constructs, the desired gene fragment (rather than the full plasmid) was amplified by PCR. Restriction sites were incorporated at the 5' and 3' ends of the gene that could then be used to insert the desired gene fragment into the appropriate pET vector (section 2.2.3).

2.3.1.1 Restriction Digestion

PCR products were purified as described in section 2.3.3. Purified DNA (1 µg) was incubated with the appropriate restriction endonuclease (10 units, table 2.8) at 37°C for 1 hour followed by heat inactivation at 80°C for 20 minutes. To prevent self-ligation the vector DNA was treated with Antarctic phosphatase (5 units) incubated at 37°C for 1 hour and heat-inactivated at 70°C for 5 minutes. Vector DNA was then digested under the same conditions as the DNA fragment to be inserted.

Construct	Plasmid Vector	Restriction Enzyme
DnaD 129-196	pET28b	NdeI, HindIII
DnaA 1-81	pET28b	NdeI, HindIII
DnaD 1-196-his	-	DpnI
DnaA 1-105	pET28b	NdeI, HindIII

Table 2.18 *Enzymes used for restriction endonuclease digestion of PCR products and the corresponding plasmid vector DNA. Reactions of 50 µL were undertaken according to the NEB protocol specific to the enzyme required. For the DnaD 1-196-his construct, the full template plasmid was amplified during PCR therefore did not require insertion into a new vector and in this case the PCR product was digested with DpnI to remove any template DNA prior to re-ligation.*

2.3.1.2 Gel Extraction

Digestion products were run on a 0.8% (w/v) agarose gel (section 2.3.8.1) and gel extraction was used to purify the desired DNA fragments (section 2.3.3).

2.3.1.3 Ligation using T4 DNA Ligase

Ligation reactions utilised T4 DNA ligase with a 3:1 ratio of insert to vector DNA incubated in ligase buffer (NEB) for 20 minutes at room temperature then 4°C overnight. Prior to ligation of the DnaD 1-196-his construct, the DNA was incubated with T4 PNK (10 units) at 37°C for 30 minutes. T4 DNA ligase was heat inactivated at 65°C for 10 minutes then DNA was transformed into *E. coli* XL-1 Blue electrocompetent cells as described in section 2.3.6.1. DNA was extracted and purified (section 2.3.3) from overnight *E. coli* cultures, of cells containing the desired plasmid, prior to being sequenced as described in section 2.3.4 to confirm success of the cloning.

2.3.2 Molecular Cloning: Q5 Site Directed Mutagesis

The NEB Q5 site directed mutagenesis kit was employed to create all DnaA domain I mutants required for this project. As part of the protocol, the whole template plasmid was amplified to contain the desired mutation then a KLD (kinase, ligase, DpnI) enzyme mix was used to remove template DNA and ligate the plasmid in one step.

Reaction Component	Final Concentration
Q5 hot start high-fidelity 2x master mix	1 x
Forward primer	0.5 µM
Reverse primer	0.5 µM
Template DNA – pET28b-DnaA-NTD-81	1-25 ng

Table 2.9 *Reaction components used during PCR amplification of DNA. Oligonucleotide primers are defined for individual reactions in table 2.11 and 2.12.*

Step	Temperature (°C)	Time	Number of Cycles
Initial denaturation	98	30 sec	1
Strand separation	98	10 sec	
Primer annealing	-	30 sec	25
Strand extension	72	30 sec / kb	
Final extension	72	2 min	1

Table 2.10 *Thermocyclic conditions used for PCR amplification of DNA. The primer annealing temperature and strand extension times were variable and are described for individual PCR reactions. The primer annealing temperature used was dependent on the specific oligonucleotide primers required for each reaction. The strand extension time was calculated by the length of DNA to be amplified and the activity of Q5 polymerase (1 unit per 50 µL reaction can process 1 kb per 20-30 seconds depending on complexity).*

Construct	Primer	Sequence 5'–3'	Ta (°C)
S75A	Forward	GGAAGAATTG**gcc**ATTAAGTTTGTCATTCC	58
	Reverse	CCGGTTAATTCATATATAGTATC	
T26A	Forward	GAGTTTTGAG**gcc**TGGATGAAGTC	58
	Reverse	GGTTTGCTCAACTTTTTTTC	
S56A	Forward	CTGGCTGGAG**gcc**AGATACTTGC	68
	Reverse	TCTCTGGCAAATTCATTGGGAGC	
T70A	Forward	ATATGAATTA**gcc**GGGGGAAGAATTG	61
	Reverse	CTGGCTGGAGGCCAGATACTTGC	
S23A	Forward	GAGCAAACCG**gcc**TTGAGACTTGG	57
	Reverse	AACTTTTTTTCGATTTGAGC	
E73N	Forward	AACCGGGGAA**aac**TTGAGCATTAAG	60
	Reverse	AATTCATATATAGTATCTGCAATCAG	
R57K	Forward	GCTGGAGTCC**aaa**TACTTGCATC	62
	Reverse	CAGTCTCTGGCAAATTCATTG	
E72N	Forward	ATTAACCGGG**aac**GAATTGAGCATTAAG	59
	Reverse	TCATATATAGTATCTGCAATCAG	
G71A	Forward	TGAATTAACC**gcc**GAAGAATTGAGCATTAAG	59
	Reverse	TATATAGTATCTGCAATCAGATG	
L54A	Forward	CAGAGACTGG**gcc**GAGTCCAGATACTTG	60
	Reverse	GCAAATTCATTGGGAGCC	
E68N	Forward	TACTATATAT**aac**TTAACCGGGGAAGAATTG	59
	Reverse	TCTGCAATCAGATGCAAG	
L74N	Forward	GGGGAAGAA**gct**AGCATTAAGTTTG	56
	Reverse	GTTAATTCATATATAGTAATCTGC	
G38A	Forward	CTCACTGCAA**gcc**GATACATTAAC	61
	Reverse	TGGGCTTTGGTTGACTTC	
ΔM1-L5	Forward	GACCTGTGGAACCAAGCC	67
	Reverse	CATATATGGCTGCCGCGC	
D_{148}YARET_{153}	Forward	**agagagacg**CTGGAGTCCAGATACTTGC	63
	Reverse	**ggcataatc**ATTGGGAGCCGTGATTGT	
W53A	Forward	TGCCAGAGAC**gcc**CTGGAGTCCAG	64
	Reverse	AATTCATTGGGAGCCGTG	
F49A	Forward	TCCCAATGAA**gcc**GCCAGAGACTG	59
	Reverse	GCCGTGATTGTTAATGTATC	
D52A	Forward	ATTTGCCAGA**gcc**TGGCTGGAGT	66
	Reverse	TCATTGGGAGCCGTGATTG	

Table 2.11 *DNA oligonucleotide sequences used for PCR mutagenesis of B. subtilis DnaA domain I. Substitution mutations are indicated in lowercase bold.*

2.3.2.1 KLD Treatment

The Q5 site directed mutagenesis kit comes with a KLD enxyme mix containing kinase, ligase and DpnI. The DpnI selectively digests methylated template DNA while the kinase and ligase circularise the synthesiszed PCR product. PCR product (1 µL) was incubated with the reagents in table 2.12 (10 µL reaction) at room temperature for 5 minutes then stored at 4 °C until transformation into NEB 5-alpha chemically competent cells (section 2.3.2.2).

Reaction Component	Final Concentration
PCR product	-
2x KLD reaction buffer	1 x
10x KLD enzyme mix	1x

Table 2.12 *Reaction components used during KLD treatment of PCR products from Q5 site directed mutagenesis.*

2.3.2.2 Transformation

NEB 5-alpha high efficiency chemically competent cells are provided with the Q5 site directed mutagenesis kit. KLD treated PCR product (5 µL) was added to the cells (50 µL) and incubated on ice for 30 minutes. The cells and DNA were heat shocked at 42 °C for 30 seconds then incubated on ice for a further 5 minutes. Transformation products were incubated in SOC outgrowth medium (950 µL) for 1 hour then plated on agar supplemented with the appropriate antibiotic and incubated overnight at 37 °C.

2.3.3 DNA Purification

DNA was extracted and purified from overnight *E. coli* cultures using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) according to the manufacturer's protocol. The GeneJET Gel Extraction and DNA Cleanup Micro Kit was used, according to the manufacturer's protocol, for the extraction and purification of DNA during molecular cloning.

2.3.4 DNA Sequencing

DNA sequencing was carried out to confirm the correct plasmid was obtained during molecular cloning. The sequencing service was provided in house, by the DNA Sequencing Facility University of Nottingham Medical School, using a 3130xl ABI PRISM Genetic Analyzer (Life Technologies). The T7 promoter (5´-TAATACGACTCACTATAGGG-3´) and T7 terminator primers (5'-GCTAGTTATTGCTCAGCGG - 3') were used for full coverage sequencing and results were analysed using ApE software.

2.3.5 Production of Competent Cells

The *E. coli* strains BL21 (DE3) and XL-1 Blue were utilised during this research. BL21 (DE3) cells were used for protein expression and XL-1 Blue cells for storage of plasmid DNA (section 2.3.7). Electrocompetent cells were produced to allow incorporation of DNA into the *E. coli* genome using electroporation (section 2.3.6.1).

2.3.5.1 Production of Electrocompetent Cells

Glycerol stocks of BL21 (DE3) and XL-1 Blue *E. coli* cells were used to inoculate LB media incubated overnight at 37°C, 180 rpm. The cultures were diluted 1:100 into LB broth and incubated at 37°C until A_{590} reached 0.7. The growths were then incubated at 4°C for 30 minutes and the cells harvested by centrifugation at 3000 *g* for 10 minutes. Cells were washed by resuspension in sterile Milli-Q water followed by centrifugation at 3000 *g* for 10 minutes. This was repeated four times using decreasing resuspension volumes until a final resuspension in LB supplemented with glycerol (30% v/v). Cells were stored as 80 μL aliquots at -80°C. Sterile technique was used throughout to prevent contamination of the cell stocks.

2.3.5.2 Production of $CaCl_2$ Chemically Competent Cells

Glycerol stocks of BL21 (DE3) *E. coli* cells were used to inoculate LB media incubated overnight at 37°C, 180 rpm. The cultures were diluted 1:100 into LB broth and incubated at 37°C until A_{590} reached 0.3. The cells were harvested by centrifugation at 3000 *g* for 10 minutes then suspended in 4 mL sterile ice cold $CaCl_2$ (50 mM) and incubated on ice for 1 hour. Cells were again harvested by centrifugation at 3000 *g* for 10 minutes and suspended in 800 µL $CaCl_2$ (50 mM). Cells were stored as 100 µL aliquots at -80°C. Sterile technique was used throughout to prevent contamination of the cell stocks.

2.3.6 DNA Transformation

2.3.6.1 Transformation using Electrocompetent Cells

Plasmid DNA (2 µL) was added to electrocompetent cells (80 µL) and incubated on ice for 2 minutes. The cells and DNA were electroporated then incubated in LB media (1 mL) for 1 hour at 37°C. The transformation products were centrifuged at 3000 rpm for 4 minutes, the pellet suspended in 200 µL supernatant, plated on agar supplemented with the appropriate antibiotic and incubated overnight at 37°C. For the transformation of ligation products during molecular cloning 5 µL of plasmid DNA was added to XL-1 Blue electrocompetent cells then the protocol described above was followed.

2.3.6.2 Transformation using $CaCl_2$ Chemically Competent Cells

Plasmid DNA (1 µL) was added to $CaCl_2$ chemically competent cells (100 µL) and incubated on ice for 30 minutes. The cells and DNA were heat shocked at 42°C for 45 seconds then incubated on ice for a further 5 minutes. The transformation products were incubated in LB media (400 µL) for 1 hour then plated on agar supplemented with the appropriate antibiotic and incubated overnight at 37°C.

2.3.7 Plasmid Storage

For long-term plasmid storage LB glycerol stocks (10 % v/v glycerol) were prepared from overnight *E. coli* XL-1 Blue cultures inoculated with a single colony grown on LB agar plates with the appropriate antibiotic selection. Glycerol stocks were stored at -80°C.

2.3.8 Gel Electrophoresis

2.3.8.1 Agarose Gel Electrophoresis

Agarose gel electrophoresis was exploited to detect reaction products during molecular cloning. Purple DNA loading dye (NEB) was added to DNA before loading onto a 50 mL agarose gel (1% w/v agarose, 40 mM TRIS pH 8, 20 mM acetic acid, 1 mM EDTA, 0.5 µg/mL EtBr). Gels were electrophoresed at 70 V, 70 mA for 1 hour in TAE running buffer (40 mM TRIS pH 8, 20 mM acetic acid, 1 mM EDTA). 2 log (0.1 – 10kb) DNA ladder (NEB) was run alongside DNA samples as an indicator of fragment size. DNA was visualised by ethidium bromide under UV transilumination (Geneflow GBOX, software GeneSnap). For purification of DNA by gel extraction 0.8% (w/v) agarose gels were prepared and used as described above for 1% (w/v) agarose gels.

2.3.8.2 SDS-PAGE

SDS polyacrylamide gel electrophoresis was used to analyse protein samples. Samples were incubated with SDS-PAGE loading dye (10% v/v glycerol, 50 mM Tris-HCl pH 6.8, 0.05% w/v bromophenol blue, 12.5 mM EDTA, 100 mM DTT, 2% w/v SDS) at 90°C for 5 minutes prior to loading on the polyacrylamide gel. The acrylamide percentage of the resolving gel was selected according to the size of the protein being purified, compositions are described in table 2.13. In all cases protein samples were run through a stacking gel (4% (v/v) polyacrylamide) before entering the resolving gel. 5 mL SDS-PAGE gels were run at 180 V, 40 mA into SDS-PAGE running buffer (25 mM Tris pH 8.3, 190 mM glycine, 0.1% w/v SDS) for a time dependent on the acrylamide percentage of the resolving gel. Protein standard markers (PageRuler Unstained Protein Ladder,

Thermo Scientific) were run alongside samples as an indicator of product size. Protein bands were visualised using GelCode Blue Safe protein stain (Thermo Scientific) according to the manufacturer's instructions and the gel imaged under white light using a Geneflow GBOX and software GeneSnap.

Reagent	Buffer Volume / mL					
	4%	7.5%	12%	15%	17.5%	20%
30% (w/v) Protogel	8 mL	25 mL	40 mL	50 mL	58.3 mL	66.7 mL
1.5 M Tris-HCl pH 8.8	-	25 mL	25 mL	25 mL	25 mL	25 mL
0.5 M Tris-HCl pH 6.8	15 mL	-	-	-	-	-
10% (w/v) SDS	0.6 mL	1 mL	1 mL	1 mL	1 mL	1 mL
MilliQ-water	36.4 mL	49 mL	34 mL	24 mL	15.7 mL	7.3 mL

Table 2.13 *Reagents used to prepare stock resolving and stacking (4% (v/v) polyacrylamide) buffers for gel electrophoresis. To set the polyacrylamide gel, buffers were supplemented with 5 µL / mL ammonium persulfate (APS, 10% w/v solution) and 0.5 µL / mL TEMED.*

2.3.9 Crosslinking of DnaA Filaments

$DnaA^{CC}$ (1.5 µM) was added to oligomer formation buffer (25 mM HEPES pH 7.6, 200 mM NaCl, 100 mM potassium glutamate, 10 mM $MgCl_2$) supplemented with 2 mM ATP and 3 nM DNA (819 bp fragment of *B. subtilis oriC* containing the DnaA box cluster downstream of the *dnaA* gene, see appendix for sequence information) the reaction was incubated for 15 minutes at 37°C. Cysteine specific cross-linker BMOE (Thermo Scientific) was added at 2 mM and the reaction proceeded for 5 minutes at 37°C. Cysteine (50 mM) was added with a further 5 minutes incubation at 37°C to quench the BMOE cross-linking reaction. Products of the crosslinking reaction were visualised using SDS-PAGE as described in section 2.3.8.2.

2.3.10 Florencent Labelling of DnaA

The flurophores Atto647N-maleimide and Cy3B-maleimide were chosen as a FRET pair suitable for investigating the DnaA nuceloptorein filament using single molecule FRET. Single cystein DnaA mutants N191C and A198C were labelled with each of the fluorescent dyes as described in section 2.3.10.1.

2.3.10.1 Maleimide Congugation

Atto647N-maleimide or Cy3B-maleimide (10 mM in DMSO) was added dropwise at 10 x molar excess to proteins $DnaA^{N191C}$ and $DnaA^{N198C}$ in 50 mM Tris, 100 mM NaCl, pH 7.5. The reaction was flushed with nitrogen and left to proceed overnight, at 4°C with rotation, prior to being quenched with 1 mM DTT. The reactive dyes were shielded from direct light throughout the labelling procedure.

2.3.10.2 Purification

Atto647N-labelled and Cy3B labelled $DnaA^{N191C}$ and $DnaA^{A19C}$ were seperated from the excess of fluorescent dye by extensive dialysis. Nickel-affinity chromatography, desalt size exclusion chromatography and seperation by gravity spin columns was trialed for purifcation of the labelled protein from the excess dye. Such methods led to considerable loss of protein due to aggregation onto column surfaces or membranes. Therefore, overnight dialysis at 4°C (into an appropriate buffer) was repeated until no further excess dye molecules were present in the exchanged buffer.

2.4 Protein Expression

The T7 bacterial expression system, using the pET vector series, was exploited for the overexpression of recombinant proteins. This system allows high-level expression from the bacteriophage T7 RNA polymerase which is tightly regulated by IPTG-dependent inhibition of the *lac* repressor.

2.4.1 Growth Media

Several growth media were selected for use dependent on the desired outcome of the growth. Luria broth was employed for small scale overnight *E. coli* cultures and the more nutrient rich 2xYT media was used for overexpression of proteins. Minimal media was exploited for overexpression of isotopically-labelled proteins for NMR analysis.

2.4.1.1 Luria Broth

Reagent	Mass (g/L)
Tryptone	10
Yeast extract	5
Sodium chloride	5

Table 2.14 *Components of Luria broth (LB). Volumes of 200 mL were made up with distilled water and autoclaved for sterilisation.*

2.4.1.2 2xYT Media

Reagent	Mass (g/L)
Casein digest peptone	16
Yeast extract	10
Sodium chloride	5

Table 2.15 *Components of 2xYT growth medium (Melford). Volumes of 500 mL were made up with distilled water and autoclaved for sterilisation in 2 L baffled flasks.*

2.4.1.3 Luria Agar

Reagent	Mass (g/L)
Agar	15
Tryptone	10
Yeast extract	5
Sodium chloride	10

Table 2.16 *Components of Luria Agar (LB Agar). Volumes of 200 mL were made up with distilled water and autoclaved to sterilise and set the agar. Sterile agar was melted, the appropriate antibiotic added at working concentration and plates of approximately 25 mL were poured and set under aseptic conditions.*

2.4.1.4 M9 Minimal Media

Minimal media was made up to 1 L with distilled water and the components in table 2.17 then autoclaved for sterilisation in a 2 L baffled flask. Ammonium chloride (1 g) and glucose (4 g) were dissolved in 20 mL Milli-Q water and sterile filtered through a 0.22 μm filter into 1 L minimal media. The media was then supplemented with biotin (10 mg), thiamine (10 mg), MEM vitamin solution (1 mL, Sigma-Aldrich), trace element

solution (1 mL, prepared according to table 2.18) and the appropriate antibiotic. For overexpression of isotopically-labelled proteins the components in table 2.19 were supplemented, as required, into minimal media in place of their unlabelled counterparts.

Reagent	Mass (g/L)
Disodium hydrogen orthophosphate	6
Potassium dihydrogen orthophosphate	3
Sodium chloride	0.5
Magnesium sulphate	0.3
Calcium chloride dihydrate	0.015

Table 2.17 *Components of M9 minimal media. The volume was made up to 1 L with distilled water.*

Reagent	Mass (mg/0.5 L)
$ZnCl_2$	40
$FeCl_3.6H_2O$	200
$CuCl_2.2H_2O$	10
$MnCl_2.4H_2O$	10
$Na_2B_4O_7.10H_2O$	10
$(NH_4)_6Mo_7O_{24}.4H_2O$	10

Table 2.18 *Components of the trace element solution. The volume was made up to 500 mL with Milli-Q water and aliquots of 1 mL stored at -20°C for use in minimal media growths.*

Isotopic Label	Reagent	Mass (g/L)
^{15}N	^{15}N-ammonium chloride	1
^{13}C	^{13}C-glucose	2

Table 2.19 *Reagents required for isotopic labelling of proteins during overexpression. Isotopically enriched media is prepared by substituting 4 g of ^{12}C glucose and 1 g of ^{14}N ammonium chloride with 2 g ^{13}C glucose and 1 g ^{15}N ammonium chloride (Goss Scientific).*

2.4.2 Expression using 2xYT media

In general, for overexpresison of *B. subtilis* DnaA and DnaD constructs, *E. coli* BL21 (DE3) cells containing the desired plasmid were inoculated into LB broth containing the appropriate antibiotic and grown overnight at 37°C, 180 rpm. The culture was diluted 1:50 into 2xYT broth supplemented with antibiotic and incubated at 30-37°C until an A_{590} 0.6-0.8 was reached. Protein expression was induced with the addition of

IPTG (1 mM), table 2.20 shows expression time and temperature specific to each recombinant protein.

Protein	Antibiotic Resistance	Expression Time / Hours	Expression Temperature / °C
DnaA	Kanamycin	4	30
DnaA domain I	Kanamycin	4	30
DnaD	Carbenicillin	8	37
DnaD 1-196	Carbenicillin	8	37
DnaD DDBH2	Kanamycin	16	20

Table 2.20 *Conditions used during overexpression of B. subtilis DnaA and DnaD recombinant proteins.*

2.4.3 Expression using M9 minimal media

Overexpression of *B. subtilis* DnaA and DnaD single domain constructs, for NMR analysis, was achieved using minimal media. *E. coli* BL21 (DE3) cells containing the appropriate plasmid were inoculated into LB broth containing kanamycin (30 μg/mL) and grown overnight at 37°C, 180 rpm. Cells from the overnight growth were harvested by centrifugation at 1000 *g* for 10 minutes, re-suspended in minimal media (section 2.4.1.4) and incubated for 1 hour at 37°C, 180 rpm. The culture was diluted 1:25 into minimal media then incubated at 30°C until A_{590} reached 0.6. Protein expression was induced with the addition of IPTG (1 mM) followed by incubation at 30°C, 180 rpm, for 16 hours.

2.4.4 Harvesting Cells

Cells were harvested from the growth culture by centrifugation at 3000 x *g* for 20 minutes at 4°C. The cell pellet was then stored at -80°C until required for purification.

2.4.5 ^{15}N and ^{13}C Labelling

For use in multidimensional NMR experiments, DnaA domain I and DnaD DDBH2 domain were expressed in M9 minimal media enriched with the desired isotope (section 2.4.1.4) to obtain isotopically labelled protein. Overexpression was achieved as

described in section 2.4.3. As the supplemented ^{13}C-glucose and ^{15}N-ammonium chloride are the only source of carbon or nitrogen, respectively given to the *E. coli* cells, this method produces uniformly labelled protein.

2.5 Protein Purification

Fast protein liquid chromatography was utilised for protein purification including; affinity, ion exchange and size exclusion techniques to ensure high purity was achieved. All purifications were performed using an AKTA Prime system (GE Healthcare) where the UV absorption was monitored at 280 nm to detect protein presence. Purity was assessed by SDS-PAGE (section 2.3.8.2) and protein concentration determined spectrophotometrically (section 2.5.4).

2.5.1 Cell Lysis and Protein Extraction

Cells pellets were suspended in the appropriate lysis buffer (20 mL per 1 L growth, table 2.21) and lysed by sonication for 8 x 30 seconds at 10 microns (MSE soniprep 150). Lysed cells were incubated with benzonase (10 units) at 37°C for 30 minutes then cell debris was removed by centrifugation at 35,000 x *g* for 30 minutes at 4°C.

Protein	Buffer Composition
DnaA domain I DnaD DDBH2 domain	50 mM potassium phosphate pH 7.4 0.5 M NaCl 20 mM imidazole 1 mM PMSF
DnaA	50 mM tris pH 8 0.5 M NaCl 20 mM imidazole 1 mM PMSF
DnaD DnaD 1-196	20 mM tris pH 7.5 0.5 M NaCl 2 mM EDTA 1 mM DTT 20% (w/v) sucrose 1 mM PMSF

Table 2.21 *Cell lysis buffer conditions for B. subtilis DnaA and DnaD recombinant proteins.*

2.5.2 Polyhistidine-tagged Constructs

The DnaA domain I and DnaD DDBH2 domain plasmids incorporated an N-terminal his_6-tag to aid purification of the recombinant proteins. When analysed by NMR spectroscopy such tags can complicate spectra by producing strong signals in the correpsonding regions. Therefore, the postion of the his_6-tag allowed for its cleavage by thrombin. The DnaA and the truncated DnaD 1-196 plasmids incorporated a permanent C-terminal his6-tag as these would not be used for NMR experiments.

2.5.2.1 Nickel-affinity Chromatography

The initial purification stage for constructs with polyhistidine tags was nickel-affinity chromotography. Soluble protein, extracted from *E. coli* cells by sonication (section 2.5.1), was filtered through a 0.22 μm syringe filter prior to loading onto a 5 mL HisTrap HP column equilibrated with binding buffer (table 2.22). The column was washed thoroughly then the bound protein was eluted using a linear gradient 0 - 50% elution buffer over 100 mL.

Protein	Chromatography Column	Buffer Composition	
		Binding	Elution
DnaA domain I, DnaD DDBH2 domain	HisTrap HP	50 mM KP pH 8 0.5 M NaCl 20 mM imidazole	50 mM KP pH 8 0.5 M NaCl 0.5 M imidazole
	Superdex 75 26/60	50 mM KP pH 7.4 100 mM NaCl	
DnaA	HisTrap HP	50 mM TRIS pH 8 0.5 M NaCl 20 mM imidazole	50 mM TRIS pH 8 0.5 M NaCl 0.5 M imidazole
	HiTrap Q HP	50 mM TRIS pH 7.5 1 mM EDTA 1 mM DTT	50 mM TRIS pH 7.5 1 M NaCl 1 mM EDTA 1 mM DTT
	Superdex 200 26/60	50 mM TRIS pH 7.5 100 mM NaCl	
DnaD 1-196	HisTrap HP	50 mM TRIS pH 8 0.5 M NaCl 20 mM imidazole	50 mM TRIS pH 8 0.5 M NaCl 0.5 M imidazole
	Superdex 75 26/60	20 mM TRIS pH 7.5 300 mM NaCl 2 mM EDTA 1 mM DTT	

Table 2.22 *The chromatography columns and buffer compositions used for purification of B. subtilis DnaA and DnaD recombinant proteins. KP corresponds to mixed potassium (K_2HPO_4/KH_2PO_4) buffer. All buffers were filtered through a 0.22 µm membrane and degassed thoroughly before use. All columns were purchased from GE Healthcare.*

2.5.2.2 Anion Exchange

$DnaA^{CC}$ protein collected from nickel-affinity chromatography was dialysed overnight into HiTrap Q HP binding buffer (table 2.22) supplemented with 100 mM NaCl. Protein was then loaded onto a 5 mL HiTrap Q HP column pre-equilibrated with binding buffer. The column was washed thoroughly and the bound protein was eluted using a linear gradient 10 – 50% elution buffer over 100 mL.

2.5.2.3 Thrombin Cleavage

DnaA domain I and DnaD DDBH2 domain protein samples collected after nickel-affinity chromatography were dialysed into thrombin cleavage buffer (20 mM tris pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$). The N-terminal his_6-tag was then cleaved by restriction grade thrombin (1-2 units per mg protein) for 16 hours at room temperature. The cleavage product was filtered through a 0.22 µm syringe filter prior to loading onto a 5 mL HisTrap HP column pre-equilibrated with HisTrap binding buffer (lacking imidazole, table 2.22). The column was washed thoroughly to collect unbound, and therefore cleaved, protein in the flow through. Un-cleaved protein, free his_6-tag and thrombin was collected with elution buffer.

2.5.2.4 Gel Filtration

The appropriate Superdex gel filtration column was pre-equilibrated in running buffer (table 2.22) prior to loading of the protein sample (at a volume >12 mL) and run between 0.5 and 1.5 mL/min. Fractions containing the desired protein were pooled, concentrated (section 2.5.5) when necessary and snap frozen for storage at -80°C.

2.5.2.5 Desalt

The DnaA domain I protein collected after gel filtration was concentrated to a volume >4 mL. The sample was then loaded onto a HiTrap Desalt column (5 x 5 mL) pre-equilibrated in milliQ-water. Fractions containing DnaA domain I were then lyophilised for storage at -80°C.

2.5.3 *Bacillus subtilis* DnaD

The DnaD full length protein is prone to agggregation under low salt conditions as well as at high protein concentration. Addition of a polyhistidine tag at the C-terminus has been shown to interfere with the DNA binding activity hence and un-tagged purification method was used with this protein. Ammonium sulfate cut was used to

exchange buffers between columns and enable re-suspension of the DnaD protein at the desired conductivity. Both the ion exchange and Heparin affinity columns require low salt for binding. Therefore, to ensure DnaD bound but did not aggregate onto the column a constant conductivity of 10 – 12mS was maintained by running a low percentage of elution buffer into the mixing chamber while loading and washing the columns. The size exclusion chromotography profile was used to determine that DnaD had been obtained in its active tetrameric form.

Protein	Chromatography Column	Buffer Composition	
		Binding	Elution
DnaD	HiTrap Q HP	20 mM glycine pH 9.8 2 mM EDTA 1 mM DTT	20 mM glycine pH 9.8 1 M NaCl 2 mM EDTA 1 mM DTT
	HiTrap Heparin HP	50 mM BIS-TRIS pH 6.7 2 mM EDTA 1 mM DTT	50 mM BIS-TRIS pH 6.7 1 M NaCl 2 mM EDTA 1 mM DTT
	Superdex 75 26/60	20 mM TRIS pH 7.5 300 mM NaCl 2 mM EDTA 1 mM DTT	

Table 2.23 *The chromatography columns and buffer compositions used for purification of B. subtilis DnaD. All buffers were filtered through a 0.22 µm membrane and degassed thoroughly before use. All columns were purchased from GE Healthcare.*

2.5.3.1 Ammonium Sulfate Cut

Soluble protein, extracted from *E. coli* cells by sonication (section 2.5.1), was precipitated using a 50% ammonium sulphate cut (2.9 g per 10 mL) then clarified at 17,500 rpm for 30 minutes at 4°C.

2.5.3.2 Anion Exchange

DnaD protein collected from ammonium sulfate cut was suspended in a mixture of HiTrap Q HP binding and elution buffer (Table 2.23) to achieve a conductivity of 40 mS. Protein was then loaded onto a 5 mL HiTrap Q HP column pre-equilibrated in Q binding

buffer. While loading, a constant 10 -12 mS conductivity was maintained by running a low percentage of elution buffer into the mixing chamber. This was required to prevent aggregation of the DnaD protein under low salt conditions. The column was washed thoroughly and bound protein was eluted using a linear gradient 10 – 50% elution buffer over 100 mL.

2.5.3.3 Heparin-affinity Chromatography

DnaD protein collected from anionic exchange was precipitated with ammonium sulfate cut (as described in section 2.5.3.1) then suspended in a mixture of HiTrap Heparin HP binding and elution buffer (Table 2.23) to achieve a conductivity of 40 mS. Protein was then loaded onto a 5 mL HiTrap Q HP column pre-equilibrated in Q binding buffer. While loading, a constant 10 -12 mS conductivity was maintained by running a low percentage of elution buffer into the mixing chamber. The column was washed thoroughly and bound protein was eluted using a linear gradient 10 – 80% elution buffer over 150 mL.

2.5.3.4 Gel Filtration

A Superdex 75 26/60 gel filtration column was pre-equilibrated in running buffer (table 2.23) prior to loading of the protein sample (at a volume >12 mL) and run between 0.5 and 1.5 mL/min. The elution profile indicated oligomeric DnaD. Fractions containing protein were pooled, supplemented with 10% (v/v) glycerol and snap-frozen for storage at -80°C.

2.5.4 Protein Quantification

The theoretic molecular weight and extinction coefficient of protein constructs were calculated from the amino acid sequence using the ExPASy ProtParam tool. These values were used for the quantification of protein along with the A_{280} recorded using a NanoDrop 2000 spectrophotometer (Thermo Scientific).

Protein	Molecular Weight / AMU	Extinction Coefficient / M^{-1} cm^{-1}	Abs $_{0.1\%}$ / AU
$DnaA^{CC}$-his	52158.5	38390	0.736
DnaA 1-81	9662.0	19480	2.016
^{15}N labelled	9772.2	19480	1.993
$^{15}N^{13}C$ labelled	10206.9	19480	1.909
DnaA 1-105	12421.0	19480	1.568
^{15}N labelled	12563.0	19480	1.551
DnaD	27638.7	35410	1.281
DnaD 1-196-his	24373.9	26930	1.105
DnaD 129-196	8334.5	13980	1.677
DnaD 129-232	12877.6	22460	1.744
^{15}N labelled	13037.5	22460	1.723

Table 2.24 *Theoretical protein molecular weight and extinction coefficient values calculated using the ExPASy online resource. The extinction coefficient is calculated from absorbance at 280 nm and* $Abs_{0.1\%}$ *corresponds to the absorbance of a 1 g/L sample of the protein.*

2.5.5 Concentrating Proteins

Following purification, proteins were concentrated as desired using centrifugal concentration by membrane ultrafiltration. Vivaspin columns (GE healthcare) with the appropriate molecular weight cut off were selected for this process, protein was applied to the column and the volume reduced to achieve the desired concentration. Buffer exchange was also achieved using this method. The desired buffer was used to dilute the protein, the sample concentrated by centrifugation and the process repeated to remove residual unwanted buffer.

2.6 Biophysical Methods

2.6.1 Nuclear Magnetic Resonance Spectroscopy

NMR experiments were recorded at 298K on a Bruker 800 MHz Avance III Spectrometer with a QCI cryoprobe. For each sample the 90° pulse and transmitter frequency were calibrated prior to running an experiment. The number of scans and points collected in each dimension for each experiment was selected subject to the protein concentration. Data acquisition and processing was carried out using Topspin 3.1.b.53 software.

2.6.1.1 Processing and Analysis

1D data sets were processed using an exponential window function with 2 Hz line broadening. For multidimensional datasets a shifted squared sine bell was used with the offset being optimized to achieve the best balance between resolution and signal to noise. All data were zero-filled by a factor of 2. Linear prediction was applied to heteronuclear dimensions. Analysis of 2D and 3D data was achieved using CCPN Analysis software.

2.6.1.2 Sample Preparation

Lyophilised DnaA proteins were suspended to the desired concentration in 600 µL NMR buffer (50 mM potassium phosphate pH 7.4, 50 mM NaCl, 5% (v/v) D_2O, 0.02% NaN_3). DnaD protein samples were buffer exchanged into 700 µL NMR Titration buffer (50 mM potassium phosphate pH 7.4, 100 mM NaCl, 5 mM DTT, 5% (v/v) D_2O, 0.02% (w/v) NaN_3) to achieve the desired concentration. Samples were clarified at 13,500 rpm for 5 minutes to remove any insoluble material prior to running on the NMR spectrometer. 3D experiments were acquired at a DnaA domain I concentration of 300 µM while during titration experiments ^{15}N labelled protein at 100 µM was used to a final ratio of 8:1 ligand:protein. DnaA and DnaD samples were dialysed (16 hours, 4°C) into the desired NMR buffer prior to running on the spectrophotometer to ensure the buffer remained consistent throughout titration experiments. For titration experiments

in the presence of ssDNA, this was supplemented into the NMR Titration buffer to achieve a 1:1 ratio with the final concentration of DnaD.

2.6.1.3 1D Experiments

One dimensional protein spectra were recorded using excitation sculpting water suppression, with a spectral width of 15 ppm. The transmitter frequency for hydrogen was centred at 4.706 ppm and 64 scans were acquired.

2.6.1.4 2D Experiments

For ^{15}N labelled and ^{14}N selectively unlabelled material 2D ^{1}H-^{15}N-heteronuclear single quantum coherence (HSQC) spectra were acquired over a spectral width of 16 ppm in the ^{1}H dimension and 45 ppm in the ^{15}N dimension. The transmitter frequency for nitrogen was centred at 119 ppm. Typically, 64 scans with 2048 complex points in the nitrogen dimension and 128 in the hydrogen dimension were acquired. Quadratic detection in the nitrogen dimension was achieved using States-TPPI.

2.6.1.5 3D Experiments

3D experiments were acquired for backbone assignment of DnaA domain I using $^{13}C^{15}N$ labelled material. These experiments included the HNCO, (HCα)CONH, CβCαNH, HNCα, and CβCα(CO)NH. The proton and nitrogen dimensions were optimised as described for 2D experiments (section 2.6.1.4). The transmitter frequency for carbon was centred at either 39 ppm for CαCβ experiments or 176 ppm for CO experiments. For the HNCO and HNCα 16 scans while for the (HCα)CONH, CβCαNH and CβCα(CO)NH 64 scans with 40 points in the nitrogen dimension and 128 points in the carbon dimension were acquired. Quadratic detection was achieved, as with 2D experiments, using States-TPPI.

2.6.2 Mass Spectrometry

Experiments were performed on a SYNAPT High Definition Mass Spectrometry system (Waters) using electrospray ionization and quadrupole time of flight mass analyser. Data analysis was performed using Mass Lynx software. Samples were diluted, or lyophilised samples suspended, into 25 mM ammonium acetate to a final concentration of 10 μM then a 100 μL sample was injected at 5 μL/min.

2.6.3 Circular Dichroism

Circular dichroism was used to asess secondary sturcture content of both DnaD DDBH2 domain and DnaA domain I (wild type and mutant) constructs. The Chirascan Plus Spectrometer (Applied Photophysics) was used for data aquistion and data analysis was performed using the CDNN software. Samples were diluted, lyophilised samples suspended, in CD buffer (50 mM potassium phosphate pH 7.4, 50 mM NaCl) to a concentration of 0.1 mg/mL in 500 μL then were loaded into a 1mm quartz cuvette and scans were collected at 190-260 nm at 25°C. Three scans were taken for each sample and the baseline (buffer only sample) was subtracted from the averaged spectra.

2.6.4 Surface Plasmon Resonance Spectroscopy

SPR spectroscopy was performed using a BIAcore 3000 (GE Healthcare) using an NTA (GE Healthcare) sensor chip with a carboxymethylated dextran surface pre-immobilsed with nitrotriacetic acid (NTA) to bind histidine-tagged biomolecules.

2.6.4.1 Surface Preparation

The NTA chip surface is activated by injecting 0.5 mM $NiCl_2$ across the surface at 10 μL/min. The reference surface does not get activated for use as a blank reference. DnaA domain I constructs were affinity captured to the NTA surface via their N-terminal his_6-tag. A ligand concentration of 0.2 μM in SPR buffer (10 mM HEPES pH 7.4, 150 mM NaCl) was injected over 10 minutes at 10 μL/min for sufficient capture. The chip

surface was regenerated, between runs, by injecting 350 mM EDTA pH 8.3 across the surface at 10 μL/min. This will remove nickel and any chelated molecules from the surface so that the NTA chip can be re-activated for use.

2.6.4.2 Data Analysis

The analyte was prepared from a stock concentration (in SPR buffer) as a dilution series with at least five concentrations and one repeated value per run. Analyte was injected over the chip surface at 10 μL/min for 2.5 minutes. The association and dissociation time was optimised to achieve steady-state equilibrium. The chip surface was regenerated and activated (as described in section 2.6.4.1) between each analyte injection as required owing to the non-covalent immoblisation of ligand to the surface.

2.6.4.3 Kinetic Analysis

The BIAevaluation software (GE Healthcare), utilised during this research, contains a curve fitting procedure that solves the differential form of the rate equation using numerical integration algorithms. The best fit model, such as separate association and dissociation phase fitting or the simultaneous Langmuir fit, is determined by an iterative process to describe the interaction and allow extraction of the kinetic rate constants from the experimental data. The process optimises initial parameter values to an algorithm that minimises the fitting value χ^2 (average squared residual) a measure of the difference between the experimental data and the fitted curve. Sensograms from the concentration series are simultaneously fit to reduce the range in the calculated values as local fits may introduce bias into the model. Stoichiometry of the interaction can also be determined using the following equation:

$$R_{max} = \frac{MW_{analyte}}{MW_{ligand}} . R_L . S_m \qquad (2.6.1)$$

Where R_{max} is substituted by R_L, the ligand response value (immobilisation level on the surface in RU) and the equation is rearranged to find S_m, the stoichiometry defined by the binding sites on the ligand. A theoretical R_{max} can be calculated and compared to the experimental value to determine activity of the ligand surface.

2.6.4.4 Steady State Analysis

An alternate approach to kinetic analysis is to use the R_{max} value for the sensograms during the steady state phase to determine a value for K_D. During steady state k_a = k_d hence this method does not provide information on the kinetics of the interaction. During this research, K_D values were calculated from saturation plots of R_{max} against analyte concentration, as illustrated in Figure 2.1. These were fitted, using non-linear regression, to the Hill Equation[108]:

$$R_{max} = \frac{B_{max}[A]^h}{K_d^h + [A]^h} \qquad (2.6.2)$$

Where [A] is the concentration of free analyte. Within the PRISM software (Graphpad) the 'one site with specific binding with Hill Slope' primary fit was utilised which provides information on stoichiometry and cooperativity of binding interactions via parameter h and the parameter B_{max} determines the maximum number of binding sites available on the surface.

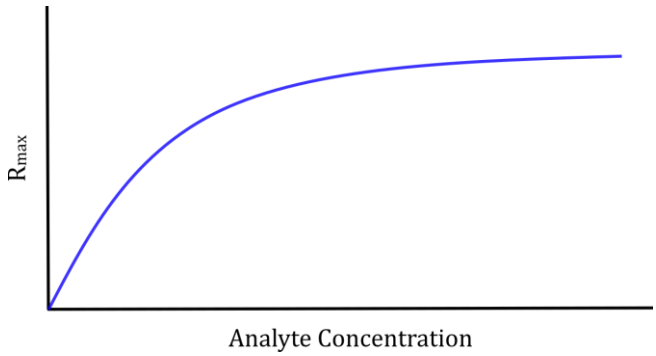


Figure 2.1 *Saturation plots of R_{max} values, obtained from sensograms at steady state, against analyte concentration allow non-linear regression analysis for determination of K_D values.*

Chapter 3

Biochemical Characterisation of the DnaA-DnaD Interaction

3.1 Introduction

DnaA is a highly conserved master replication initiator protein across the bacterial species. The mechanism of action in *E. coli* has been well documented but little is known of the additional role of essential initiator protein DnaD in *B. subtilis* and related gram-positive organisms. It has been established that an interaction occurs between DnaA and DnaD[106], these initiator proteins form high order oligomeric complexes[98] and both display DNA remodelling activity[99]. It has been reported that DnaD may act as a negative regulator of DNA replication initiation[65, 66], however, its requirement for viability[103], abundance within the cell[103] and involvement in replication restart[89] suggest that DnaD plays a more cooperative role during DNA replication initiation.

The interaction of DnaD with DnaA was investigated to further understand the role of DnaD during DNA replication initiation and reveal the mechanism of DNA melting at the origin of replication. This chapter describes the expression and purification of the DnaA and DnaD constructs required for this study. Preliminary study of the DnaA nucleoprotein filament revealed that DnaD can alter its oligomeric conformation and provided evidence for the direct interaction between DnaD and DnaA domain I. Ishigo-Oka *et al.* reported that the interaction between DnaA and DnaD occurs through the DnaD DDBH2 domain[106]. Thus, providing the grounds for further

characterisation of the interaction between the two individual domains in subsequent chapters.

3.2 Protein Expression

For the purpose of this research, it was necessary to produce a variety of recombinant protein constructs for the *B. subtilis* initiators DnaA and DnaD. The well-established *E. coli* recombinant overexpression technique was employed for this work. *E. coli* has the ability to grow quickly, and to a high density, using low cost substrates hence it is a popular choice for expression of recombinant proteins. This system also facilitates the isotopic labelling technique required for NMR studies, the details of which are described in Chapter 4. The BL21 (DE3) *E. coli* strain, in combination with pET expression vectors, was used to produce the DnaA and DnaD recombinant constructs. The conditions required for the soluble expression of individual constructs are described in sections 2.4.2 and 2.4.3. The domain boundaries for each of the targets of interest were chosen after consultation with the literature, an X-ray crystal structure of DnaA domain I (in complex with SirA) defined a sequence for stable folded protein. The sequence corresponding to the DnaD DDBH2 domain along with the truncation for disrupting DNA binding activity of DnaD was based on the work of Marston *et al.*[90].

3.2.1 DnaA Expression

The wild type DnaA and the DnaA^{CC} constructs were supplied by Geoff Briggs (Soultanas group) in the pET-28b vector between cloning sites *NcoI* and *NotI*. This vector provides a C-terminal polyhistidine tag (AAALEHHHHHH) immediately prior to the target protein to aid purification. The single cysteine mutants (at positions A198C and N191C) were produced by site-directed mutagenesis of the wild type construct. Expression in the pET vector is under control of the *T7* promoter and *lac* operator and is induced by the addition of isopropyl β-D-thiogalactoside (IPTG), the non-hydrolysable allolactose mimic. The vector also contains the *lacI*^{q} gene so that expression can be tightly regulated until IPTG displaces the lac repressor at the operator region.

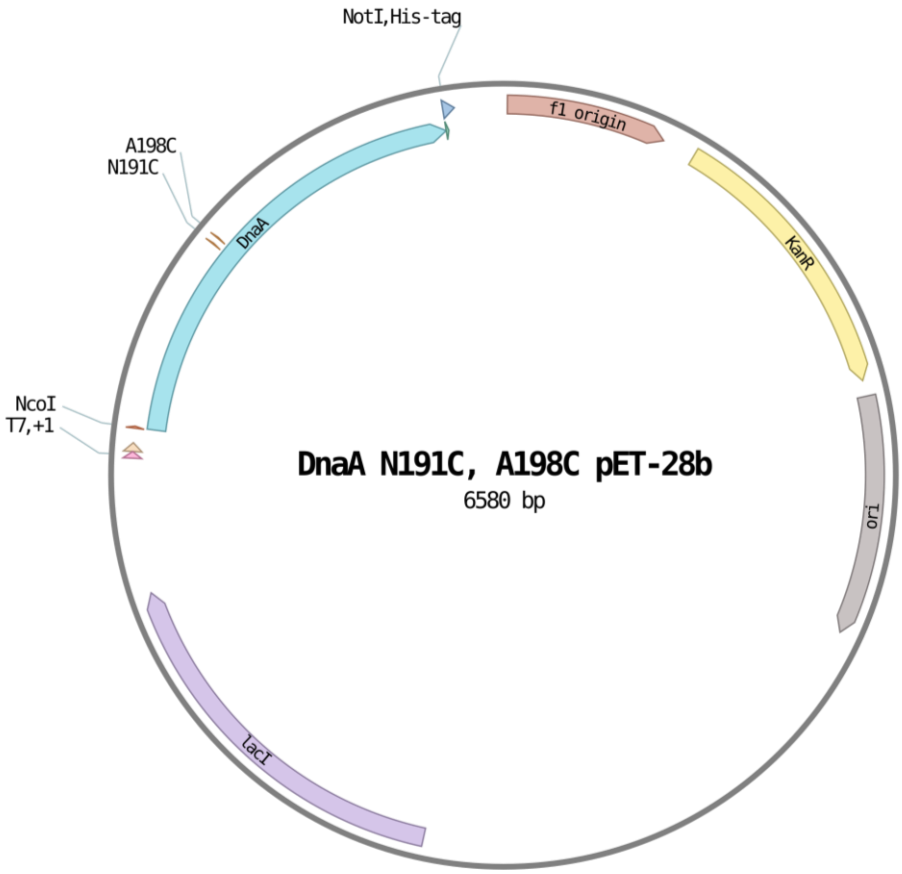


Figure 3.1 *The annotated plasmid map for pET-28b showing inserted DnaA gene. The pET-28b vector infers kanamycin resistance. Expression of the target gene is under control of the T7 promoter, induced by addition of IPTG.*

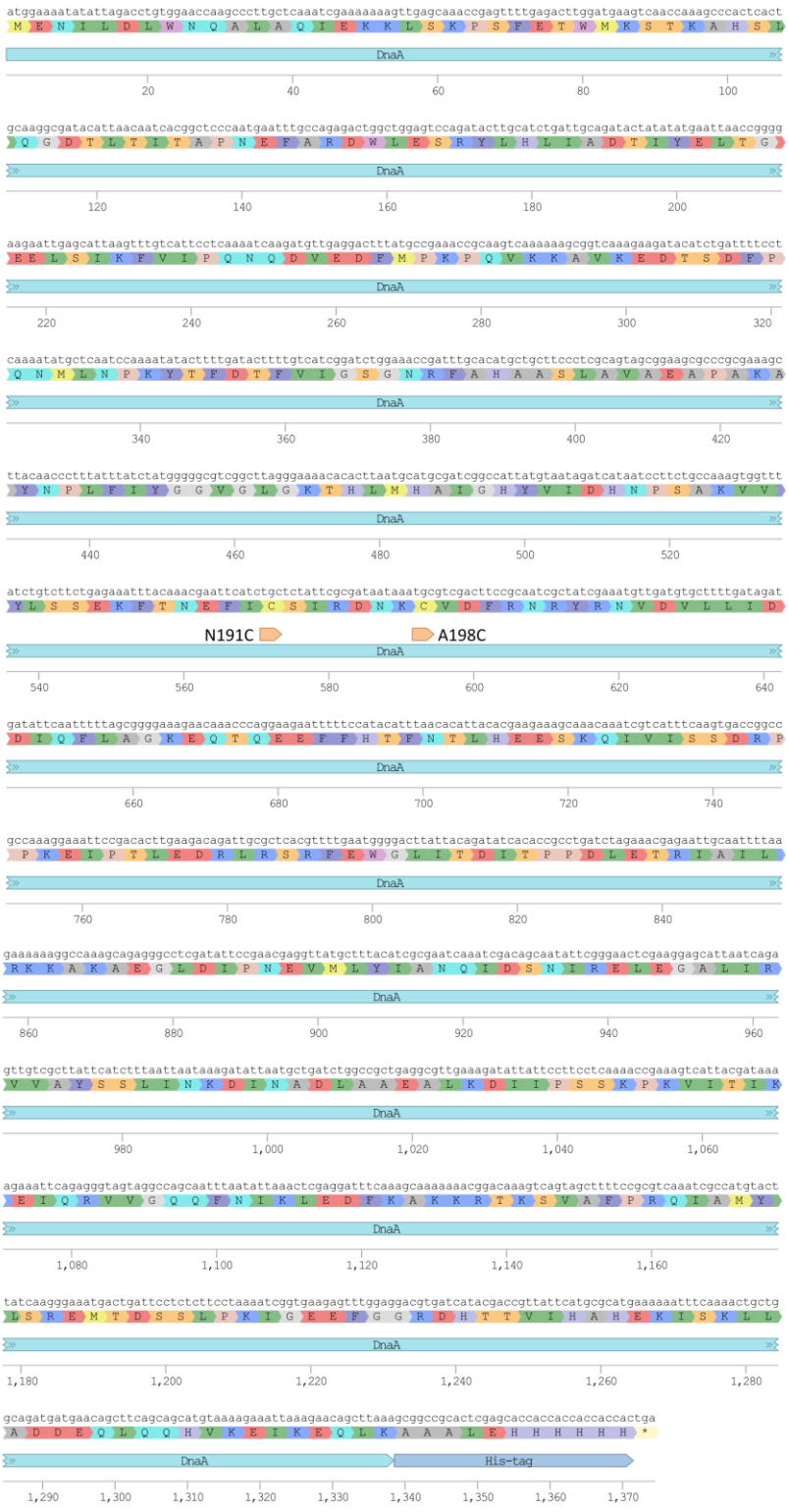


Figure 3.2 *The coding sequence for the DnaA^{CC} construct in pET-28b.*

3.2.2 DnaD Expression

The wild type DnaD construct was supplied by William Grainger (Soultanas group) in the pET-22b vector between cloning sites *NdeI* and *HindIII*. The truncated DnaD construct, DnaD 1-196, was created by deletion of the region Glutamate 197 up to and including the *XhoI* restriction site. This provided a C-terminal polyhistidine tag (AAALEHHHHHH) to aid purification.

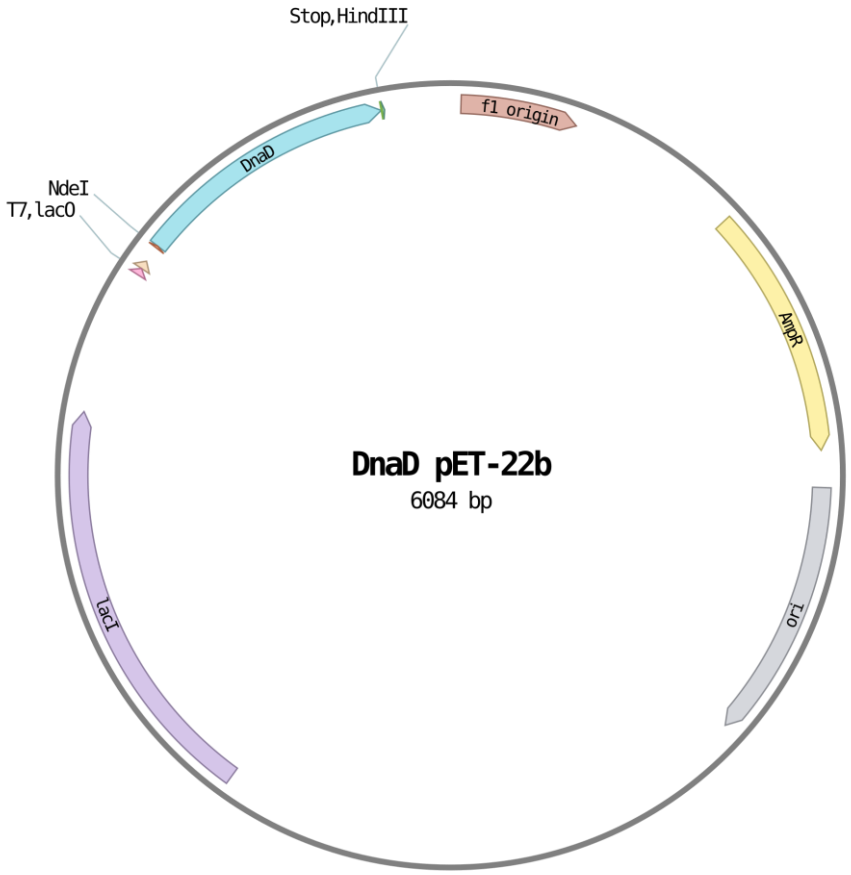


Figure 3.3 *The annotated plasmid map for pET-22b showing inserted DnaD gene. The pET-22b vector infers ampicillin resistance. Expression of the target gene is under control of the T7 promoter, induced by addition of IPTG.*

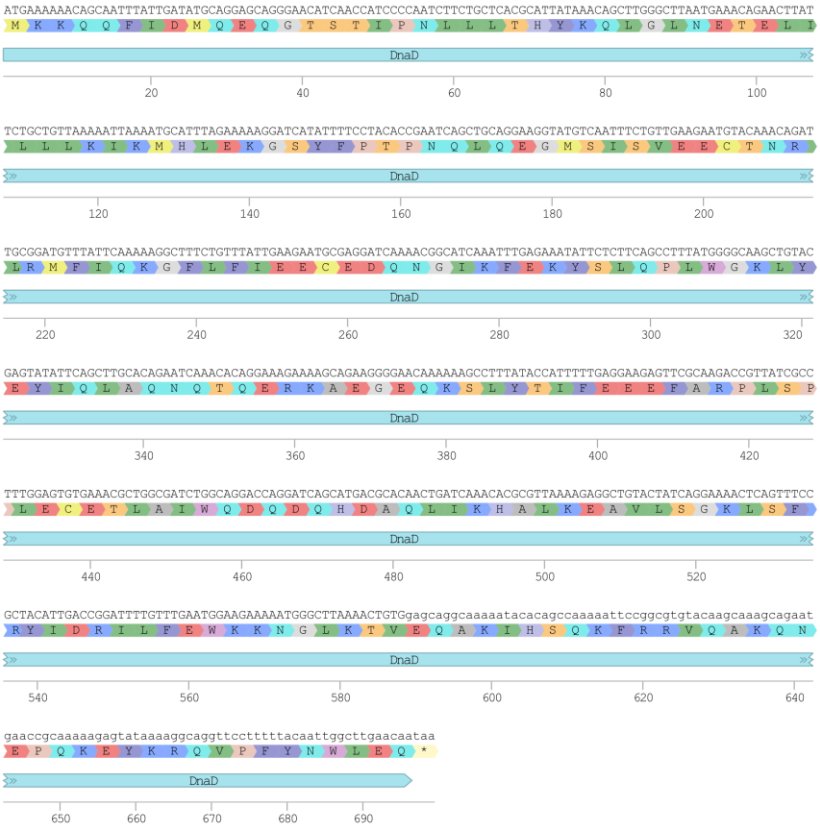


Figure 3.4 *The coding sequence for DnaD in the pET-22b vector.*

3.2.3 DnaA Domain I and DnaD DDBH2 Domain Expression

The pET-28b vector was also utilised for the DnaA domain I and DnaD DDBH2 domain constructs. In these instances, the target gene was cloned between the *NdeI* and *HindII* sites to provide an N-terminal polyhistidine tag separated from the protein by a thrombin cleavage site. The DnaD DDBH2 domain construct was kindly supplied by Farhat Marston (Craven Group, Univeristy of Sheffield) in the pET-28b vector. The truncated 129-196 construct was created by inserting a stop site (TAA) immediately after Valine 196.

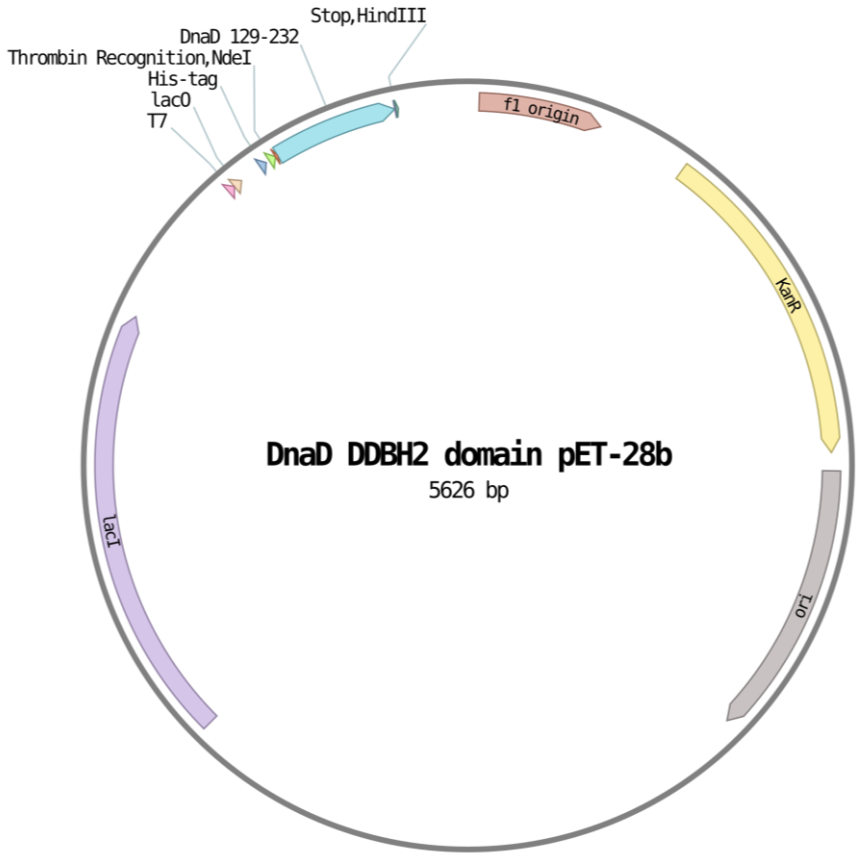


Figure 3.3 *The annotated plasmid map for pET-28b showing inserted DnaD DDBH2 gene. The pET-28b vector infers kanamycin resistance. Expression of the target gene is under control of the T7 promoter, induced by addition of IPTG.*

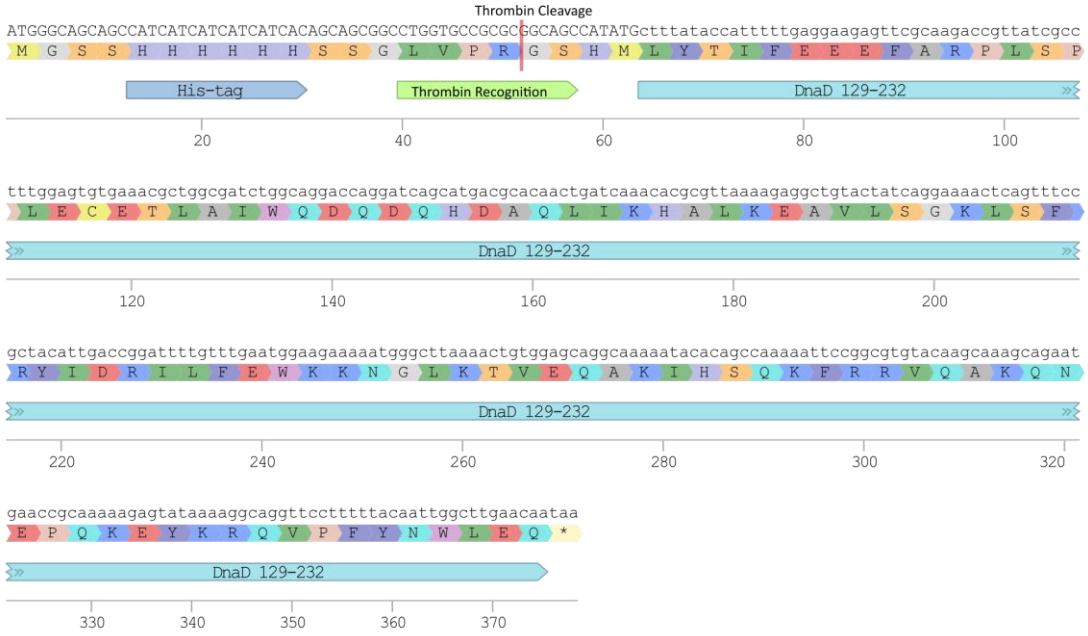


Figure 3.6 *The coding sequence for the DnaD DDBH2 construct in pET-28b. The thrombin cleavage site is highlighted, post cleavage a short leader sequence of GSHM precedes DnaD residue L129.*

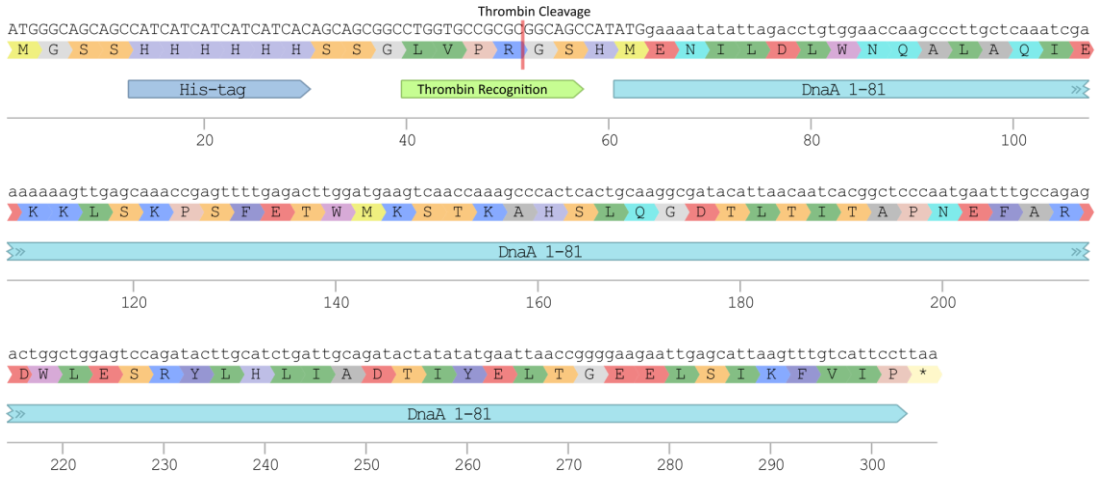


Figure 3.7 *The coding sequence for the DnaA domain I construct in pET-28b. The thrombin cleavage site is highlighted, post cleavage a short leader sequence of GSH precedes DnaA residue M1.*

3.3 Protein Purification

Fast protein liquid chromatography (FPLC) was utilised for protein purification during this research. Liquid chromatography specific to the purification of proteins is characterised by both a mobile and stationary phase, provided by an aqueous buffer and solid resin matrix respectively. Separation relies on the affinity of individual proteins for the two phases, and this can be manipulated to separate the desired component from the contaminated mixture. The pH and conductivity of the mobile phase can be altered, while a range of resins are available to manipulate the stationary phase (typically a cross-linked agarose bead which is modified by attachment of additional groups to the surface). The various forms of FPLC employed throughout this project are discussed below, all purifications were performed using an AKTA Prime system (GE Healthcare) where the UV absorption was monitored at 280 nm to detect protein presence. Purity was assessed by SDS-PAGE (section 2.3.8.2) and protein concentration determined spectrophotometrically (section 2.5.4).

3.3.1 Affinity Chromatography

3.3.1.1 Nickel-affinity

Nickle-affinity columns are a commonly used form of immobilised metal ion affinity chromatography (IMAC), in this case the matrix is functionalised with the nitrilotriacetic acid chelator. Most biological molecules do not have a high affinity for metal ions, hence these columns are used in partnership with recombinant proteins genetically engineered to contain a polyhistidine-tag[109]. The his-tag has a micromolar affinity for the metal ion functionalised matrix and the fusion products are typically eluted using a gradient of imidazole (within the liquid phase) to out-compete the protein.

3.3.1.2 Heparin-affinity

Heparin-sepharose columns are utilised for separation of biomolecules with an affinity to heparin. Heparin is a highly sulfonated glycosaminoglycan that can bind a wide range of biomolecules including; DNA-binding proteins, coagulation factors, serine protease inhibitors, growth factors, hormone receptors and lipoproteins[110]. For its interaction with DNA binding proteins, heparin mimics the polyanionic structure of the nucleic acid. The selectivity of heparin can be modified by altering the pH or ionic strength of the liquid phase and typically a salt gradient is used to elute the target protein.

3.3.2 Ion Exchange

The principle of ionic affinity between the applied protein sample and the solid phase is the basis of ion exchange. The surface charge of the protein target is dictated by the pH and conductivity of the liquid phase, and hence these are important to the strength of the ionic interaction required to bind the target protein to the solid phase[111]. The matrix for anionic exchange columns carry a positive charge via functionalised groups such as quaternary ammonium (Q), diethylaminoethyl (DEAD), and quaternary aminoethyl (QAE) which are bound by negatively charged proteins (under the conditions of the liquid phase). Cationic exchange columns use a negatively charged matrix, surfaces include sulfopropyl (SP) and carboxymethyl (CM), to bind positively charged proteins. Typically, during ion exchange, the analyte is loaded in a low salt buffer then a salt gradient is used to elute bound proteins. Proteins show great variety in their surface charge making this process ideal for separation.

3.3.3 Size Exclusion

Size exclusion chromatography, also known as gel filtration, separates the components of a mixture based on their molecular size. The stationary phase consists of a porous solid matrix where smaller molecules can enter the pores while larger molecules are excluded. Therefore, elution of separated products occurs largest to smallest according to the time spent interacting with the matrix. This technique is suited to removing either high or low molecular weight contaminants from a sample as a significant difference in molecular weight (minimum 10%) is essential for good separation[112].

3.3.4 DnaA Purification

The DnaA, DnaA^{CC}, DnaA^{N191C} and DnaA^{A198C} recombinant proteins were purified following the same methodology. The first stage of purification was the application of the soluble fraction of the cell lysate onto a HisTrap nickel-affinity column. The His-tagged protein target was captured from the crude mixture and eluted from the column over a 0.5 M imidazole gradient. The eluted sample was dialysed overnight into a low salt (100 mM) buffer to allow binding onto a HiTrap Q anion exchange column. The target protein was eluted from the column using a 1 M salt gradient. As a final purification stage, the sample was concentrated to a volume less than 12 mL for loading onto a Superdex 200 26/60 size exclusion column. The purity of the sample was confirmed by SDS-PAGE analysis at each stage of purification (Figure 3.8).

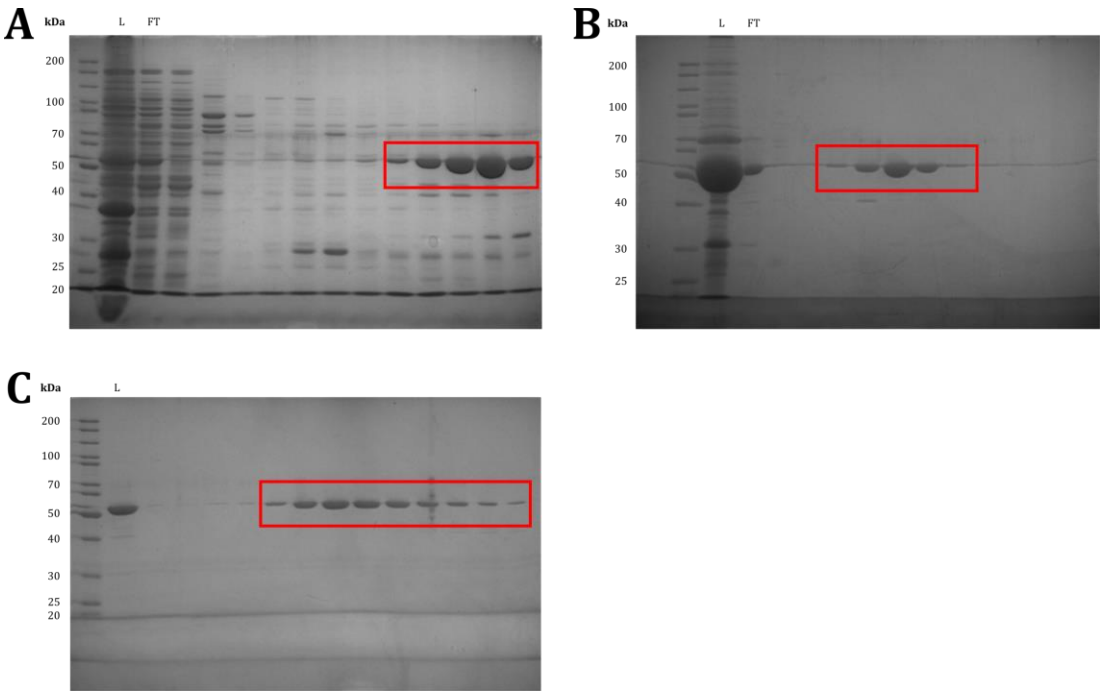


Figure 3.8 *Purification of DnaA^{CC}. SDS-PAGE analysis is shown for each purification stage. The first lane in each SDS-PAGE analysis contains a molecular weight marker, with values displayed to the left, the lane represented by 'L' contains a sample of the loaded material for each column and 'FT' represents the column flow though. Fractions taken forward are highlighted.* ***A.*** *Elution of DnaA^{CC} from the nickel-affinity column over an imidazole gradient. B. Elution of DnaA^{CC} from the anion exchange column over a salt gradient.* ***C.*** *Size exclusion purification of DnaA^{CC}, pure protein is observed at 52 kDa.*

3.3.5 DnaD Purification

The DnaD recombinant protein was purified using an un-tagged methodology. Ammonium sulfate cut (50%, 2.9 g per 10 mL), followed with clarification by centrifugation at 17, 500 rpm, was used to initially capture DnaD from the crude cell lysate. The sample was resuspended to a conductivity of 40 mS to allow binding onto a HiTrap Q anion exchange column. The target protein was eluted from the column using a 1 M salt gradient and the eluted sample was treated by a second ammonium sulfate cut. This allowed buffer exchange to 40 mS for binding onto a HiTrap Heparin-affinity column. The target protein was eluted over a 1 M salt gradient. As a final purification stage, the sample was loaded onto a Superdex 75 size exclusion column from which the

elution profile indicated oligomeric DnaD. The purity of the sample was confirmed by SDS-PAGE analysis at each stage of purification (Figure 3.9).

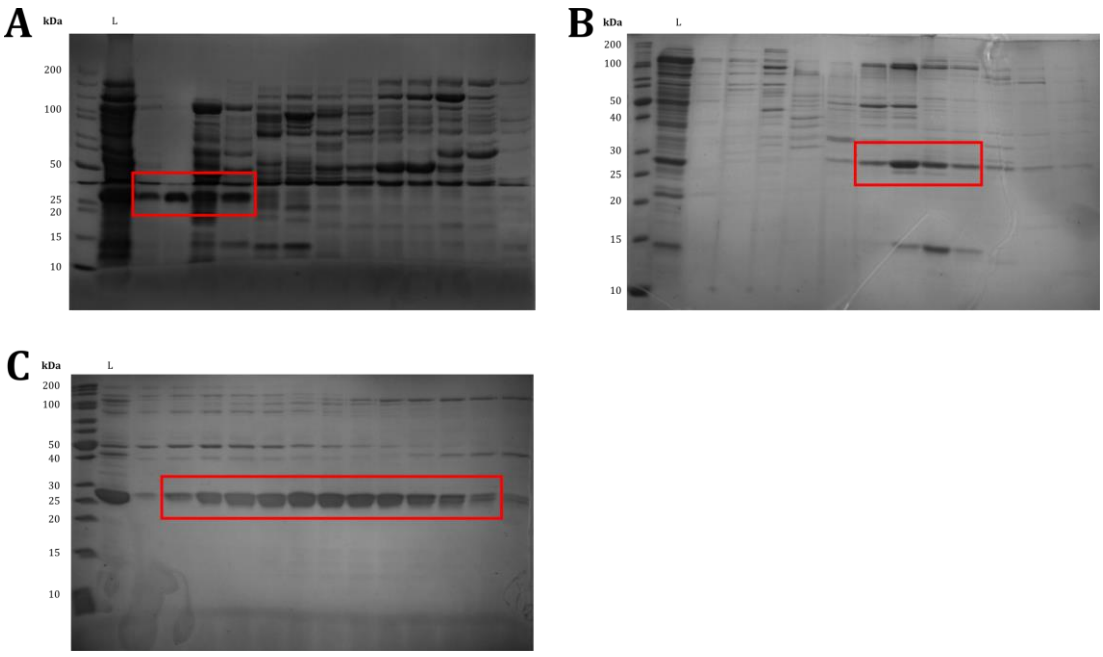


Figure 3.9 *Purification of DnaD. SDS-PAGE analysis is shown for each purification stage. The first lane in each SDS-PAGE analysis contains a molecular weight marker, with values displayed to the left, and the lane represented by 'L' contains a sample of the loaded material for each column. Fractions taken forward are highlighted.* ***A.*** *Elution of DnaD from the anion exchange column over a salt gradient.* ***B.*** *Elution of DnaD from the Heparin-affinity column over a salt gradient.* ***C.*** *Size exclusion purification of DnaD, monomeric protein is observed at 27.6 kDa.*

3.3.6 DnaD 1-196 Purification

The DnaD 1-196 recombinant protein was purified following a similar methodology as that used for DnaA purification (section 3.3.4). The first stage of purification was the application of the soluble fraction of the cell lysate onto a HisTrap nickel-affinity column. The His-tagged protein target was captured from the crude mixture and eluted from the column over a 0.5 M imidazole gradient. The eluted sample was then concentrated to a volume less than 12 mL for loading onto a Superdex 75 26/60 size exclusion column. The purity of the sample was confirmed by SDS-PAGE analysis at each stage of purification (Figure 3.10).

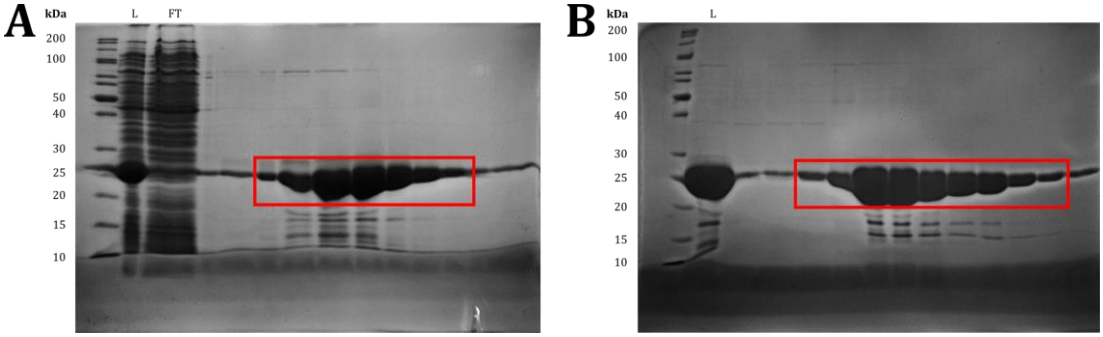


Figure 3.10 *Purification of truncated DnaD (DnaD 1-196). SDS-PAGE analysis is shown for each purification stage. The first lane in each SDS-PAGE analysis contains a molecular weight marker, with values displayed to the left, the lane represented by 'L' contains a sample of the loaded material for each column and 'FT' represents the column flow though. Fractions taken forward are highlighted.* ***A.*** *Elution of DnaD 1-196 from the nickel-affinity column over an imidazole gradient.* ***B.*** *Size exclusion purification of DnaD 1-196, the protein is observed at 24.4 kDa.*

3.3.7 DnaA Domain I and DnaD DDBH2 Domain Purification

The DnaA domain I and the DnaD DDBH2 domain recombinant proteins were purified following the same methodology. The first stage of purification was the application of the soluble fraction of the cell lysate onto a HisTrap nickel-affinity column. The His-tagged protein target was captured from the crude mixture and eluted from the column over a 0.5 M imidazole gradient. The eluted sample was dialysed into thrombin cleavage buffer (see section 2.5.2.3) over 4 hours then incubated with restriction grade thrombin (1-2 units per mg protein) overnight to cleave the N-terminal his-tag. The sample was passed through the HisTrap column for a second time to separate the cleaved his-tag from the target protein. As a final purification stage, to remove high molecular weight contaminants, the sample was concentrated to a volume less than 12 mL for loading onto a Superdex 75 size exclusion column. The purity of the sample was confirmed by SDS-PAGE analysis at each stage of purification (Figure 3.11 and 3.12).

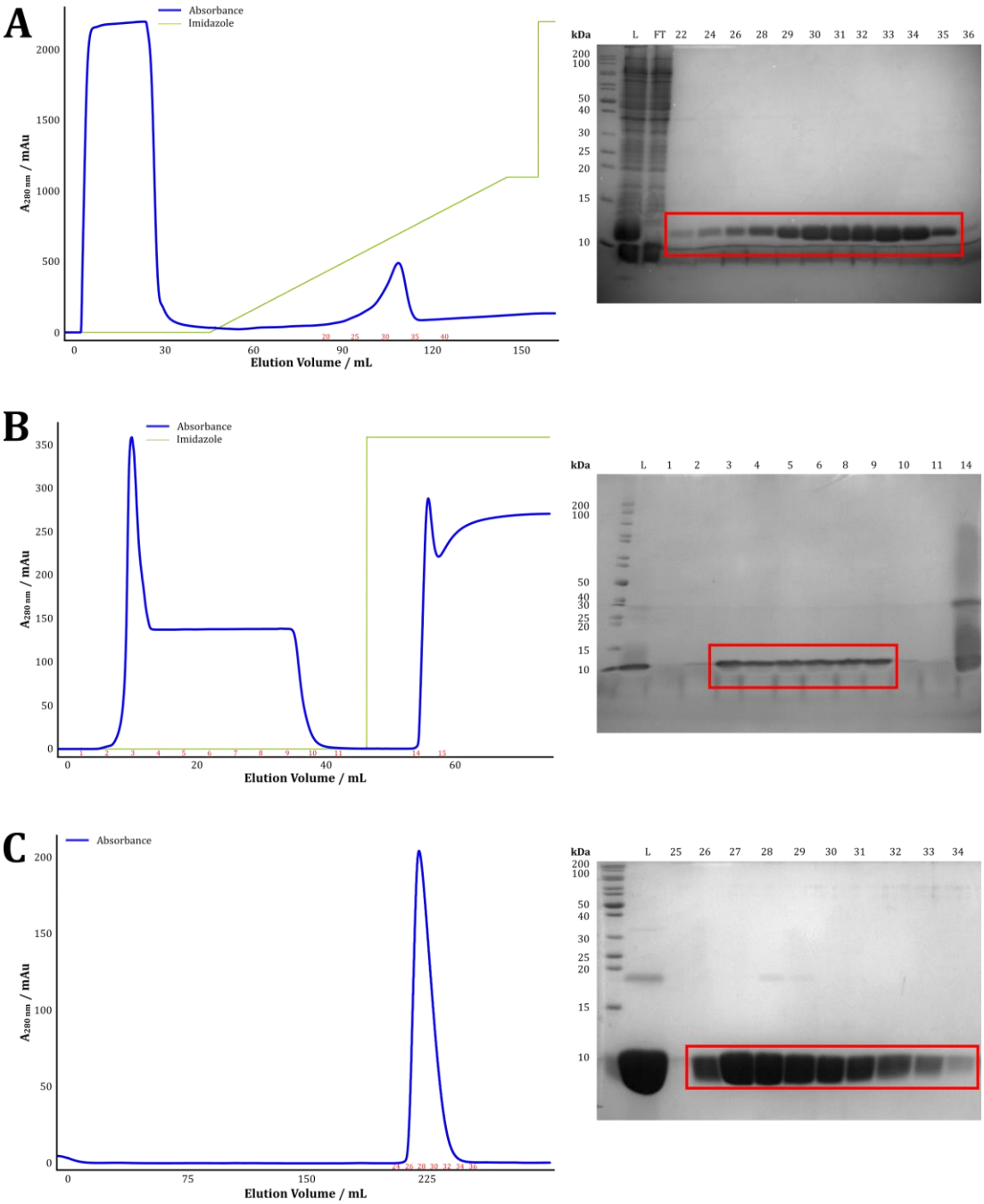


Figure 3.11 *Purification of DnaA domain I. The chromatogram is shown on the left and SDS-PAGE analysis on the right for each purification stage. Absorbance at 280 nm is displayed in blue while imidazole concentration is displayed in green. The first lane in each SDS-PAGE analysis contains a molecular weight marker, with values displayed to the left, the lane represented by 'L' contains a sample of the loaded material for each column and 'FT' represents the column flow though. Fraction numbers are indicated above the lanes and those taken forward are highlighted.* ***A.*** *Elution of DnaA domain I from the nickel-affinity column over an imidazole gradient.* ***B.*** *Separation of products after thrombin cleavage of the N-terminal polyhistidine tag. Cleaved DnaA domain I does not bind to the nickel-affinity column while uncleaved protein, thrombin and cleaved his-tag are eluted from the column with a 500 mM imidazole wash.* ***C.*** *Size exclusion purification of DnaA domain I. Pure protein is observed at this stage, running at 9.7 kDa.*

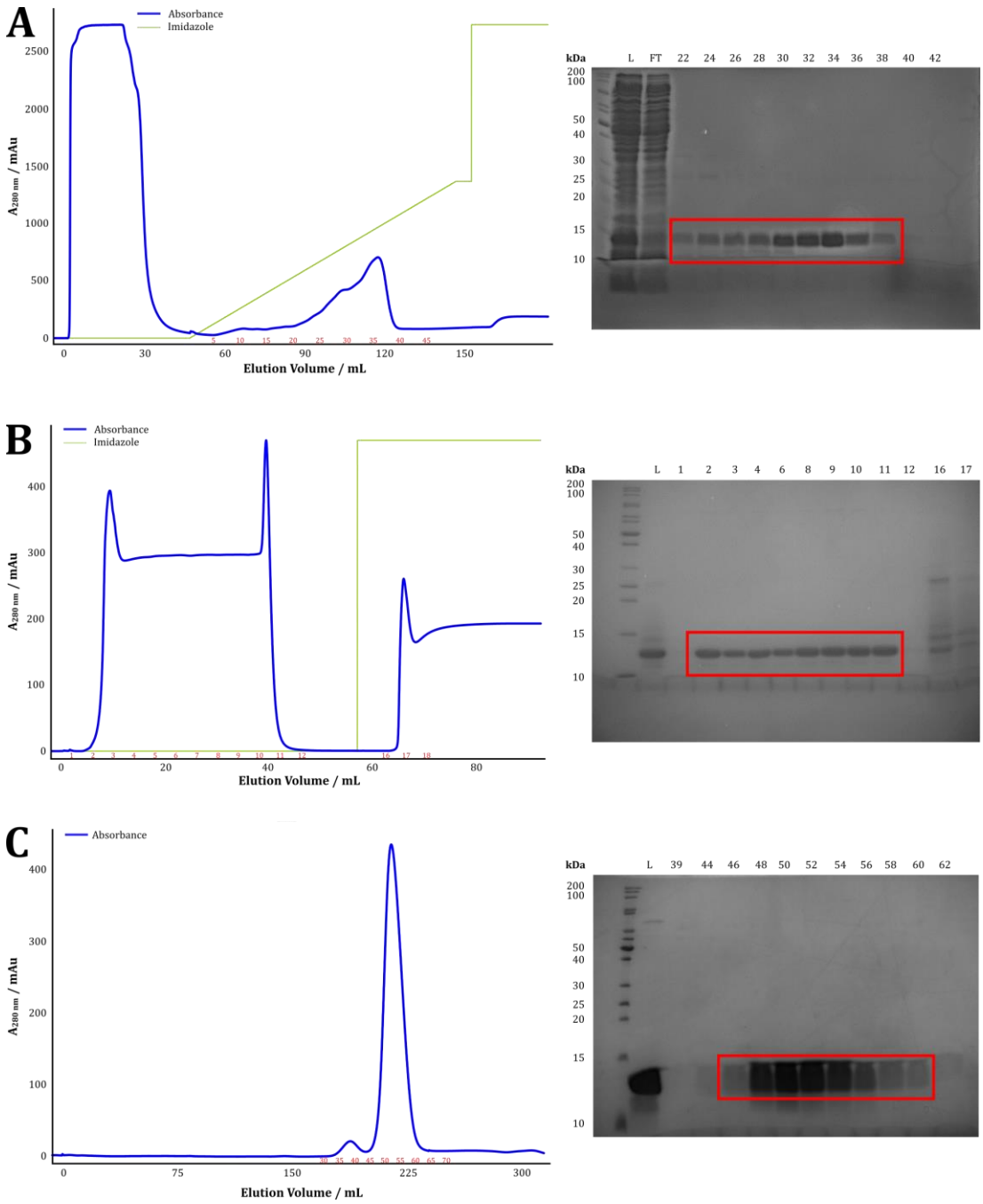


Figure 3.12 *Purification of the DnaD DDBH2 domain. The chromatogram is shown on the left and SDS-PAGE analysis on the right for each purification stage. Absorbance at 280 nm is displayed in blue while imidazole concentration is displayed in green. The first lane in each SDS-PAGE analysis contains a molecular weight marker, with values displayed to the left, the lane represented by 'L' contains a sample of the loaded material for each column and 'FT' represents the column flow though. Fraction numbers are indicated above the lanes and those taken forward are highlighted.* ***A.*** *Elution of DnaD DDBH2 domain from the nickel-affinity column over an imidazole gradient.* ***B.*** *Separation of products after thrombin cleavage of the N-terminal polyhistidine tag. Cleaved DDBH2 domain does not bind to the nickel-affinity column while uncleaved protein, thrombin and cleaved his-tag are eluted from the column with a 500 mM imidazole wash.* ***C.*** *Size exclusion purification of the DnaD DDBH2 domain. Pure protein is observed at this stage, running at 12.5 kDa.*

3.4 Characterisation of the DnaA Nucleoprotein Filament

A site-specific cross-linking assay was developed by Scholefield and Murray[66] to detect DnaA filament assembly on both double stranded and single stranded DNA. This technique involves the introduction of two cysteine residues into domain III of the *B. subtilis* DnaA protein at positions N191 and A198, the product referred to as DnaA^{CC}. Within the DnaA right handed helical filament structure, the residue N191 of one DnaA molecule is within 9 Å proximity of the residue A198 of a second adjacent molecule (Figure 3.13). Introduced cysteines at these positions can be cross-linked by the sulfhydryl specific cross-linker bis(maleimido)ethane (BMOE) with a spacer of 8 Å. The cross-linked protein complexes can then be visualised by SDS-PAGE electrophoresis. This technique is advantageous for visualisation of the DnaA filament as the high molecular weight species formed during the oligomerisation reaction are too large to enter the gel matrix under native conditions.

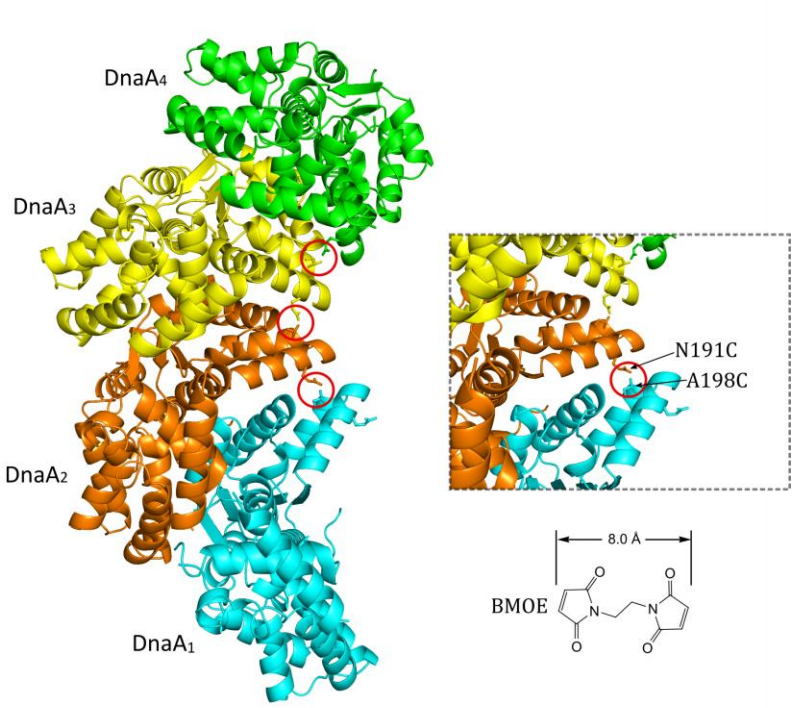


Figure 3.13 *Cartoon representation of the helical DnaA structure from A. aeolius (PDB – 2HCB). DnaA monomers are identified by individual colours. The residues representing introduced cysteines at position N191 and A198 (for crosslinking assays with B. subtilis) are highlighted. BMOE conjugates residue C191 of the first DnaA monomer with C198 of the second monomer, as indicated by the red circle.*

Using the crosslinking method for detection of DnaA filaments, it was reported that DnaD inhibits DNA-dependent DnaA filament formation but has no effect on pre-formed DNA filaments[65, 66]. However, the recombinant DnaD protein used in this research contained a C-terminal polyhistidine-tag which is known to diminish the DNA binding activity of DnaD[90]. These results are also based on the proposal that the degree of BMOE cross-linking observed corresponds with the level of DnaA filament formation achieved. It should not be interpreted strictly that the inhibition of BMOE cross-linking corresponds to the inhibition of filament formation as this method of detection does not allow for altered conformations of the DnaA filament to be recognised. We propose that DnaD acts cooperatively with DnaA during DNA replication initiation and a true understanding of any conformational changes to the DnaA filament in the presence of DnaD are of importance to this research. Therefore, the recombinant proteins $DnaA^{CC}$, DnaD and DnaD 1-196 were obtained for use in DnaA nucleoprotein filament studies.

3.4.1 Optimisation of the DnaA Filament Formation Reaction

DnaA filament formation was induced using an oligomerisation reaction and adjacent monomers within the filament were then cross-linked, using the cysteine specific maleimide conjugation reaction with BMOE, as described in section 2.3.9. SDS-PAGE was used to visualise the products of the crosslinking reaction. The oligomerisation reaction was repeated under various reaction conditions, as described below, to initially characterise DnaA filament formation in *B. subtilis*.

3.4.1.1 *oriC* DNA Fragments

As mentioned previously, there are two DnaA box clusters proximal to the *oriC* on the *B. subtilis* bacterial chromosome. For investigation of DnaA filament formation, double stranded linear *oriC* DNA fragments were utilised (provided by Dr Geoff Briggs, Soultanas Group). Four regions of the bacterial chromosome were chosen, consisting of; an 846 bp fragment containing the DnaA-box cluster upstream of the *dnaA* gene, an 819 bp fragment containing the DnaA-box cluster downstream of the *dnaA* gene, a 2436

bp fragment containing both upstream and downstream DnaA-box clusters, and an 811 bp region of the chromosome containing no DnaA binding sites (see Appendix 8.1 for sequence specific information). The latter was amplified for use as a control to measure non-specific binding of DnaA to DNA. Figure 3.14 displays the various levels of crosslinked DnaA molecules observed using the different *oriC* fragments within the oligomerisation reaction. Extensive cross-linking of high molecular weight DnaA species was observed in the presence of DNA fragments containing DnaA-box clusters. Low levels of cross-linking were observed in the absence of DNA or presence of non-specific DNA. During the oligomerisation reaction, the DnaA protein was in a molar excess of 500-fold to DNA (protein excess to the number of binding sites being dependent on the specific fragment).

Electrophoretic mobility shift assays (EMSA) were additionally used to study DnaA binding to the various *oriC* DNA fragments. In this case, BMOE cross-linking was not necessary prior to electrophoresis due to the native conditions used. At high concentrations of DnaA, large species were formed during the oligomerisation reaction which could not enter the agarose gel, therefore, lower concentrations of DnaA (187 nM – 1.5 μM) were used for analysis by EMSA. It was revealed that DnaA actively bound to all fragments of DNA (Figure 3.14) including the region amplified without DnaA-box clusters. The greatest shift was seen with the DNA fragment containing the DnaA-box cluster downstream of the *dnaA* gene. This region is located just ahead of the DUE in the *oriC* and contains 4 DnaA-box sites. Subsequent experiments, therefore, utilised this particular *oriC* DNA fragment.

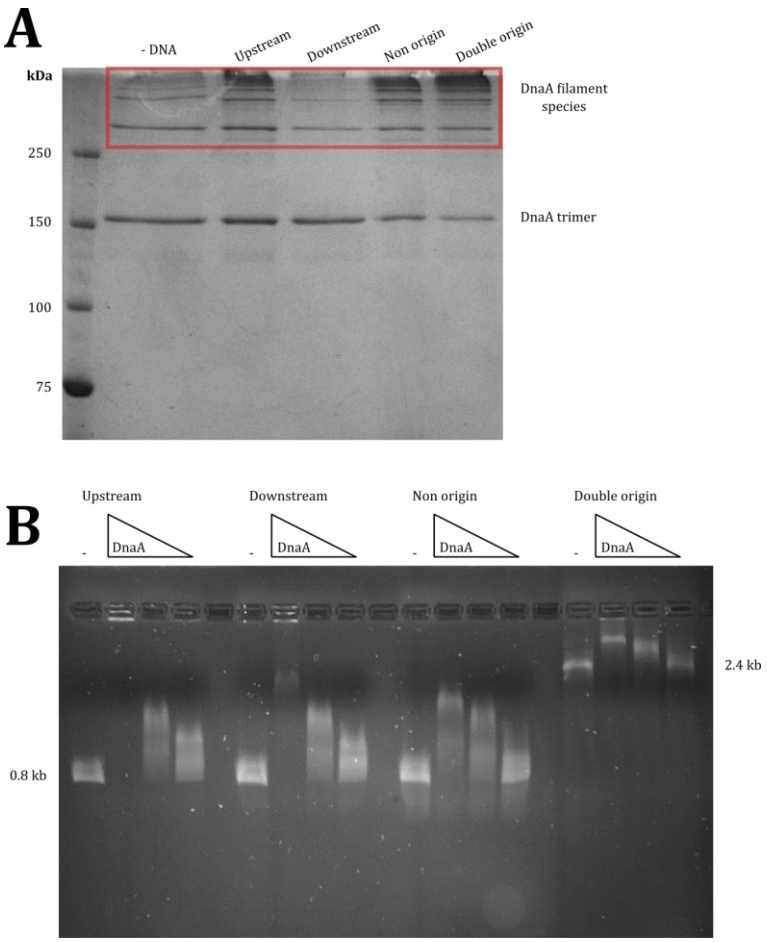


Figure 3.14 *Visualisation of DnaA nucleoprotein filaments formed using various DNA fragments.* ***A.*** *SDS-PAGE analysis of BMOE crosslinked DnaA after oligomerisation. The first lane on the gel contains a molecular weight marker, with values displayed to the left. Each of the remaining lanes contains the product of the oligomerisation reaction using no DNA or varied DNA fragments (DNA upstream of the dnaA gene, DNA downstream of the dnaA gene, non-origin DNA, and DNA containing both the upstream and downstream DnaA-box clusters). To enhance visualisation of the high molecular weight crosslinked DnaA species (highlighted), monomeric DnaA was run off the gel shown above. The gel was run in tandem with a higher percentage of bisacrylamide to visulaise monomeric DnaA ensuring DnaA concentration is consistent throughout.* ***B.*** *EMSA analysis of DnaA nucleoprotein filaments formed using various DNA fragments. Triangles above the lanes represent the decreasing DnaA concentration used (1.5 µM, 750 nM, 375 nM, 187 nM respectively) and lanes indicated by a '-' represent the absence of DnaA.*

3.4.1.2 DnaA Concentration and ATP Dependent Oligomerisation

Having established DnaA filament formation could be achieved using the DNA fragments amplified by PCR, other elements in the reaction were investigated to find

the optimum conditions for DnaA filament formation *in vitro.* The oligomerisation reactions were performed using various concentrations of the DnaA^{CC} protein to reveal that increasing concentrations of DnaA led to higher extents of BMOE cross-linking and the observation of many high order oligomeric species by SDS-PAGE. A concentration of 1.5 µM DnaA^{CC} was found to be sufficient for qualitative visualisation of the DnaA filament species using the crosslinking technique (as displayed in Figure 3.15).

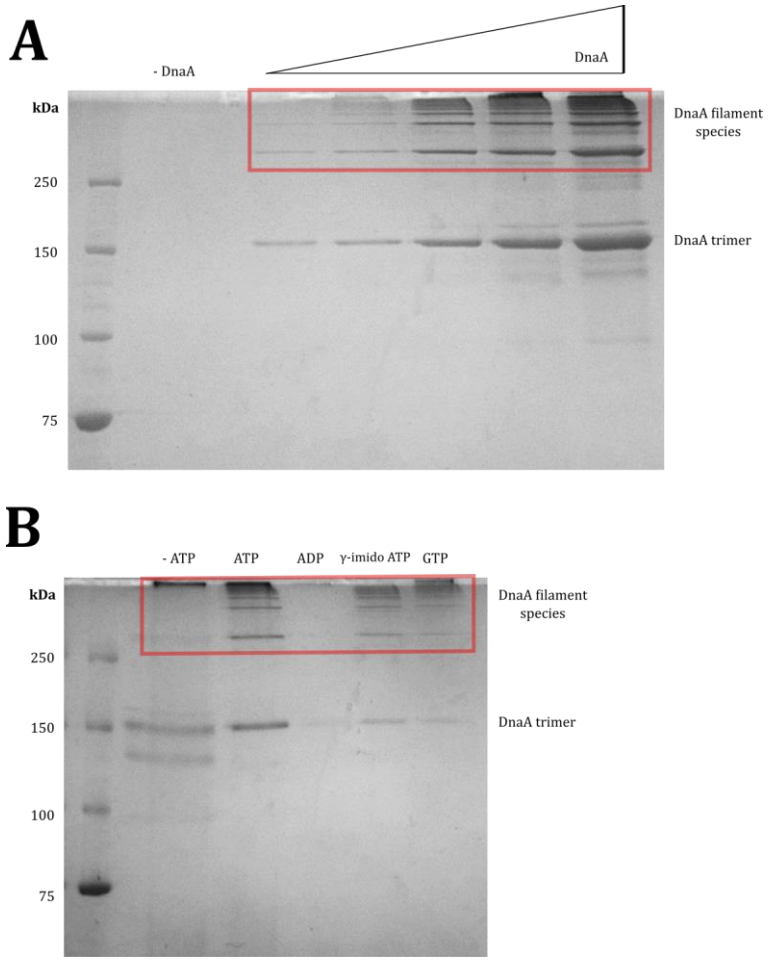


Figure 3.15 *SDS-PAGE visualisation of BMOE crosslinked DnaA nucleoprotein filaments. The first lane of each gel contains a molecular weight marker, with values displayed to the left. To enhance visualisation of the high molecular weight crosslinked DnaA species (highlighted), monomeric DnaA was run off the gels shown above. These were run in tandem with a higher percentage of bisacrylamide to visulaise monomeric DnaA ensuring DnaA concentration was consistent throughout.* ***A.*** *Analysis of DnaA filaments after oligomerisation using varied DnaA concentration. The triangle above the lanes represents the increasing DnaA concentration used (0.75 µM, 1.5 µM, 3 µM, 6 µM, 12 µM respectively)* ***B.*** *Analysis of DnaA filaments after oligomerisation using various ATP analogues.*

As described in Chapter 1, ATP binding with domain III of DnaA is essential for oligomerisation of monomers into the filament species. Engagement of the γ-phosphate of ATP, via an arginine in the sensor II helix of subdomain IIIb, induces a conformation that allows an arginine from a second DnaA monomer, via the box VII helix of subdomain IIIa, to interact *in-trans* with the same γ-phosphate. In the presence of ADP, dimerisation is prevented by steric clash between the sensor II arginine and box VII helix of the second DnaA molecule. The dependence of ATP on DnaA filament formation was investigated by performing the oligomerisation reaction in the absence of ATP and in addition the presence of ADP, non-hydrolysable γ-imido ATP and GTP. It was determined that presence of ATP or an analogue of ATP is required for observation of DnaA filaments using the crosslinking technique, and that presence of ADP inhibited observation of these high molecular weight DnaA oligomers. These results are displayed in Figure 3.15.

3.4.2 DnaD Inhibits Crosslinking of DnaA Filaments

Having characterised DnaA filament formation in *Bacillus subtilis,* the next phase of this research was to investigate the effect of initiator protein DnaD on stability of the DnaA filament. To achieve this, oligomerisation reactions were undertaken in the presence of DnaD. There are two situations in which stability of the DnaA filament may be altered by presence of ancillary proteins; during oligomerisation of DnaA into the filament, and after DnaA filament formation has occurred. The filament species may be affected differently in each state, therefore, order of addition reactions were utilised where DnaA was incubated with DnaD (15 minutes, 37°C) either prior to or post oligomerisation.

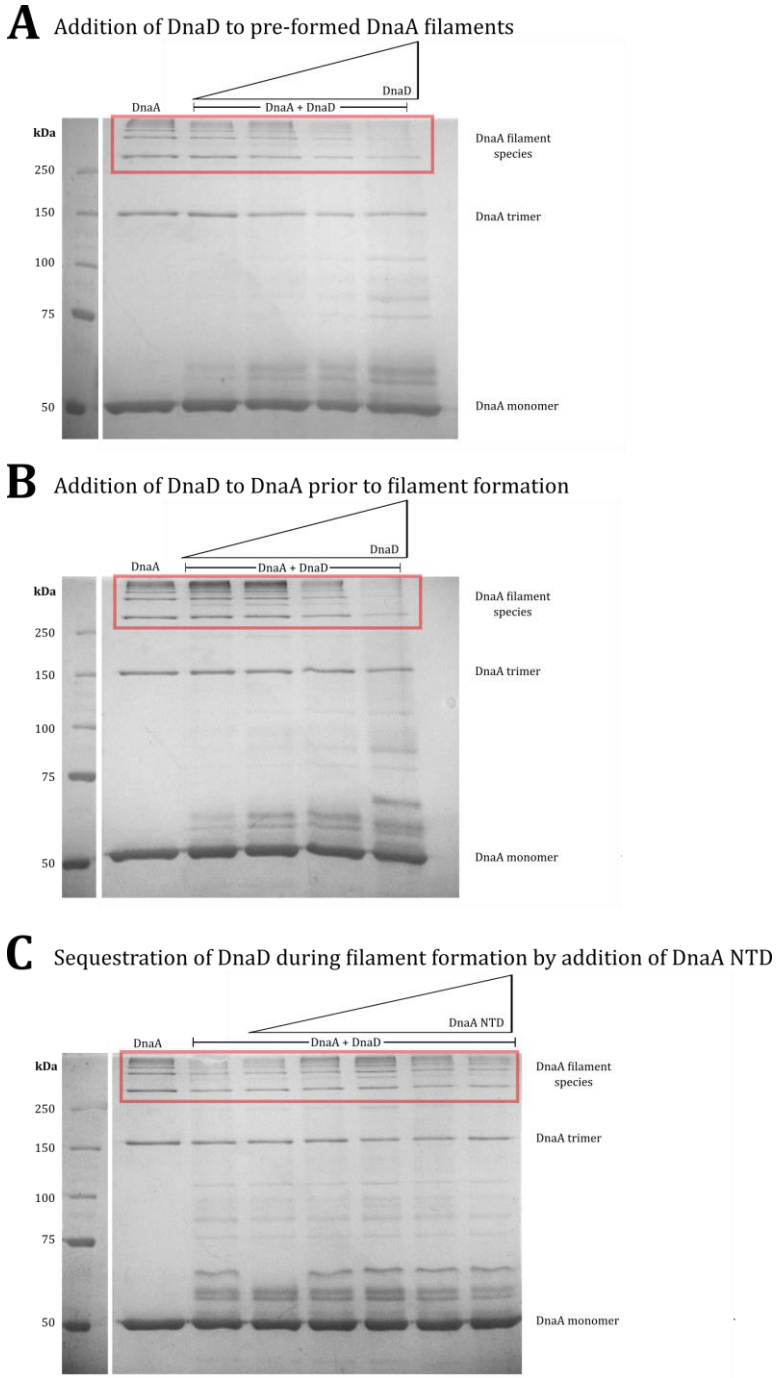


Figure 3.16 *DnaD inhibits crosslinking of DnaA nucleoprotein filaments. SDS-PAGE visualisation of BMOE crosslinked DnaA nucleoprotein filaments in the presence of DnaD. The first lane of each gel contains a molecular weight marker, with values displayed to the left. High molecular weight DnaA species are highlighted. Bands observed between 50 and 150 kDa correspond to impurities within the DnaD stock.* ***A.*** *Analysis of DnaA filaments when incubated with DnaD post oligomerisation. The triangle above the lanes represents the increasing DnaD concentration used (1.5 μM, 3 μM, 6 μM, 12 μM respectively)* ***B.*** *Analysis of DnaA filaments when incubated with DnaD (at various concentrations) prior to oligomerisation.* ***C.*** *DnaA domain I sequesters DnaD from inhibiting crosslinking of DnaA nucleoprotein filaments. DnaA was incubated with DnaD (6 μM) and various concentrations of DnaA domain I (NTD) prior to oligomerisation. The triangle above the lanes represents the increasing DnaA domain I concentration.*

It was revealed that the presence of DnaD inhibited cross-linking of the DnaA filament in a concentration dependent manner (as displayed in Figure 3.16). This observation occurs in the case where DnaD is present during assembly of the DnaA filament and where DnaD is added to the pre-formed DnaA filament. As we cannot assume that inhibition of cross-linking equates to the absence of the DnaA filament species, it is not definitive that DnaD prevents filament formation or promotes disassembly of the DnaA filament. It was reported by Scholefield and Murray[68] that DnaD inhibits formation of the DnaA helical filament and that pre-formed DnaA filaments are resistant to DnaD activity. Our results could support the statement that DnaD prevents filament formation, however, they conflict in the case of DnaD activity on pre-formed DnaA filaments. These experiments were repeated with the addition of *B. subtilis* replication initiator protein DnaB in the place of DnaD, in this case no inhibition of crosslinking was observed over various concentrations of DnaB (data not shown) indicating that the affect observed in the presence of DnaD is specific.

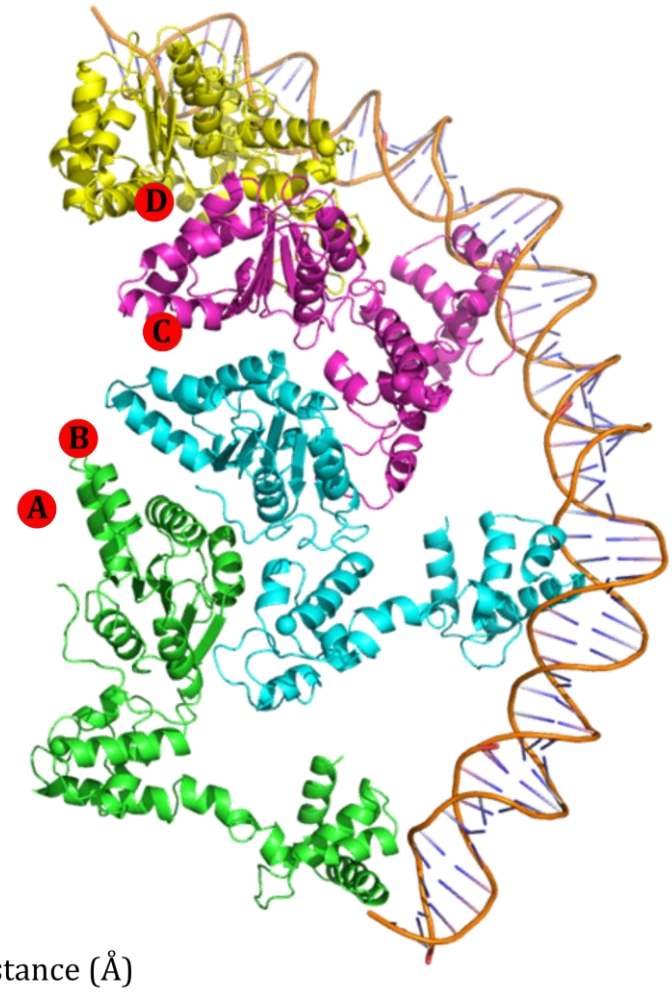
As discussed in Chapter 1, the N-terminal domain of DnaA is involved in the interaction with the DnaC helicase and may be an interaction 'hub' for multiple regulatory interactions. It was proposed that DnaD interacts with DnaA via this domain. The oligomerisation reactions were consequently repeated in the presence of DnaD and the addition of DnaA domain I. This experiment revealed that DnaD dependant inhibition of DnaA crosslinking was reduced by the addition of DnaA NTD (as shown by the level of high molecular weight DnaA species present on the gel) as displayed in Figure 3.16. This suggests that DnaA domain I sequesters DnaD from the DnaA Filament, thereby preventing DnaD induced inhibition of filament cross-linking. These results led us to propose that a direct interaction occurs between DnaD and domain I of DnaA, this was subsequently investigated by high resolution NMR as described in Chapter 5.

3.4.3 Single Molecule FRET

As an explanation for the results discussed in section 3.4.2, we propose that DnaD alters the DnaA filament conformation, increasing the distance between the introduced cysteine residues, rather than preventing assembly or promoting disassembly of the DnaA filament. This is consistent with a model of collaboration between DnaA and DnaD rather than an antagonistic role of DnaD during replication initiation. In collaboration with Timothy Craggs (University of Sheffield) fluorescence resonance energy transfer (FRET) analysis was used to investigate the DnaA nucleoprotein filament stability in the presence of DnaD.

Duderstadt *et al.* reported a model of the DNA-DnaA mini-filament containing four DnaA molecules[82], this was used to assess the interatomic distances of residues A198 and N191 within the filament and determined that labelling at position N191C would be appropriate for FRET experiments. The FRET pair Cy3B and Atto647N were selected for use in these experiments, Cy3B with an excitation wavelength of 559 nm and an emission wavelength of 570 nm, and Atto647N with an excitation wavelength of 644 nm and emission wavelength of 669 nm. For the purpose of this research, a single cysteine residue was introduced at position N191 of DnaA. The $DnaA^{N191C}$ protein was labelled with the fluorescent dyes Cy3B-maleimide and Atto647-maleimide at the introduced cysteine via a conjugation reaction (as described in section 2.3.10 and displayed in Figure 3.17).

The $DnaA^{CC}$, Cy3B-$DnaA^{N191C}$ and Atto647N-$DnaA^{N191C}$ proteins were mixed, prior to the oligomerisation reaction, in a 1:1:2 molar ratio respectively (providing an average of 2 labels per 4 DnaA molecules). The single molecule FRET experiments revealed the presence of a high FRET species (E*=0.7) corresponding to the DnaA nucleoprotein filament. On the addition of DnaD, the FRET signal was shifted to a lower value of E*=0.45 (Figure 3.18), this was consistent when DnaA was incubated with DnaD either prior to or post oligomerisation. These data indicated that a high FRET species was still present under these conditions (corresponding to the DnaA filament) but an altered distance between the fluorophore-labelled molecules was being observed. Hence, these results align with our proposal that on binding of DnaD, the DnaA filament undergoes a conformational change (potentially untwisting of the nucleoprotein filament) and this DnaD-mediated change can be induced on pre-formed DnaA filaments or during their formation in the presence of DnaD.



Distance (Å)

N191	A	B	C	D
A		28.2	50.0	64.3
B			24.2	41.1
C				20.6

Figure 3.17 *The model of the DNA-DnaA mini filament was used to assess the interatomic distances of residues N191 within the filament (courtesy of Dr Timothy Craggs, University of Sheffield). Labelling at position N191 gave distance ranges deemed appropriate for the FRET experiments. Residue N191 within 4 adjacent DnaA molecules are represented by red circles. Model based on X-ray crystal structures of E. coli DnaA domain IV complexed to DnaA-box DNA [PDB - 1J1V] and A. aeolicus DnaA domain III [PDB - 1L8Q].*

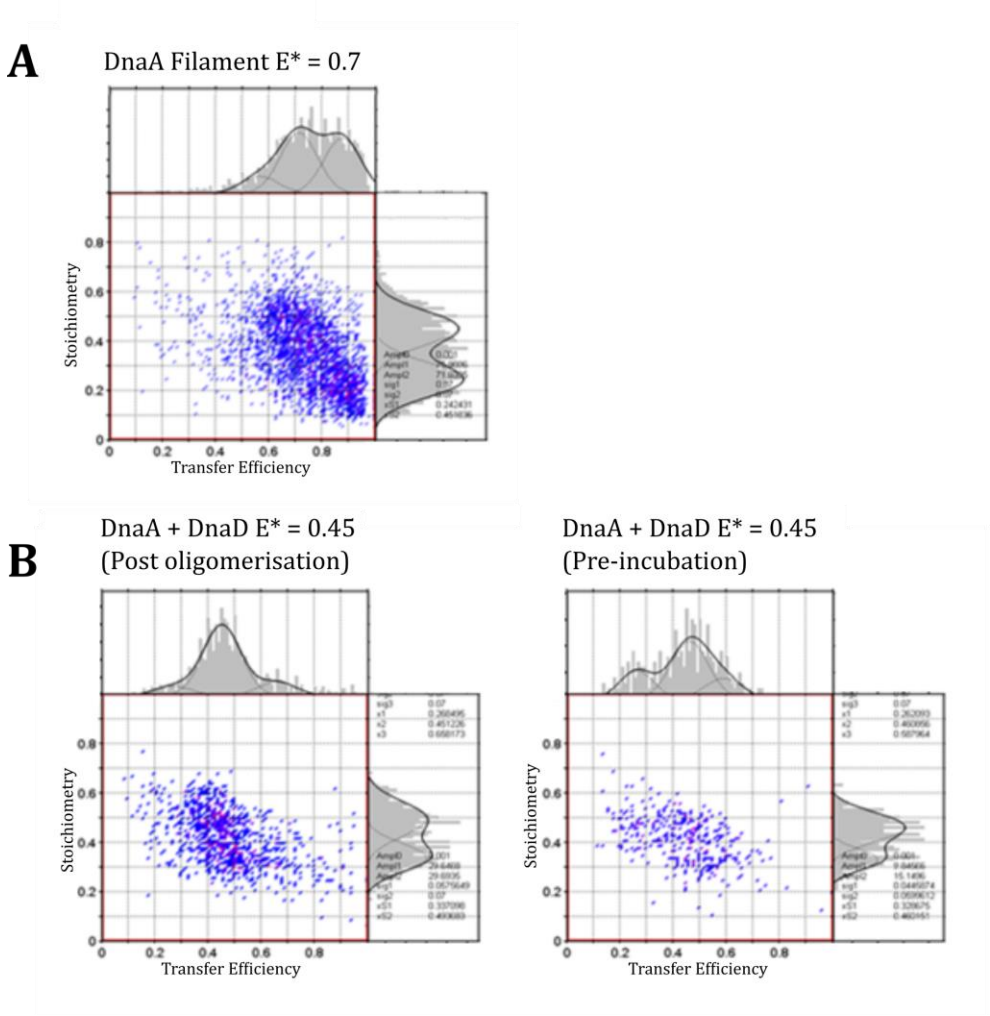


Figure 3.18 *Single molecule FRET analysis (courtesy of Dr Timothy Craggs, University of Sheffield) of the DnaA nucleoprotein filament. Cy3B and Atto647N labelled DnaA^{N191C} was mixed with unlabelled protein (1:1:2 respectively) prior to the oligomerisation reaction. FRET-stoichiometry histograms are displayed with data fitted to Gaussian kernel density estimation.* ***A.*** *The DnaA filament produces a high-FRET species with the signal of E*=0.7.* ***B.*** *In the presence of DnaD the FRET signal is shifted from E*=0.7 to E*=0.45 indicating a conformational change within the DnaA filament. This response is observed when DnaA is incubated with DnaD prior to oligomerisation (left) and on addition of DnaD to pre-formed DnaA filaments (right).*

DnaD has the potential to interact simultaneously with both DnaA and the dsDNA wrapped around the outside of the DnaA nucleoprotein filament. It is unclear and of interest whether both interactions are required to alter the conformation of the DnaA filament. Therefore, the FRET experiments were repeated using the truncated DnaD construct (DnaD 1-196) lacking the C-terminal region essential to DNA binding. Similar results were obtained using this construct indicating a conformational change within the DnaA filament occurs independently of DNA binding by DnaD (Figure 3.19). The same FRET shift was seen (from E*=0.7 to E*=0.45) in the presence of truncated

DnaD suggesting that a similar extent of conformational change, or untwisting, is observed as for the full-length protein.

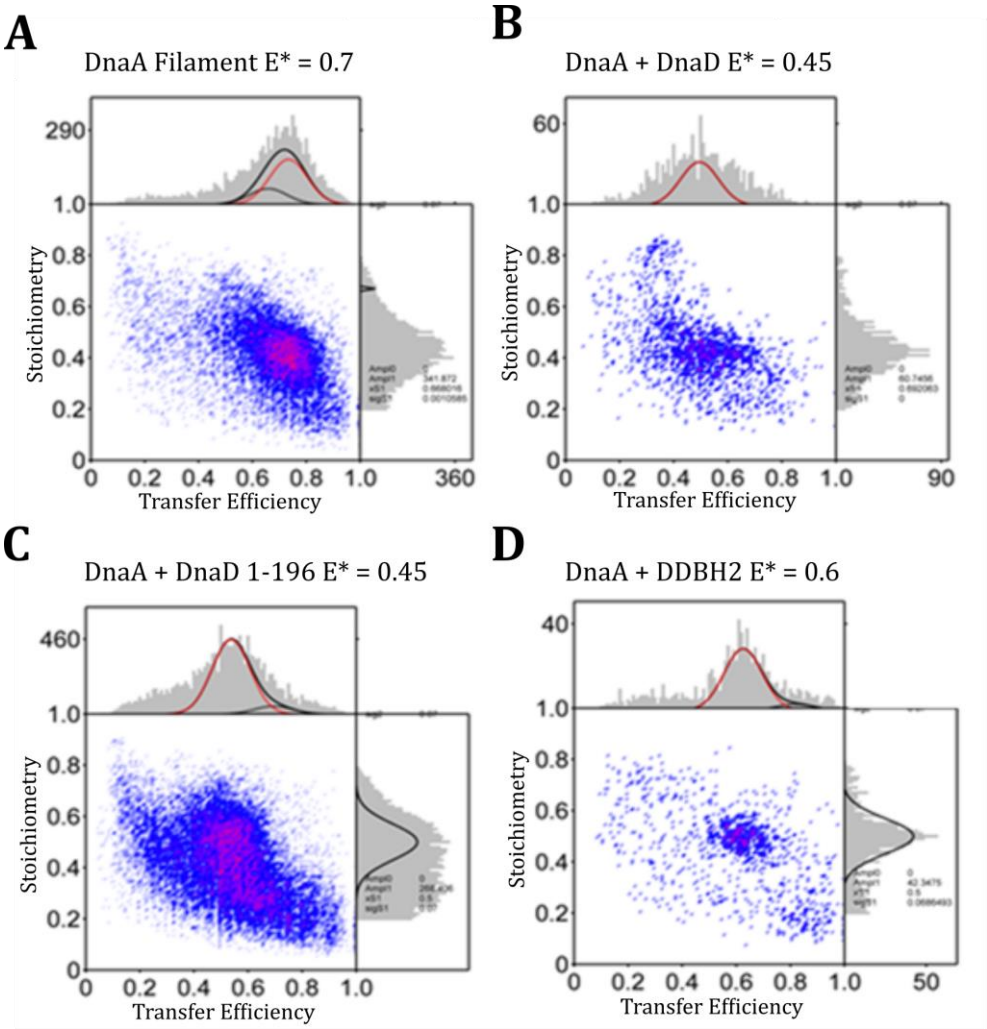


Figure 3.19 *Single molecule FRET analysis (courtesy of Dr Timothy Craggs, University of Sheffield) of the DnaA nucleoprotein filament. Cy3B and Atto647N labelled DnaA^{N191C} was mixed with unlabelled protein (1:1:2 respectively) prior to the oligomerisation reaction. FRET-stoichiometry histograms are displayed with data fitted to Gaussian kernel density estimation. Data displayed in A and B (originally displayed in Figure 3.18) are included as a comparision to the results displayed in C and D.* ***A.*** *The DnaA filament produces a FRET signal of E*=0.7.* ***B.*** *In the presence of DnaD the FRET signal is shifted from E*=0.7 to E*=0.45 indicating a conformational change within the DnaA filament.* ***C.*** *In the presence of the truncated DnaD 1-196 construct the FRET signal is shifted from E*=0.7 to E*=0.45 indicating the conformational change is independent of DNA binding by DnaD.* ***D.*** *In the presence of isolated DnaD DDBH2 domain the FRET signal is shifted from E*=0.7 to E*=0.6 indicating a conformational change (to a much smaller extent) can be induced without the presence of oligomeric DnaD.*

As described in Chapter 1, yeast-two-hybrid experiments identified the interaction between DnaA and the C-terminal region of DnaD (the DDBH2 domain)[106]. The FRET experiments were repeated with addition of the isolated DDBH2 domain (in place of full length DnaD) to reveal a shift to E*=0.6 from E*=0.7 (Figure 3.18), this suggests that a conformational change within the DnaA filament can be induced by the isolated DDBH2 domain but to a much smaller extent than with the full length DnaD or the DnaD 1-196 construct. The N-terminal DDBH1 domain of DnaD is required for self-oligomerisation of DnaD, into scaffold like formations[100], thus it is of interest that a conformation change within the DnaA filament can also be induced in the absence of oligomeric DnaD. It is of importance to note that a limitation of the preliminary FRET data presented is the lack of negative control such as addition of DnaB or BSA in the place of DnaD as this would remove any doubt that the affect on the DnaA filament is non specific.

3.5 Summary

A reliable expression and purification strategy has been developed for *B. subtilis* DnaA and DnaD recombinant proteins, for both the full-length constructs and individual target domains. This has provided material of a high quality for the biochemical and biophysical analysis of the DnaA-DnaD interaction described in section 3.4 and to follow in subsequent chapters.

Investigation of the DnaA nucleoprotein filament has shown that these high molecular weight species can be obtained *in vitro* and observed using a cross-linking technique combined with SDS-PAGE. DnaD reduces crosslinking observed on addition to pre-formed DnaA filaments as well as when present during filament formation. The inhibitory effect of DnaD can be reduced by addition of DnaA domain I, to the oligomerisation reaction, through sequestration of the DnaD molecules. This confirmed that an interaction occurs between DnaD and domain I of DnaA.

Single molecule FRET experiments revealed that DnaD induces an altered confirmation of the DnaA nucleoprotein filament and that the conformational change is independent of the DNA binding activity of DnaD. These results align with the inhibition of DnaA crosslinking observed in the presence of DnaD, as distortion of the conformation will disrupt the location of the introduced cysteine residues within the

DnaA filament. We propose that the observed conformational change corresponds to untwisting of the DnaA nucleoprotein filament, by DnaD, as a mechanism to aid DnaA mediated melting of the *oriC* DUE.

DnaD forms large structures (DDBH1 domain activity) through self-oligomerisation. As discussed in Chapter 1, we proposed DnaD acts as a structural scaffold potentially involved in bridging of the two DnaA nucleoprotein filament complexes located either side of the *dnaA* gene within the chromosomal origin of replication. The looped DNA structure has been observed by electron microscopy[102] and distortion of the DnaA filament structure would be expected to form such a conformation.

Chapter 4

NMR Assignment of DnaA Domain I

4.1 Introduction

NMR spectroscopy is a well-established tool for determining protein structure that is capable of solving structures to atomic resolution, has the ability to accurately measure the dynamic properties of proteins, and can be used to probe the process of protein folding[113, 114]. The expansion of the range of applications in the field of nuclear magnetic resonance originated from the development of pulsed Fourier transform NMR spectroscopy by Ernst and Anderson[115] and the conception of multidimensional NMR spectroscopy[116, 117]. Further development followed the introduction of three-dimensional NMR experiments in the 1990s. Addition of a third dimension provided far greater resolution compared to the crowded proton spectral window accessible by two dimensional experiments. Although 3D experiments are inherently less sensitive than analogous 2D experiments, genetic engineering techniques allow the production of proteins labelled with NMR observable stable isotopes (^{13}C and ^{15}N) to overcome this issue[118]. Chemical shifts for the Cα and Cβ nuclei provide additional structural information relating to backbone angles, and therefore secondary structure content, a further advantage to 3D experiments[119, 120]. Low inherent sensitivity was formerly a limiting factor for probing the structure of biological macromolecules. This was due to a size limitation cause by two technical barriers; large molecules have slower tumbling rates and thereby shorter signal relaxation times, and increased molecular weight introduces more complexity to the spectrum. Advances in hardware, such as high field magnets (>600 MHz) and

sensitivity increasing cryoprobes, along with the expansion of theoretical and experimental capabilities have allowed the study or larger proteins up to 50 kDa size[121].

Currently, NMR spectroscopy rivals X-ray crystallography for determining the three-dimensional structure of biological macromolecules at atomic resolution. NMR was the method of choice to study the DnaA-DnaD interaction via use of the isolated domains involved (DnaA domain I and the DnaD DDBH2 domain). The structure of *B. subtilis* DnaA domain I, in complex with inhibitor of DNA replication SirA, was previously determined by X-ray crystallography[107]. Utilising NMR spectroscopy, comprehensive assignment of the DnaA domain I amide backbone has enabled investigation of DnaD binding on an atomic scale and facilitated characterisation of the DnaA-DnaD binding interface.

4.2 ^{1}H NMR

One-dimensional NMR techniques are extremely useful for providing structural information and conformational analysis of small molecules (<1000 kDa). However, these have limited application in the complex spectra of biological macromolecules, which have several hundred ^{1}H signals that appear highly overlapped in a limited spectral region. This experiment, having excellent reproducibility, is useful for identifying contaminants within the protein sample from either unwanted proteins or buffer components. The spectrum of a folded protein will produce chemical shifts that are widely dispersed, due to ring-current effects on methyl groups, and downfield shifted amide and Cα signals. In contrast, an unfolded protein spectrum will exhibit signals clustered around the amide, Cα and methyl proton regions with characteristic random coil shifts[122]. These represent an average value of all possible conformations that an amino acid can adopt when the protein state is highly dynamic and lacking secondary structure[123]. Figure 4.1 shows the ^{1}H NMR spectrum for DnaA domain I. Good dispersion of signals and the upfield shifted methyl signal (the result of packing against aromatic side-groups) indicate the protein is folded.

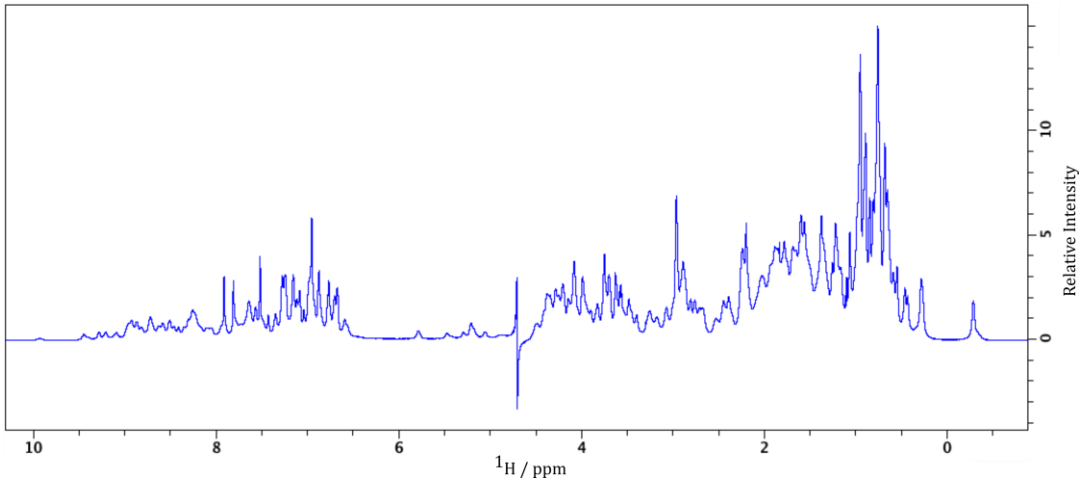


Figure 4.1 *^{1}H NMR spectrum for ^{15}N DnaA domain I. Dispersion of peaks and the upfield shifted methyl signal indicate the protein is folded. As indicated by the doublet just below 0 ppm the protein contains two unique spin ½ nuclei. The water signal lies at approximately 4.6 ppm. No indication of contamination is apparent in the spectrum.*

4.3 ^{15}N-HSQC

The ^{15}N HSQC (heteronuclear single quantum coherence) spectrum is the most standard 2D NMR experiment. During the experiment, magnetisation is transferred from the amide proton to the attached nitrogen nuclei via J-coupling and using an INEPT (insensitive nuclei enhanced by polarisation transfer) sequence. The chemical shift is evolved on the nitrogen, during the time delay t_1, then magnetisation is transferred back to the amide proton for detection (Figure 4.2). The acquisition time is designated t_2 and during the experiment the indirect evolution time t_1 is increased incrementally[124]. Fourier transform with respect to the two time domains generates a two-dimensional spectrum with two independent frequency dimensions. All N-H correlations are shown, these consist mainly of the backbone amide groups but the HSQC also contains signals from the amides of the asparagine, glutamine, tryptophan and in principal arginine side chains. Generally the first heteronuclear experiment to be undertaken, the ^{15}N HSQC provides an amide backbone 'fingerprint' for the protein characterised by it amide resonances. The chemical shift dispersion in ^{1}H and ^{15}N is a good indicator of folded structure and hence can provide an indication of the level of structural disruption caused by mutagenesis when compared to the wild type spectrum.

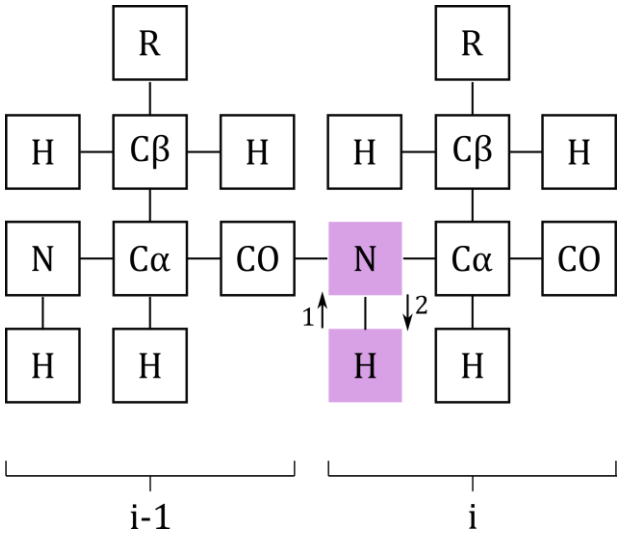


Figure 4.2 *The ^{1}H-^{15}N HSQC Experiment. Magnetisation is transferred from the amide proton to the attached nitrogen, the chemical shift is evolved on the nitrogen then magnetisation is transferred back to the amide proton for detection.*

4.3.1 Optimisation of NMR Conditions

The ^{15}N HSQC experiment is useful to establish suitability of the selected protein for analysis by NMR. It is used to determine if the protein is folded and the quality of the spectrum is assessed to ensure adequate coverage (in terms of resolution, dispersion and lineshape) of the amide backbone signals is achieved. It is also required that the protein remains stable under NMR conditions for several weeks while collecting three-dimensional data. These factors are considered before proceeding on to three-dimensional experiments, where significantly more expensive isotopic labelling is required. Optimisation of NMR conditions was undertaken using the ^{15}N HSQC experiment.

Full coverage of the expected amide backbone signals for DnaA domain I was not initially achieved in the ^{15}N HSQC experiment. Titration experiments for salt concentration, temperature and pH of the protein solution were undertaken (individually and in combination) to achieve the highest visibility in terms of the number of backbone amide signals observed. However, a small number of signals remained particularly weak and 89 of the expected 92 signals (including asparagine and glutamine side chain amide peaks) were observed for this protein construct. The final buffer conditions chosen for observation of the DnaA domain I construct consisted of; 50 mM potassium phosphate (K_2HPO_4/KH_2PO_4), 25 mM NaCl, pH 7.4. Protein

stability was monitored over a period of 1 to 6 weeks where no apparent protein degradation (doubling of peaks and the appearance of new peaks in the lower-mid region of the spectrum is indicative of protein degradation) was visible within the ^{15}N HSQC spectrum.

4.4 Triple Resonance Experiments

Though they are inherently less sensitive than two-dimensional NMR experiments, the addition of a third dimension provides greater resolution and additional information on secondary structure can be gained using ^{13}C chemical shifts[119, 120]. The three-dimensional experiments defined here rely on magnetisation transferred primarily though the J-coupling of adjacent atoms and correlation of the chemical shifts of nuclei through scalar coupling. The J-coupling parameter is valuable for magnetisation transfer due to the uniformity of the magnitude and relative freedom from effects of conformation within the backbone. The transverse relaxation time (T_2) determines the linewidth and is dependent on factors relating to protein structure such as tumbling rate, internal mobility and local conformation. This makes the T_2 a crucial parameter as the one-bond J-coupling must be significantly higher than the linewidths for efficient magnetization transfer, as is the case for most globular proteins up to 30 kDa.

The backbone resonance assignment methodology relies on ^{13}C, ^{15}N labelled protein and requires the CβCαNH and CβCα(CO)NH experiments. However, insufficient quality of spectra often requires additional experiments for full assignment such as the HNCα, HN(CO)Cα, HNCO and HN(Cα)CO experiments. Larger proteins, greater than 150 residues, give lower quality spectra due to the slower tumbling rate observed and for proteins greater than 250 residues deuteration of the protein may be necessary to obtain spectra of sufficient quality for backbone resonance assignment.

An inherent sensitivity problem was initially found with some of the triple resonance experiments on DnaA domain I. This problem was likely associated with fast relaxation of the Cα nuclei and subsequently the direct pathway pulse sequences, as opposed to the 'out and back' sequences, were used to acquire the CβCαNH, CβCα(CO)NH and (HCα)CONH experiments. Analysis of triple resonance experiments enables sequential assignment of the amide backbone residues as these are linked by

magnetisation transfer pathways, amino acid types also show characteristic Cα and Cβ chemical shifts.

4.4.1 CβCαNH

During the CβCαNH experiment, the ^{1}H magentisation is evolved on the H-Cα and H-Cβ simultaneously and transferred to the respective attached ^{13}C atom then from Cβ to the Cα. From the Cα magnetisation can be transferred simultaneously to the N_i or N_{i-1} and finally to the amide proton (N-H). As magnetisation is transferred to N_i from both $C\alpha_i$ and $C\alpha_{i-1}$ for each NH peak there will be two Cα and Cβ peaks visible, relating to [$C\alpha_i$, $C\beta_i$] and [$C\alpha_{i-1}$, $C\beta_{i-1}$]. However, coupling to the directly bonded Cα and Cβ is stronger giving these peaks a higher intensity in the spectrum.

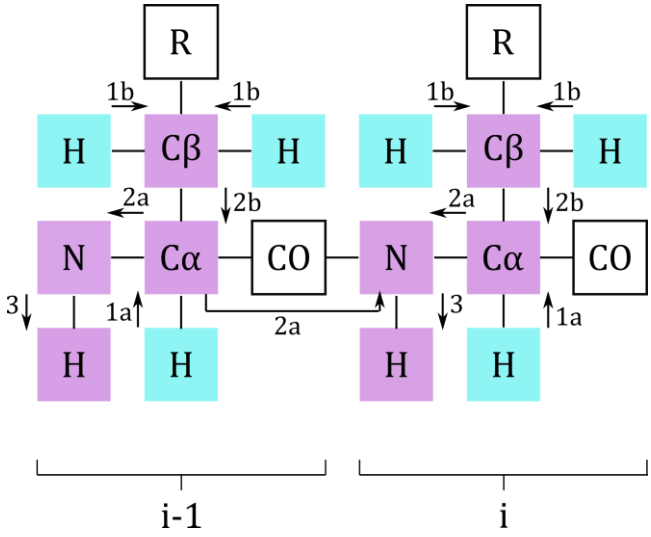


Figure 4.5 *The CβCαNH experiment. Arrows indicate magnetisation route used for acquisition of the NMR experiment. Detected nuclei are shown in purple while nuclei that transfer magnetisation are highlighted in blue[125].*

Using direct pathway pulse sequences improved spectra obtained compared to the 'out and back' method, however, poor magnetisation transfer was still apparent by missing signals in the CβCαNH spectrum. Therefore, additional experiments were used to overcome this issue. These included the HNCα experiment (signal is evolved on H-N, magnetisation passed to the N and then via J-coupling to $C\alpha_i$ or $C\alpha_{i-1}$ and back via the same route for detection) and the CβCαNH experiment optimised for Cβ signals.

4.4.2 CβCα(CO)NH

For the CβCα(CO)NH experiment, the signal is again evolved on the H-Cα and H-Cβ simultaneously and transferred to the respective carbon atoms then from Cβ to the Cα. From the Cα magnetisation is transferred to the CO (but not detected) through to the nitrogen and finally to the amide proton (N-H). As the signal is not detected on the CO, for each NH peak there will be a Cα and Cβ peaks visible relating to [$C\alpha_{i-1}$, $C\beta_{i-1}$].

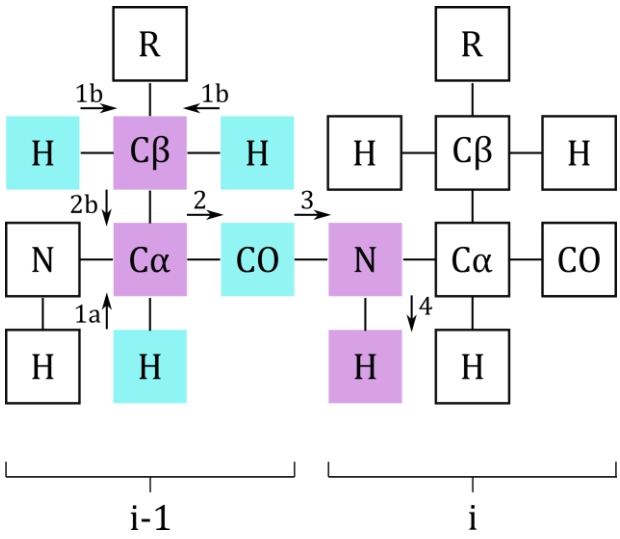


Figure 4.6 *The CβCα(CO)NH experiment. Arrows indicate magnetisation route used for acquisition of the NMR experiment. Detected nuclei are shown in purple while nuclei that transfer magnetisation are highlighted in blue[126, 127].*

4.4.3 HNCO

During the HNCO experiment, the signal is evolved on the H-N and magnetisation transferred to the nitrogen then through to the attached CO_{i-1}. Magnetisation is returned to the amide proton via the same route and the chemical shift is evolved on all three nuclei for detection. The HNCO is the most sensitive triple resonance experiment. In addition to backbone CO-NH correlations, asparagine and glutamine correlations are also visible in the spectrum.

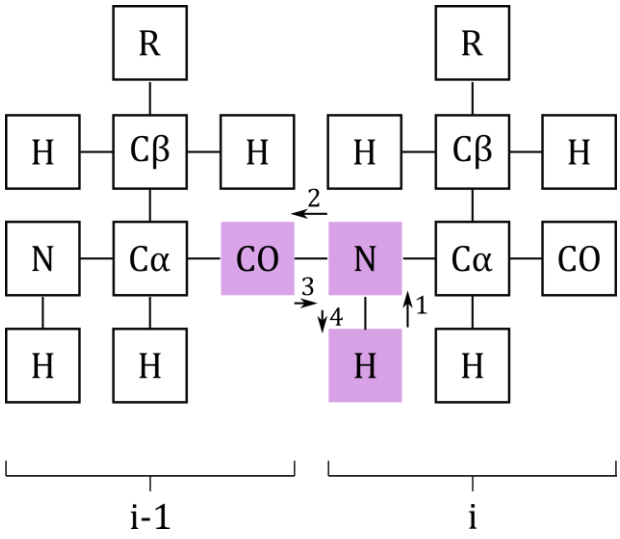


Figure 4.3 *The HNCO experiment. Arrows indicate magnetisation route used for acquisition of the NMR experiment[128-130].*

4.4.4 (HCα)CONH

For the (HCα)CONH experiment, the signal is evolved on the H-Cα and transferred to the respective carbon (but not detected). From the Cα magnetisation is transferred simultaneously to the CO_i or CO_{i-1} through to the nitrogen and finally to the amide proton (NH). As magnetisation is transferred to N_i from both CO_i and CO_{i-1} for each NH peak there will be two carbonyl peaks visible, relating to [CO_i] and [CO_{i-1}].

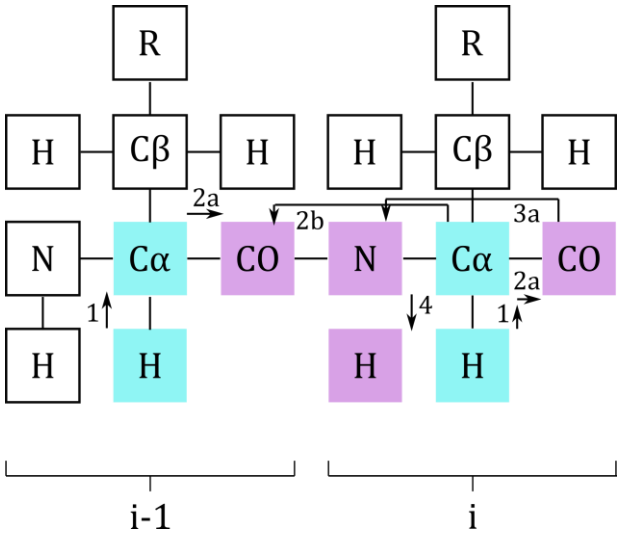


Figure 4.4 *The (HCα)CONH experiment. Arrows indicate magnetisation route used for acquisition of the NMR experiment. Detected nuclei are shown in purple while nuclei that transfer magnetisation are highlighted in blue[130, 131].*

Experiment	Observed Carbon Nuclei
CβCαNH	Cα(i-1), Cβ(i-1), Cα(i) Cβ(i)
CβCα(CO)NH	Cα(i-1), Cβ(i-1)
HNCO	CO(i-1)
(HCα)CONH	CO(i-1), CO(i)

Table 4.1 *Summary of the carbon nuclei observed during each of the triple resonance experiments described in sections 4.4.1 – 4.4.4. Notations 'i' and 'i-1' are indicative of the atom position in relation to the detected amide proton within the amino acid sequence.*

4.5 Sequential Backbone Assignment

Standard triple resonance backbone assignment is based on the CβCαNH and CβCα(CO)NH spectra, where the CβCαNH correlates each NH group with the Cα and Cβ of residue i and i-1 within the protein sequence and the CβCα(CO)NH correlates the NH group to the Cα and Cβ of the i-1 residue only. These matched pairs of experiments can be overlaid for any given NH to link the Cα and Cβ of the i and i-1 residues and thereafter trace connectivity of the amide backbone. The same principle can be applied to the HNCO and (HCα)CONH, or equivalent experiments, where the HNCO correlates each NH group to the CO of residue i-1 only and the (HCα)CONH correlates the NH group to the CO of the i and i-1 residue. The Cα and Cβ chemical shifts are more dispersed than the CO shifts and using the corresponding spectra an increased number of links can be made between residues (up to 4 compared to 2) therefore greater resolution is achieved using these spectra. Cα and Cβ chemical shifts are also more characteristic of residue type than the carbonyl chemical shift.

The first stage of sequential assignment involves eliminating ^{15}N HSQC peaks corresponding to side chain amides of asparagine and glutamine residues. These are identified by pairs of peaks at the same nitrogen, but different proton, chemical shifts. Navigating from each NH peak in the ^{15}N HSQC spectrum to the corresponding nitrogen and proton frequency in the CβCαNH, CβCα(CO)NH, HNCO and (HCα)CONH spectra identifies peaks representing the Cα and Cβ of the i and i-1 residue (CβCαNH), Cα and Cβ of the i-1 residue (CβCα(CO)NH), CO of the i-1 residue (HNCO) and CO of the i and i-1 residue ((HCα)CONH). These signals can be assigned to the appropriate spin systems, arbitrarily given to each NH peak in the ^{15}N HSQC spectrum, then sequential linking of the amide peaks is achieved both manually and using the semi-automated protein sequence assignment tool within the CCPN Analysis software[132]. Residues glycine, alanine, serine and threonine can be identified by their characteristic Cα and Cβ

chemical shifts. Assigning these residue types identifies possible locations within the protein sequence for the linked fragments by elimination, along with use of the prediction tool within CCPN Analysis[132]. Several cycles of analysis are involved working from the most confident assignments to narrow down options for the less confident regions which can then be assigned by process of elimination.

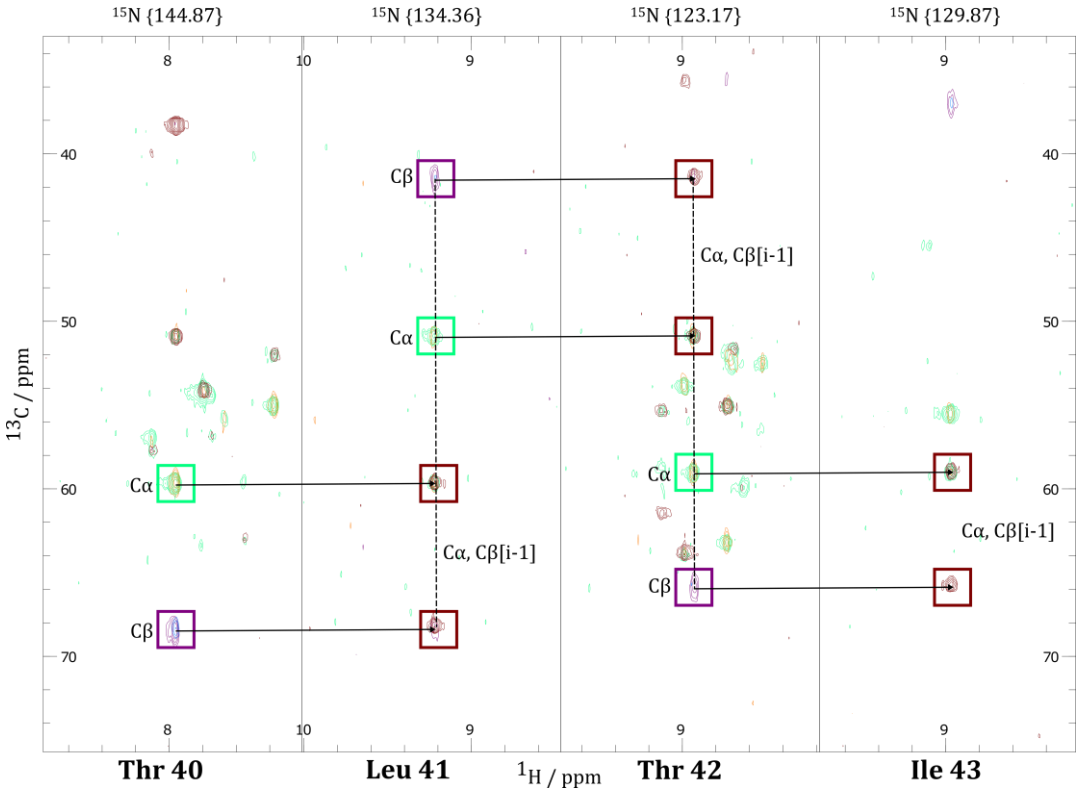


Figure 4.7 *Overlaid spectra of CβCα(CO)NH (brown), CβCαNH (blue and orange), HNCα (green) and Cβ optimised CβCαNH (purple). Strips represent various ^{15}N chemical shifts that contain corresponding peaks used to establish amide backbone connectivity.*

Figure 4.7 shows the sequential links made between the backbone NH signals from the ^{15}N HSQC spectrum and the Cα and Cβ signals from the CβCαNH, CβCα(CO)NH, HNCα and the CβCαNH experiment optimised for Cβ signals. For the majority of assigned residues links are made from the amide frequency to the Cα and Cβ of the proceeding residue within the protein sequence. The HNCO and (HCα)CONH connectivity between carbonyl signals also supported assignment. In theory using overlaid pairs of experiments should allow total assignment of the amide backbone resonances but signal overlap and insufficient magnetisation transfer result in the

requirement for multiple experiments and additional techniques before full assignment can be achieved.

4.6 ^{14}N Selective Unlabelling

The first step towards three-dimensional structure determination of proteins by NMR involves sequence specific resonance assignments of its backbone and sidechain ^{1}H, ^{13}C and ^{15}N nuclei, as described in section 4.5. However, in the case of spectral overlap or larger molecular weight proteins, it often becomes necessary to identify spin systems corresponding to the different amino acid types selectively to aid resonance assignments. The traditional approach to selective labelling of a specific amino acid type is achieved by feeding the host organism with the desired isotopically labelled (^{13}C/^{15}N/^{2}H) amino acid, while supplying the rest of the amino acids in unlabelled form[133, 134]. The drawback of this method is the use of expensive enriched amino acids resulting in undesirably high costs. An alternative approach is the method of amino acid selective 'unlabelling' (or reverse labelling) which involves selective unlabelling of specific amino acid types against a uniformly ^{13}C/^{15}N labelled background. This technique can also be used to aid backbone connectivity through linking of the amide resonance for the selectively unlabelled residue with its C-terminal neighbour using HN-detected double and triple resonance spectra[135].

To aid assignment of the DnaA domain I ^{15}N HSQC spectrum, the amino acid specific unlabelling technique was utilised[135, 136]. ^{14}N selective unlabelling is suitable only for amino acids not subject to cross-metabolism during biosynthesis due to isotope scrambling (as displayed in Table 4.2). Standard expression for a ^{15}N enriched sample was followed but in addition the media was supplemented with a source of the ^{14}N L-amino acid selected for unlabelling. Uniform unlabelling can be achieved using this technique as *E. coli* cells will use the supplemented amino acid source rather than synthesize their own. This method aided assignment of ambiguous peaks within the ^{15}N HSQC spectrum in addition to verification of the existing assignment.

Amino Acid	Amino Acid Cross Metabolised With
Arginine	-
Asparagine	-
Aspartate	Phenylalanine
	Tyrosine
	Asparagine
	Lysine
Glutamate	Proline
Histidine	-
Isoleucine	Leucine
	Valine
Leucine	Isoleucine
	Valine
Lysine	-
Phenylalanine	Tyrosine
Valine	Leucine
	Isoleucine
Tyrosine	Phenylalanine

Table 4.2 *Isotope scrambling observed when selectively unlabelling amino acids due to cross metabolism. Information taken from Krishnarjuna et al.[135].*

Comparison of the ^{15}N enriched HSQC and the ^{14}N unlabelled spectra allowed identification of unlabelled residues, this being due to the absence of signals corresponding to ^{14}N atoms in the HSQC spectrum (Figure 4.8). Identification of residue type for an amide resonance significantly reduces ambiguity when assigning peaks. This technique was applied to the naturally occurring amino acids that do not undergo cross-metabolism, lysine, arginine, asparagine and histidine, to allow uncomplicated interpretation of the corresponding spectra[135]. These four amino acid types account for approximately 18% of the DnaA domain I sequence therefore identifying these in the ^{15}N HSQC would be highly beneficial to assignment. The spectra corresponding to ^{14}N Lys, ^{14}N Arg, ^{14}N Asp and ^{14}N His DnaA domain I samples are shown in Figures 4.9, 4.10, 4.11 and 4.12 respectively where identified peaks are labelled in red.

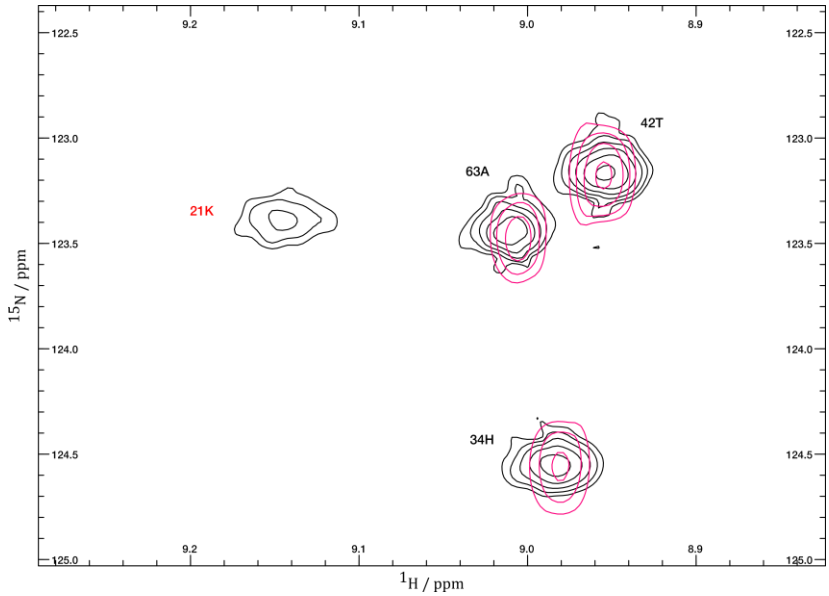


Figure 4.8 *The result of ^{14}N Lysine incorporation into a ^{15}N labelled sample of DnaA domain I. The ^{14}N Lys incorporated spectrum (pink) overlaid with the uniformly labelled ^{15}N HSQC spectrum (black). Peaks corresponding to lysine residues are identified by their absence when the spectra are compared.*

Residue Type	Unlabelled HSQC Peaks	Assigned Peaks	Missing Peaks
Lysine	6	6	0
Arginine	2	2	0
Asparagine	3	3	0
Histidine	2	2	1

Table 4.3 *Summary of ^{14}N selective labelling experiments used for residue identification in the ^{15}N HSQC spectrum. A histidine residue within the protein sequence remains unaccounted for due to absence of the corresponding backbone amide peak observed from the HSQC experiment.*

The combination of ^{14}N leucine, isoleucine and valine was also used for a selective unlabelling experiment as they corresponded to a significant portion of unassigned peaks in the ^{15}N HSQC spectrum. These amino acids will undergo cross metabolism to each other, leading to isotope scrambling, which complicates analysis. This experiment worked only to a degree as the backbone amide signals for isoleucine, leucine and valine residues were still present in the spectrum, however, they appeared as noticeably weaker signals compared to the fully ^{15}N enriched spectrum. These results were useful in aiding assignment by reducing ambiguity, these three amino acid types accounting for approximately 24% of the DnaA domain I sequence. The spectrum corresponding to the triply unlabelled protein with ^{14}N Leu, ^{14}N Ile, ^{14}N Val is shown in figure 4.13.

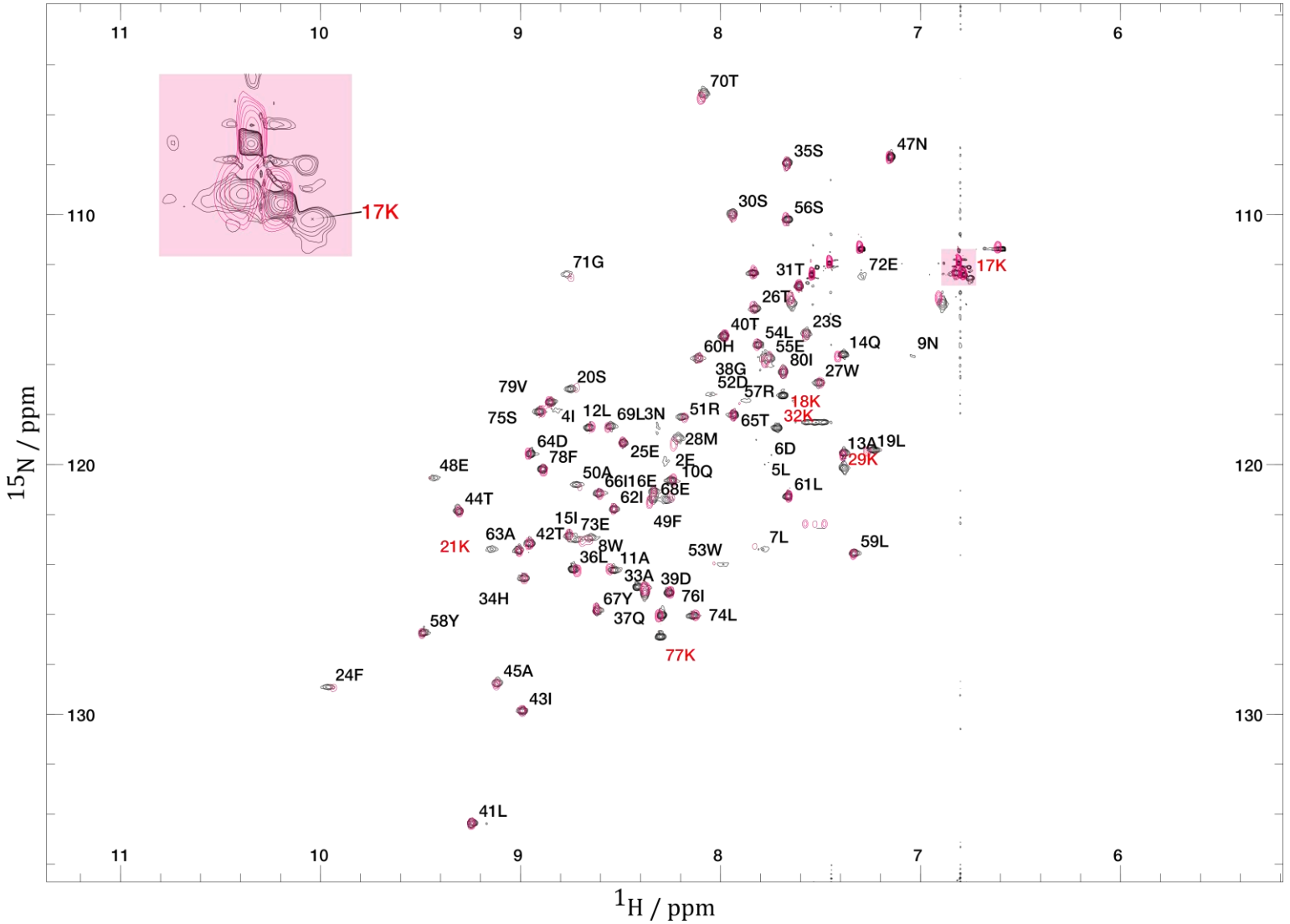


Figure 4.9 *^{15}N HSQC spectral overlay of fully enriched ^{15}N DnaA domain I (black) and ^{14}N Lys selectively labelled DnaA domain I (pink). Lysine residues are highlighted in red, all expected peaks were observed and assigned.*

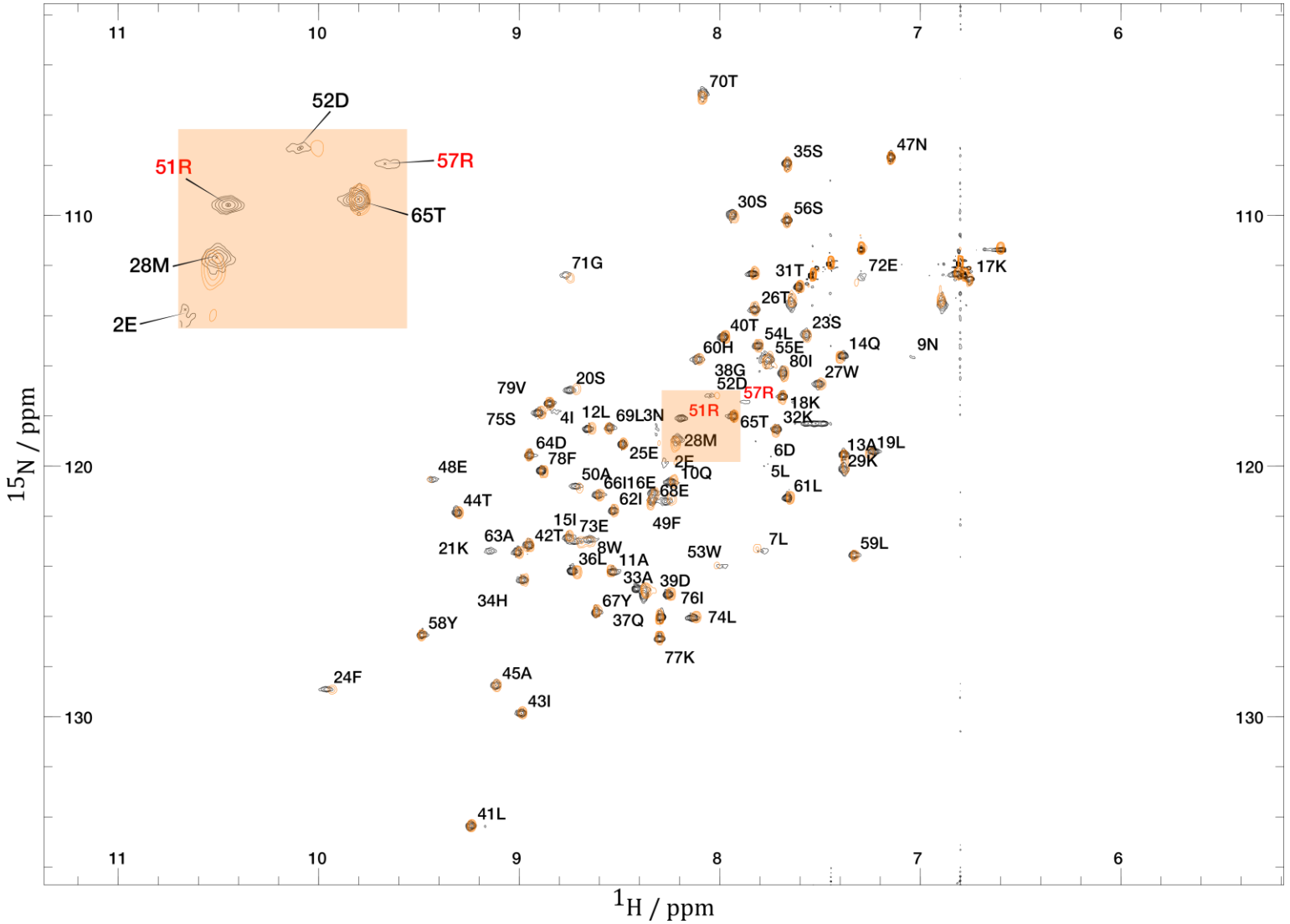


Figure 4.10 *^{15}N HSQC spectral overlay of fully enriched ^{15}N DnaA domain I (black) and ^{14}N Arg selectively labelled DnaA domain I (orange). Arginine residues are highlighted in red, all expected peaks were observed and assigned.*

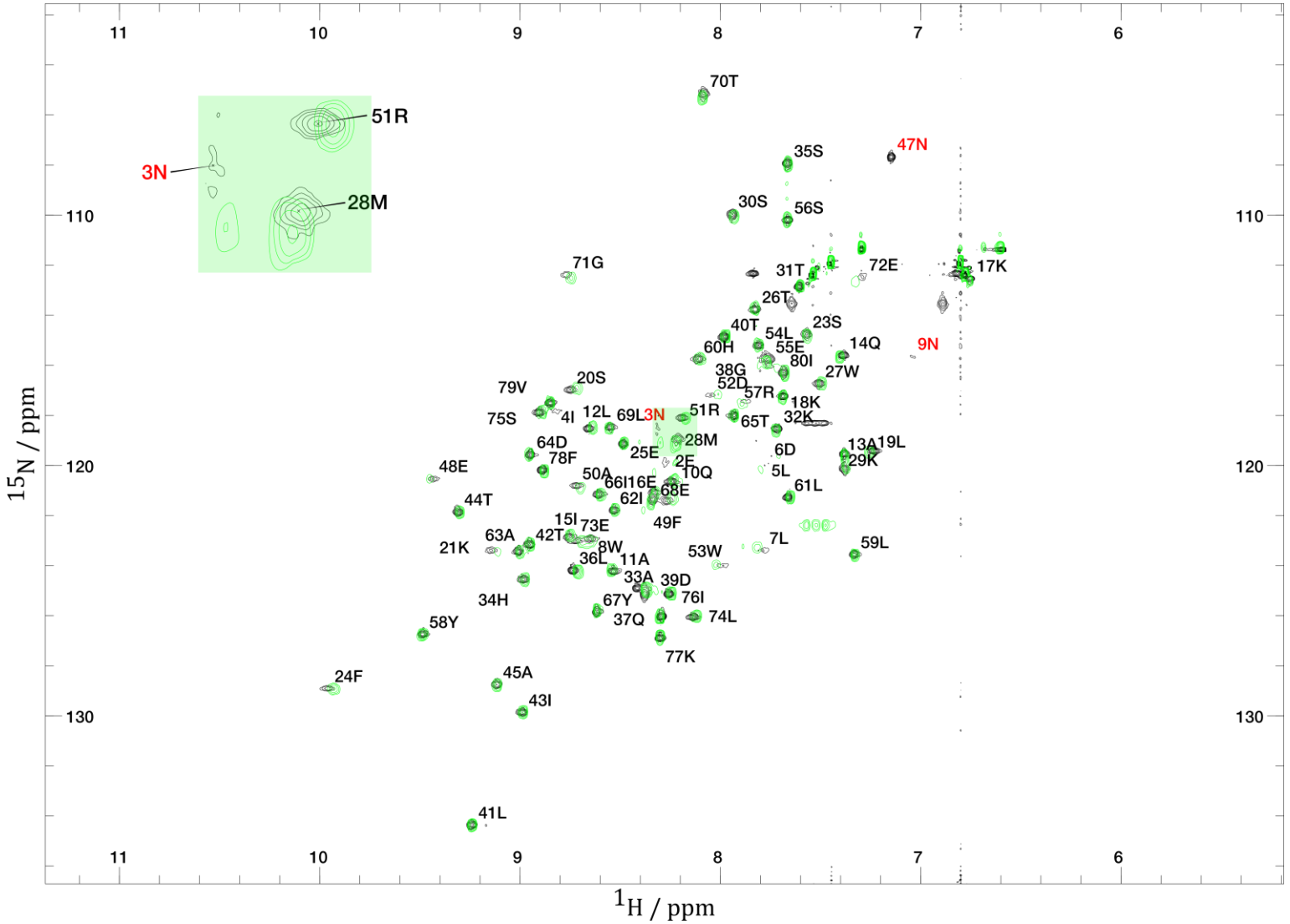


Figure 4.11 *^{15}N HSQC spectral overlay of fully enriched ^{15}N DnaA domain I (black) and ^{14}N Asn selectively labelled DnaA domain I (green). Asparagine residues are highlighted in red, all expected peaks were observed and assigned.*

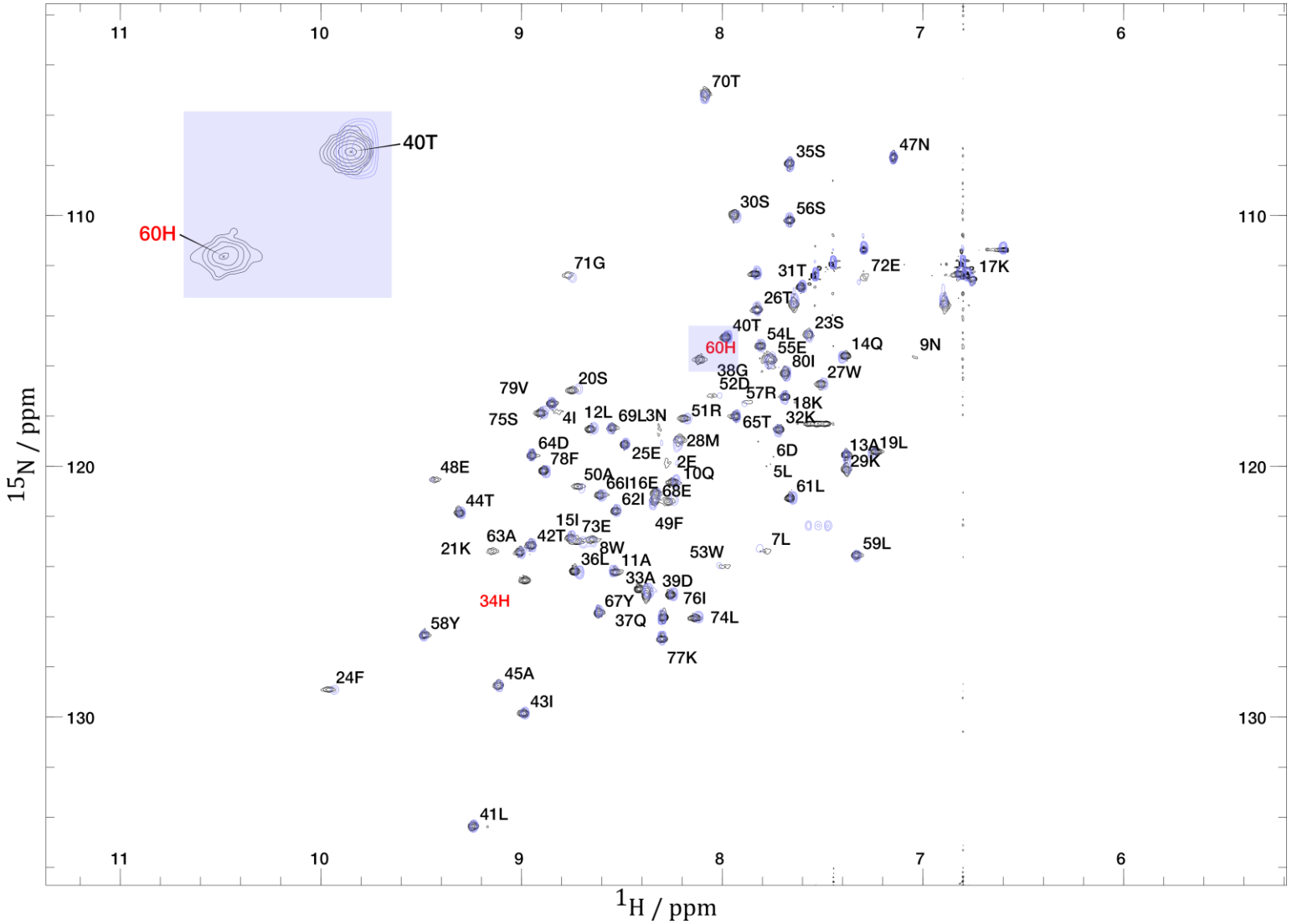


Figure 4.12 *^{15}N HSQC spectral overlay of fully enriched ^{15}N DnaA domain I (black) and ^{14}N His selectively labelled DnaA domain I (blue). Histidine residues are highlighted in red, one histidine peak is not observed in the ^{15}N HSQC experiment.*

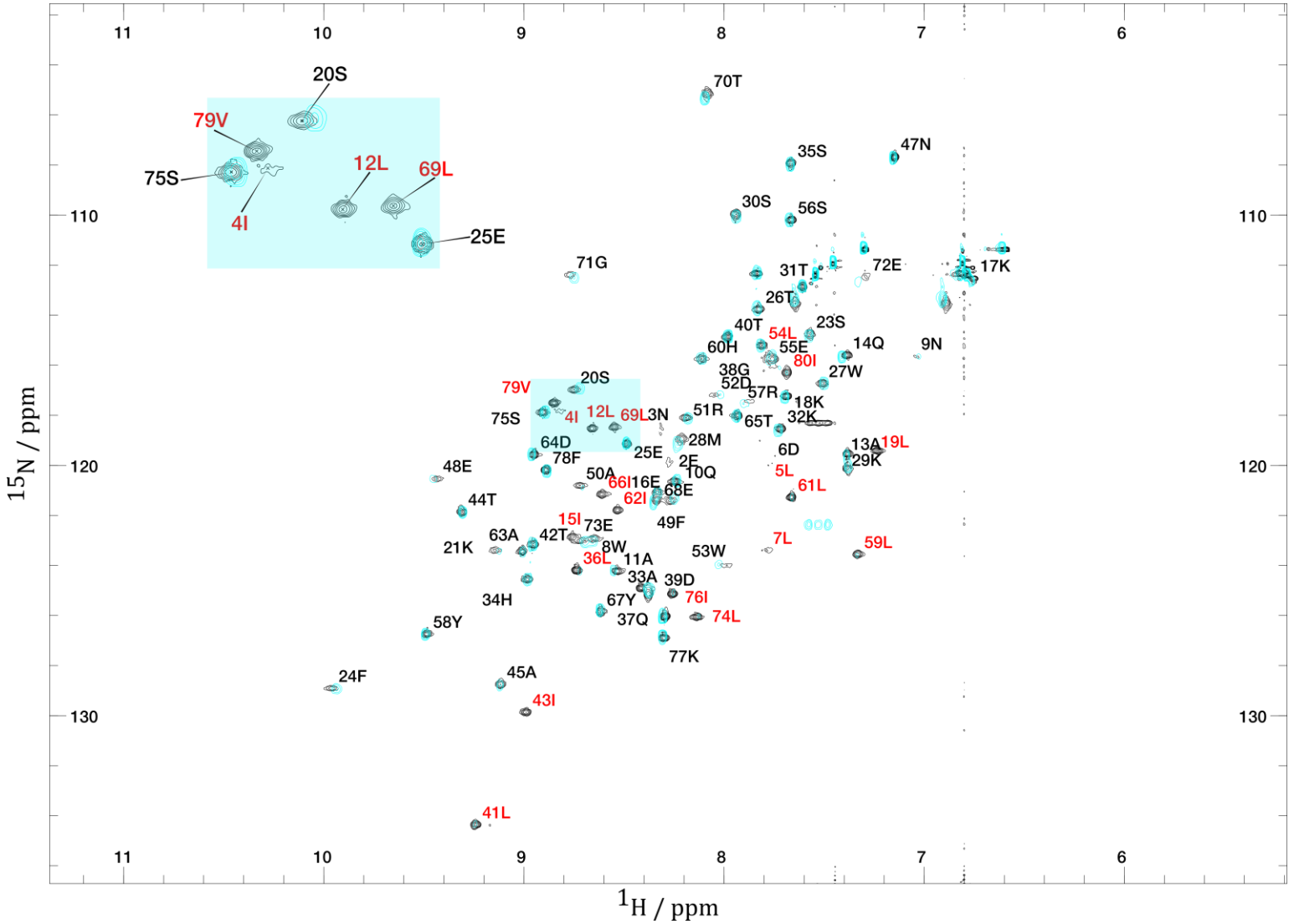


Figure 4.13 *^{15}N HSQC spectral overlay of fully enriched ^{15}N DnaA domain I (black) and ^{14}N Leu, ^{14}N Ile, ^{14}N Val selectively labelled DnaA domain I (cyan). Leucine, isoleucine and valine residues are highlighted in red, peak intensities were noticeably reduced but not absent as expected.*

4.7 Site Directed Mutagenesis

Using the methods previously described, regions of ambiguity still remained within the ^{15}N HSQC spectrum. In particular, the unstructured loop regions were difficult to assign using the sequential technique (described in section 4.5) due to missing amide Cα, Cβ and carbonyl resonances. Therefore, site directed mutagenesis was used to locally perturb specific amide resonances within the spectrum and allow residue identification. On visual analysis of the overlaid wild type and mutant ^{15}N HSQC spectra, the altered residue will produce a signal relating to its own chemical environment that should be noticeably different to the original residue at that particular location. This will result in relocation, and permit identification, of the amide backbone resonance corresponding to this position within the protein sequence. Minor shifting of several peaks can occur during these experiments relating to perturbations to the amino acids surrounding the point mutation in both sequence and space hence these shifts can also be informative to the assignment process.

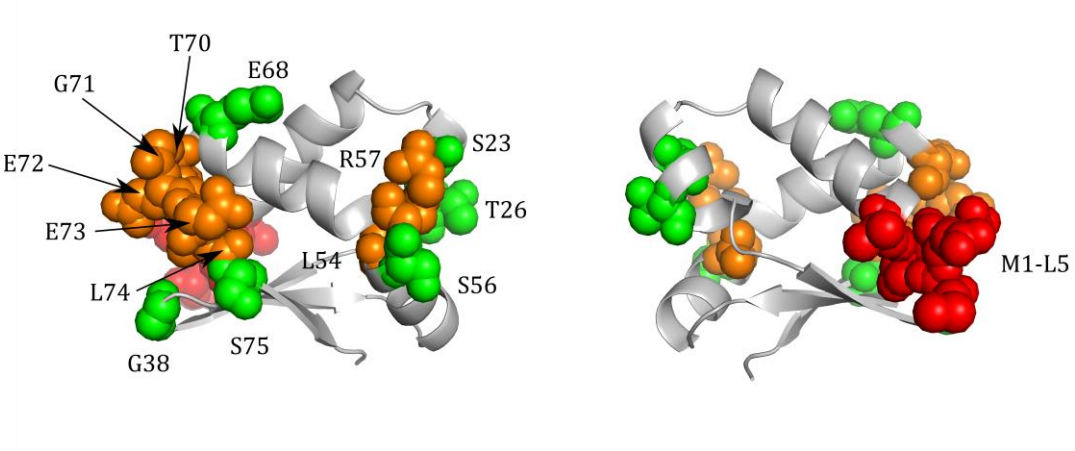


Figure 4.14 *Selected point mutations produced to facilitate assignment of specific residues and surrounding regions of the DnaA domain I ^{15}N HSQC spectrum. Highlighted in red is the region truncated in the ΔM1-L5 mutation. Displayed in green are the mutation sites that were used in the HSQC experiment, those in orange did not produce a sufficient quantity of soluble protein for further use and the region highlighted in red produced unstable unfolded protein.*

Residues selected for mutagenesis were positioned in unassigned regions of the protein sequence consisting of areas in and around the loop regions with less defined secondary structure (Figure 4.14). Initially, surface exposed serine and threonine

residues (S23, T26, S56, T70, S75) were selected as these could be mutated to alanine residues, without disrupting secondary structure of the domain, to distinctly alter the backbone amide chemical shift observed at these positions. With the exception of the T70A mutant these experiments were successful for assigning the corresponding resonances within the ^{15}N HSQC spectrum. The spectra corresponding to the S23A, T26A, S75A and S65A DnaA domain I mutants are shown in Figures 4.15, 4.16, 4.17 and 4.18 respectively where identified peaks are labelled red.

Although glycine residues should have characteristic Cα shifts within the CβCαNH spectrum both glycine residues within the domain sequence were ambiguous to assign due to lack of connectivity. For this reason, glycine 38 and glycine 71 were mutated to alanine to provide confidence in their assignment within the ^{15}N HSQC. The quantity of protein obtained for the G71A mutant was insufficient for use in the HSQC experiment due to low expression and solubility issues. The spectrum corresponding to the G38A DnaA domain I mutant is shown in Figure 4.19 where the identified peak is labelled red. The result of this experiment was not as clear as with the previous mutations as the signal corresponding to glycine 38 appeared overlapped and the peak showed relatively low resolution within the spectrum, however it did allow assignment by elimination for both glycine residues within the DnaA domain I sequence.

Further mutations included; L54A, R57K, E68N, E72N, E73N, L74A. These positions span two loop regions as highlighted in Figure 4.14. Based on the crystal structure of DnaA domain I, these positions seemed suitably surface exposed and conservative mutations (rather than all to alanine mutations) were selected to avoid complication by disruption of secondary structure. Insufficient protein quantity and quality (due to insolubility and degradation) was obtained from many of these mutants, with the exception of E68N, likely due to destabilisation of the secondary structure. The spectrum corresponding to the E68N DnaA domain I mutant is shown in Figure 4.20 where the identified peak is labelled red. A reduction in yield was observed for the E68N mutant, compared to the wild type, which impaired the resolution of the corresponding ^{15}N HSQC. Several amide resonances were consequently reduced to near the noise region complicating analysis of the spectrum. However, reviewing the Cα and Cβ shifts from the three-dimensional spectra in combination with the mutant data allowed assignment of glutamate 68.

The N-terminus of DnaA domain I also contained high ambiguity due to low resolution backbone amide resonances. To aid confidence of assignment in this area, a

truncation mutation ΔM1-L5 was created as it was suspected flexibility of this region may have been affecting the resolution of these residues. Overlaying the wild type and mutant ^{15}N HSQC spectra would then highlight resonances corresponding to this region. However, the truncated protein showed severe degradation during purification highlighting the importance of this region to the integrity of the domain.

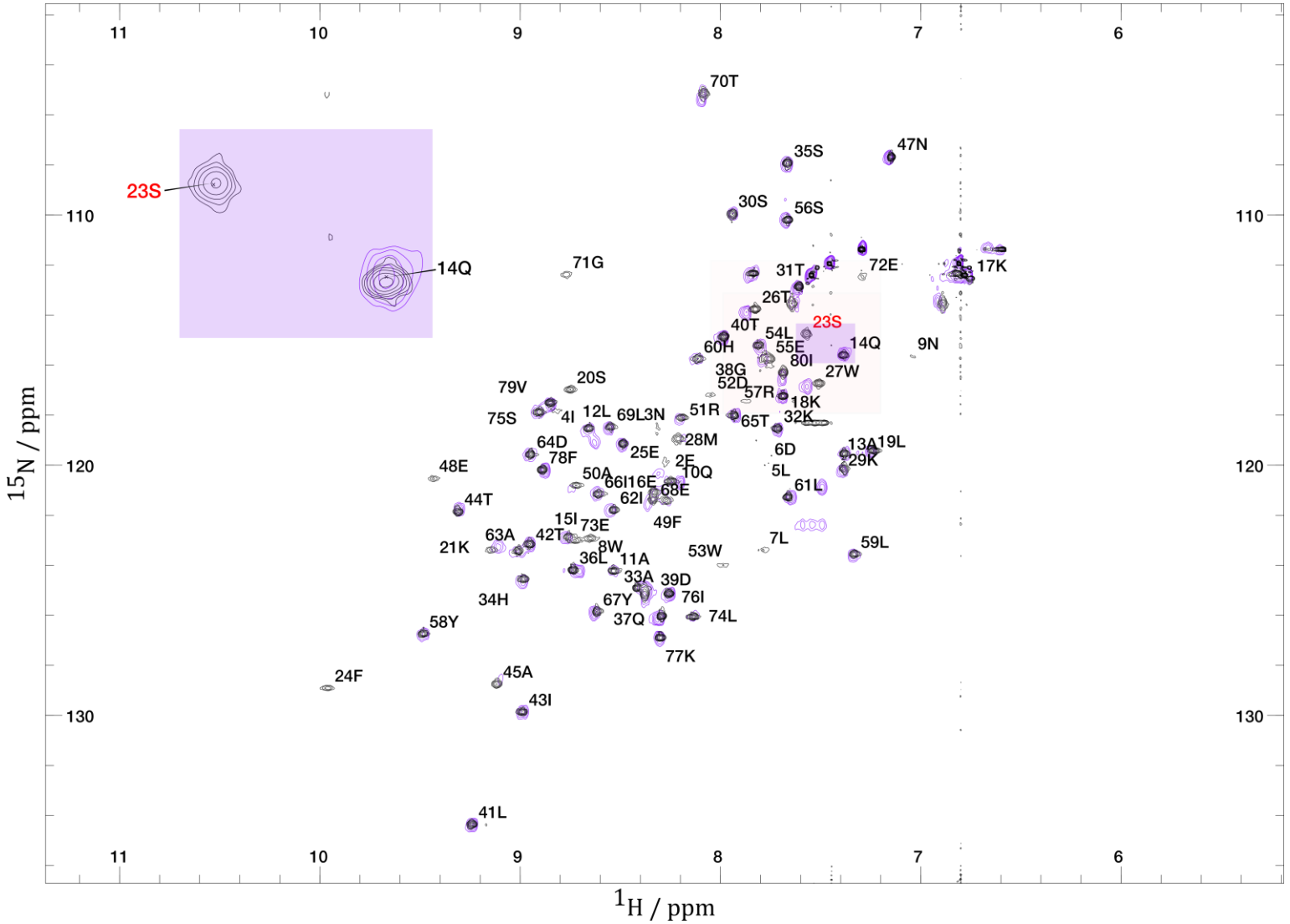


Figure 4.15 *^{15}N HSQC spectral overlay of DnaA domain I (black) and S23A mutant (purple). Serine 23 is highlighted red.*

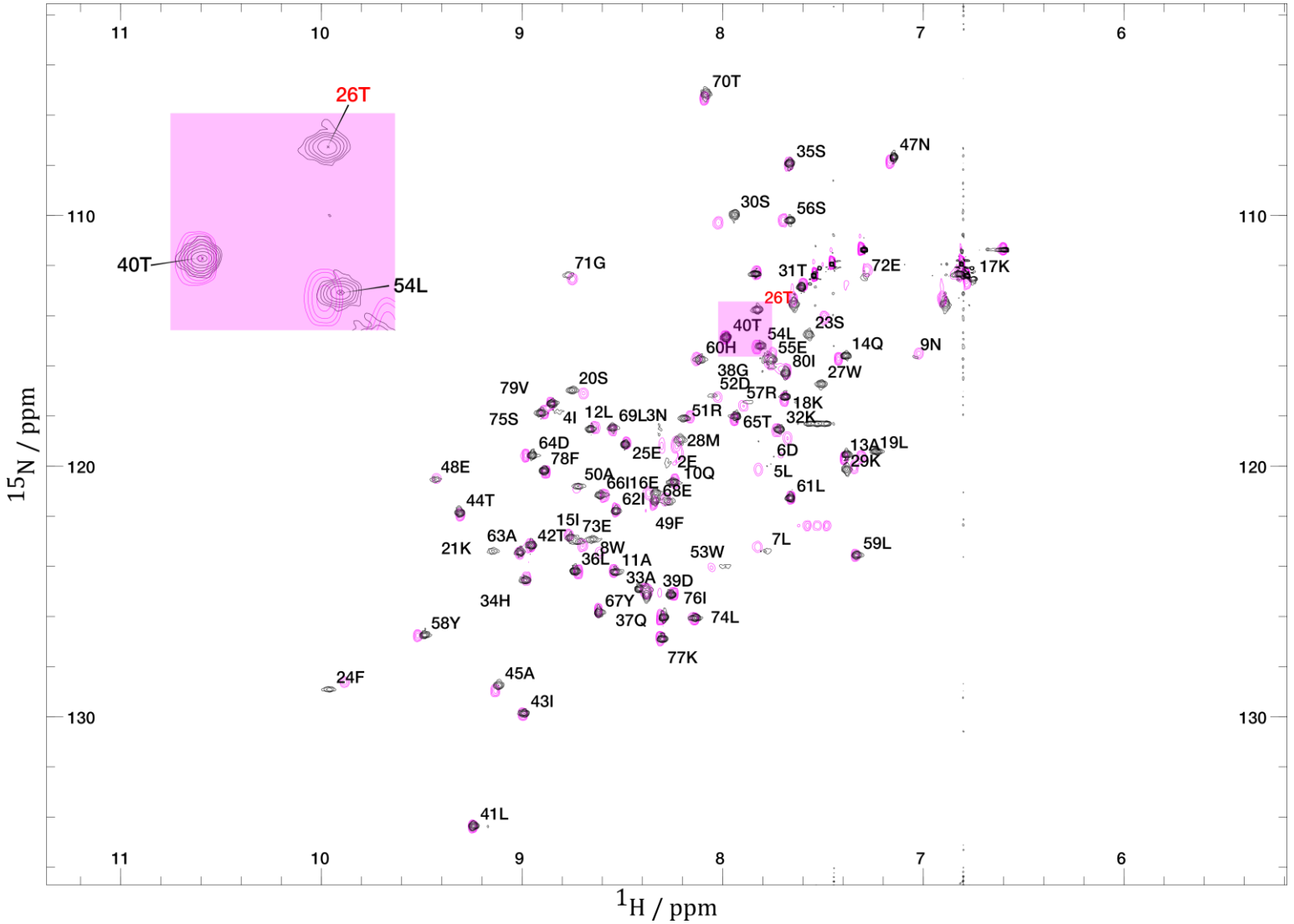


Figure 4.16 *^{15}N HSQC spectral overlay of DnaA domain I (black) and T26A mutant (magenta). Threonine 26 is highlighted red.*

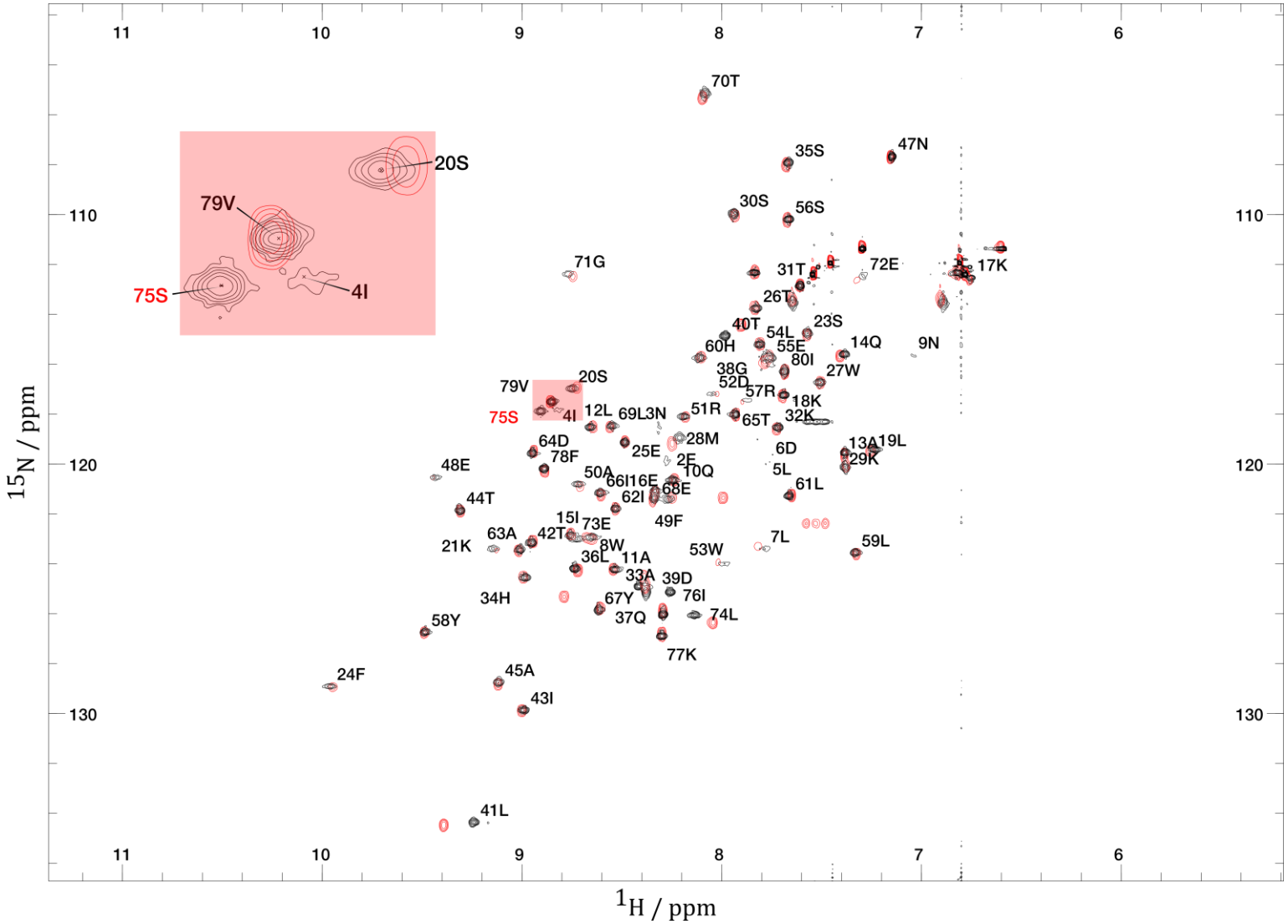


Figure 4.17 *^{15}N HSQC spectral overlay of DnaA domain I (black) and S75A mutant (red). Serine 75 is highlighted red.*

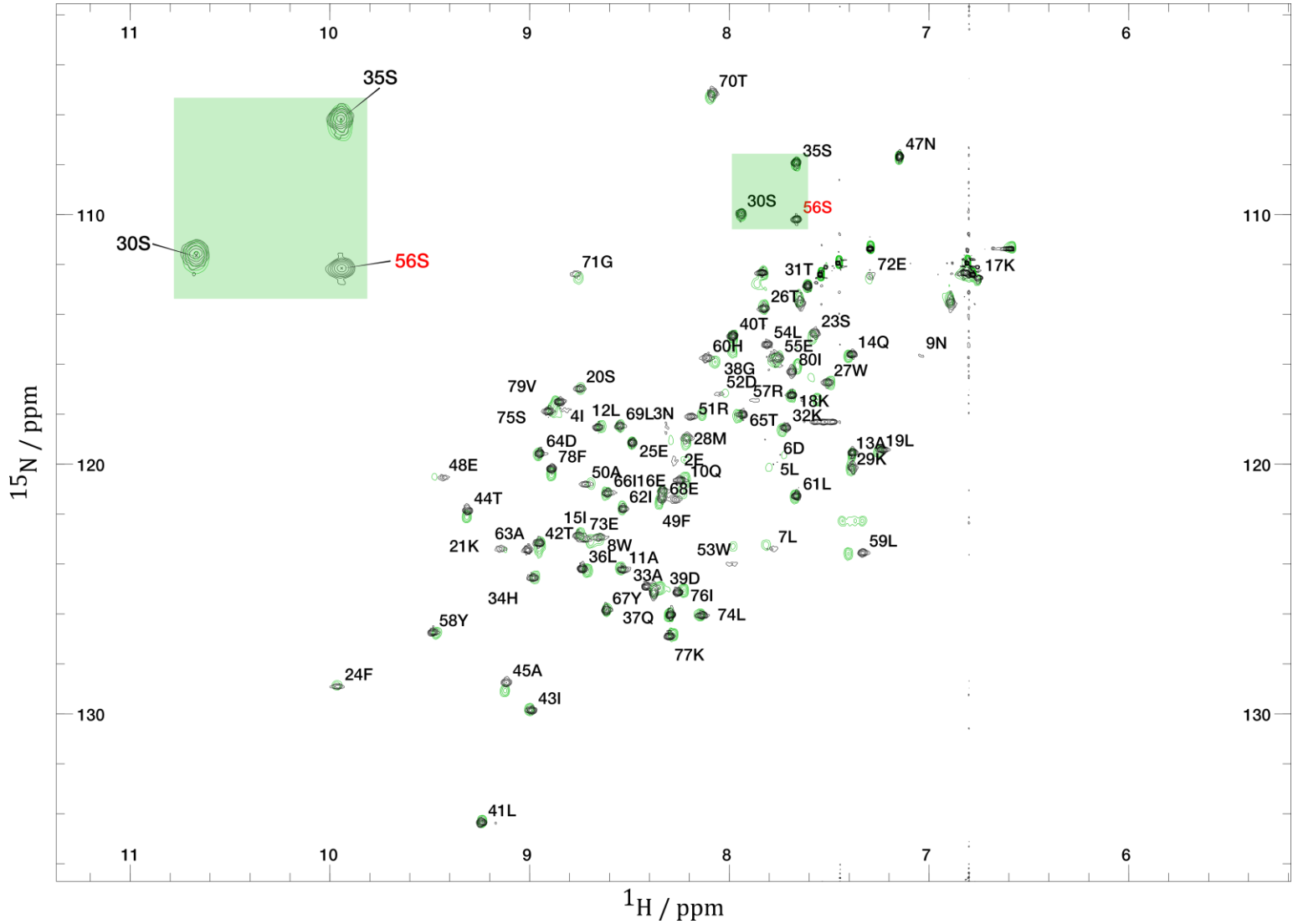


Figure 4.18 *^{15}N HSQC spectral overlay of DnaA domain I (black) and S56A mutant (green). Serine 56 is highlighted red.*

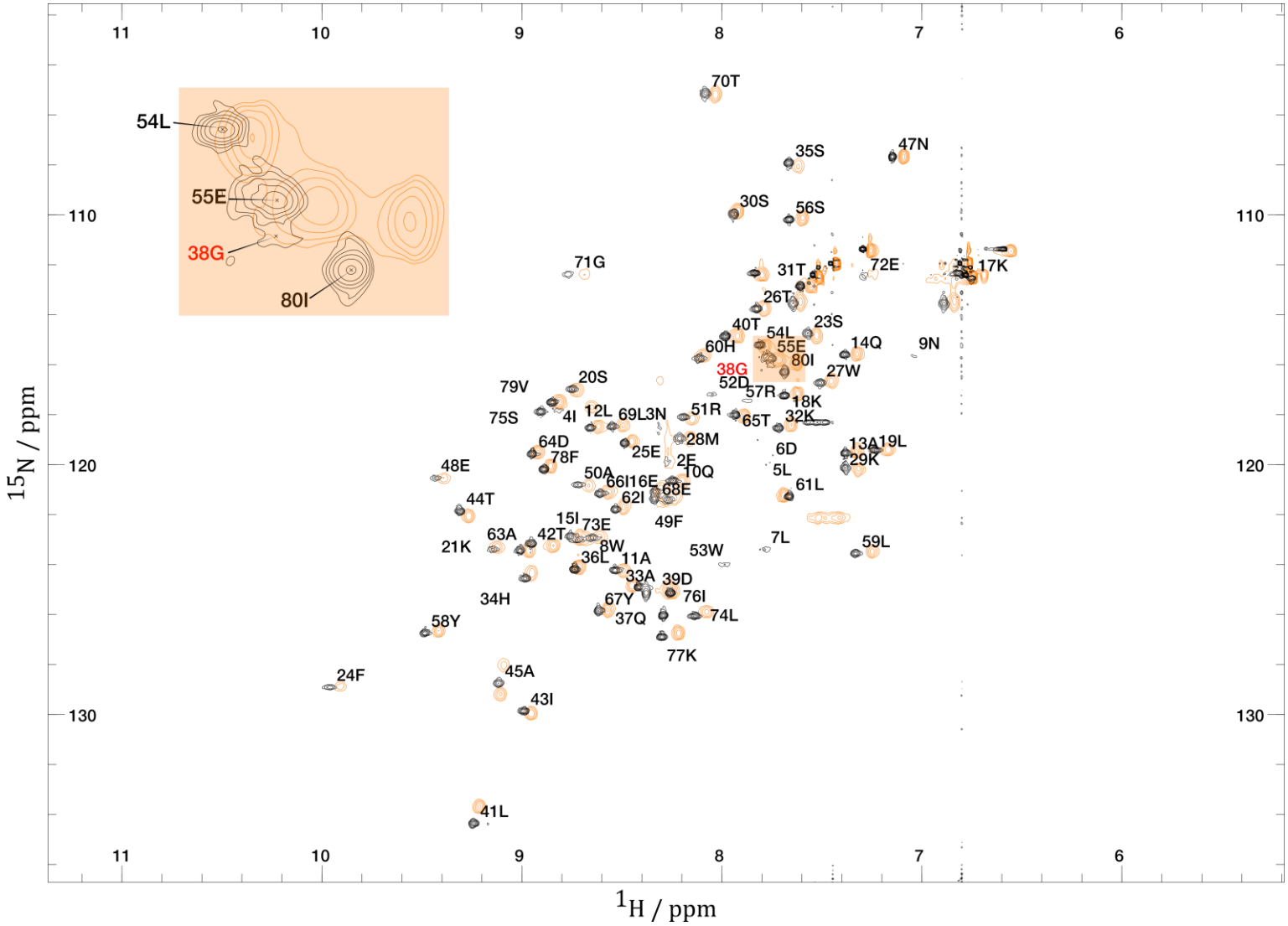


Figure 4.19 *^{15}N HSQC spectral overlay of DnaA domain I (black) and G38A mutant (orange). Glycine 38 is highlighted red where the corresponding peak appears overlapped and low resolution within the spectrum. Assignment was aided by the characteristic Cα signal, related to the amide resonance at this position, observed in the HNCα spectrum.*

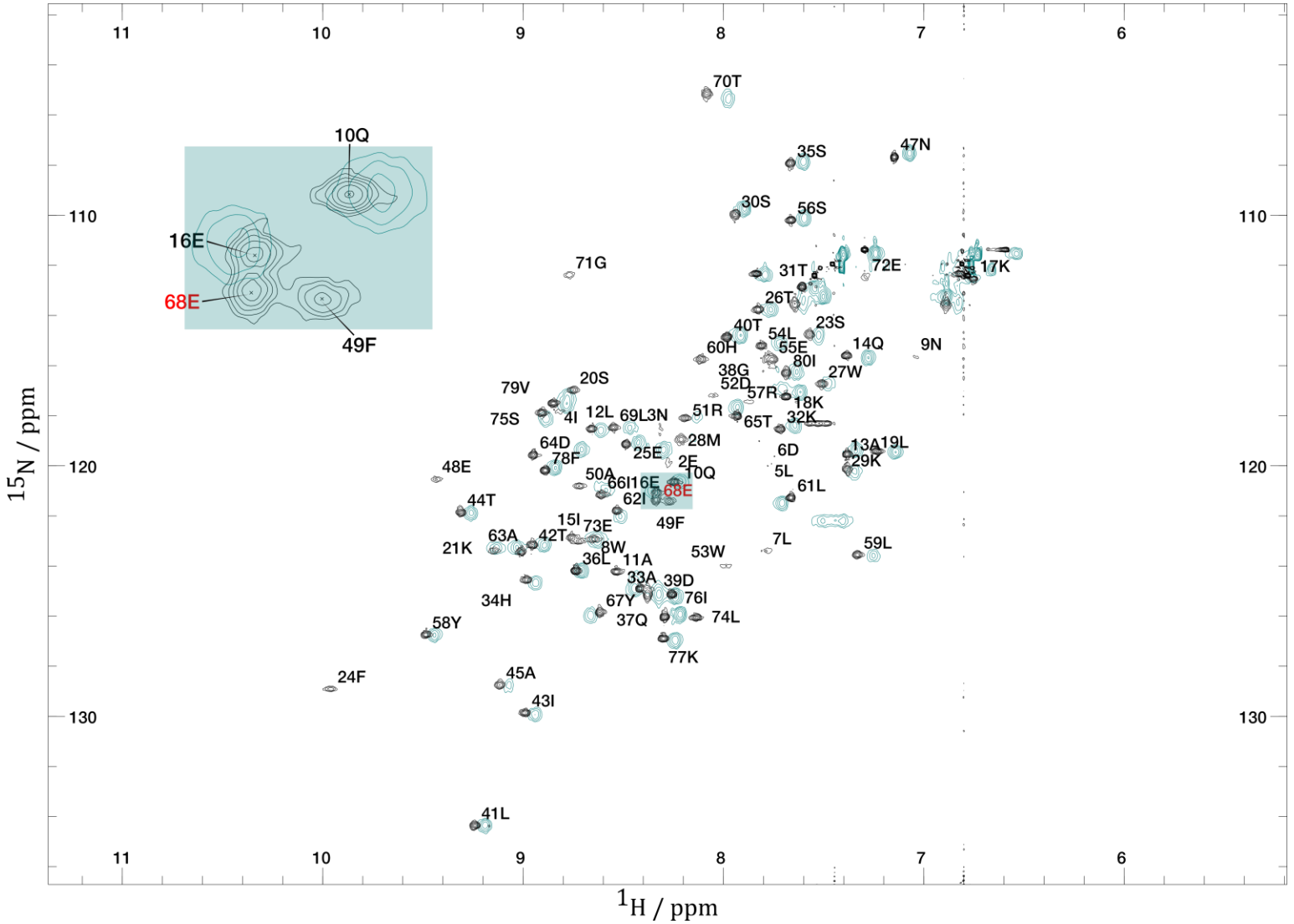


Figure 4.20 *^{15}N HSQC spectral overlay of DnaA domain I (black) and E68N mutant (teal). Glutamate 68 is highlighted red. Lower resolution was observed in the mutant spectrum where a few signals were reduced to the noise region. Assignment was aided by corresponding peaks within the CβCαNH spectrum.*

4.8 Discussion

The methods described in this chapter have allowed 97% assignment of the DnaA domain I amide backbone (excluding proline residues which are not observable in the ^{15}N HSQC experiment). Multiple three-dimensional experiments were required to provide sequential amide backbone connectivity. Using this technique based on backbone amide Cα, Cβ and carbonyl resonances approximately 57% of the amide backbone signals were assigned. Side chain assignment using HCCH-TOCSY and HNHα experiments proved inconclusive. Consequently, two-dimensional ^{14}N selective unlabelling and site directed mutagenesis experiments proved vital to reducing ambiguity within the assignment thereby allowing the further 40% assignment to be acieved.

In the ^{15}N HSQC spectrum 89 peaks were identified (with an additional 15 side chain peaks from asparagine, glutamine and tryptophan residues) of which all were confidently assigned, as displayed in Figure 4.21. Taking into account that proline residues are not observed in the ^{15}N HSQC experiment, 3 unique peaks are missing from the spectrum. These correspond to a short leader sequence Gly-Ser-His resulting from thrombin cleavage of the N-terminal polyhistidine-tag, incorporated into the recombinant protein to aid purification of the DnaA domain I construct. This unstructured region allows flexibility affecting the resolution of these residues and hence the signal may be broadened beyond detectable resolution under the conditions of the experiment.

The assignment obtained provides a foundation to further NMR characterisation of the DnaA domain I structure. Demonstrated in the subsequent chapter, characterisation in relation to the essential binding interaction with associated initiator protein DnaD has been undertaken.

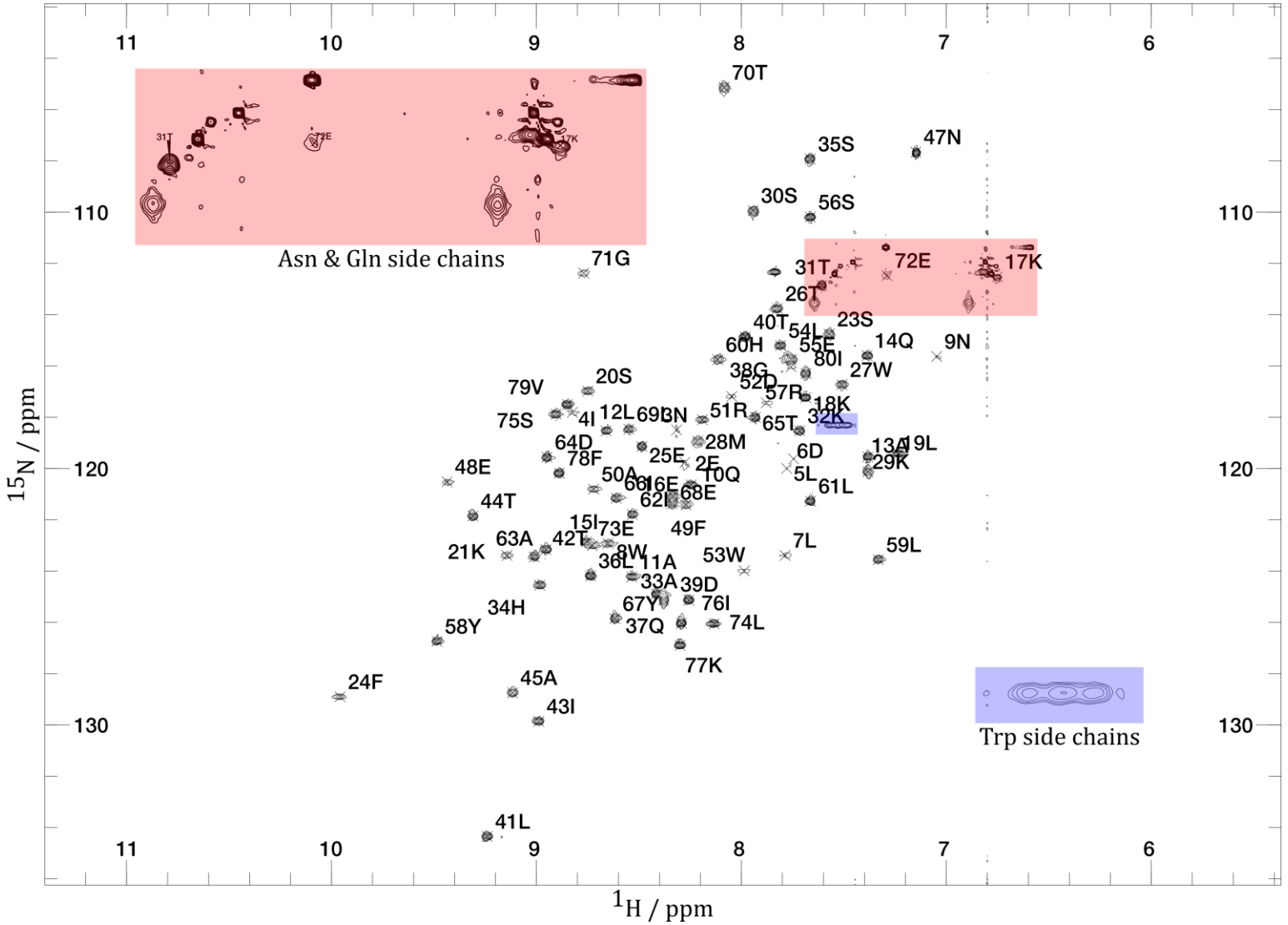


Figure 4.21 *^{15}N HSQC spectrum of DnaA domain I. The backbone amide peaks are labelled with the residue-specific assignment (97% excluding proline residues) that was achieved using a combination of the sequential assignment method, ^{14}N amino acid selective unlabelling and mutagenesis techniques. The locations of the side chain amide peaks corresponding to asparagine, glutamine and tryptophan have been highlighted.*

Chapter 5

Biophysical Characterisation of the DnaA-DnaD Interaction

5.1 Introduction

The interaction between the *Bacillus subtilis* initiator proteins DnaA and DnaD is essential for local melting of the duplex DNA at the *oriC* and subsequent loading of the replisomal machinery required for DNA replication. This interaction provides an interesting antimicrobial target for intervention in bacterial DNA replication, particularly gram-positive pathogens such as *Clostridium* and *Staphylococcus*, as well as advancing the understanding of bacterial DNA replication. However, beyond identification of the interaction between DnaA and DnaD there is no characterisation of the binding interface or detail of the kinetic profile for this interaction.

Two different methods were exploited to probe the interaction between DnaA domain I and the DnaD DDBH2 domain; high-field NMR spectroscopy and surface plasmon resonance (SPR) spectroscopy, both of which have the ability to detect weak and transient biomolecular interactions. NMR spectroscopy was used to structurally characterise the binding interface between the two protein domains, subsequently, SPR spectroscopy was utilised to quantitatively investigate the residues directly involved in the binding interaction.

5.1.1 Chemical Shift Perturbation

Chemical shifts are particularly sensitive to the electronic environment of a nucleus. Perturbations can be caused through covalent changes to the molecular structure but also by non-covalent interactions with binding partners and solvents. Subsequently, CSPs are an excellent probe for identification of interaction surfaces within protein complexes. Chemical shift mapping can provide information on the location as well as the strength of a binding interaction while having the advantage of being a straight forward method. It is important, however, to take into consideration the strength of the interaction when applying this method as an interaction weaker than 3μM is required, the method being best suited to interactions in the mM range. These conditions give rise to fast exchange where the k_{off} of the interaction is greater than the chemical shift difference between the free and bound states. Consequently, the observed chemical shift is a population weighted average of the respective free and bound states until the binding partner has been added to excess[137].

The ^{15}N HSQC experiment is classically used for CSP analysis as it is highly sensitive, the peaks are generally well resolved and it provides a probe for each non-proline amino acid within the protein. The chemical shifts of amide protons are influenced largely by through-space effects such as ring currents (from nearby aromatic groups) and the formation of hydrogen bonds. The chemical shifts of carbon atoms (Cα and Cβ), in contrast, are predominately affected by through-bond effects[137]. These atoms are typically not surface exposed and therefore further removed from the effects of ligand binding making them less informative for such studies. Ligand binding encompasses through-space effects where hydrogen bonding is the major contribution to the observed difference in amide chemical shift, consequently, the ^{15}N-HSQC or ^{15}N-TROSY experiments are most useful for observation of a ligand binding interface. Chemical shifts are monitored over a series of experiments where increasing concentrations of ligand are titrated into a sample of ^{15}N labelled protein of interest. Chemical shifts perturbed on addition of the interaction partner can be used to identify amino acids involved in the interaction surface. However, it is important to consider that chemical shift perturbation may not necessarily indicate a direct interaction with the binding partner but rather indicates a change in the magnetic environment of a nucleus. Widespread chemical shift perturbation can be observed where a substantial structural rearrangement occurs upon complex formation, resulting in perturbations distal to the ligand interaction site.

The respective scale of chemical shift ranges for nitrogen and hydrogen are significantly different, therefore, their relative contribution must be weighted to the calculated CSP value for each peak. Throughout this work, the CSPs were calculated according to the equation below[137]:

$$CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (0.14\ x\ \delta_N^2)]} \qquad (5.1.1)$$

Where δ_H is the amide proton chemical shift and δ_N is the amide nitrogen chemical shift. The scaling coefficient of δ_N^2 is derived from the relative ratio of amide proton to amide nitrogen chemical shift range. Where the amide chemical shift range for ^{15}N is considered to be approximately 22 ppm and for ^{1}H to be 3 ppm[138]. Applied threshold limits[137] were used to determine the significance of the calculated CSP values:

$$Lower\ Threshold\ \geq\ \mu + \sigma, \quad Higher\ Threshold\ \geq\ \mu + 2\sigma \qquad (5.1.2)$$

This takes into account the population mean (μ) and standard deviation (σ). Residues characterised by a CSP exceeding one or both of these thresholds were considered to have a significant conformational difference on addition of ligand. Residues identified by this method were then mapped onto the structure (X-ray crystal or NMR) to identify the interaction surface.

5.1.2 Surface Plasmon Resonance

Surface plasmon resonance has become a widely used technique for the study of biomolecular interactions since its development in the 1990s[139, 140]. The technique facilitates observation of binding in real-time, label free, providing a kinetic profile of the interaction where the optical method of detection is independent of the analyte. The method can be used to study a range of interactions; protein-protein[141], DNA-RNA[142], DNA-protein[143], antibody-antigen[144, 145] and small molecule drug candidates[146].

5.1.2.1 SPR Theory

Surface plasmon resonance is a process that occurs when plane-polarised light hits a thin metal film under total internal reflection (TIR) conditions. A surface plasmon is an electro-magnetic surface wave that propagates parallel to a metal-dielectric interface as illustrated in Figure 5.1. An incident light beam is used to excite the surface plasmon in a resonant manner for instance, plane-polarised light passing through a glass prism at a given wavelength and angle allows the incoming beam to match the momentum of the surface plasmon. The light beam hits the conducting metal layer (commonly gold) at the interface of the glass sensor chip, with a high refractive index (RI), and a low RI material such as liquid and then undergoes TIR.

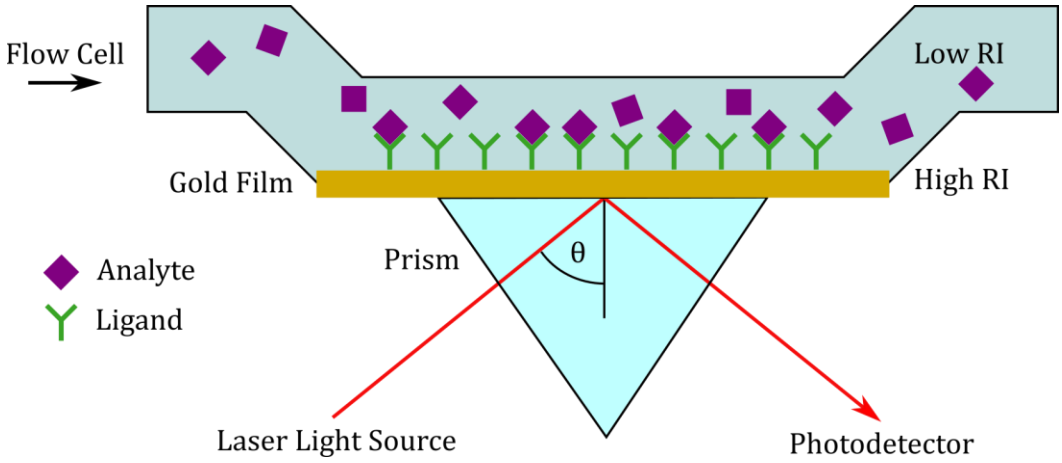


Figure 5.1 *Schematic showing the generation of surface plasmons via total internal reflection of incident light at the interface of the sensor chip and flow cell.*

The incident light is completely reflected at TIR, however, before the critical angle is reached the beam is partly refracted and therefore altering the incidence angle (θ) changes the out-coming light until this critical angle is reached (Figure 5.2). At the angle of incidence (θ_{TIR}) the incident light photons interact with the free-electrons in the metal layer to generate plasmons and this induced resonance decreases the intensity of the detected reflected light. The angle at which the resonance occurs is highly sensitive to the RI of the medium at the glass interface. This allows changes in the surface, such as analyte binding, to be measured by a change in light intensity when the angle of resonance is perturbed.

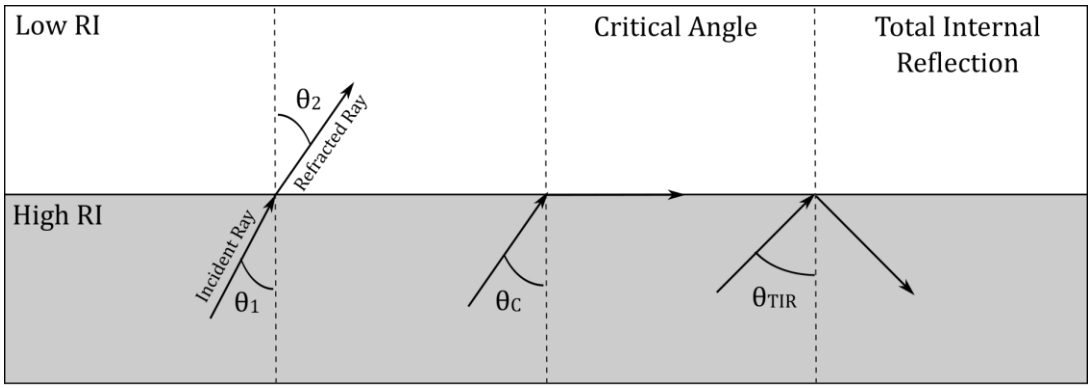


Figure 5.2 *The angle of incidence determines if a beam of light is reflected or refracted. Before the critical angle is met the incident beam is refracted at an angle θ_2, at the critical angle the θ_2 is equal to 90° and total internal reflection (TIR) occurs for angles greater than the critical angle θ_C.*

5.1.2.2 Sensor Surface

There are two methods for immobilising a ligand to a sensor surface; the permanent immobilisation of the ligand via a covalent bond such as with CM5 surfaces consisting of a carboxymethylated dextran matrix, or transient affinity capture of the ligand such as with SA (streptavidin) and NTA (nitrilotriacetic acid) surfaces. A number of sensor surfaces are available, utilising these methods, for use with SPR spectroscopy instrumentation. CM5 is one of the most widely used surface types where direct immobilisation can be achieved via the amine coupling of surface exposed lysine residues through ε-amino groups[147]. This sensor surface was initially used to immobilise the proteins in this study, however, both immobilisation of DnaA domain I and DnaD DDBH2 failed to produce a response on injection of analyte. Using this method, it is possible for biological activity of the ligand to be altered due to a covalent bond within the active site or restricted free movement of the ligand may have prevented binding to the analyte.

Consequently, the NTA surface was selected where transient immobilisation is achieved through affinity capture of polyhistidine-tagged ligand. The binding of histidine relies on the nickel chelated NTA surface where the affinity of this interaction (ca. K_D 10^{-6} M) is sufficient for analysis of subsequent analyte binding[148]. The advantage of this method for immobilisation is the homogeneous orientation of the ligand molecules. SPR measurements are recorded in real-time as sensograms (Figure

5.3) where the angle of incidence is affected by changes at the glass-metal interface. This is convenient for monitoring ligand binding as a shift in the angle is directly proportional to a change in mass at the surface.

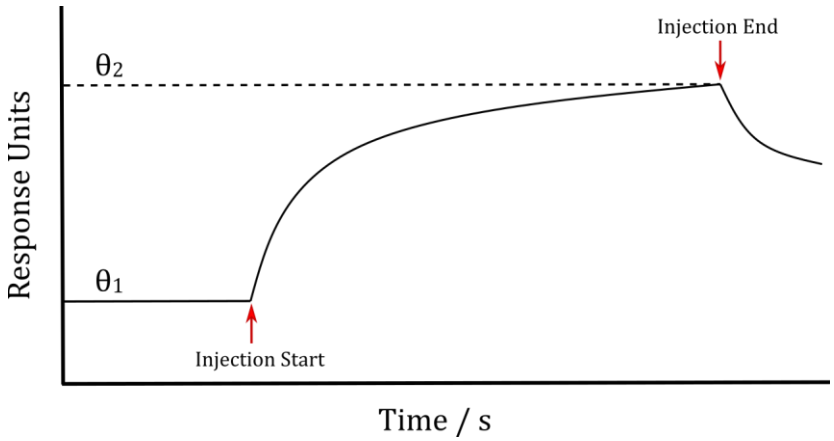


Figure 5.3 *A sensogram representing the adsorption of analyte onto a ligand surface. The start and end points of the analyte injection are indicated by arrows. The angle of resonance θ_1 is constant prior to analyte injection. The adsorption of analyte at the ligand surface alters the angle of resonance to θ_2 which is observed as a change in light intensity response at the detector. Post injection the analyte dissociates from the surface and the decrease in the response is observed as the angle of resonance returns to θ_1.*

5.1.2.3 Experimental Design

Multi-cycle kinetic assays were used during this work. This method uses single analyte injections over a series of varied concentrations where the ligand surface is regenerated between injections. As the K_D value for the interaction between DnaA and DnaD was unknown, the analyte concentrations were determined by initial calibration injections. Typically, a range of 0.1x – 10x the K_D of an interaction is used to determine analyte concentrations required[141]. The lowest analyte concentration that gave a detectable response was selected for the start point and the range was scaled to the highest concentration achievable under solubility limits of the DnaD DDBH2 domain. A repeat of at least one concentration was included within each series to validate the results.

To prevent re-association occurring, the transfer of analyte to the immobilised ligand should be faster than the binding interaction and diffusion rate, hence, a lower limit of 30 mL/min is recommended[149]. However, during this work, a transfer speed of 10 mL/min was used as speeds greater than this did not produce SPR curves of sufficient quality. This can be accounted for by the rebinding effects[148, 150] observed when transiently immobilising ligands to an NTA surface. During injection of ligand, a maximum association is followed by slight dissociation as rebinding is required for stability at low binding affinity (Figure 5.4). Hence, the binding appears more stable at partial occupation of ligand binding sites than at higher concentrations and additionally the dissociation rate is increased at higher flow rates.

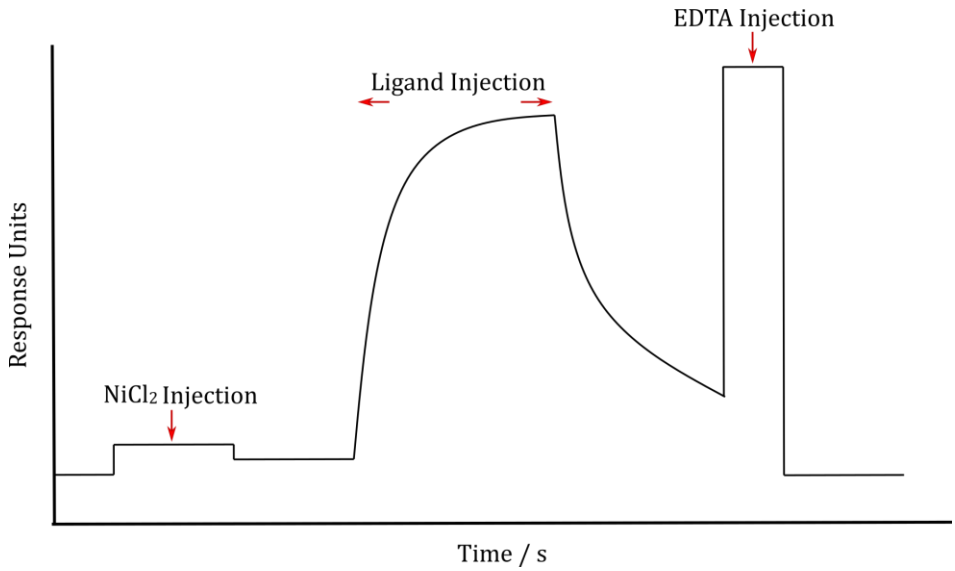


Figure 5.4 *A sensogram representing the affinity capture of histidine-tagged proteins onto the NTA chip surface. The surface is charged with $NiCl_2$ for which a slight increase in response (typically ± 40 RU) is observed. With injection of the ligand a maximum association is observed followed by dissociation due to rebinding effects[148, 150]. The injection of EDTA strips the nickel and histidine-tagged protein from the sensor surface hence regenerating the chip.*

The amount of ligand immobilised onto the surface is also of importance to reduce limiting effects of mass transfer and minimise artefacts from over-crowding of molecules on the sensor surface. Mass transfer is the first stage of binding analyte to the ligand where the analyte is transferred out of the bulk of solution towards the sensor chip surface. If diffusion from the bulk solution is slower than the rate of binding to the ligand then a shortage of analyte at the interaction surface will occur and

calculated kinetic constants under these conditions will reflect the mass transfer rate rather than binding kinetics. Mass transfer is minimised by immobilising less ligand to the surface. Throughout these experiments typically 700 arbitrary response units (RU) were maintained during ligand immobilisation which, after dissociation due to rebinding effects, resulted in approximately 200 RU immediately prior to analyte injection.

5.1.2.4 Data Analysis

Interaction kinetics describe the interaction between one or more components which remain unchanged post interaction, as opposed to enzyme kinetics, and can be grouped into three distinct phases; association - two or more molecules bind to each other, steady state – the number of molecules binding equals the number of bonds breaking, and dissociation – the bonds between molecules are broken. Each of the phases provides information for the interaction between molecules in terms of how fast the association and dissociation occurs and the strength of the interaction. Kinetic and steady state analysis methods are described in section 2.6.4.

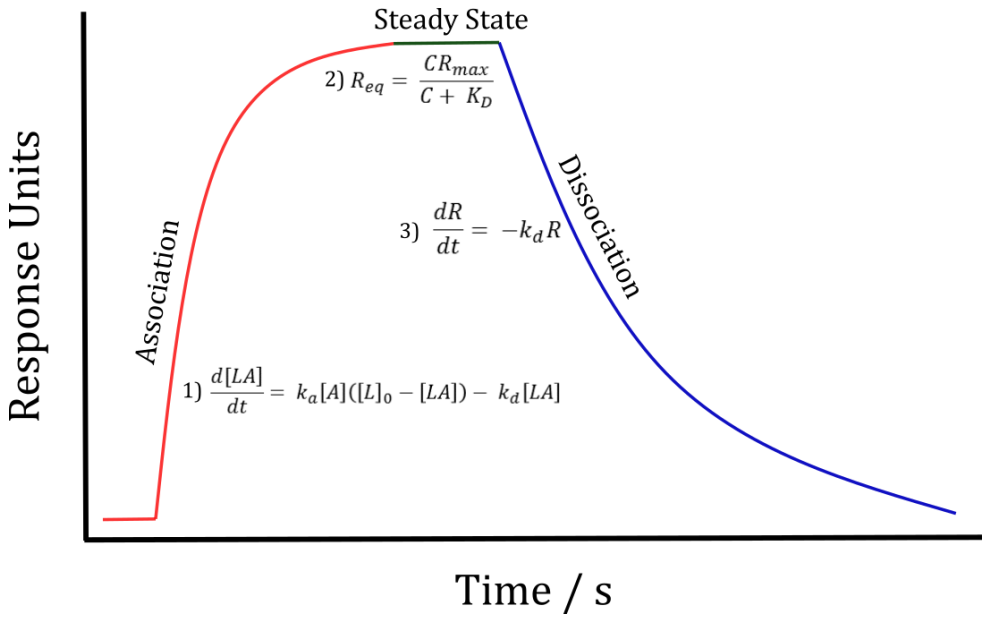


Figure 5.5 *A sensogram showing the phases of interaction kinetics. The association phase is displayed in red, the steady state phase in green and the dissociation phase in blue. The equations used to determine the rate and equilibrium constants for the 1:1 binding model are displayed; 1) derives the association rate constant, 2) derives the equilibrium rate constant and 3) derives the dissociation rate constant.*

5.2 Chemical Shift Perturbation Analysis

5.2.1 Identifying the DnaD Interaction Patch of DnaA Domain I

The titration of unlabelled DnaD DDBH2 domain (final concentration 800 μM) into a ^{15}N labelled sample of 100 μM DnaA domain I was performed incrementally to an 8:1 excess. Figure 5.6 illustrates selected peaks that were subject to perturbation during the experimental series.

The DDBH2 domain of DnaD is involved in DNA binding. The conserved motif YxxxIxxxW ($Y_{180}IDRI_{184}LFEW_{188}$) along with a region of the unstructured C-terminus (residues 206-215) are essential for DNA binding activity[90]. It is therefore of interest to investigate the effect of DNA binding on the DnaA-DnaD interaction as at the origin of replication both proteins bind to each other and function to remodel the DNA for replication initiation. DnaD can bind both double stranded and single stranded DNA with a higher affinity for the latter, therefore, a short 10 mer of ssDNA (5' GTTATTGCTC 3') previously used in DnaD-DNA binding studies[90] was selected for this investigation. The titration of unlabelled DnaD DDBH2 domain (final concentration 800 μM) into a ^{15}N labelled sample of 100 μM DnaA domain I was repeated, in the presence of ssDNA (800μM), incrementally to an 8:1 excess. Figure 5.7 illustrates selected peaks that were subject to perturbation during the experimental series.

A truncation of the DnaD DDBH2 domain, residues 129 – 196, was created which lacked the putative DNA binding motif and was predicted to have no DNA binding activity. As mentioned previously, the unstructured C-terminal region containing residues 206 – 215 is essential to the DNA binding activity of DnaD. This construct was used to investigate whether DnaD maintained its DnaA binding activity without the ability to bind DNA. The titration of unlabelled DnaD DDBH2 domain truncation (final concentration 800 μM) into a ^{15}N labelled sample of 100 μM DnaA domain I was performed incrementally to an 8:1 excess. Figure 5.8 illustrates selected peaks that were subject to perturbation during the experimental series.

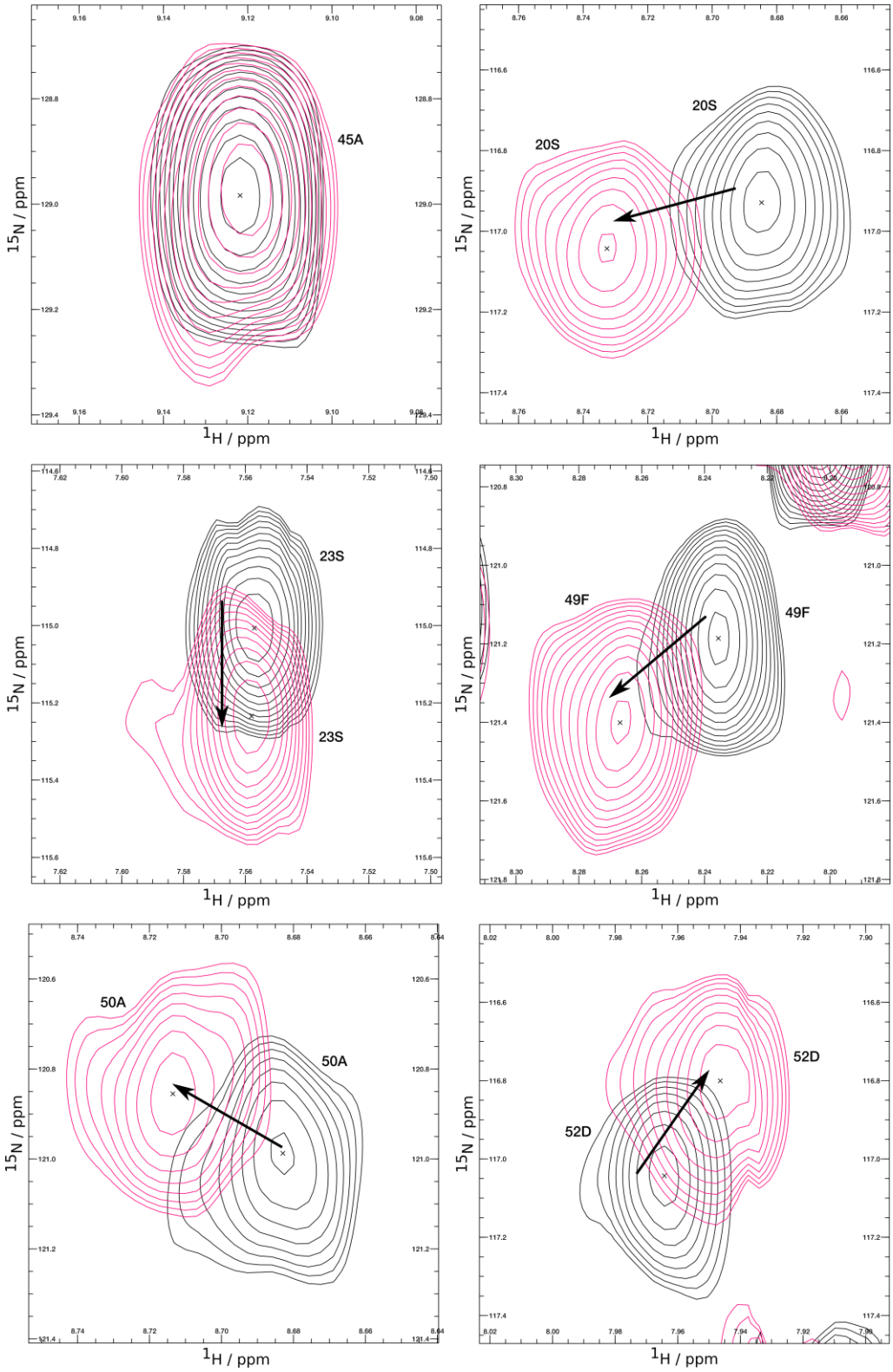


Figure 5.6 *Overlay of spectral peaks representing DnaA domain I (100 µM) ^{15}N HSQC in black and the 8:1 excess of DnaD DDBH2 domain (800 µM) to DnaA domain I spectrum is shown in pink. The arrows indicate direction of movement of the peak during the titration series. The alanine 45 peak (top left) has been included to illustrate the behaviour of peaks not involved in the binding interface.*

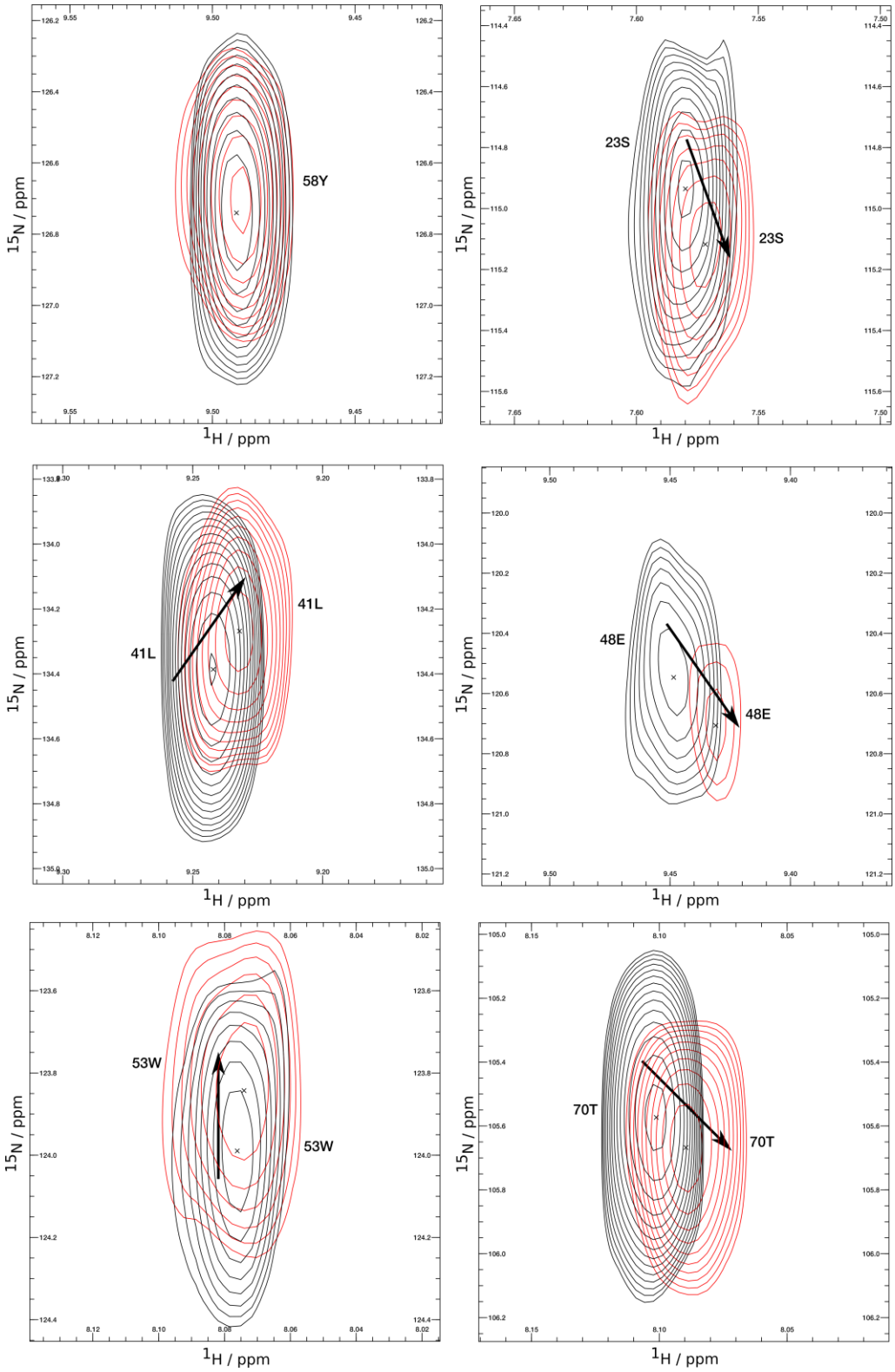


Figure 5.7 *Overlay of spectral peaks representing DnaA domain I (100 μM) ^{15}N HSQC in black and the 8:1 excess of DnaD DDBH2 domain (800 μM) to DnaA domain I spectrum is shown in red, both spectra were run in the presence of ssDNA (800μM). The arrows indicate direction of movement of the peak during the titration series. The tyrosine 58 peak (top left) has been included to illustrate the behaviour of peaks not involved in the binding interface.*

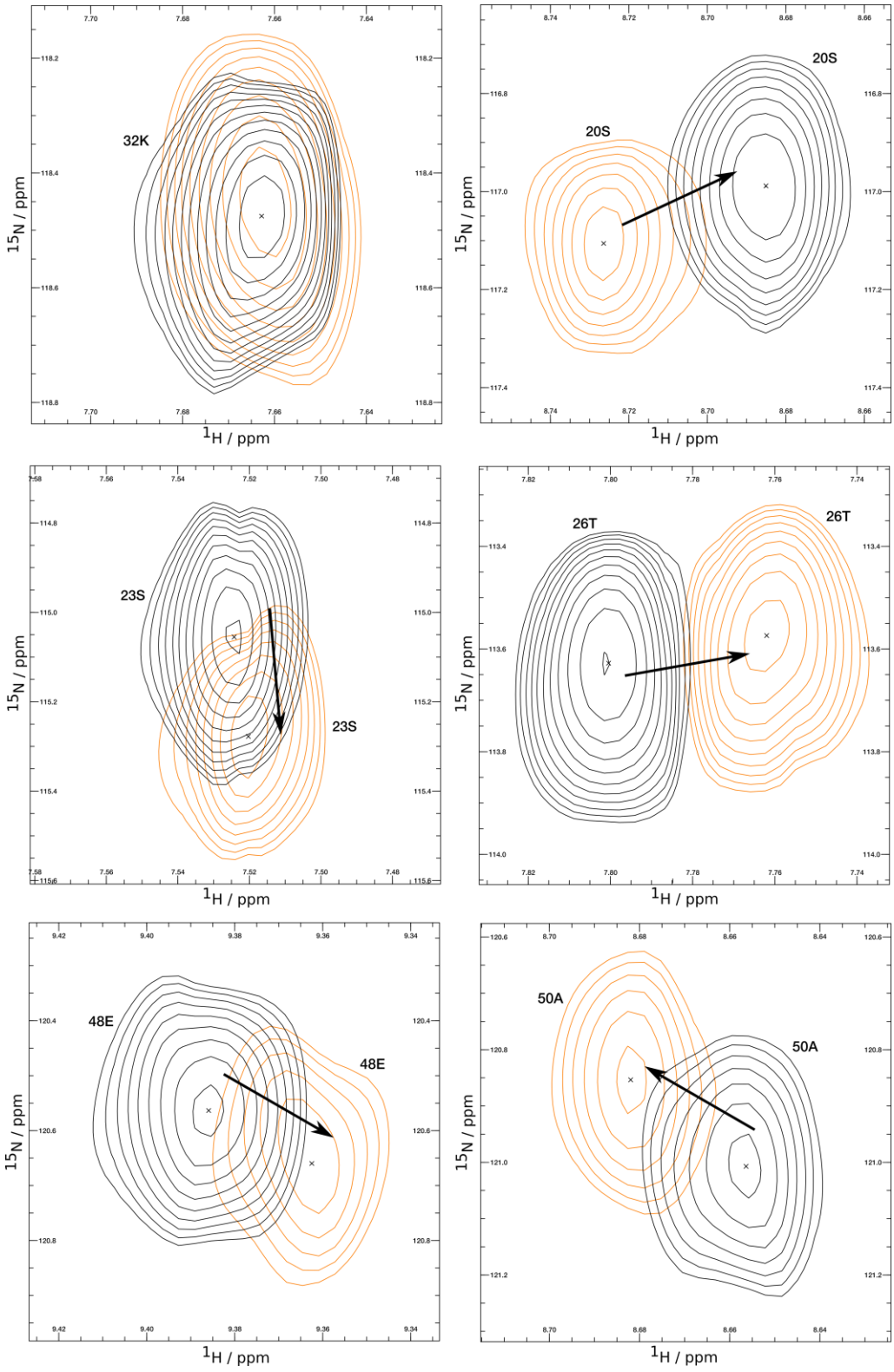


Figure 5.8 *Overlay of spectral peaks representing DnaA domain I (100 µM) ^{15}N HSQC in black and the 8:1 excess of DnaD DDBH2 domain truncation (residues 129 – 196, 800 µM) to DnaA domain I spectrum is shown in orange. The arrows indicate direction of movement of the peak during the titration series. The lysine 32 peak (top left) has been included to illustrate the behaviour of peaks not involved in the binding interface.*

CSP analysis was performed on the resulting ^{15}N HSQC spectra from each of the titration experiments and plots showing the chemical shift difference vs residue number are displayed in Figure 5.9. Threshold values to determine significantly shifted resonances were determined as described in section 5.1.1, residues exceeding these thresholds are likely directly involved in the DnaA-DnaD interaction though conformational changes can lead to disruption in the environment of residues away from the binding interface. It is apparent when comparing the results from the separate titration experiments that there is a high degree of overlap between the residues that are observed above the threshold of significance. These results confirm that the interaction between DnaA domain I and the DnaD DDBH2 domain is independent of the unstructured C-terminal region of DnaD (residues 206 - 215) essential to DNA binding.

The residues identified as significant to the DnaA-DnaD binding interface were mapped onto the crystal structure of DnaA domain I to reveal a well-defined binding surface (as displayed in Figure 5.10). The interface is positioned on one face of the domain and can be clustered into three groups; Lys17, Ser20, Ser23 and Thr26 located at the interface between the C-terminus of the α1 helix and N-terminus of the α2 helix; Glu48, Phe49, Ala50, Asp52, Trp53, Leu54, Glu55 and S56 located throughout the α3 helix; and His60 located at the N-terminus of the α4 helix. The lack of significant CSPs within the β-sheet regions of the domain demonstrates that the interface from DnaA domain I is extensively α-helical. Additional CSPs of significance were observed for Leu69 and Thr70, at the C-terminus of the α4 helix, and Leu41 located on the β2-sheet. These residues are positioned away from the main binding interface identified and may indicate an allosteric effect experienced by the structure in response to DnaD DDBH2 binding at the α2 and α3 helix.

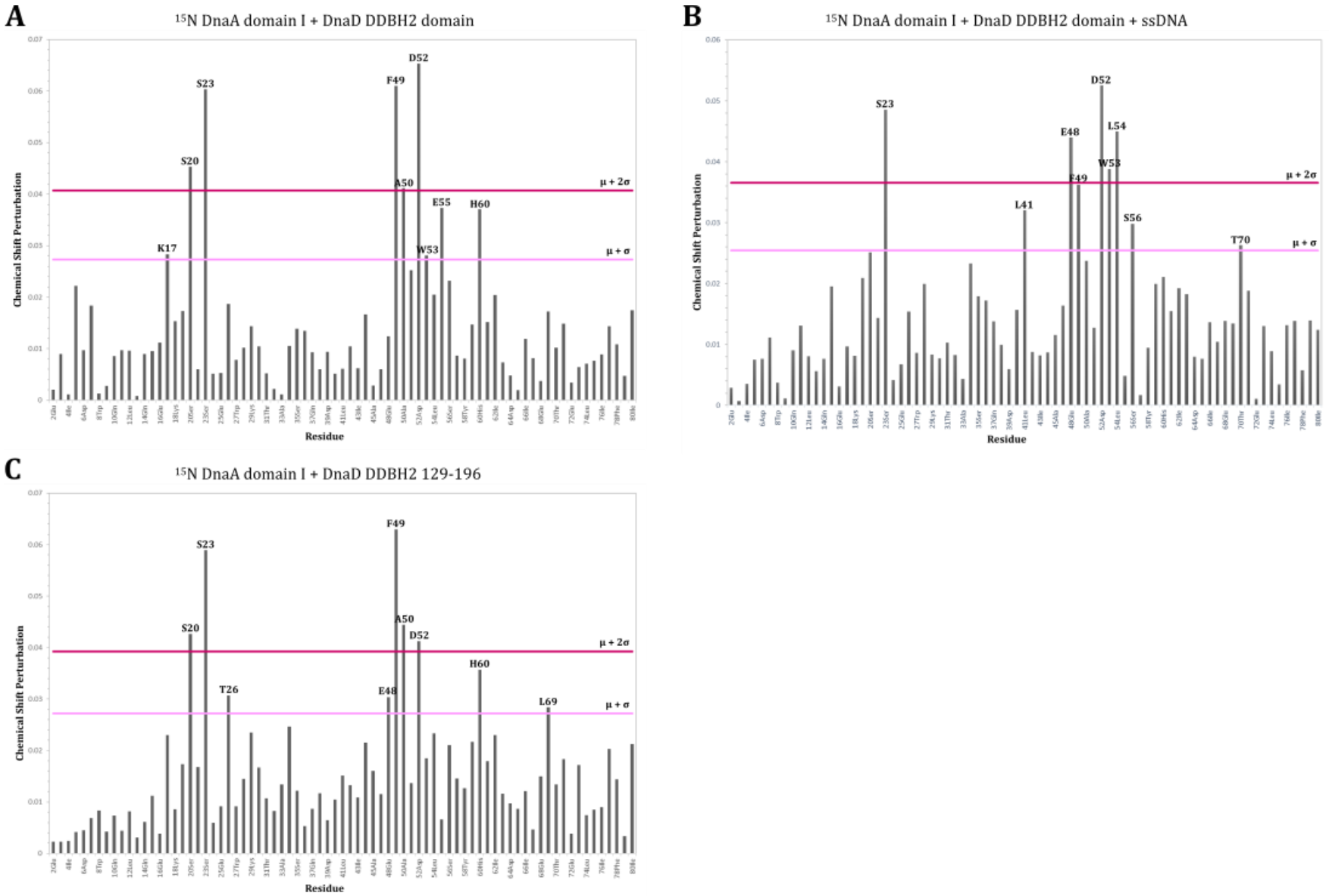


Figure 5.9 A. *CSP analysis of the DnaA domain I titration with the DnaD DDBH2 domain. Threshold values of significance were calculated using μ = 0.0139 and σ = 0.0134.* ***B.*** *CSP analysis of the DnaA domain I titration with the DnaD DDBH2 domain in the presence of ssDNA (800 μM). Threshold values of significance were calculated using μ = 0.0143 and σ = 0.0111.* ***C.*** *CSP analysis of the DnaA domain I titration with the DnaD DDBH2 domain truncation (residues 129 - 196). Threshold values of significance were calculated using μ = 0.0152 and σ = 0.0120.* ***A, B & C.*** *The μ + 2σ and μ + σ limits are highlighted in red and pink respectively.*

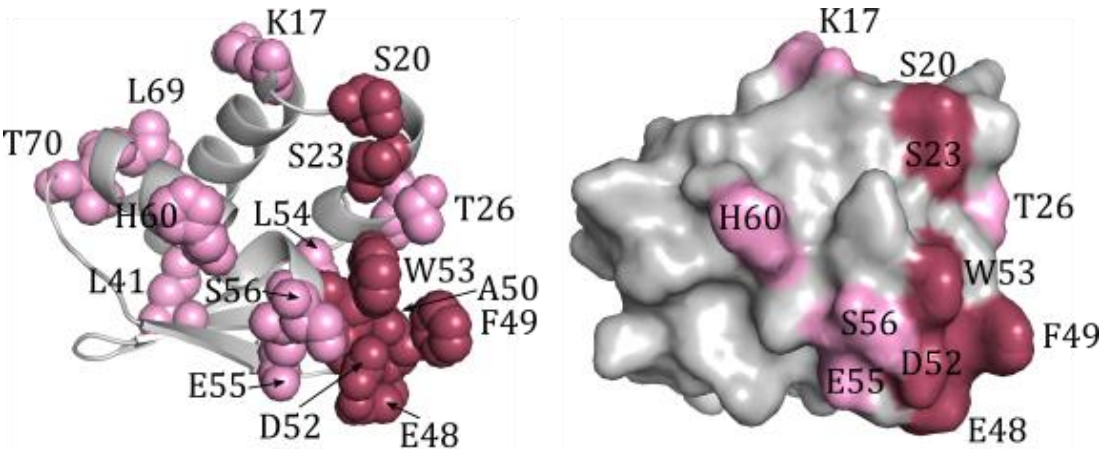


Figure 5.10 *Residues exceeding the CSP threshold limits of the DnaA domain I titrations with DnaD DDBH2 domain (DDBH2 truncation, and in the presence and absence of ssDNA) are mapped onto the X-ray crystal structure of DnaA domain I 9PDB – 4TPS).*

The chemical shift differences observed during the titration were also used to calculate a binding affinity between DnaA domain I and the DnaD DDBH2 domain. The following equation was used to fit shift differences observed at each ligand concentration:

$$y = A((B + x) - \sqrt{(B + x)^2 - 4x})) \qquad (5.2.1)$$

This equation is available as a fitting function within the CCNP Analysis software[132]. Where A = (maximum chemical shift difference)/2, B = 1 + K_D / [DnaA domain I], x = [DnaD DDBH2 domain] / [DnaA domain I] and y = chemical shift difference. The value for affinity can be extracted from a plot of the chemical shift difference against the ligand:protein ratio. Experimental data fitted in this way, for DnaA domain I amide backbone signals in the ^{15}N HSQC, are displayed in Figure 5.11 and table 5.1. The residues shown were selected from those that experienced the largest CSP during the titration experiments and provided the best fit to equation 5.2.1, determined by the error for the calculated K_D value. For dissimilar isotopes CCPN uses a scaling factor of 1 to compare shifts in the first dimension (^{1}H) and a scaling factor of 0.15 to compare shifts in the second dimension (^{15}N). Note that the scaling factor used in equation 5.1.1 is 0.14 for the nitrogen dimension, hence, there are minior differences between the shift differences reported between Figures 5.9 and 5.11. As discussed by Williamson[137] there is no theory to provide a weighting for the relative chemical shifts of different nuclei (numerous values between 0.1 and 0.45 have been used)

though the standard practise of the Euclidean weighting of approximately 0.14 seems to be optimal.

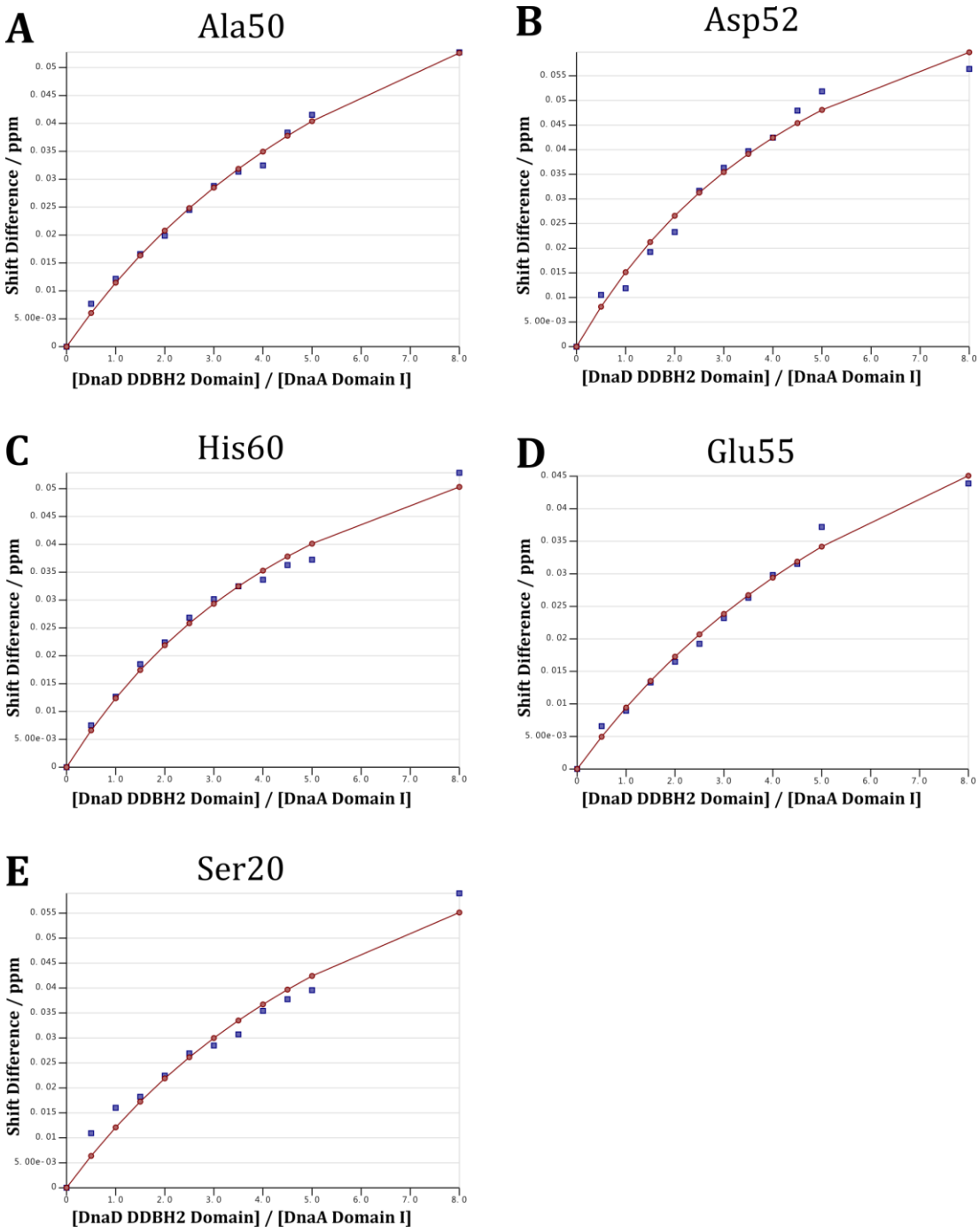


Figure 5.11 *The non-linear fit for the chemical shift changes of residues* ***A.*** *Ala50,* ***B.*** *Asp52,* ***C.*** *His60,* ***D.*** *Glu55 and* ***E.*** *Ser20 for the DnaD DDBH2 domain titration against DnaA domain I (100 μM). Experimental data are shown in blue and the fit shown in red.*

This method should ideally be used to fit data where saturation of the binding site has been achieved, to provide an accurate value for the A parameter dependent on the maximum chemical shift difference. However, from the plots in Figure 5.11 it is apparent that the NMR titration did not reach saturation at the binding site, as demonstrated by the continuing upward trajectory of the chemical shift difference. The limited stability of the DnaD DDBH2 domain at a higher ligand excess meant that saturation could not be achieved during these experiments. Despite this, the values calculated are of a similar magnitude (within the hundred μM rage) to those determined by SPR (section 5.3, below).

Peak	K_D / μM	Error K_D / μM
Ala50	701	85
Asp52	454	99
His60	491	78
Glu55	789	138
Ser20	686	197

Table 5.1 *The K_D parameters determined using the chemical shift differences measured in the DnaA domain I ^{15}N HSQC spectrum for the peaks corresponding to residues Ala50, Asp52, His60, Glu55 and Ser20. The mean plus standard deviation across the calculated K_D has a value of 768 μM with an error of 168 μM.*

5.2.2 Identifying the DnaA Interaction Patch of the DnaD DDBH2 Domain

The backbone assignment of the DnaD DDBH2 domain was previously completed by Dr Farhat Marston[90] (Craven Group, University of Sheffield) whose NMR data were kindly supplied to us for use in assignment of the ^{15}N-HSQC spectrum for this domain. This information removed the need for acquisition of three-dimensional NMR experiments on $^{13}C^{15}N$ labelled protein, as described for DnaA domain I in Chapter 4. The HSQC spectra observed for the DnaD DDBH2 domain (Figure 5.12) are relatively poor quality compared to the well resolved spectra obtained for DnaA domain I. This is possibly a consequence of the unstructured C-terminal region of the domain or the effect of transient self-association. The assignment shows a degree of overlap within the amide backbone resonances which would potentially complicate the observation of chemical shift perturbations in these regions.

The titration of unlabelled DnaA domain I (final concentration 800 μM) into a ^{15}N labelled sample of 100 μM DnaD DDBH2 domain was performed incrementally to

an 8:1 excess. Figure 5.13 illustrates selected peaks that were subject to perturbation during the experimental series. As described in section 5.2.1 the effect of DNA binding on the DnaA-DnaD interaction was investigated using ssDNA (5' GTTATTGCTC 3') that has been shown to be bound by the DnaD DDBH2 domain[90]. The titration of unlabelled DnaA domain I (final concentration 800 μM) into a ^{15}N labelled sample of 100 μM DnaD DDBH2 domain was repeated, in the presence of ssDNA (800μM), incrementally to an 8:1 excess. Figure 5.14 illustrates selected peaks that were subject to perturbation during the experimental series.

CSP analysis was performed on the resulting ^{15}N HSQC spectra from both of the titration experiments and plots showing the chemical shift difference vs residue number are displayed in Figure 5.15. Threshold values to determine significantly shifted resonances were determined as described in section 5.1.1, residues exceeding these thresholds are likely directly involved in the DnaA-DnaD interaction. The results from the separate titration experiments display a high degree of overlap between the residues that are observed above the threshold of significance. However, the addition of ssDNA to the experiment produced poorer quality HSQC spectra with weaker and broadened peaks displaying increased overlap between backbone amide residues. This is likely due to the increased tumbling time of the DnaD-DNA complex and complicates observation of chemical shift changes in certain regions of the ^{15}N HSQC spectrum. In particular the peaks corresponding to the unstructured C-terminal region of the domain were affected, and consequently certain amide resonances were excluded from the CSP analysis.

The residues identified as significant to the DnaA-DnaD binding interface were mapped onto the crystal structure of DnaA domain I as displayed in Figure 5.16. These residues correspond to a defined binding surface within the structured N-terminal region of the domain and can be clustered into two groups; Leu129, Tyr130, Ile132, Phe133, Glu134, and Glu135 located on the α1 helix and N-terminal region of the loop between the α1 and α2 helices; and Lys164, His165, Glu169 and Val171 located throughout the α3 helix. The residues Phe133 and Glu169 are not surface exposed on the identified interaction face and hence the CSP observed at these positions is likely due to allosteric effects experienced in these regions during DnaA domain I binding.

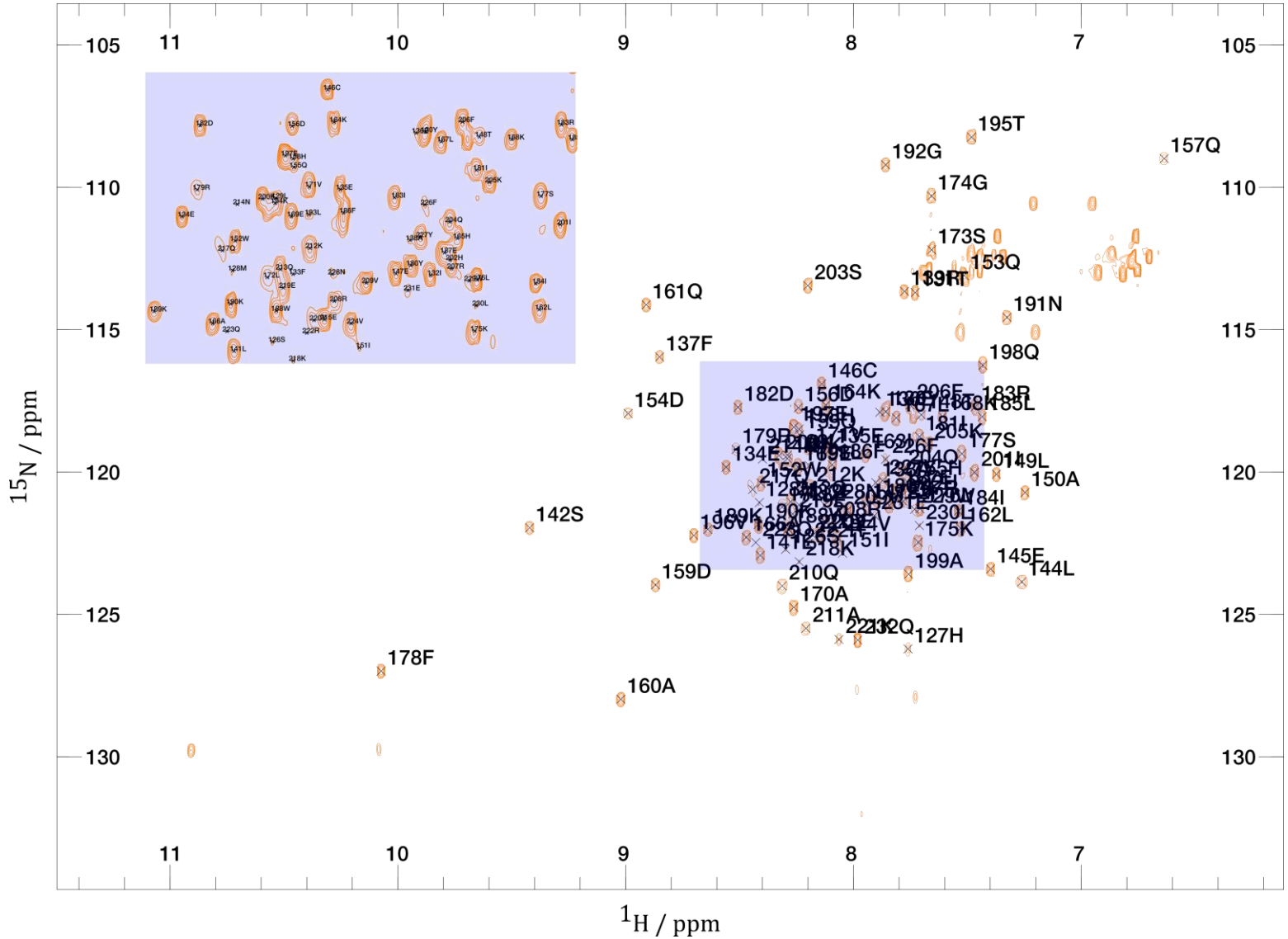


Figure 5.12 *The ^{15}N HSQC spectrum of the DnaD DDBH2 domain (100 μM). The backbone resonance peaks are labelled with the residue-specific assignment that was achieved using data supplied by Dr Jeremy Craven (University of Sheffield).*

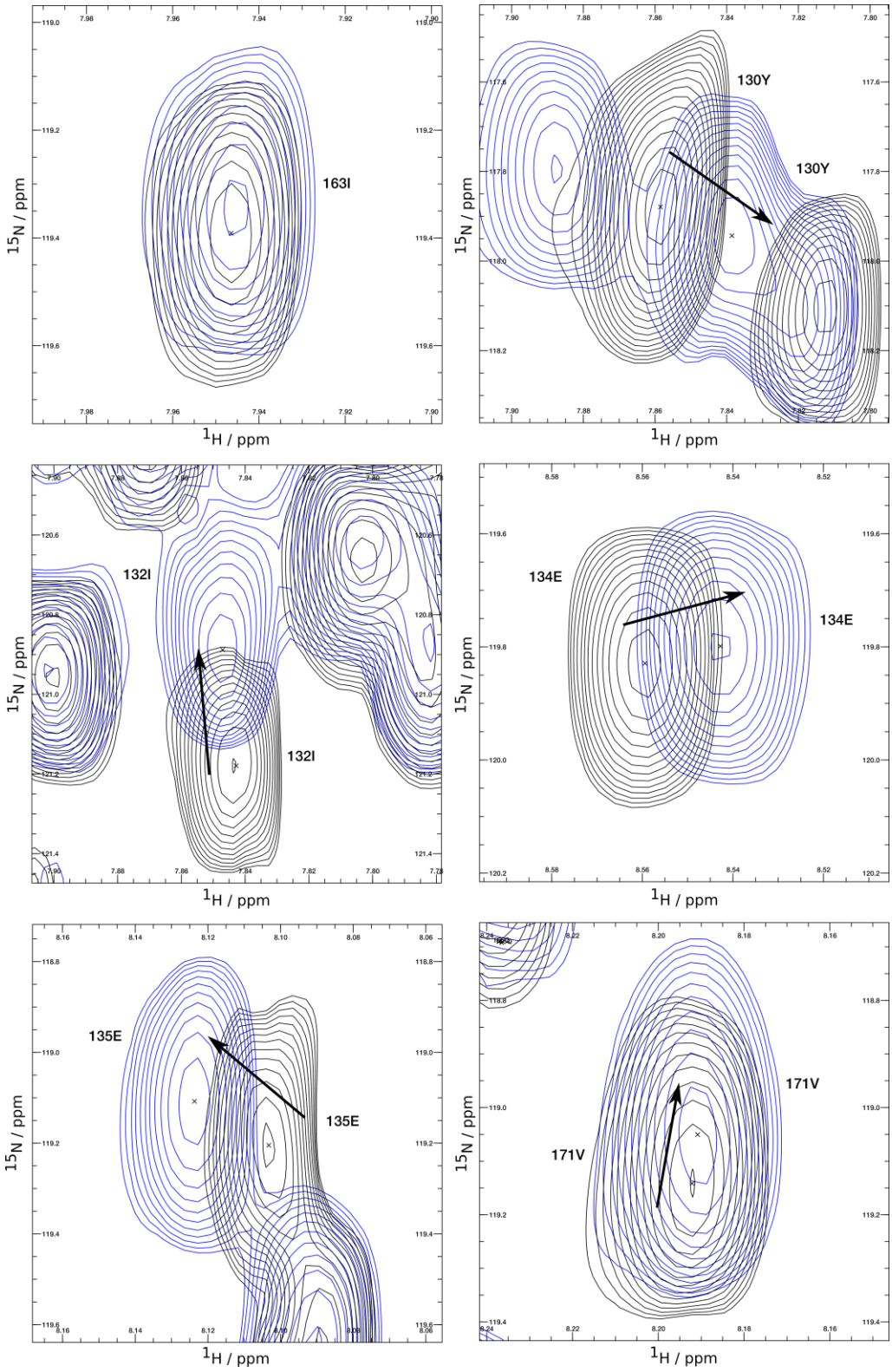


Figure 5.13 *Overlay of spectral peaks representing DnaD DDBH2 domain (100 μM) ^{15}N HSQC in black and the 8:1 excess of DnaA domain I (800 μM) to DnaA domain I spectrum is shown in blue. The arrows indicate direction of movement of the peak during the titration series. The isoleucine 163 peak (top left) has been included to illustrate the behaviour of peaks not involved in the binding interface.*

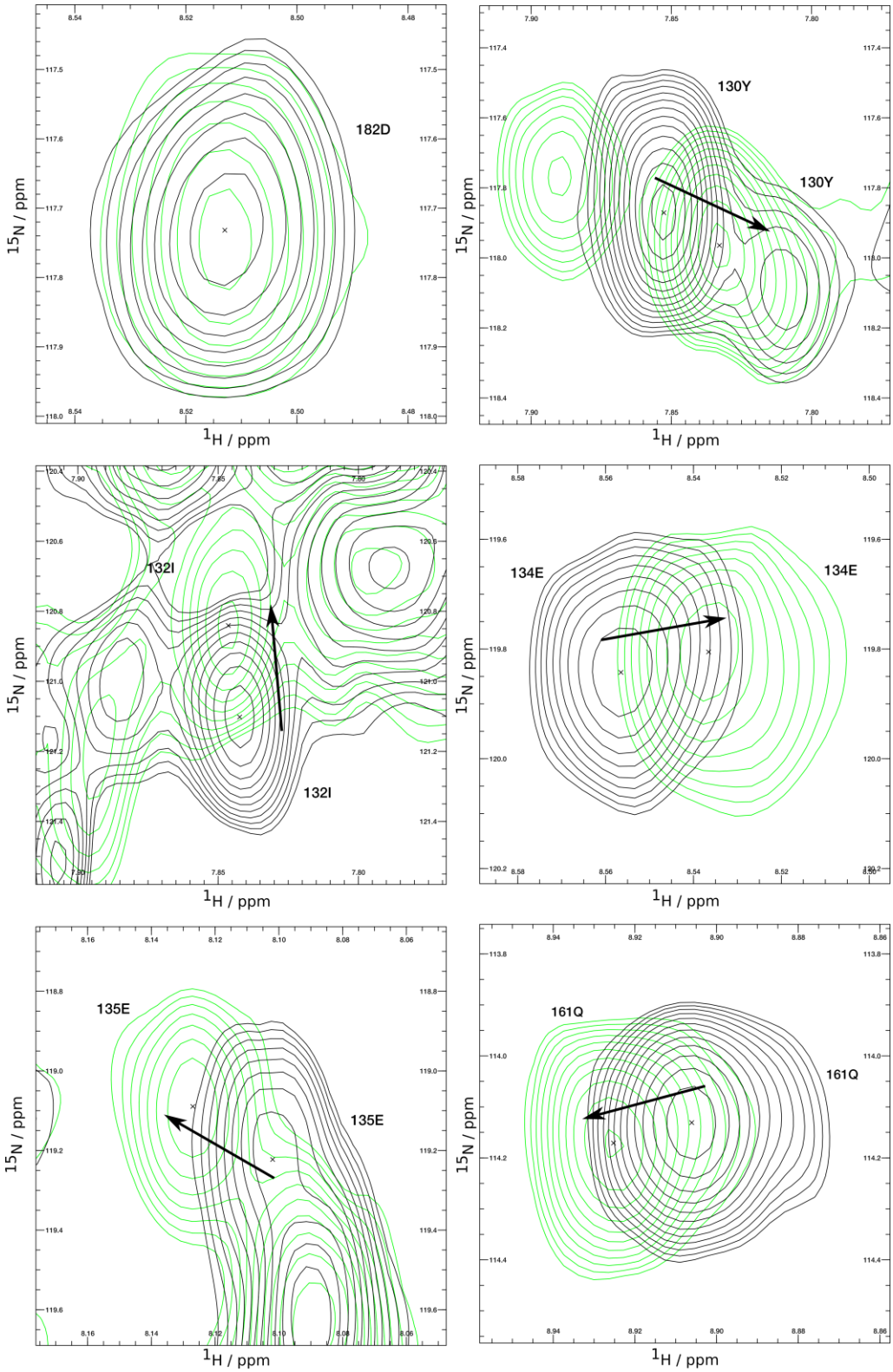


Figure 5.14 *Overlay of spectral peaks representing DnaD DDBH2 domain (100 µM) ^{15}N HSQC in black and the 8:1 excess of DnaA domain I (800 µM) to DnaA domain I spectrum is shown in green, both spectra were run in the presence of ssDNA (800µM). The arrows indicate direction of movement of the peak during the titration series. The aspartate 182 peak (top left) has been included to illustrate the behaviour of peaks not involved in the binding interface.*

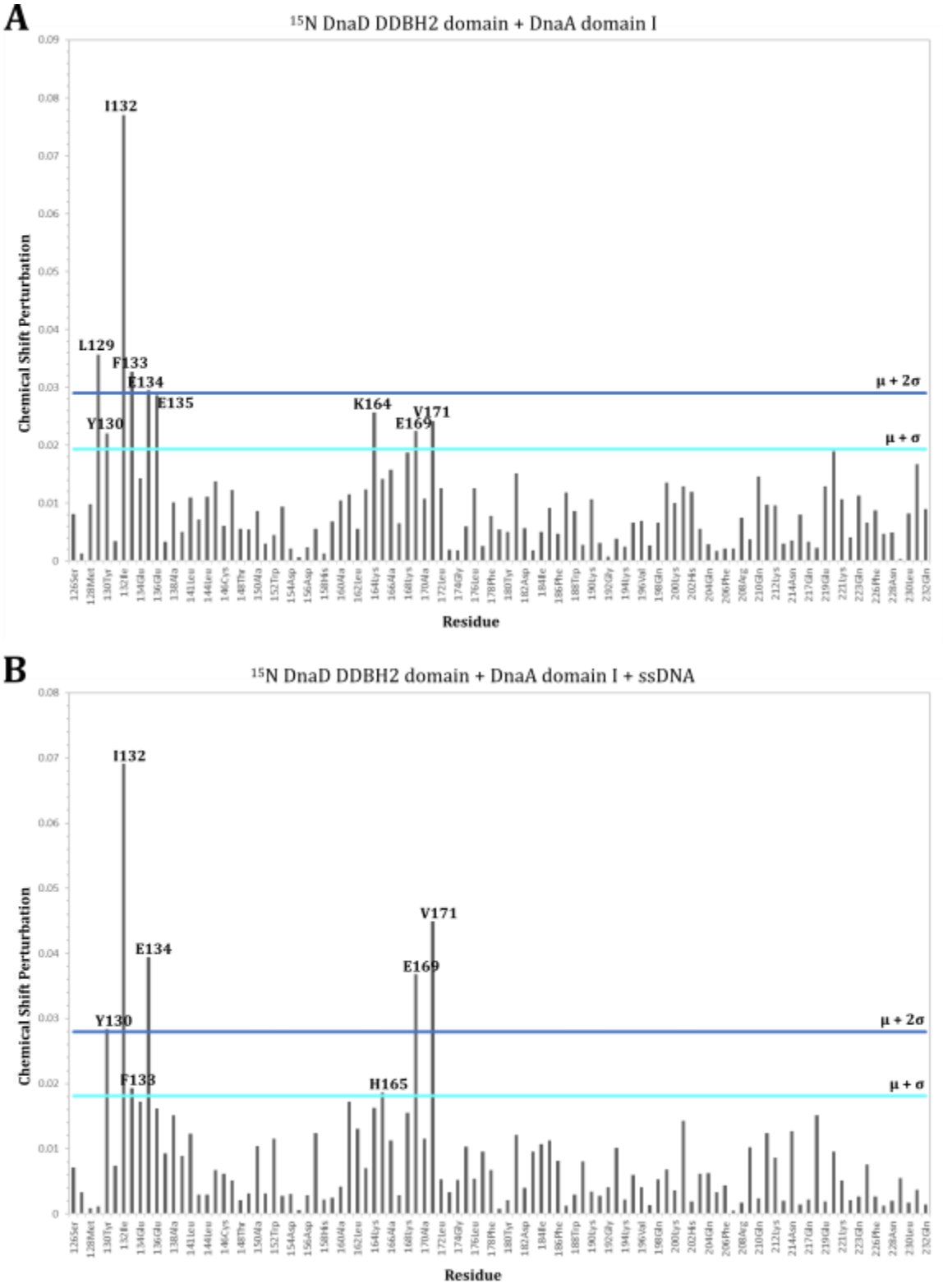


Figure 5.15 A. *CSP analysis of the DnaD DDBH2 domain titration with DnaA domain I. Threshold values of significance were calculated using μ = 0.00956 and σ = 0.00974.* ***B.*** *CSP analysis of the DnaD DDBH2 domain titration with DnaA domain I in the presence of ssDNA (800 μM). Threshold values of significance were calculated using μ = 0.00831 and σ = 0.00982.* ***A & B.*** *The μ + 2σ and μ + σ limits are highlighted in blue and cyan respectively.*

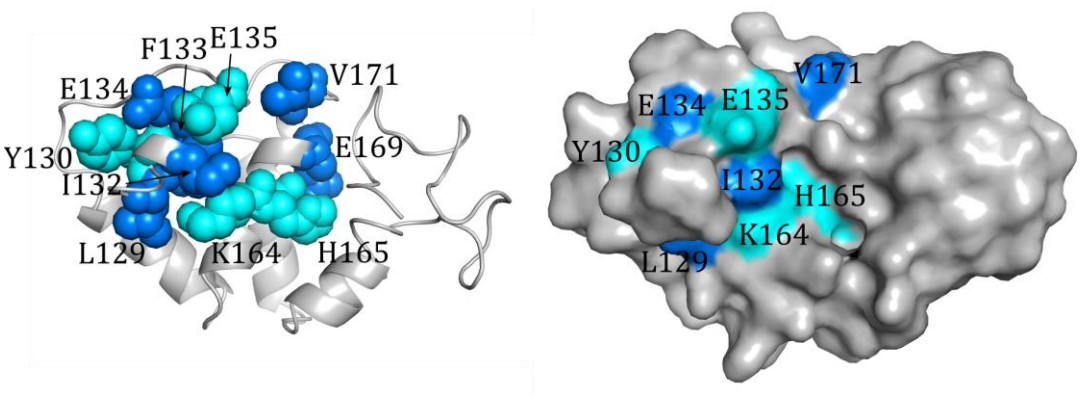


Figure 5.16 *Residues exceeding the CSP threshold limits of the DnaD DDBH2 domain titrations with DnaA domain (in the presence and absence of ssDNA) are mapped onto the modelled structure of the DnaD DDBH2 domain. Structure supplied by Dr Jeremy Craven (University of Sheffield) based on NMR coordinates and the crystal structure of the DnaD-like replication protein from Streptococcus mutans PDB – 2ZC2).*

The chemical shift differences observed during the titration were used to calculate a binding affinity between the DnaD DDBH2 domain and DnaA domain I. As described for section 5.2.1, equation 5.2.3 was used to fit shift differences observed at each ligand concentration. Where A = (maximum chemical shift difference)/2, B = 1 + K_D / [DnaD DDBH2 domain], x = [DnaA domain I] / [DnaD DDBH2] and y = chemical shift difference. Experimental data were fitted, as a plot of the chemical shift difference against the ligand:protein ratio, for DnaD DDBH2 domain amide backbone signals in the ^{15}N HSQC, as displayed in Figure 5.17 and table 5.2 shows the extracted binding affinity values for individual residues. The residues shown were selected from those that experienced the largest CSP during the titration experiments and provided the best fit to equation 5.2.1, determined by the error for the calculated K_D value.

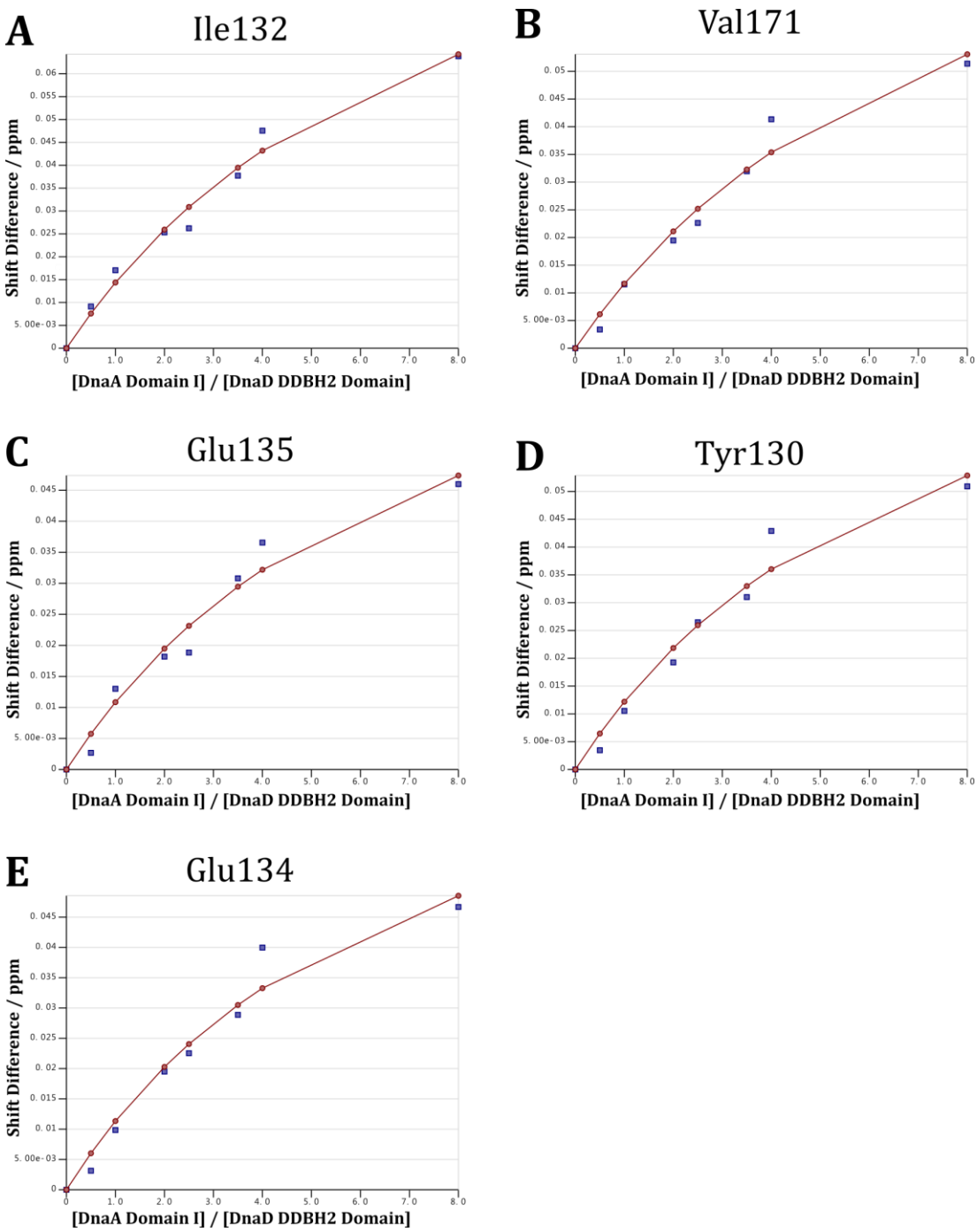


Figure 5.17 *The non-linear fit for the chemical shift changes of residues* ***A.*** *Ile132.* ***B.*** *Val171,* ***C.*** *Glu135,* ***D.*** *Tyr130 and* ***E.*** *Glu134 for the DnaA domain I titration against the DnaD DDBH2 domain (100 μM). Experimental data are shown in blue and the fit shown in red.*

As with DnaA domain I, the NMR titration did not reach saturation at the binding site, as demonstrated by the continuing upward trajectory of the chemical shift difference. As a consequence, the extracted binding affinity values are an estimate calculated based on extrapolated data. The plots in Figure 5.17 also show significant

differences between the experimental and fitted data and hence the K_D values obtained from this data produced high errors. The weak nature of the interaction between the single domains of DnaA and DnaD would align with only a portion of the population forming the bound complex, where the observed peaks represent a weighted average of the free and bound states, as indicated by the small chemical shift differences observed. However, the calculated values are of a similar magnitude to those calculated for DnaA domain I (section 5.2.1).

Peak	K_D / μM	Error K_D / μM
Ile132	648	188
Val171	684	248
Glu135	603	239
Tyr130	590	237
Glu134	568	230

Table 5.2 *The KD parameters determined using the chemical shift differences measured in the DnaD DDBH2 domain ^{15}N HSQC spectrum for the peaks corresponding to residues Ile132, Val171, Glu135, Tyr130 and Glu134. The mean plus standard deviation across the calculated K_D has a value of 665 μM with an error of 251 μM.*

5.3 SPR Kinetic Analysis

A ligand surface of DnaA domain I was generated through affinity capture of an N-terminal his_6-tag onto a nickel chelated NTA chip surface. Kinetic study of the interaction with the DnaD DDBH2 domain was conducted as a dilution series against the DnaA domain I surface. The series contained seven concentration points (48.12 – 770 μM) with one repeat, as displayed in Figure 5.18.

The data from the concentration series were initially fitted using the Langmuir isotherm (available witin the BIAevaluation software) and equilibrium constants were derived using recommendations by O'Shannessy *et al.* [151]. Both the separate and simultaneous fitting of the association and dissociation phases were investigated (Appendix 8.3). However, there were clear systematic deviations in the association and dissociation phases indicating the Langmuir 1:1 model is not optimum for these experimental data. This may be a consequence of limitations inherent within the technique, steady state analysis is typically considered the best treatment for SPR data.

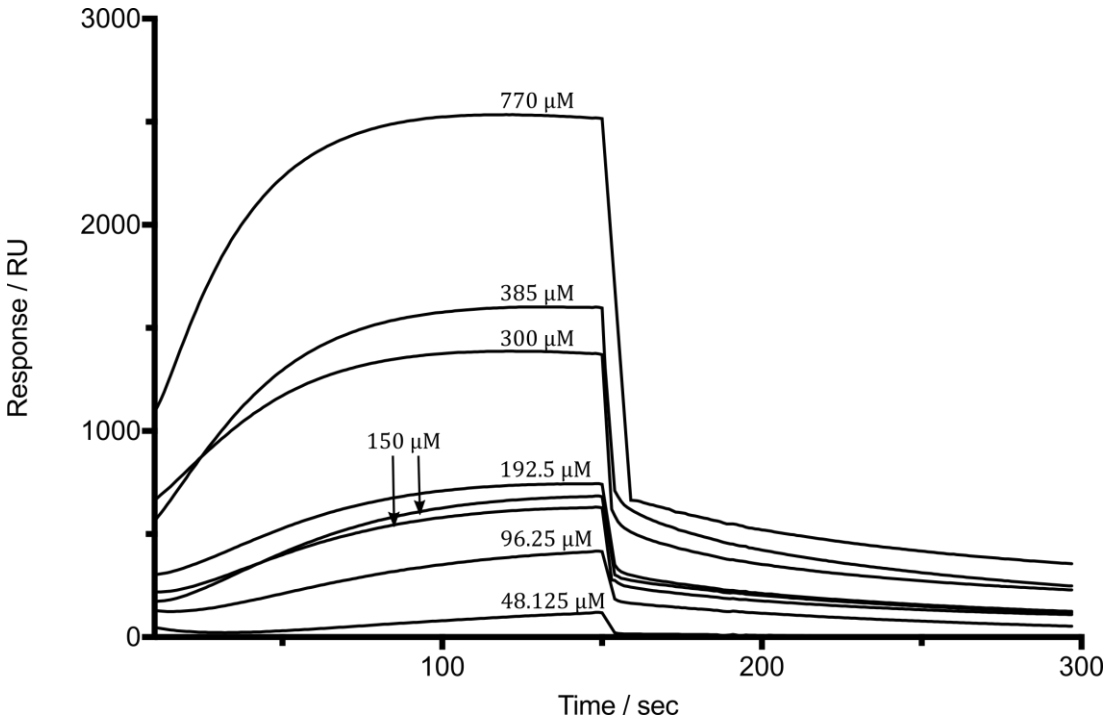


Figure 5.18 *The DnaD DDBH2 domain concentration series (48.125 – 770 µM) for the DnaA domain I immobilised surface. A repeat for the 150 µM concentration is included and a flow rate of 10 µL/min was maintained throughout. The sensogram profile indicates the binding interaction between DnaA domain I and the DnaD DDBH2 domain.*

Subsequently, a saturation plot (Figure 5.19) was generated from the end of the association phase, close to steady state, for non-linear regression Hill Slope analysis. The results are summarised in Table 5.3. The calculated K_D values using this method are of a similar magnitude to those estimated from the NMR titration data (section 5.2). For a 1:1 binding interaction the Hill parameter *h* is expected to be close to 1, however, the calculated value from the experimental data is 1.441. This value corresponds to a receptor or ligand that has multiple binding sites with positive cooperativity and is consistent with the slight sigmoidal shape to the plot. If the interaction between DnaA domain I and the DnaD DDBH2 domain is not a 1:1 interaction this would explain the poor fit of the experimental data to the Langmuir 1:1 binding models. DnaA domain I was immobilised to the sensor surface in one orientation as dictated by the N-terminal his_6-tag used for affinity capture, therefore it is more probable that the analyte (DnaD DDBH2 domain) contains the multiple binding sites.

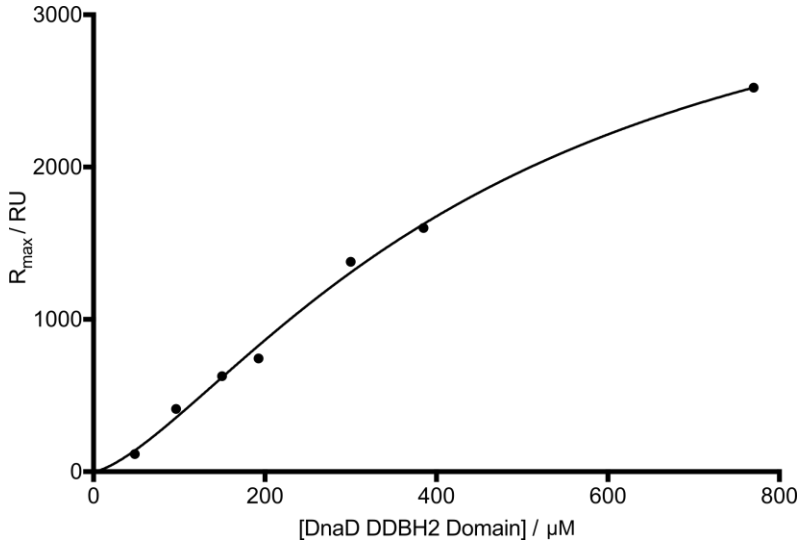


Figure 5.19 *A saturation plot with non-linear regression Hill Slope analysis was used to calculate the DnaD DDBH2 domain binding affinity to the DnaA domain I surface.*

Construct	K_D / μM	B_{max}	h	R^2
DnaD DDBH2 Domain	458.2 ±95.14	3715	1.441	0.996

Table 5.3 *The binding affinity of the DnaD DDBH2 domain for the DnaA domain I surface were calculated from the non-linear regression Hill slope analysis of the saturation plot.*

The bivalent analyte model, available within the BIAevaluation software, was investigated as a consequence of steady state analysis results. This model describes binding of a bivalent analyte to immobilised ligand, where one analyte molecule can bind one or two ligand molecules, and is often relevant to signalling molecules where dimerisation of the receptor is common or the study of intact antibodies binding to an immobilised antigen. However, as with the Langmuir 1:1 binding model clear systematic deviations were observed between the experimental and fitted data indicating an inadequate fit (Appendix 8.3).

5.3.1 DnaA domain I Mutants

The results of the chemical shift perturbation analysis identified the involvement of the DnaA domain I α2, α3 and α4 helices in the interaction with the DnaD DDBH2 domain. In particular, the α3 residues E48 – W53 (Highlighted in Figure 5.20) displayed consistently significant levels of perturbation during the titration experiments and interestingly this region overlaps with the SirA binding patch[107]. Therefore, this region was selected for further investigation into the involvement in the DnaA-DnaD interaction.

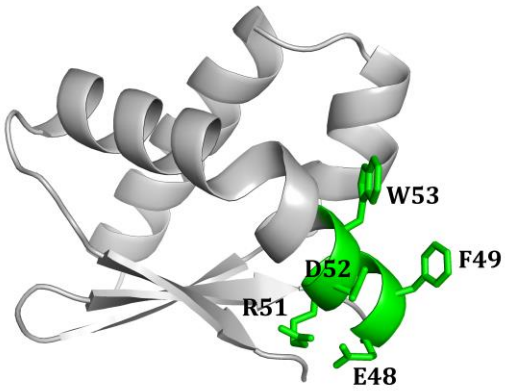


Figure 5.20 *The DnaA domain I crystal structure with region Glu48 – Trp53 highlighted in green with residue side chains displayed as sticks. Significant chemical shift perturbation was observed in this region on addition of the DnaD DDBH2 domain. Consequently, residues within this region were selected for mutation experiments to further investigate the DnaA-DnaD interaction.*

Initially four simultaneous point mutations were generated with conservative substitutions (based on Hidden Markov Model results) to prevent disruption to the secondary structure of the domain. Within the region E48 – W53 the following mutations were incorporated; E^{48} to Aspartate, F^{49} to Tyrosine, D^{52} to Glutamate and W^{53} to Threonine, the mutant product is denoted DYARET. A ligand surface of the DnaA domain I mutant DYARET was generated through affinity capture of an N-terminal his_6-tag onto a nickel chelated NTA chip surface. Kinetic study of the interaction with the DnaD DDBH2 domain was conducted as a dilution series against the DnaA domain I surface. The series contained five concentration points (96.25 – 385 μM) with one repeat, as displayed in Figure 5.21.

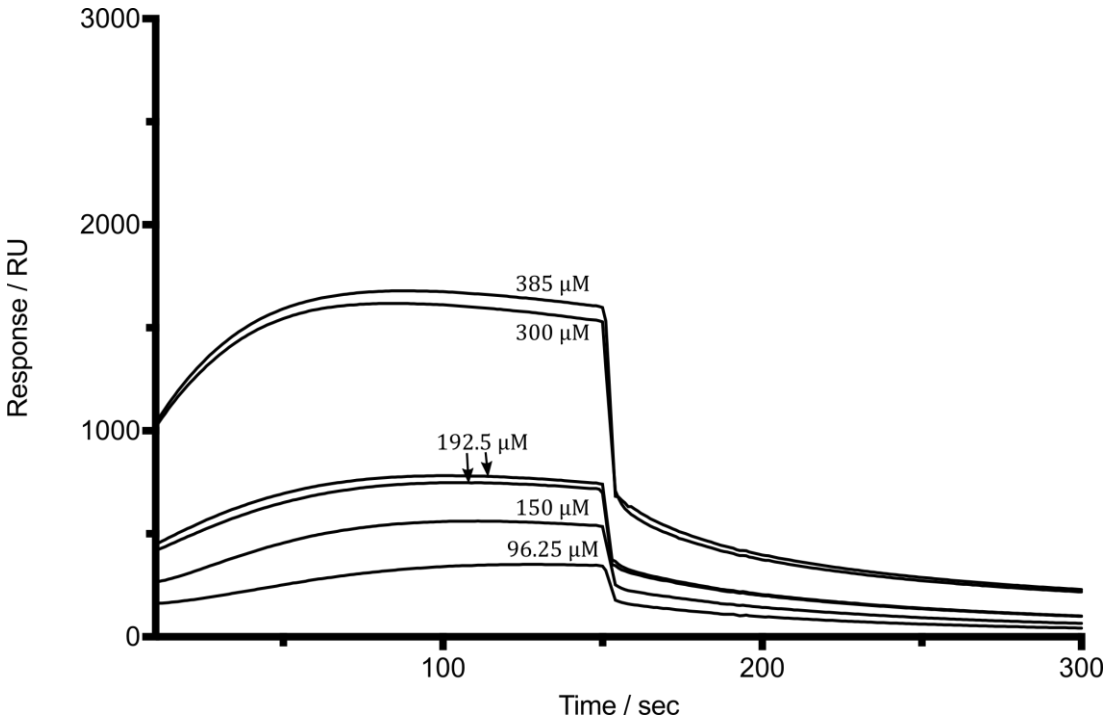


Figure 5.21 *The DnaD DDBH2 domain concentration series (96.25 – 385 µM) for the DnaA domain I mutant DYARET immobilised surface. A repeat for the 192.5 µM concentration is included and a flow rate of 10 µL/min was maintained throughout.*

As previously noted, the kinetic analysis of the interaction between DnaA domain I and the DnaD DDBH2 domain, using the models within the BIAevaluation software, did not produce good quality fits to the experimental data and hence kinetic values generated using this method may not provide true representations for the interaction. Therefore, steady state analysis was chosen to analyse data from the mutant DnaA domain I experiments. A saturation plot (Figure 5.22) was generated from the end of the association phase, close to steady state, for non-linear regression Hill Slope analysis. The results are displayed in Table 5.4.

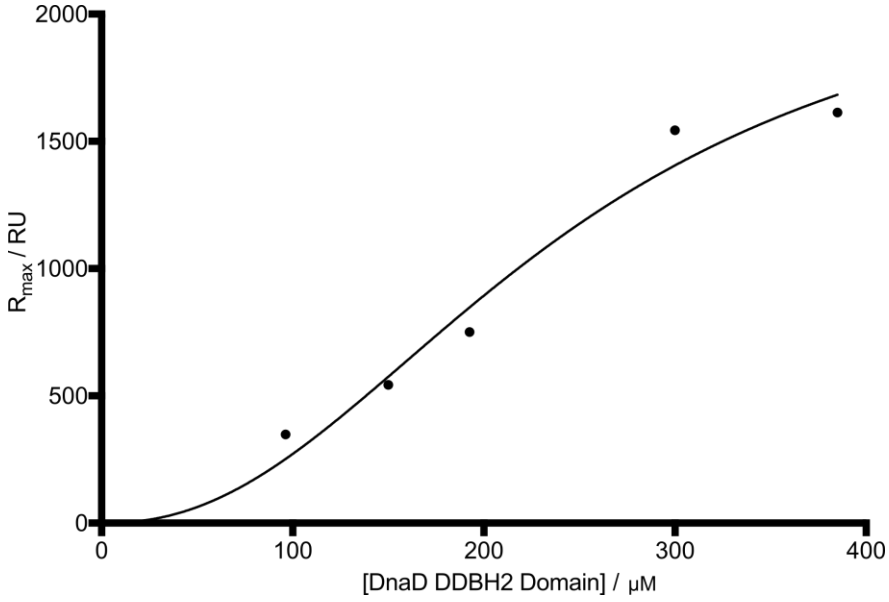


Figure 5.22 *A saturation plot with non-linear regression Hill Slope analysis was used to calculate the DnaD DDBH2 domain binding affinity to the DnaA domain I mutant DYARET surface.*

Construct	K_D / μM	B_{max}	h	R^2
Wild Type	458.2 ±95.14	3715	1.441	0.996
DYARET	243.1 ±118.1	2281	2.252	0.9679

Table 5.4 *The binding affinity of the DnaD DDBH2 domain for the DnaA domain I, wild type and mutant DYARET surface were calculated from the non-linear regression Hill slope analysis of the saturation plot.*

The Hill Slope analysis indicates that the kinetics of the binding interaction have been altered. However, contrary to the aim of the mutation, the binding affinity appears to have increased according to the calculated K_D value. This aligns with the slight increase in response observed during the concentration series compared to the wild type domain (Figure 5.18). Also of interest is the h value of 2.252 indicating and interaction greater than 1:1 with positive cooperativity, although this was also observed for the wild type domain the value has significantly increased for the DYARET mutant. The R^2 value (a fraction between 0.0 and 1.0) quantifies goodness of fit, where a higher value indicates the model fits the data better. The lower R^2 value compared to the wild type plot, as well as the higher error in the calculated values, may account for the observed differences in the K_D and h values for this mutant.

The conservative mutation of residues within the DnaA domain I α4 helix did not disrupt the binding interaction with the DnaD DDBH2 domain. Consequently, residues within this region were selected for mutation to Alanine with an aim to disrupt the binding interaction. Single residue mutations at Phe49, Asp52 and Trp53 were chosen to minimise disruption of the secondary structure within the domain. A ligand surface of the DnaA domain I mutant (W53A, D52A or F49A) was generated through affinity capture of an N-terminal his_6-tag onto a nickel chelated NTA chip surface. Kinetic study of the interaction with the DnaD DDBH2 domain was conducted as a dilution series against the mutant DnaA domain I surface. The series contained six concentration points (96.25 – 770 μM) with one repeat, as displayed in Figures 5.23 and 5.24. The F49A mutant surface failed to produce a consistent response during the concentration series. Absence of clear association and dissociation phases and lack of concentration dependence was observed for this mutant and subsequently the data from the experimental series could not be used to obtain kinetic data for the interaction.

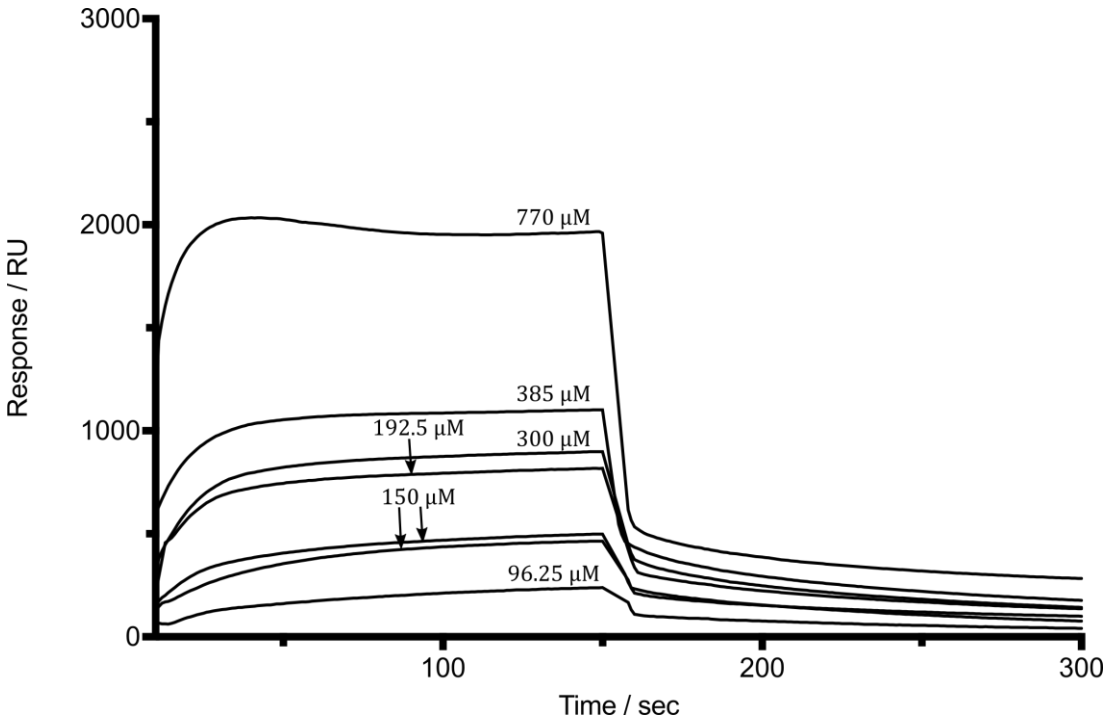


Figure 5.23 *The DnaD DDBH2 domain concentration series (96.25 – 770 μM) for the DnaA domain I mutant W53A immobilised surface. A repeat for the 150 μM concentration is included and a flow rate of 10 μL/min was maintained throughout.*

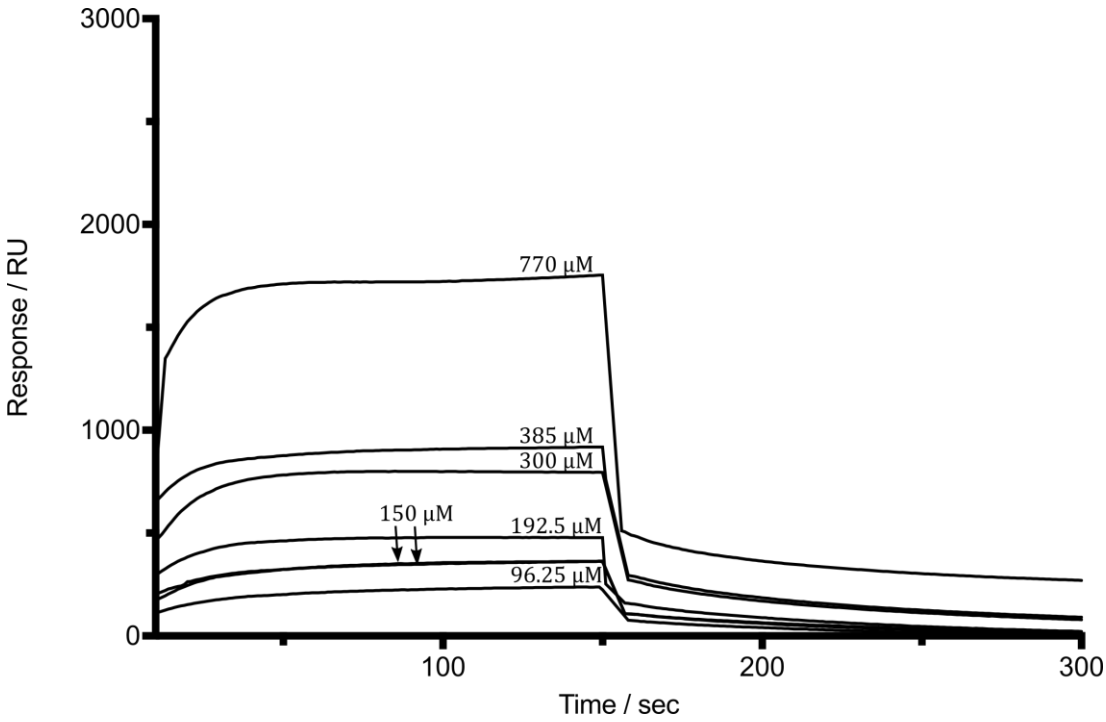


Figure 5.24 *The DnaD DDBH2 domain concentration series (96.25 – 770 µM) for the DnaA domain I mutant D52A immobilised surface. A repeat for the 150 µM concentration is included and a flow rate of 10 µL/min was maintained throughout.*

The interaction with the DnaD DDBH2 domain is maintained with the D52A and W53A mutant surfaces, however, the magnitude of response compared to the wild type domain indicate a weaker interaction is being observed. At the 300 µM concentration point there is 1.5x and 1.7x decrease in response for mutants W53A and D52A respectively. A noticeable different in the shape of the sensogram profiles is also observed, particularly at the 770 µM concentration, indicating the mutations have causes a change in the kinetic profile of the interaction. Saturation plots for the mutants D52A and W53A (Figure 5.25) were generated from the end of the association phase, close to steady state, for non-linear regression Hill Slope analysis. The results are displayed in Table 5.5.

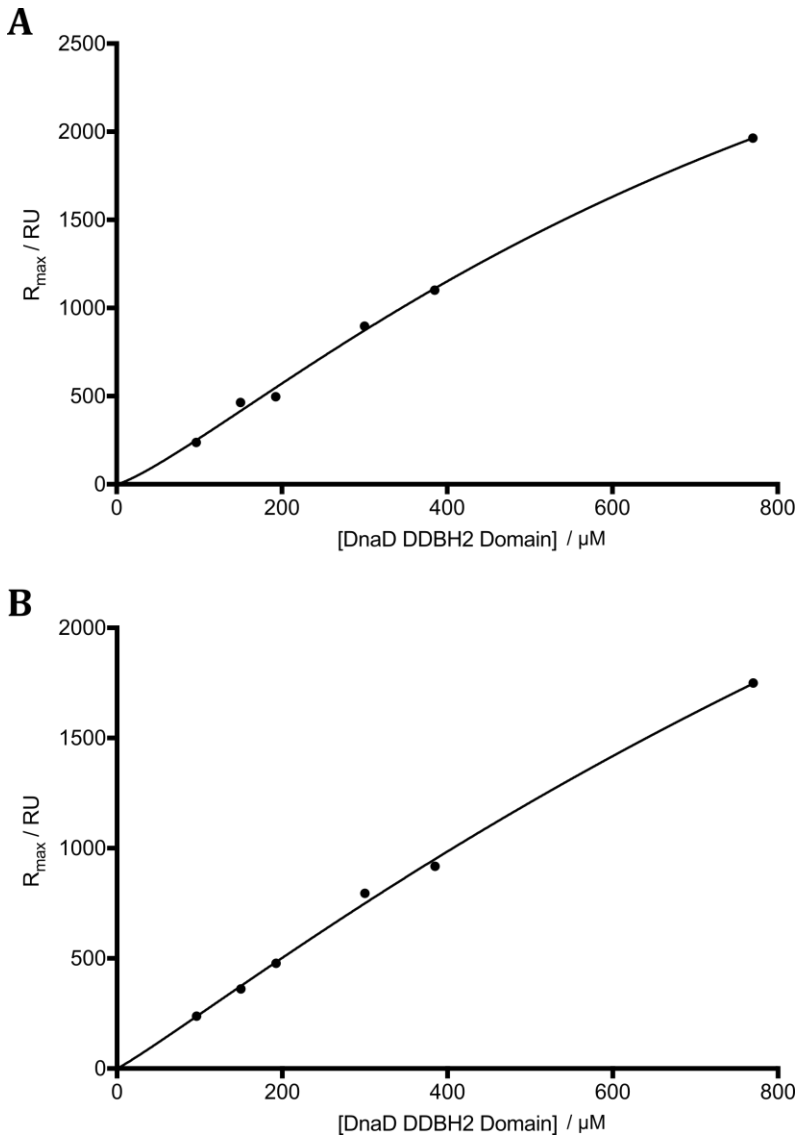


Figure 5.25 *Saturation plots with non-linear regression Hill Slope analysis were used to calculate the DnaD DDBH2 domain binding affinity to the DnaA domain I mutant surfaces* ***A.*** *W53A and* ***B.*** *D52A.*

Construct	K_D / µM	B_{max}	h	R^2
Wild Type	458.2 ±95.14	3715	1.441	0.996
W53A	945.2 ±483.2	4496	1.241	0.997
D52A	2084 ±2108	6897	1.086	0.997

Table 5.5 *The binding affinity of the DnaD DDBH2 domain for the DnaA domain I, wild type and mutant, surface were calculated from the non-linear regression Hill slope analysis of the saturation plot.*

The K_D values, generated using steady state analysis, are significantly higher than the wild type with a 2-fold increase for the W53A mutant surface and almost 5-

fold increase with the D52A mutant. The Hill Slope parameter *h* value of 1.086 calculated for the D52A mutant indicates a 1:1 interaction rather that the multiple binding sites indicated by the wild type domain data. This suggests that the mutation has destabilised the sphere for ligand binding within DnaA domain I and thus reduced affinity for the DnaD DDBH2 domain. Although the error in calculated K_D values for the DnaA mutants is very high this is often a consequence of measuring such weak interactions, and is likely to be close to the limit of what can be measured by SPR.

5.4 Discussion

5.4.1 Mapping the DnaA-DnaD Binding Interface

Chemical shift perturbation analysis, of the titrations of the DnaD DDBH2 domain into DnaA domain I, highlights three regions within the domain where residues exceed the CSP threshold of significance. Mapped onto the X-ray structure (PDB - 4TPS, SirA in complex with domain I of DnaA) it is observed that two main regions of the structure are perturbed, at the higher threshold, located within the α2 and α3 helices with the third region, perturbed to a lesser extent, located on the α4 helix. Together these form a defined binding surface within DnaA domain I. Interestingly, this overlaps with the DnaA-SirA binding interface identified by X-ray crystallography[107] as shown in Figure 5.26. SirA is an inhibitor of DNA replication that prevents re-initiation of replication, by binding to DnaA, in *B. subtilis* cells committed to sporulation[152]. The overlap between these binding interfaces within DnaA domain I suggests that DnaD interacts at the interface while DnaA is active and that the SirA binding interaction could sterically hinder the DnaA-DnaD interaction, as a negative regulator of replication initiation, by competing for the same binding site.

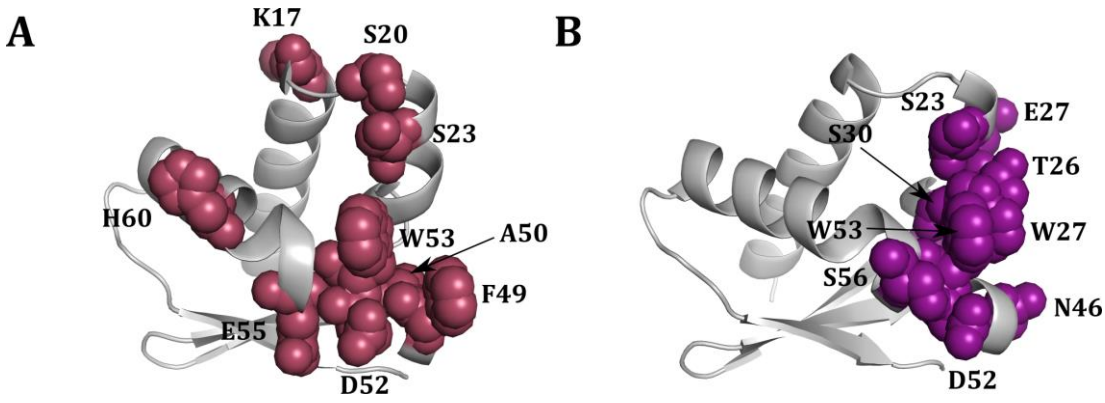


Figure 5.26 A. *Residues of the DnaD binding interface, identified by CSP analysis, mapped onto the X-ray crystal structure of DnaA domain I.* ***B.*** *Residues of the SirA binding interface, identified by X-ray crystallography, mapped onto the structure of DnaA domain I [PDB – 4TPS].*

As displayed in Figure 5.27, the DnaD binding interface also overlaps with the corresponding α2 and α3 helices identified as the interaction surface between *H. pylori* DnaA with, regulator of DNA replication, HobA[69]. As discussed in section 1.5.2, we hypothesize that the *B. subtilis* DnaD tetramer resembles the HobA tetramer found in *H. pylori* (Figure 1.15) and these results may support this theory. DnaD and HobA may have comparable roles during DNA replication initiation whereby HobA forms a structural scaffold for DnaA self-oligomerisation.

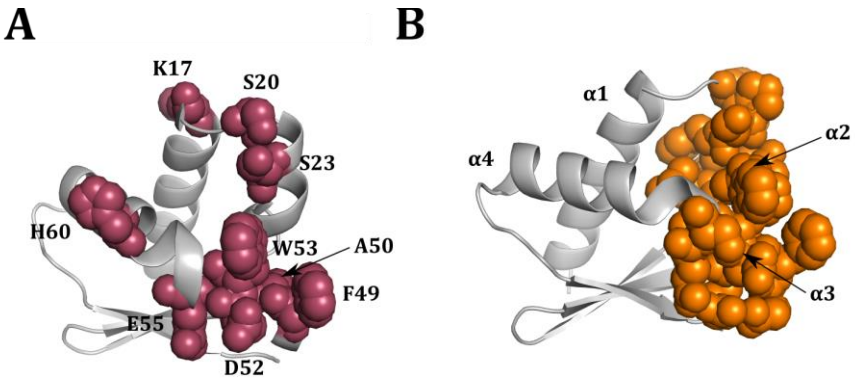


Figure 5.27 A. *Residues of the DnaD binding interface, identified by CSP analysis, mapped onto the X-ray crystal structure of DnaA domain I (PDB – 4TPS).* ***B.*** *The α2 and α3 helices of DnaA domain I, highlighted orange, correspond to the H. pylori HobA-DnaA interface. Based on the X-ray crystal structure of H. pylori HobA bound to DnaA (PDB – 2WP0)[69].*

Chemical shift perturbation analysis, for the titrations of DnaA domain I into the DnaD DDBH2 domain, highlights two main regions within the DDBH2 domain where residues exceed the CSP threshold of significance. Mapped onto the modelled structure it is observed that these regions are located within the α1 and α3 helices to form a defined binding surface within the DnaD DDBH2 domain. This region is consistent with previous yeast-two-hybrid data[106] identifying residues 134 – 140 as essential to the interaction with DnaA. The DNA binding interface on the DDBH2 domain was identified using NMR titration experiments and gel shift analysis of DnaD mutants[90]. The conserved YxxxIxxxW motif located within the α4 helix along with residues 206 – 215 located in the unstructured C-terminal region of the domain are essential to the DNA binding activity of DnaD. It is of interest that the DnaA binding interface, identified by CSP analysis, is observed on the opposite face to that involved in DNA binding (Figure 5.20). This suggests that DnaD can bind both DNA and DnaA simultaneously. However, examining the orientation of a DnaA filament bound to DNA it is unlikely that a single DnaD DDBH2 domain could interact with both DnaA domain I and DNA at the same time and it is more probable that within the DnaD tetramer separate monomers undergo the two interactions.

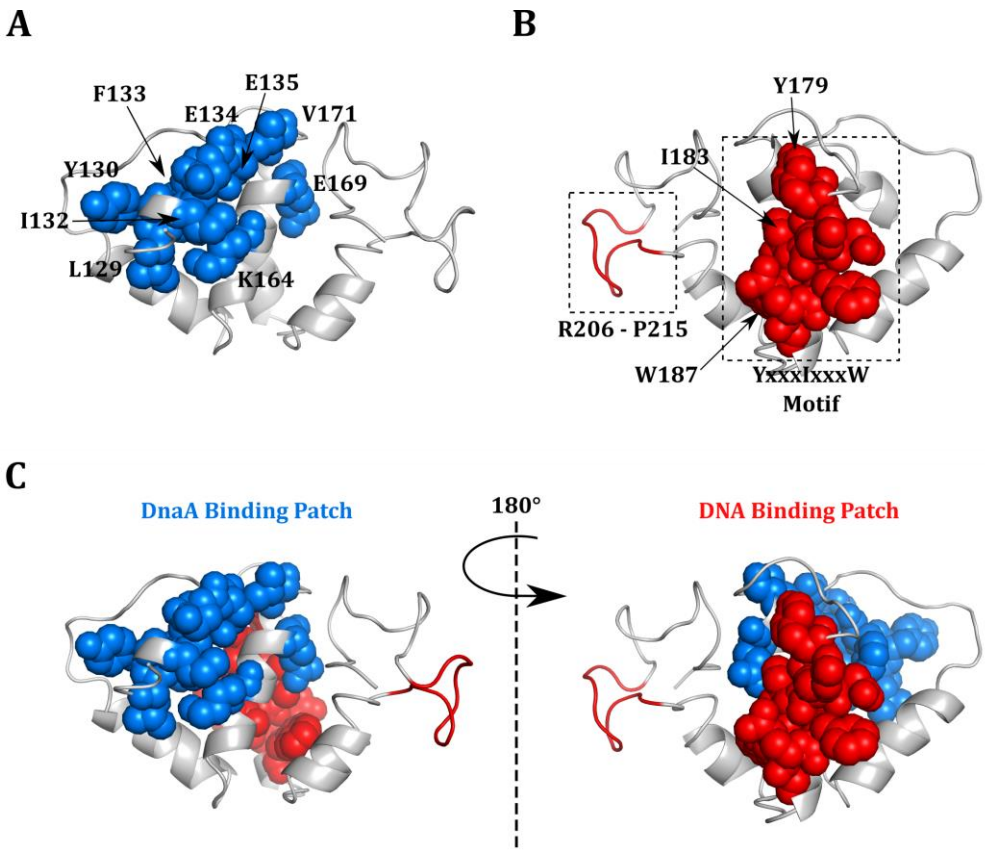


Figure 5.20 A. *Residues of the DnaA binding interface, identified by CSP analysis, mapped onto the NMR structure of the DnaD DDBH2 domain.* ***B.*** *The conserved YxxxIxxxW motif and residues 206 -215, essential for DNA binding activity, highlighted on the DnaD DDBH2 domain.* ***C.*** *Residues of the DnaA binding interface (blue) and the DNA binding interface (red) highlighted on the NMR structure of the DnaD DDBH2 domain. Interface regions form distinct surfaces found on opposite faces of the domain.*

5.4.2 Kinetic Analysis of the DnaA-DnaD Interaction

ITC was originally investigated for study of the binding affinity of the DnaA-DnaD interaction. However, no response could be obtained using this method under a range of conditions, possibly due to the very weak nature of the binding interaction and limited supply of material. Hence, the more versatile technique SPR was utilised for investigation of the interaction. The CM5 (carboxymethylated dextran matrix) surface failed to capture the ligand, DnaA domain I or DnaD DDBH2 domain, in an orientation that produced a response for binding of the analyte. This is not unexpected as lysine residues, used for coupling to the surface, are located throughout both domains with some residing within the binding interface identified by CSP studies. Consequently, the

NTA (carboxymethylated dextran pre-immobilised with nitrotriacetic acid) surface was chosen as both constructs contained an N-terminal his_6-tag (cleaved by thrombin prior to NMR studies) utilised for the purification process. No response was observed for the DnaA DDBH2 domain surface on addition of analyte, likely due to the proximity of the interaction surface to the N-terminal region of the domain. However, this method proved successful for capture of the DnaA domain I surface.

The experimental data obtained for the concentration series, of DnaD DDBH2 domain against the DnaA domain I surface, confirmed the binding interaction between the two domains and that the interaction occurs in the absence of additional supporting interactions such as DNA binding. The K_D values calculated for this interaction are the first to be reported and indicate the interaction is very weak, occurring in the μM - mM range. The value calculated by steady state analysis, using saturation plots with Hill Slope analysis, showed a reasonable agreement with those calculated for individual residues, using NMR titration data, indicating a reliable value despite the high associated errors. There is an indication from Hill Slope analysis that the interaction may not be 1:1 but instead involve multiple binding sites with positive cooperativity. This model is not unlikely as multiple points of contact could be expected *in vivo* where a scaffold of DnaD tetramers would interact with a helical DnaA nucleoprotein filament. It also aligns with the weak nature of the interaction, where many low affinity interactions combine to produce a significant conformational change in the duplex DNA at the *oriC* during DNA replication initiation.

Chemical shift perturbations are highly sensitive to changes in the environment and observed significant CSPs for individual residues may arise from both direct and allosteric effects. SPR was used as a method to aid characterisation of the binding interface between DnaA and DnaD. Mutations within the α3 helical region of DnaA domain I were chosen as this region is significant to both the interaction with DnaD and with the inhibitor of replication initiation SirA[107]. The conservative mutation of four residues within the helix (E48D, F49Y, D52E, W53T) did not inhibit the interaction with DnaD leading to individual residues being substituted for alanine (F49A, D52A, W53A). The F49A surface did not produce a reliable response on addition of analyte which prevented interpretation of the kinetics for this interaction. This is likely due to a marked decrease in the binding affinity outside of the limits for observation by SPR. Nevertheless, the D52A and W53A surfaces produced interesting results during the concentration series. The interaction with the DnaD DDBH2 domain is still supported by these DnaA domain I mutants. It was evident that the mutations altered the kinetic

profile of the interaction as differences were observed in the curve of the association phase compared to the wild type surface. The response, for both mutants, was decreased in comparison to the wild type surface and this aligned with the K_D values calculated from steady state analysis which indicated a clear decrease in affinity for the DnaD DDBH2 domain. Hill Slope analysis indicated that the interaction is closer to 1:1 for both mutants, particularly D52A, suggesting that if multiple binding sites are present then one of these could be dependent on the α3 region mutated corresponding to the observed decrease in affinity for DnaD without elimination of the interaction.

Further optimisation of the SPR experimental conditions is necessary to produce a more accurate kinetic model for the DnaA-DnaD interaction. The data fitting poorly to the available BIAevaluation models is an important limiting factor for these results. Ideally each residue identified by CSP analysis, both within DnaA domain I and the DnaD DDDBH2 domain, would be mutated for use in the SPR experiments to gain information on their exact involvement in the DnaA-DnaD interaction. However, the data from these preliminary experiments highlight the significance of the individual residues to the DnaA-DnaD interaction while indicating the importance of the binding interface as a whole, rather than individual contacts, to the DnaA-DnaD interaction.

Chapter 6

Modelling the DnaA-DnaD Interface

The most accurate structures of protein complexes are deduced experimentally using X-ray crystallography and NMR spectroscopy and are deposited in the Protein Data Bank (PDB). As of April 2018, the PDB contained 139555 structures (note this does not represent unique structures as multiple entries are entered for the same protein determined under differing conditions) with approximately 90% from X-ray and 9% from NMR experimental methods[153]. There is a large gap between the number of unique sequences known and the structures solved, and in recent years computational methods for the prediction of 3D models of proteins have become invaluable. There are two distinct methods for the computational modelling of protein-protein complex structures; ridged-body protein docking, and template-based modelling (also referred to as homology-modelling). Although the structure of DnaA domain I was solved by X-ray crystallography[107], the DnaD DDBH2 domain has a largely unstructured C-terminus rendering this method unsuitable for determination of the DnaA-DnaD complex structure. Therefore, computational methods were explored for modelling the DnaA-DnaD binding interface, providing further structural insights into this essential interaction.

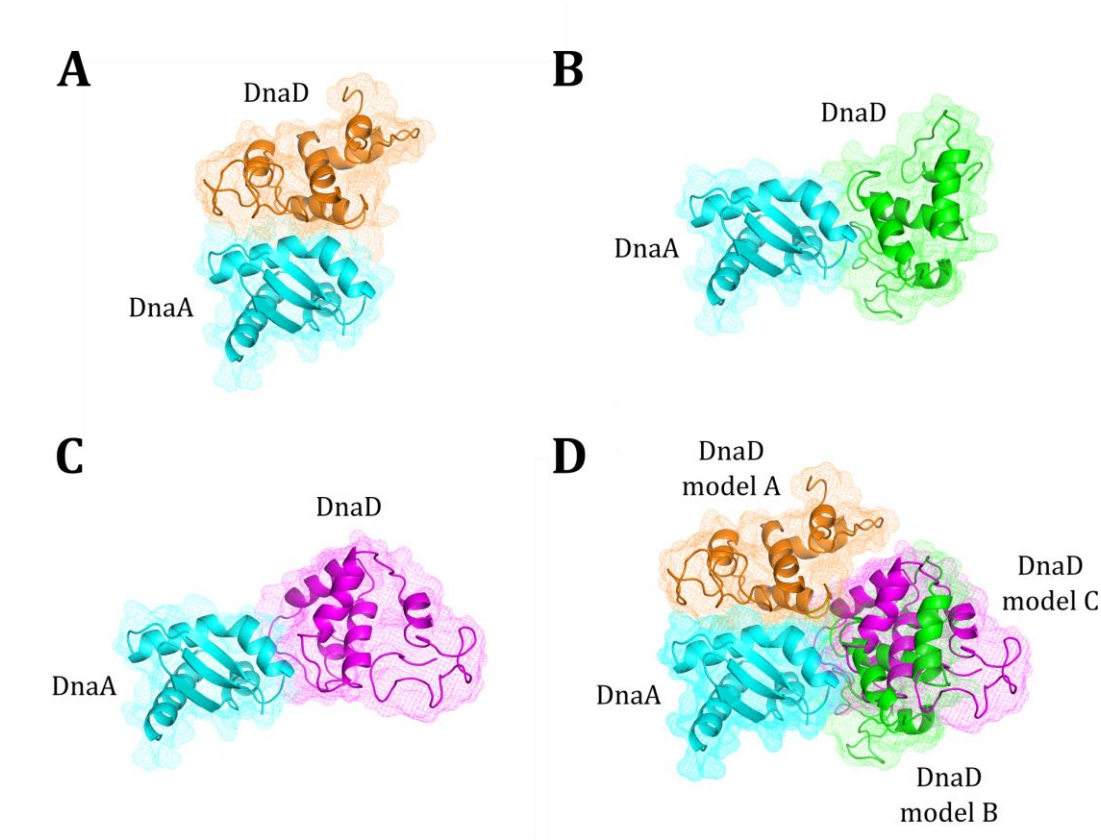
6.1 Template-based Modelling

Rapid advances in template-based modelling of protein-protein complexes follows the trend in structure prediction of individual proteins. Such methods can provide beneficial structural information on the interface location on one or both proteins and the nature of the interactions stabilising the interface. Physical characterisation of a protein complex, even with low accuracy, is valuable compared to its absence. Previous studies[154] have shown that docking of low resolution protein models can still yield structurally meaningful results to predict binding interfaces and these can then serve as a starting point for further structural analysis. The low-resolution modelling of protein-protein interactions is also central to efforts of modelling large protein-protein interaction networks[155].

The structure of a protein-protein complex can be constructed using known complex structures as a template. The templates are detected either by homology-based sequence alignment or structure-based comparison. An advantage of this method is that the structure of the monomer components are not pre-required for this method (models are in principle constructed from amino acid sequences). Also, that complex templates are used from the bound form (compared to unbound structures used in ridged-body docking) which are not usually sensitive to conformational change induced upon binding. There are four stages to template-based modelling; finding known structures related to the sequences to be modelled, aligning the target sequence to the template sequence, copying aligned regions of template structures to construct structural frameworks, and constructing the unaligned regions (looped regions) and adding side-chain atoms. The quality of the predicted output model is therefore dependent on the accuracy of the template identifications. Various resources are available for homology-based prediction of protein-protein complexes including; SWISS-MODEL[156], MODELLER[157], and 3D-JIGSAW[158] amongst others.

The InterEvDock2 online resource[159] was used to predict the structure of the DnaA-DnaD interaction. InterEvDock predicts the structure of protein-protein interactions using evolutionary information, hence, is well suited where homologous sequences are available for both binding partners. InterEvDock requires the input of PDB files for each binding partner (either uploaded to the server or retrieved via the PDB accession code) from which the sequence is automatically extracted. Exhaustive rigid-body docking between the two input structures is performed using FRODOCK (fast rotational docking)[160] which combines a search algorithm based on spherical harmonics and an energy-based scoring function (including van der Waals, electrostatics and desolvation terms). The best 10,000 FRODOCK clusters are then re-scored using InterEvScore[161] and SOAP-PP potentials[162]. InterEvScore is a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to co-evolutionary information (evolutionary information is scored only for residues belonging to apolar patches). SOAP-PP is a statistically optimised atomic potential which captures orientation-dependent interactions (such as hydrogen bonds). The top 1000 models for each of the three scoring functions are clustered using the FCC (fraction of common contacts) method[163] which groups structural models according to common contacts at the interface. Clusters are then ranked according to the mean score of the top 30% models and the top 10 predictions from each scoring method are provided by the server.

The top 3 models based on FRODOCK score[160], SOAP-PP score[162], and InterEvScore[161] are displayed in Figure 6.1. The models identified by physics based potential and statistically optimised atomic potential show a degree of overlap that is not observed with the model based on co-evolutionary information which is distinct. The top models predict a purely α-helical interface from DnaA domain I and a majority unstructured, with some α-helical, contribution by the DnaD DDBH2 domain.



Model	FRODOCK	SOAP	InterEv
A	2162.21	-7719.90	149.02
B	987.39	-8172.7	171.48
C	1341.06	-7104.97	1162.27

Figure 6.1 *Predicted models of the DnaA domain I - DnaD DDBH2 domain interaction using the InterEvDock2 online resource. DnaA domain I and DnaD DDBH2 domain are represented as mesh and cartoon where DnaA domain I is displayed in cyan. The top 3 models are displayed based on the FRODOCK score, SOAP score and InterEv score (results tabulated).* ***A.*** *Top model based on FRODOCK score, DnaD DDBH2 domain displayed in orange.* ***B.*** *Top model based on SOAP score, DnaD DDBH2 domain displayed in green.* ***C.*** *Top model based on InterEv score, DnaD DDBH2 domain displayed in pink.* ***D.*** *Overlay of top 3 models from InterEvDock.*

InterEvDock also provides the top 5 residues on each chain predicted to be involved in contacts based on the consensus top 10 models from each method. These are displayed in table 6.1. The ranked residues from DnaA domain I correspond closely with residues identified by CSP analysis of the DnaA-DnaD interaction (Chapter 5), this

is likely due to the overlap with the DnaA-SirA binding interface for which the complex structure is deposited in the PDB[107]. The ranked residues within the DnaD DDBH2 domain however are not consistent with the CSP experimental data. Of those predicted R179, R183 and K212 are within the DNA binding interface[90] while the remaining residues lie within the unstructured C-terminal region. The results from the template-based modelling highlights the advantages and disadvantages of this technique, while it proved successful in predicting the binding interface within DnaA domain I, it somewhat missed the mark with the DnaD DDBH2 domain, instead highlighting a region required for an alternative interaction.

Residue Score	DnaA Domain I	DnaD DDBH2 Domain
1	W53	R179
2	F49	F226
3	D52	R183
4	S56	Y227
5	R57	K212

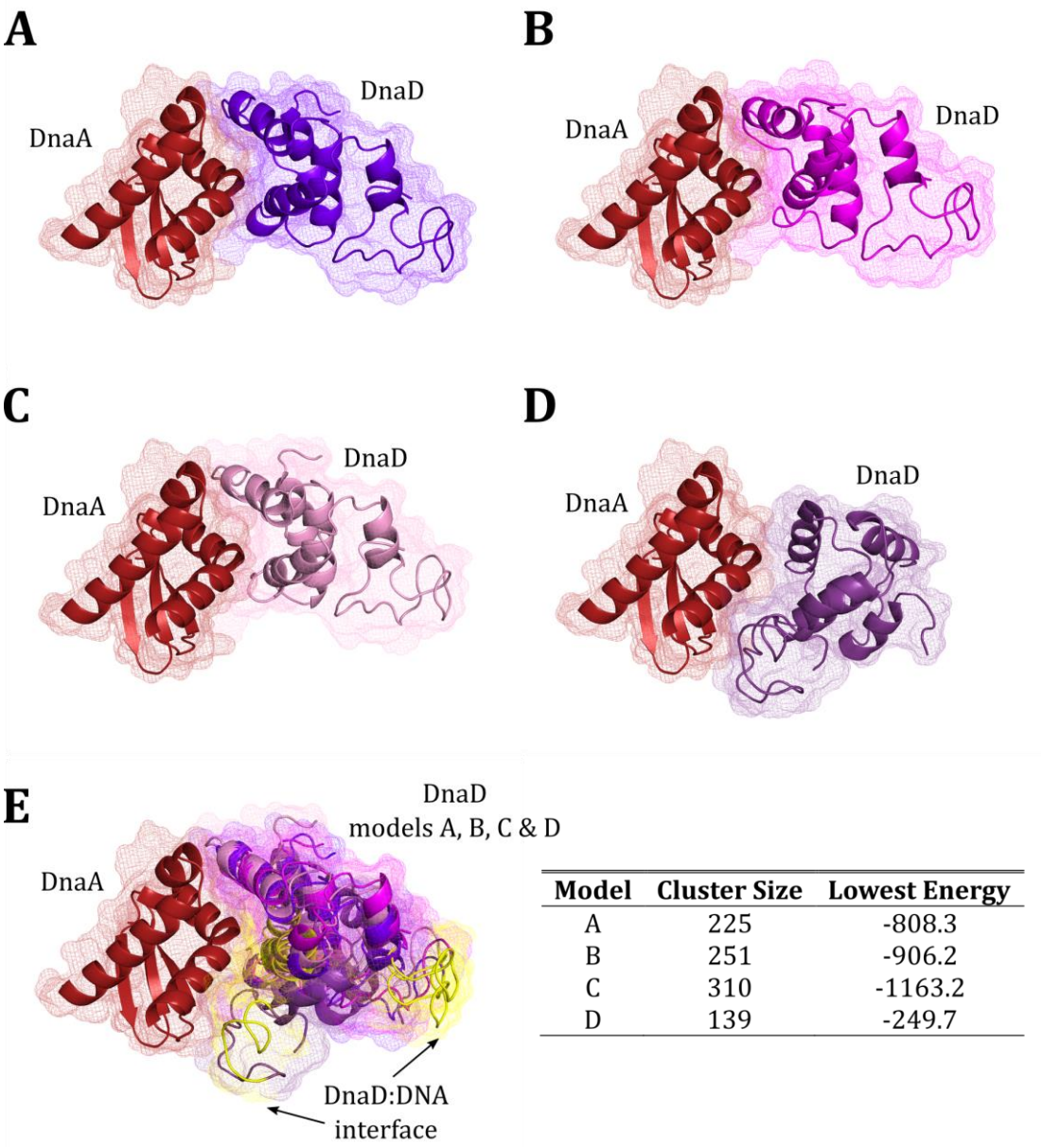
Table 6.1 *InterEvDock top 5 residues on each chain (DnaD domain I and DnaD DDBH2 domain) predicted to be involved in contacts based on the consensus top 10 models from each method (FRODOCK, SOAP and InterEv).*

7.2 Rigid-body Docking

Rigid-body docking using known protein structures is an alternative method to homology-based prediction of protein-protein complexes. The applicability of template-based docking is limited to coverage of template structures, hence direct methods can be more useful to many applications. Direct rigid-body docking is based on thermodynamics and attempts to identify the complex structure with the lowest conformational Gibbs free energy. Numerous resources are available for computational prediction of protein-protein complexes using rigid-body docking such as;

FRODOCK[160], DOCK/PIERR[164], PRISM[165], DOT[166], ZDOCK[167] and GRAMM[168].

The online resource ClusPro[169] was used to predict the structure of the DnaA-DnaD interaction, requiring the PDB file input for each binding partner. ClusPro performs three computational stages; rigid-body docking using the PIPER (Fast Fourier Transform based) docking algorithm [170], root-mean-square deviation (RMSD) based clustering of the 10,000 lowest energy structures, and refinement of structures using energy minimization. The top models obtained from ClusPro are displayed in Figure 6.2, in general the models are ranked by cluster size (clustered with 9 Å Cα RMSD radius) and ClusPro recommends that models are not judged based on the PIPER scores (electrostatics, hydrophobics, Van der Waals). Considerable overlap is observed between the top models, with a consistent interacting face within DnaA domain I that is purely α-helical (as found with the template based modelling). A predominantly α-helical interface is also observed for the DnaD DDBH2 domain, with the exception of the Van der Waals plus electrostatic top model which incorporates some of the unstructured tail of the domain within the interface. However, each of the top models identified by ClusPro produce an interface that overlaps with the DNA binding patch (YxxxIxxxW motif) for DnaD and do not compliment the CSP data.



Model	Cluster Size	Lowest Energy
A	225	-808.3
B	251	-906.2
C	310	-1163.2
D	139	-249.7

Figure 6.2 *Predicted models of the DnaA domain I – DnaD DDBH2 domain interaction using the ClusPro online resource. DnaA domain I and DnaD DDBH2 domain are represented as mesh and cartoon, DnaA domain I is displayed in red. The top 4 models are displayed based on the electrostatic-favoured score, hydrophobic-favoured score, Van de Waals plus electrostatic score and the balanced score ranked by cluster size. Values for the cluster size and lowest energy weighted score[169] for each model are tabulated.* ***A.*** *Top model based on ClusPro balanced score, DnaD DDBH2 domain displayed in purple.* ***B.*** *Top model based on electrostatic score, DnaD DDBH2 domain displayed in magenta.* ***C.*** *Top model based on hydrophobic score, DnaD DDBH2 domain displayed in pink.* ***D.*** *Top model based on Van der Waals plus electrostatic score, DnaD DDBH2 domain displayed in violet.* ***E.*** *Overlay of top 4 models from ClusPro. The DNA binding patch within the DnaD DDBH2 domain (YxxxIxxxW motif and F206-E215) is highlighted in yellow.*

7.3 HADDOCK Restraint Driven Docking

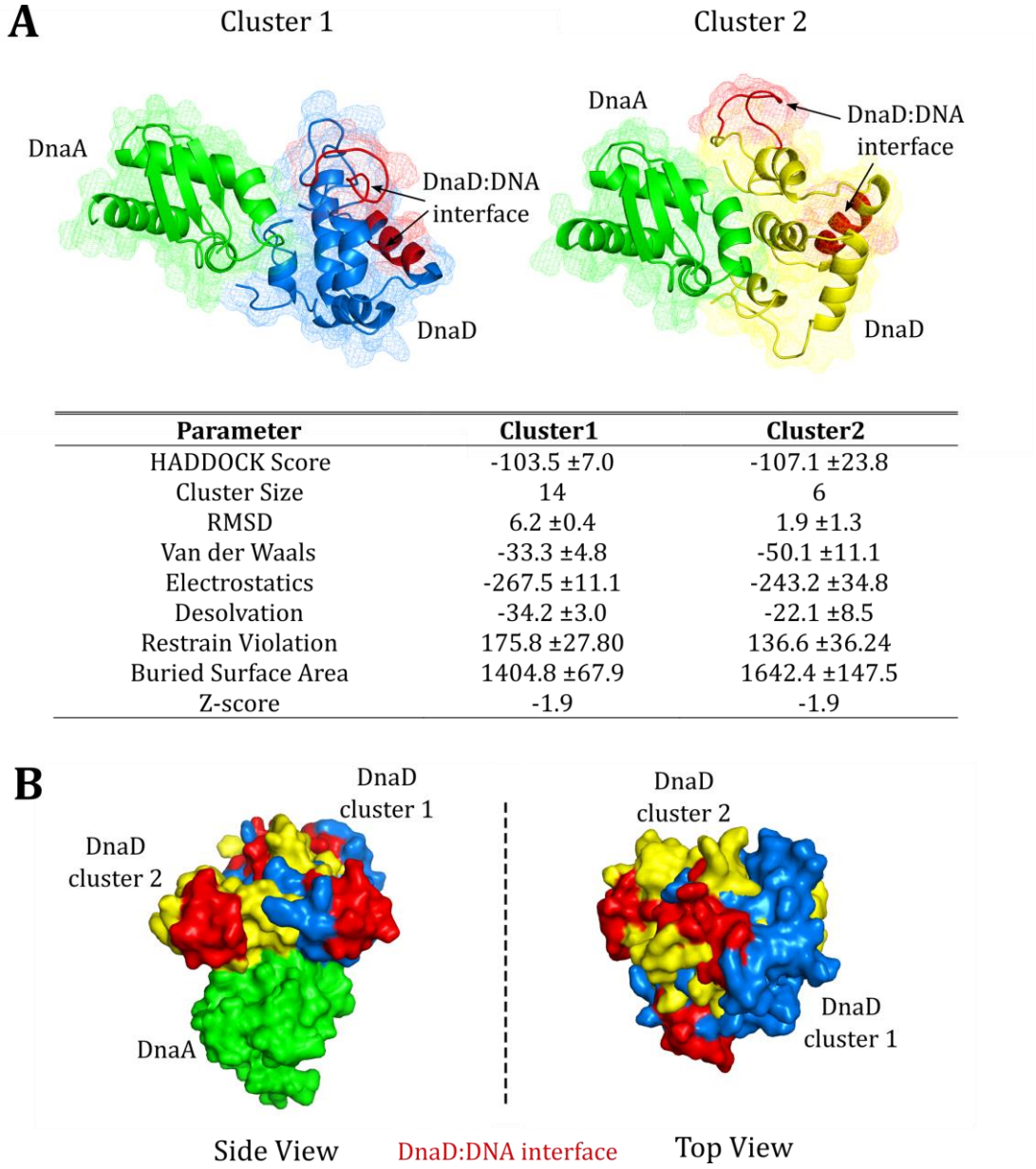
The information generated from biochemical and biophysical experiments on protein-protein interactions can in principle be used to model the structure of the corresponding protein complex. Such methods are valuable when conventional NMR and crystallographic approaches fail. Docking approaches model the structure of a complex based on the structure of the constituents (as described for rigid-body docking, section 6.1.1). Constraint-driven docking utilises interface restraints during rigid-body-docking where platforms require at least one interacting amino acid residue identified on each protein subunit within in a complex. HADDOCK[171] among all the docking methods participating in the Critical Assessment of Prediction of Interactions (CAPRI) challenge, was found to be the only true data-driven strategy[172].

HADDOCK 2.2 accessible via the WeNMR[173] server was used to model the DnaA-DnaD interface using the residue restraints obtained from CSP analysis of NMR data (described in Chapter 5). The HADDOCK docking protocol consists of; rigid-body docking, semi-flexible refinement, and final refinement in explicit solvent. Ambiguous interaction restraints (AIRs) are generated prior to running HADDOCK, these use experimental data such as NMR chemical shift perturbation (CSP) data, to define 'active' residues (experimentally derived and solvent accessible), and 'passive' residues (solvent accessible residues neighbouring active residues). To account for errors in the definition of active and passive residues, HADDOCK allows the random deletion of a fraction of the restraints for each docking run. Subsequently, multiple runs using the same AIR input are generally used to prevent bias. The HADDOCK score given to output models is a weighted sum of intermolecular electrostatics, van der Waals, desolvation, and AIR restraints. A z-score is also given which represents how many standard deviations the HADDOCK score of a given cluster is away from the mean of all clusters.

Multiple HADDOCK runs were undertaken using the AIR restraints displayed in Table 6.2. Initially restraint inputs were varied to confirm output models were not purely driven by energy forces (electrostatics, hydrophobics, Van der Waals). Using identical AIR inputs for 10 HADDOCK runs, 162 structures in 13 clusters were generated. The clusters were then analysed to select the best models to fit with experimental restraints provided by the CSP data. The two best-fit clusters are displayed in Figure 6.3.

AIR Active Residues	
DnaA Domain I	**DnaD DDBH2 Domain**
K17	L129
S20	Y130
S23	I132
T26	E134
E48	E135
F49	K164
D52	V171
W53	
E55	
S56	
H60	

Table 6.2 *Ambiguous interaction restraints used with the HADDOCK program to model the DnaA domain I – DnaD DDBH2 domain interaction. NMR chemical shift perturbation data was analysed to select significant residues (as described in Chapter 5) and solvent exposed residues from this set were designated active residues. Passive residues were determined automatically by the HADDOCK program (solvent exposed neighbours of active residues).*



Parameter	Cluster1	Cluster2
HADDOCK Score	-103.5 ±7.0	-107.1 ±23.8
Cluster Size	14	6
RMSD	6.2 ±0.4	1.9 ±1.3
Van der Waals	-33.3 ±4.8	-50.1 ±11.1
Electrostatics	-267.5 ±11.1	-243.2 ±34.8
Desolvation	-34.2 ±3.0	-22.1 ±8.5
Restrain Violation	175.8 ±27.80	136.6 ±36.24
Buried Surface Area	1404.8 ±67.9	1642.4 ±147.5
Z-score	-1.9	-1.9

Figure 6.3 *Predicted models of the DnaA domain I – DnaD DDBH2 domain interaction using HADDOCK restraint-driven docking. The top clusters are displayed based on HADDOCK scoring and compliance with experimental restraints.* ***A.*** *Top two clusters represented as mesh and cartoon, DnaA domain I is displayed in green and the DnaD DDBH2 domain is displayed in blue (cluster 1) and yellow (cluster 2), the DNA binding patch within the DDBH2 domain (YxxxIxxxW motif and F206-E215) is highlighted in red. The HADDOCK scores for each cluster are displayed in the table below the models.* ***B.*** *Overlay of the top two clusters displayed as surface representations. DnaA domain I is displayed in green and the DnaD DDBH2 domain is displayed in yellow and blue for cluster 1 and 2 respectively. The DNA binding patch within each DDBH2 domain is highlighted in red.*

Cluster 1		Cluster 2	
DnaA Domain I	**DnaD DDBH2 Domain**	**DnaA Domain I**	**DnaD DDBH2 Domain**
S20	**Y130**	**S20**	**L129**
K21	T131	P22	T131
P22	**E134**	**S23**	**I132**
S23	**E135**	E25	**E134**
E25	E136	**T26**	**E135**
T26	A138	W27	E136
W27	Q161	S30	A138
N47	**K164**	P46	D159
E48	H165	N47	Q161
F49	K168	**F49**	**K164**
R51	V170	A50	H165
D52	**L171**	R51	K168
W53	V196	**D52**	E197
S56	Q204	**W53**	K200
R57	Q217	**E55**	Q204
Y58	R222	**S56**	Q217
	W229	R57	W229
	L230	Y58	L230
	E231	L59	E231
	Q232	**H60**	Q232
		L61	

Table 6.3 *Residues of contact within the top two clusters from the HADDOCK predicted models of the DnaA domain I – DnaD DDBH2 domain interface. The residue interface was identified using a PyMOL script[174, 175] that applies a 1 Å² cutoff above the difference between the complex-based surface area and chain-only surface area to select for interface residues. AIR input residues (Table 7.2) are displayed in bold.*

Both clusters provide a DnaA binding interface distinct to the DNA binding patch within the DnaD DDBH2 domain. Of the 11 AIR restraints input for DnaA domain I, 8 and 9 were identified as interface residues (Table 6.3) within clusters 1 and 2 respectively, and of the 7 AIR restraints input for the DnaD DDBH2 domain, 5 were identified as interface residues for each cluster. The interaction surfaces of the individual clusters show overlap and the difference between the models can be accounted for by an approximated movement, for the DnaD DDBH2 domain, of 30 Å distance along an axis of rotation. The DnaA-DnaD interface is predominately α-helical, using polar side chain packing in a hydrophobic interface, as displayed in Figure 6.4. Current scoring functions are generally not accurate enough for reliable model

discrimination and in addition the top clusters produced similar parameter scores. Further experimental information (such as site-directed mutagenesis) would be required to confirm an exact model from those obtained through the HADDOCK restraint-driven rigid-body docking.

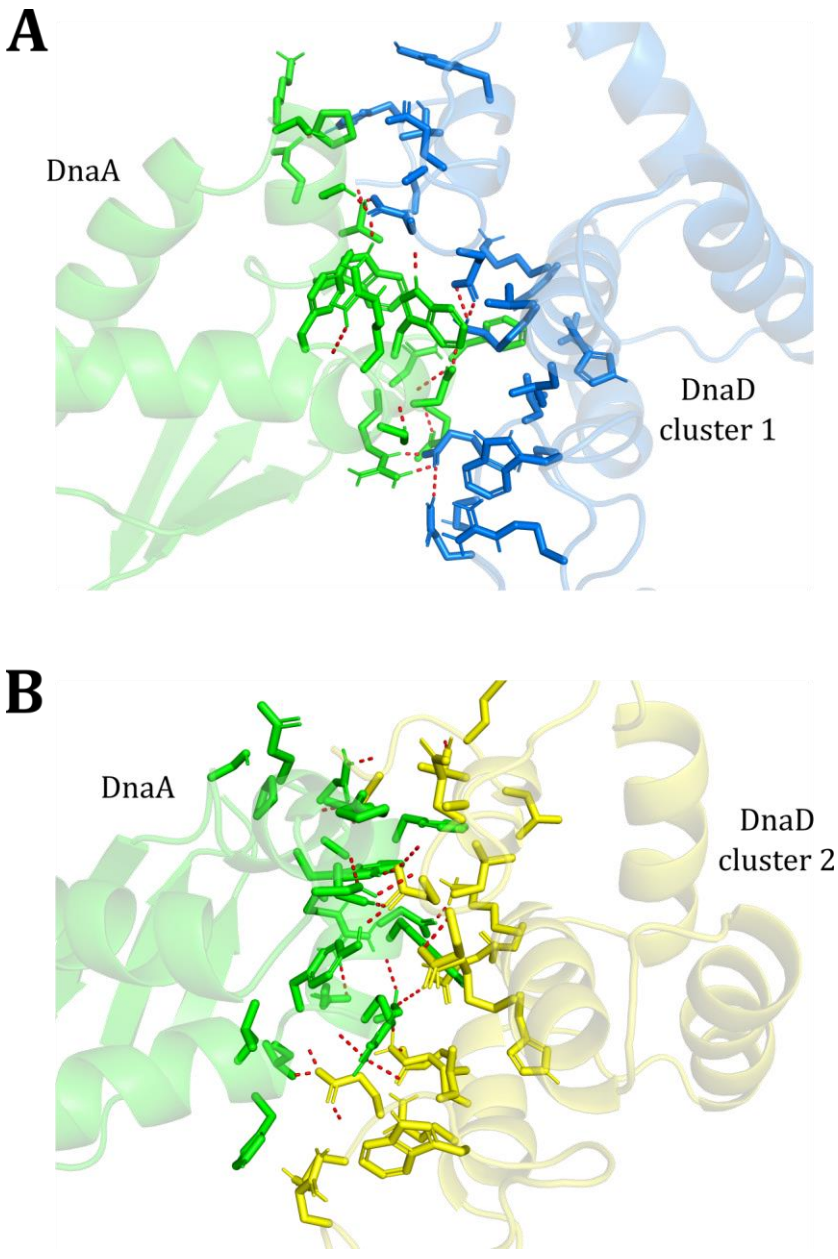


Figure 6.4 *Contact residues within* ***A****. clusters 1 and* ***B.*** *cluster 2 from the HADDOCK predicted models of the DnaA domain I – DnaD DDBH2 domain interface. DnaA domain I is displayed in green and the DnaD DDBH2 domain is displayed in blue (cluster 1) and yellow (cluster 2). Interface residue side chains are displayed as sticks and the red dashed lines represent polar contacts. Residues of contact shown are as listed in table 6.3. The interface is purely α-helical from DnaA, in both clusters, with polar side chain packing in a hydrophobic interface.*

6.4 Summary

Predicted modelling (template-based and rigid-body) and restraint driving docking methods have been used to structurally characterise the DnaA-DnaD binding interface. The predicted computational modelling techniques proved relatively accurate for identifying the DnaD binding interface within DnaA domain I. However, restraint driven docking was required to output models with a DnaA binding interface, within the DnaD DDBH2 domain, that complimented our NMR experimental data. A number of potential complex models are output by docking techniques and although compliance with CSP derived residue restraints was used to aid selection of the best models, a single complex structure could not be arrived at. The predicted models obtained from HADDOCK will provide a basis for mutagenesis experiments to confirm involvement of specific residues within the DnaA – DnaD binding interface. Subsequent low-resolution modelling of the entire complex structure (DNA, DnaA filaments and DnaD oligomers) may then be possible to provide further structural information on how the interaction between these essential proteins initiates DNA replication.

Chapter 7

Understanding the Mechanism of DNA Replication Initiation

The work presented in this thesis, entailing the biophysical characterisation of the *B. subtilis* DnaA–DnaD interaction (in solution by NMR and SPR spectroscopy) has provided a greater structural understanding of their mechanism of action during DNA replication initiation. A conformational change within the DnaA nucleoprotein filament was observed on binding of DnaD which is proposed to aid DnaA mediated melting of the *oriC* DUE during DNA replication initiation. The DnaD DDBH2 domain was found to form a direct interaction with domain I of DnaA. Chemical shift perturbation analysis revealed the specific residues involved in the DnaA-DnaD binding interface and computational methods we used to model the structure of the complex. The interface within DnaD was found to be distinct to its DNA binding patch, suggesting that DnaD can bind both DnaA and DNA simultaneously. Our findings are consistent with the original hypothesis that DnaD acts a positive regulator of DNA replication initiation and have providing further structural insights into the essential interaction between *B. subtilis* DnaA and DnaD.

7.1 Competetive Regulation of DNA Replication Initiation

It is of particular interest that the DnaA-DnaD binding interface was found to overlap with the DnaA-SirA interface determined by X-ray crystallography[107]. SirA is a negative regulator of DNA replication, it inhibits initiation of replication in diploid cells committed to sporulation. Sporulation occurs in response to nutrient starvation to produce an endospore that lies dormant until nutrient conditions improve[176]. The process of sporulation involves the asymmetric division of the cell into a larger mother cell and a smaller forespore compartment where each require one copy of the genetic information. Hence, sporulation typically follows the completion of DNA replication and is the result of a regulatory feedback mechanism involving Spo0A[177]. SirA is produced under phosphorylated-Spo0A regulation and has been identified as an inhibitor of replication, that acts by binding to DnaA, with a specific role in preventing re-initiation in cells committed to sporulation[152]. SirA overexpression mimics DnaA depletion from the origin suggesting that the mechanism of action inhibits the co-operative binding of DnaA to *oriC*[178].

It was speculated by Jameson *et. al.*[107] that SirA may inhibit the DnaA-DnaD interaction arresting assembly of the initiation complex. We tested this hypothesis by aligning the modelled HADDOCK structures (as described in Chapter 6) with the DnaA-SirA complex X-ray crystal structure, displayed in Figure 7.1. DnaD and SirA interact with domain I of DnaA and their interaction patch is highly overlapped with both requiring residues Ser20, Thr26, Glu48, Phe49, Asp52, Trp53 and Ser56 within the α2 and α3 helices of DnaA. The high degree of overlap suggests that both proteins use the same interaction patch on DnaA, thus, SirA binding could sterically hinder the DnaA-DnaD interaction to prevent re-initiation of DNA replication in *B. subtilis* cells committed to sporulation. This further suggests that DnaD plays a positive role during replication initiation, as opposed to the negative regulatory role suggest by Bonilla and

Grossman[65], and that SirA and DnaD achieve opposing regulatory functions via interaction with the same structural site on DnaA.

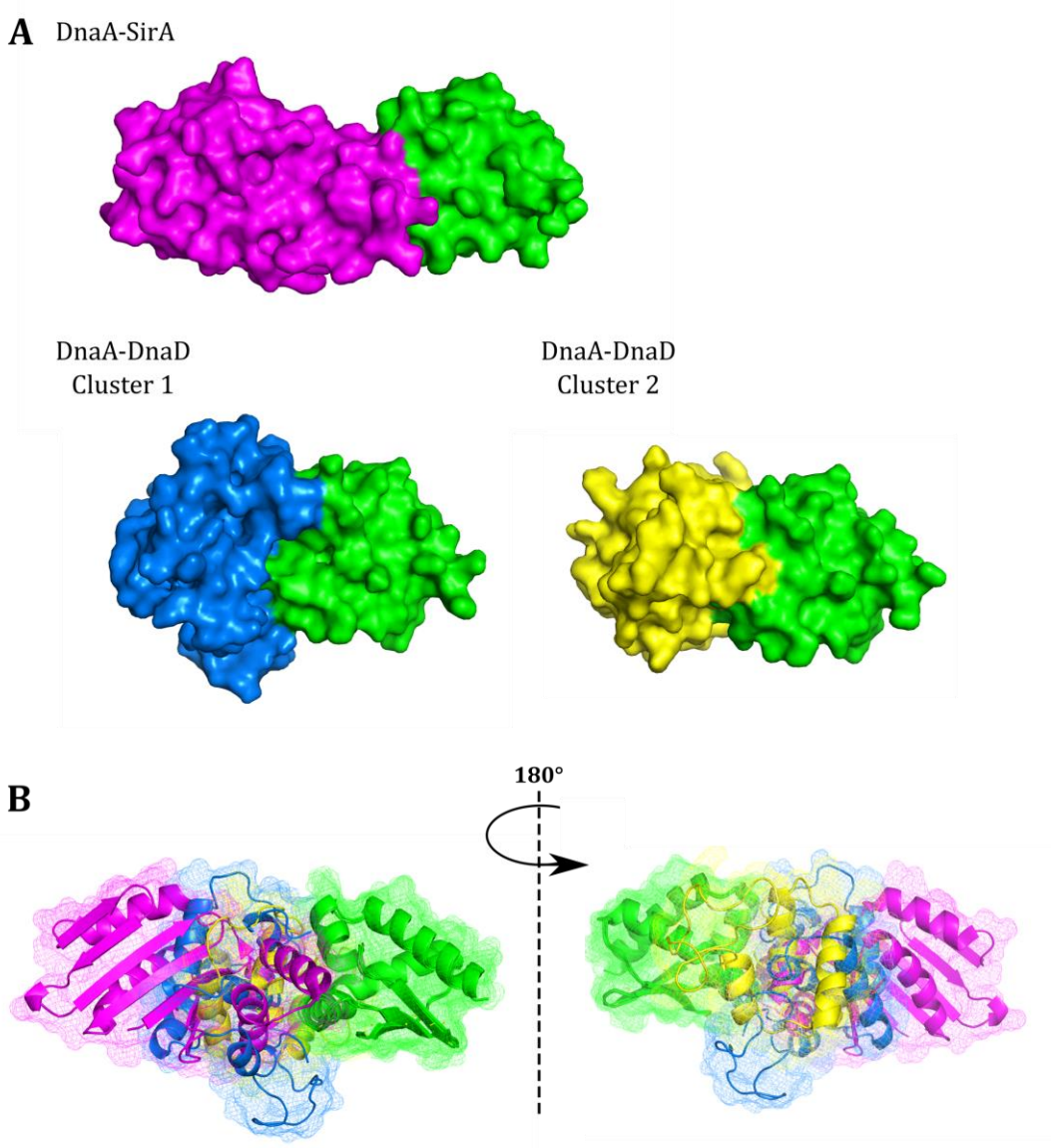


Figure 7.1 *The DnaA-DnaD interface overlaps with the DnaA-SirA binding interface.* ***A.*** *The structure of SirA – DnaA domain I complex (left), and predicted models of the DnaA domain I – DnaD DDBH2 domain interaction Cluster 1 (centre) and Cluster 2 (right) represented as surface models. DnaA domain I is displayed in green and kept ion the same orientation, SirA displayed in purple and the DnaD DDBH2 domain is displayed in blue and yellow for cluster 1 and cluster 2 respectively.* ***B.*** *Overlay of the DnaA-SirA complex with the two predicted DnaA-DnaD complex models displayed as cartoon and mesh representation. Individual domains are coloured as described for A. The structure of the DnaA-SirA binding interface was determined by X-ray crystallography[107] (PDB – 4TPS).*

7.2 Unwinding of the *oriC* DNA

Local melting of the *oriC* DNA is thought to be directly mediated by the DnaA nucleoprotein oligomer. Positive writhe is introduced into the dsDNA wrapped around the outside of the DnaA filament (through domain IV interaction) which is compensated by introducing negative supercoiling just ahead of the DnaA-DNA filament. The DNA double helix is destabilised facilitating melting at the AT rich DNA unwinding element (DUE) within the *oriC*. The ATPase domain of DnaA can also bind unwound ssDNA inside the DNA filament to stabilise strand separation.

The mechanism behind the additional requirement of initiator protein DnaD in low GC content *Firmicutes* is not yet fully understood. We propose that DnaD functions, as a positive regulator of DNA replication initiation, similarly to *H. pylori* HobA acting as a structural scaffold to aid DnaA binding to DnaA boxes within the *oriC*. HobA is an activator of DNA replication initiation, it forms tetramers to promote DnaA oligomerisation (in the same manner as the *E. coli* DiaA)[179]. The structure of the HobA-DnaA domain I complex compliments that of SirA-DnaA domain I, and therefore also compliments the DnaD-DnaA complex structure. DnaD is active as a tetramer, like HobA and DiaA, and can form high order oligomers. The architectural assistance of DnaD may be crucial to facilitate remodelling of *oriC* by DnaA, as displayed in Figure 7.2.

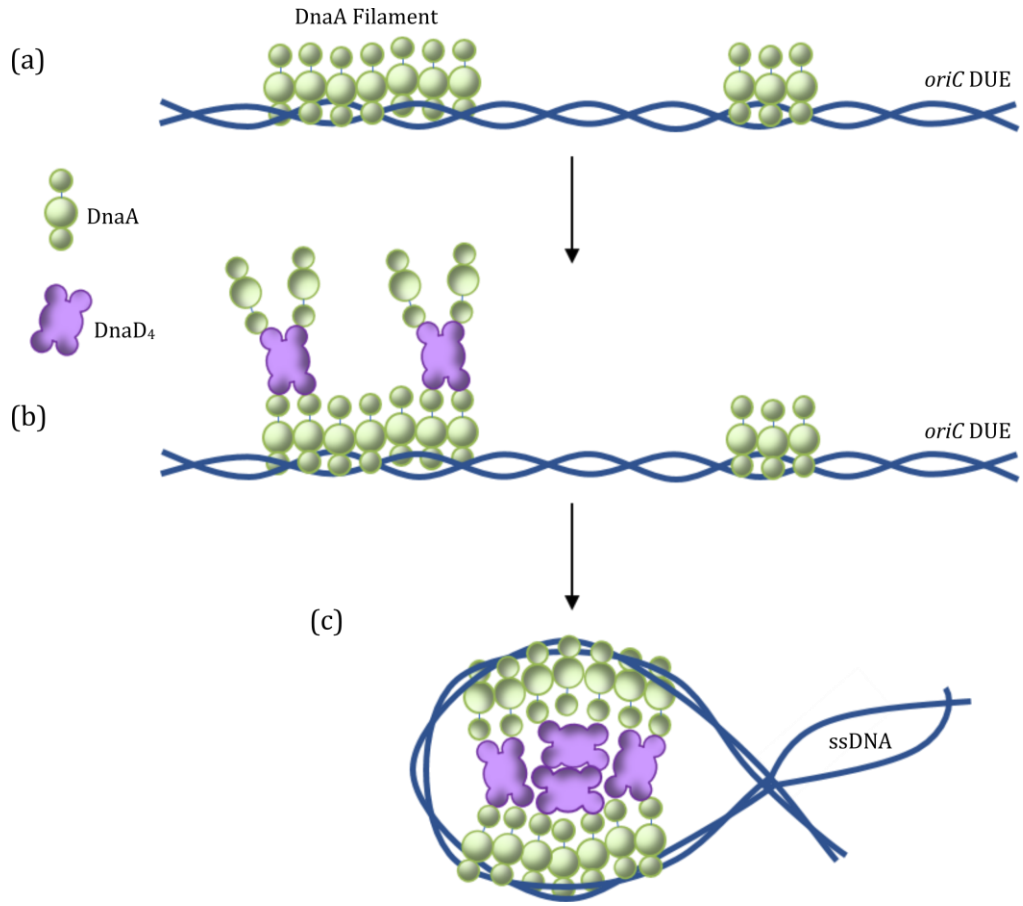


Figure 7.2 *Speculative model of the DnaA-DnaD complex and its assembly onto the B. subtilis oriC during DNA replication initiation. (a) Schematic showing DnaA binding at the DnaA box clusters of the split origin (two DnaA box cluster regions are separated by the dnaA gene) through domain IV interaction with dsDNA. Oligomerisation occurs via domain III interaction, in an ATP dependent manner, to form the DnaA nucleoprotein filament. (b) The DnaD tetramer interacts with domain I of DnaA via its exposed DDBH2 domains and induces a small conformational change within the helical structure. DnaA-bound DnaD may recruit additional DnaD to form a higher order DnaD oligomer. The proximity of the adjacent DnaA filament could extend the DnaD scaffold into this region (via DnaA-DnaA and DnaA-DnaD interaction) causing a bend in the DNA. (c) DnaD can act as a supporting scaffold to stabilise the looped DNA structure. DnaA mediated DUE melting may be assisted by DnaD-induced conversion of negative twist into writhe.*

In addition to a structural role, DnaD has global DNA remodelling activity and it has been shown to convert writhe into twist and circularise linear DNA[87, 99]. Zhang *et. al.* proposed that DnaD opens up supercoiled chromosomal DNA in a concentration-dependent manner to expose *oriC* to DnaA[102]. DnaA could then bind the DnaA-boxes to cause local unwinding of the duplex DNA. Our identification of the DnaA-DnaD binding interface suggests that DnaD may be able to interact with both DnaA and DNA simultaneously (though it is likely that within a DnaD tetramer separate DDBH2 domains are employed for these two interactions to avoid steric hinderance). DnaD may mediate DUE melting directly through its DNA remodelling activity and indirectly through assisting DnaA mediated duplex unwinding. It's possible that the conformational change observed within the DnaA filament on binding of DnaD (as described in Chapter 3) adds to the positive writhe created by the DnaA filament and allows the conformation to adopted whereby ssDNA can bind inside the filament. Further investigation of the DnaA-DNA-DnaD interactions is required to fully understand the mechanism of strand separation at the *oriC* during DNA replication initiation.

7.3 DNA Replication Initiation as a Target for Antimicrobials

DNA replication is an underexploited antimicrobial target. It is essential for cell viability, hence, inhibiting any protein required for this process would be detrimental to cell survival. Current therapies targeting DNA replication are limited to topoisomerase II inhibitors targeting DNA gyrase and topoisomerase IV which act to modify the topology of DNA during replication. Topoisomerase II inhibitors can be divided into two groups; aminocoumarins (ATPase inhibitors), and fluoroquinolones (poisons that interfere with the catalytic DNA cleavage and joining process).

Aminocoumarins are limited by poor pharmacological properties and there is major concern over the rise in fluoroquinolone resistant pathogens. With the challenge of antimicrobial resistance, exploration of new targets is essential for development of novel antimicrobials to avoid existing resistance mechanisms. Most components of the bacterial replisome are significantly different to their eukaryotic counterpart and therefore can be exploited to develop antimicrobial compounds without causing cytotoxicity.

The highly conserved replication initiator protein DnaA contains an 'interaction hub' within domain I responsible for a number of interactions (HobA, DiaA, SirA, DnaD) and is subsequently an attractive target. DnaA is strictly regulated so that DNA replication is co-ordinated with cell growth and differentiation. Disrupting this regulation could be a potential mechanism for promoting cell death. Another advantage of targeting the bacterial replicative machinery is that certain components are well conserved between bacteria, such as DnaA, while others are restricted to particular species. This allows the development of both broad and narrow range antimicrobials. A complete understanding of the mechanisms involved in DNA replication is required before effective therapies can be produced. The research presented here has provided structural information on the essential interaction between DnaA and DnaD within *B. subtilis* and related gram-positive organism, providing a basis for developing antimicrobial therapies targeting these proteins.

7.4 Final Remarks

The objectives of this thesis were to observe the interaction between *B. subtilis* DnaA and DnaD and characterise the interaction surfaces. The interaction between the single domains, DnaA domain I and the DnaD DDBH2 domain, was detected by NMR which was subsequently used to identify the specific residues involved within each protein and characterise the nature of these interfaces. The experimental data were first used to predict a restraint-free model of the DnaA domain I – DnaD DDBH2 domain complex and then using restraints from NMR CSP analysis to drive docking, which provides a basis for additional experimental techniques to confirm its accuracy. The kinetics of the interaction were investigated by SPR to obtain a binding affinity of approximately 500 μM which complimented the values calculated using NMR titration data and are the first recorded for this interaction. Single molecule FRET revealed that DnaD does not inhibit DnaA nucleoprotein filament assembly or promote their disassembly, as proposed by Scholefield et. al. 2013[66]. These data suggest a conformational change within the DnaA filament in the presence of DnaD which aligns with our proposal that DnaA and DnaD work cooperatively to unwind duplex DNA. This structural characterisation of the interaction will aid the understanding of the roles of DnaA and DnaD during DNA replication initiation and will facilitate further targeted investigations of the interaction with a particular interest to designing antimicrobial therapies in the future.

Chapter 8

Appendix

8.1 *B. subtilis oriC* sequence

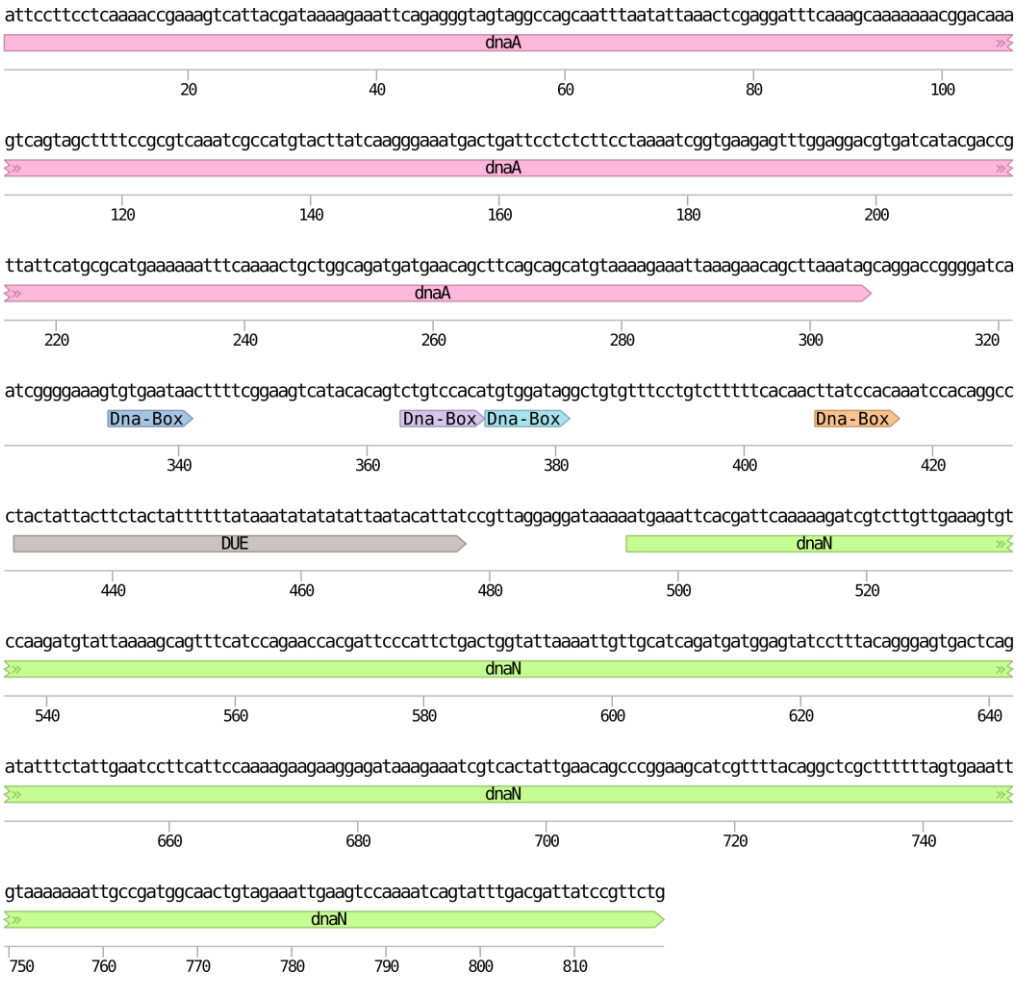


Figure 8.1 *Sequence of the B. subtilis origin fragment containing the DnaA-box cluster downstream of the dnaA gene used for DnaA filament formation reactions (Chapter 3). The dnaA gene, DnaA-boxes, DUE and dnaN gene are highlighted.*

8.2 NMR Spectra

Included in the following section are the full ^{15}N HSQC spectra of the assignment of *B. subtilis* DnaA domain I (Chapter 4). Alongside the spectrum of the DnaD DDBH2 domain and the spectral overlays used to characterise the DnaA-DnaD binding interface (Chapter 5). The ^{1}H, ^{15}N and ^{13}C chemical shifts for the amide backbone assignment of DnaA domain I are listed in Tables 8.1 and 8.2. The ^{1}H and ^{15}N chemical shift for the amide backbone assignment of DnaD DDBH2 domain are listed in Table 8.3.

- 8.2 – DnaA domain I
- 8.3 – DnaD DDBH2 domain
- 8.4 – DnaA domain I titration against DnaD DDBH2
- 8.5 – DnaA domain I titration against DnaD DDBH2 in the presence of ssDNA
- 8.6 – DnaA domain I titration against truncated DnaD DDBH2 domain
- 8.7 – DnaD DDBH2 domain titration against DnaA domain I
- 8.8 – DnaD DDBH2 domain titration against DnaA domain I in the presence of ssDNA

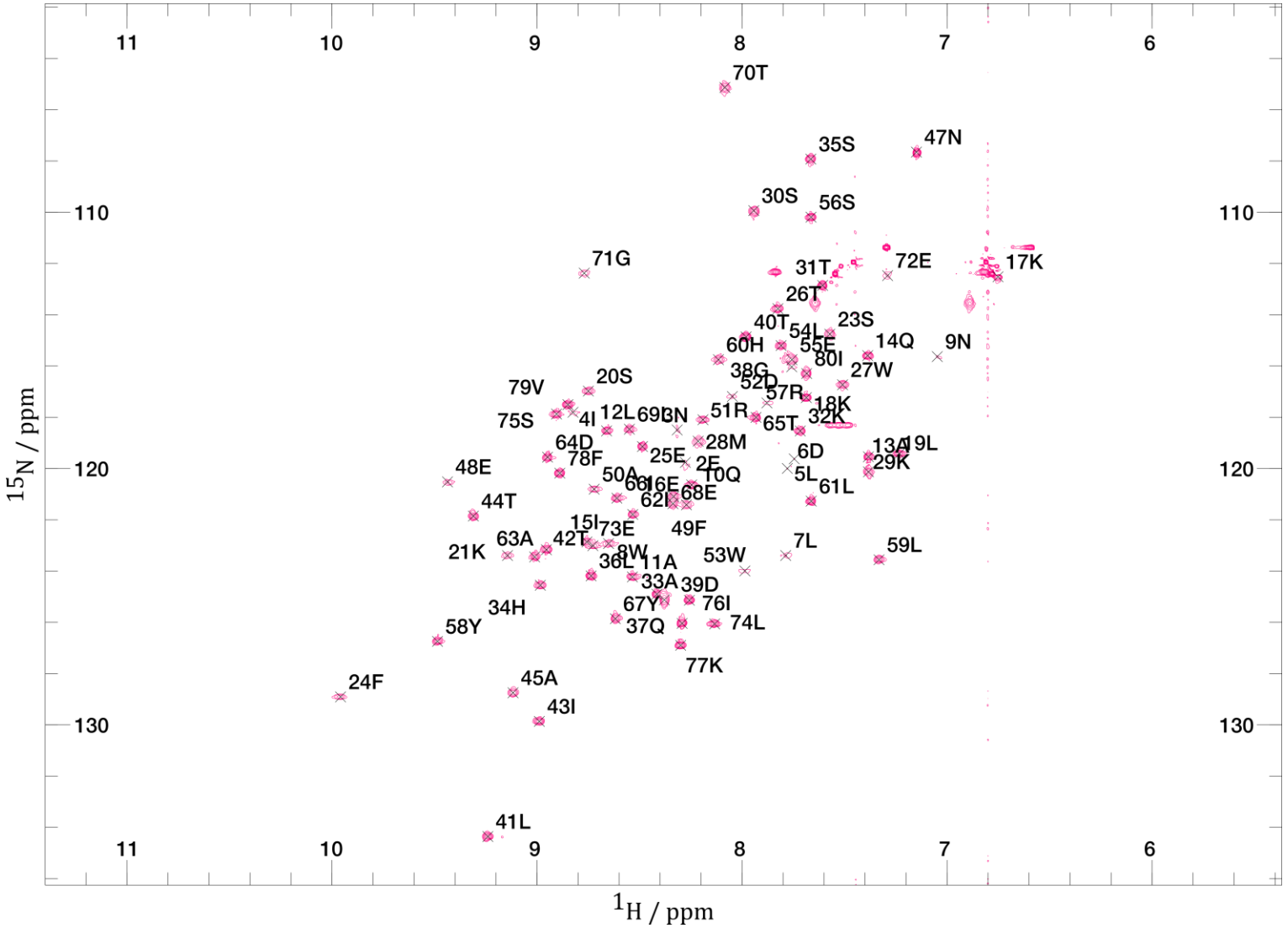


Figure 8.2 *The ^{15}N HSQC spectrum of DnaA domain I.*

Position	Residue	^{1}H/ppm	^{15}N/ppm	Position	Residue	^{1}H/ppm	^{15}N/ppm	Position	Residue	^{1}H/ppm	^{15}N/ppm
2	Glu	8.27635	119.762	29	Lys	7.38332	120.128	56	Ser	7.66342	110.193
3	Asn	8.31644	118.495	30	Ser	7.94273	109.937	57	Arg	7.87877	117.439
4	Ile	8.82464	117.806	31	Thr	7.60675	112.855	58	Tyr	9.48555	126.744
5	Leu	7.77913	119.996	32	Lys	7.71593	118.533	59	Leu	7.33177	123.559
6	Asp	7.74535	119.623	33	Ala	8.41544	124.902	60	His	8.1151	115.752
7	Leu	7.78706	123.392	34	His	8.98387	124.554	61	Leu	7.66262	121.272
8	Trp	8.65123	122.927	35	Ser	7.66444	107.927	62	Ile	8.53303	121.792
9	Asn	7.04753	115.632	36	Leu	8.73621	124.188	63	Ala	9.01167	123.449
10	Gln	8.24587	120.639	37	Gln	8.29228	126.042	64	Asp	8.94856	119.557
11	Ala	8.53478	124.228	38	Gly	7.75757	116.03	65	Thr	7.93182	118.004
12	Leu	8.6606	118.532	39	Asp	8.37905	125.125	66	Ile	8.61053	121.156
13	Ala	7.38361	119.542	40	Thr	7.98033	114.865	67	Tyr	8.61793	125.859
14	Gln	7.38702	115.594	41	Leu	9.23992	134.357	68	Glu	8.33698	121.371
15	Ile	8.75647	122.863	42	Thr	8.95423	123.168	69	Leu	8.54945	118.467
16	Glu	8.33372	121.093	43	Ile	8.98882	129.865	70	Thr	8.08363	105.134
17	Lys	6.75139	112.517	44	Thr	9.31033	121.854	71	Gly	8.76972	112.377
18	Lys	7.68636	117.233	45	Ala	9.11705	128.748	72	Glu	7.28995	112.473
19	Leu	7.23269	119.421	47	Asn	7.1493	107.655	73	Glu	8.72761	123.008
20	Ser	8.74964	116.977	48	Glu	9.43664	120.519	74	Leu	8.13344	126.063
21	Lys	9.14403	123.393	49	Phe	8.27117	121.416	75	Ser	8.90488	117.873
23	Ser	7.57202	114.757	50	Ala	8.72132	120.806	76	Ile	8.25686	125.141
24	Phe	9.95849	128.908	51	Arg	8.1906	118.096	77	Lys	8.29942	126.897
25	Glu	8.48574	119.142	52	Asp	8.04719	117.191	78	Phe	8.88864	120.181
26	Thr	7.82494	113.763	53	Trp	7.98616	123.994	79	Val	8.84919	117.507
27	Trp	7.50955	116.73	54	Leu	7.80942	115.201	80	Ile	7.6846	116.294
28	Met	8.21391	118.933	55	Glu	7.75656	115.752				

Table 8.1 *^{1}H and ^{15}N chemical shifts for the DnaA domain I amide backbone assignment.*

Position	Residue	Cα [i] ^{13}C/ppm	Cβ [i] ^{13}C/ppm	CO [i] ^{13}C/ppm	Cα [i-1] ^{13}C/ppm	Cβ [i-1] ^{13}C/ppm	CO [i-1] ^{13}C/ppm	Position	Residue	Cα [i] ^{13}C/ppm	Cβ [i] ^{13}C/ppm	CO [i] ^{13}C/ppm	Cα [i-1] ^{13}C/ppm	Cβ [i-1] ^{13}C/ppm	CO [i-1] ^{13}C/ppm
2	Glu							42	Thr	59.00623	65.80369	170.67176	50.83218	41.36162	170.8429
3	Asn							43	Ile	55.49474	36.96275	171.77595	58.95771	65.72578	170.66484
4	Ile							44	Thr	59.06465	66.77748	171.26885	55.44952	37.11935	171.82246
5	Leu				54.21364			45	Ala	45.47912	17.50562	173.25379	59.12585	66.54761	171.31889
6	Asp	54.35454	33.99764		55.39943	38.41024	176.76876	47	Asn	49.51949	36.97485	171.83478	62.08948	29.64204	174.01122
7	Leu	55.54323			61.38221			48	Glu	57.63284			49.57491	37.03695	171.86953
8	Trp	59.93362	25.86369	172.8519			173.20826	49	Phe	58.10375		175.84999	57.58611	27.04451	174.02855
9	Asn							50	Ala	52.10388	15.77309	175.74355	58.27108	36.07868	175.98239
10	Gln	56.4256	25.8035	176.1379	53.68801	34.97193	176.8193	51	Arg	57.77366	26.87214	174.09341	52.11838	15.89409	175.75894
11	Ala	52.51945	14.9322	176.92977	56.43044	26.06653	176.19367	52	Asp	54.25107	37.52621		57.80689		174.15647
12	Leu	55.27966	39.07714	175.83755	52.53655	15.15471	176.97487	53	Trp	57.98889		175.07567	54.27039	37.75519	176.10304
13	Ala	52.04585	15.00389	177.7615	55.38007	39.1665	175.93624	54	Leu	54.20577		175.47326	54.08799	28.19884	175.54527
14	Gln	55.04335	25.74194	176.69006	51.97825	15.20396	177.8015	55	Glu		41.50972		56.79995	26.67899	
15	Ile	63.16017	35.65962	174.89375	55.05713	25.99929	176.75666	56	Ser				56.90654	27.56773	
16	Glu	56.65141	26.27966	174.73308	63.09105	35.46787	174.9413	57	Arg	54.04949			57.51647	60.99891	
17	Lys	54.56408	29.73506	174.90589	56.42575	26.35325	174.74338	58	Tyr						181.82883
18	Lys	53.96341	31.94171	173.08893	54.47261	29.97348	174.97407	59	Leu	57.70291	39.96496	175.25125	37.52184		172.84113
19	Leu	50.71153	43.56287	174.19566	53.912	32.16839	173.14342	60	His	57.18583	26.21492	174.1899	57.69181	40.09149	175.30116
20	Ser	55.35918	61.37984	171.87315	50.73731	43.51401	174.25869	61	Leu	55.3197	39.11404	177.99827	56.87356	26.35765	174.27929
21	Lys	55.35313			55.32898	61.51587	171.9216	62	Ile	63.81559	35.65229	174.45666	55.3202	39.24646	178.04781
23	Ser	59.54429			62.96017		177.52259	63	Ala	53.79127	15.33392	178.37259	63.86103	35.59181	174.51527
24	Phe						181.01941	64	Asp	54.75463	37.81812	176.08486	53.80679	15.55222	178.40836
25	Glu	56.01297	27.59768	174.38649	55.67482	38.97921	175.32101	65	Thr	64.6396	66.01002	173.48927	54.87012	37.89152	176.13825
26	Thr	63.36459	66.24896	172.55631	56.65004	27.1381	175.15108	66	Ile	59.8806	32.85607	176.78828	64.73846	65.95142	173.54425
27	Trp	54.71074		173.75689	63.33057	65.98634	172.59861	67	Tyr	58.38405	35.25648	176.0372	59.87887	32.79555	176.85274
28	Met	51.73264			43.0086		170.2734	68	Glu	58.32001	35.36052	175.8347	58.35703		
29	Lys	57.04432		173.49079	51.90159	30.17588		69	Leu	56.07773	38.6959	176.26812	58.00178		175.05969
30	Ser	55.16673	60.43883	172.46052	56.90605	30.35083	173.56695	70	Thr	59.76329	70.15705				
31	Thr	59.25907	68.95911	171.49865	60.31318	55.20674	172.50053	71	Gly	42.98112			59.69374	70.06683	173.5705
32	Lys	51.55462	33.67499	171.91574	59.29428	68.83618	171.53505	72	Glu			175.73564			175.87137
33	Ala	49.9478	15.48646	172.99041	51.54232	33.78305	171.96142	73	Glu	52.52676			51.58387	29.5166	172.04461
34	His	50.01483	29.84551	172.83116	49.94717	15.80754	173.02053	74	Leu	50.91474	40.24	175.04386	52.72769	26.50307	172.87986
35	Ser	55.13367	61.92613	168.80838	55.8026	29.95979	172.84747	75	Ser	55.35876	60.95851	169.96231	50.95704	40.34198	175.05143
36	Leu	52.16069	41.97934	172.01189	55.11275	61.87414	168.86141	76	Ile	55.45084			55.41201	61.00361	173.47313
37	Gln	51.68603	28.02254	173.23584	52.03115	42.16608	172.04407	77	Lys	51.02201	33.31675	171.16907	55.42283	34.61343	172.46985
38	Gly		26.32124					78	Phe	53.79133	39.95267	173.58999	51.03225	33.44625	171.22163
39	Asp	50.79425			44.98269		171.16071	79	Val	56.16243	32.59057	171.15137	53.84359	39.99493	173.64356
40	Thr	59.60358	68.31842	170.64384	50.93996	38.28121	172.25144	80	Ile	55.7845	35.44075	171.17989	56.2127	32.82044	171.1842
41	Leu	59.64512	41.44592	170.75503	68.24153	59.60886	170.68113								

Table 8.2 *^{13}C chemical shifts for the DnaA domain I amide backbone assignment. The data shown was accumulated from the following 3D experiments: CβCαNH, CβCα(CO)NH, HNCO, HN(Cα)CO, (HCα)CONH, HNCα and a Cβ optimized CβCαNH.*

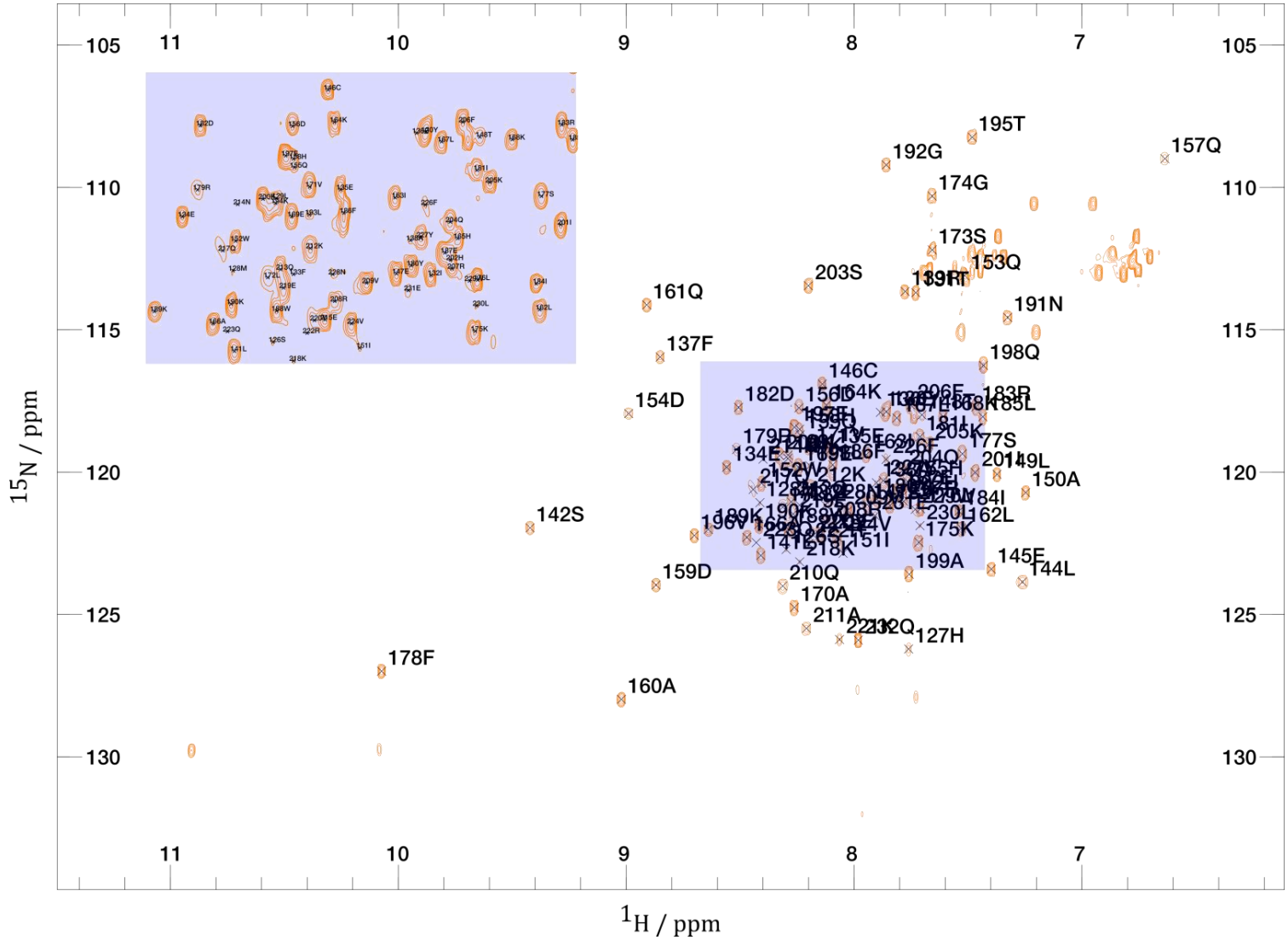


Figure 8.3 *The ^{15}N HSQC spectrum of the DnaD DDBH2 domain.*

Position	Residue	^{1}H/ppm	^{15}N/ppm	Position	Residue	^{1}H/ppm	^{15}N/ppm	Position	Residue	^{1}H/ppm	^{15}N/ppm
126	Ser	8.32	122.5	154	Asp	9.03	117.89	180	Tyr	7.92	120.91
127	His	7.65	125.34	155	Gln	8.3	118.57	181	Ile	7.74	118.57
128	Met	8.39	120.94	156	Asp	8.29	117.67	182	Asp	8.54	117.68
129	Leu	8.31	119.42	157	Gln	6.68	108.94	183	Arg	7.51	117.78
130	Tyr	7.86	117.82	158	His	8.3	118.49	184	Ile	7.59	121.2
131	Thr	7.75	113.75	159	Asp	8.89	123.9	185	Leu	7.48	117.93
132	Ile	7.89	121.14	160	Ala	9.05	127.83	186	Phe	8.12	119.89
133	Phe	8.29	120.95	161	Gln	8.94	114.14	187	Glu	7.85	120.47
134	Glu	8.59	119.74	162	Leu	7.57	121.95	188	Trp	8.34	121.84
135	Glu	8.15	119.07	163	Ile	7.99	119.31	189	Lys	8.65	121.94
136	Glu	7.9	117.7	164	Lys	8.17	117.7	190	Lys	8.46	121.73
137	Phe	8.89	116.04	165	His	7.82	120.22	191	Asn	7.38	114.55
138	Ala	7.91	120.18	166	Ala	8.51	122.35	192	Gly	7.89	109.17
139	Arg	7.82	113.64	167	Leu	7.86	118.08	193	Leu	8.13	119.68
141	Leu	8.43	122.86	168	Lys	7.66	118.06	194	Lys	8.31	119.52
142	Ser	9.49	122.08	169	Glu	8.26	119.59	195	Thr	7.51	108.25
144	Leu	7.46	124.25	170	Ala	8.29	124.66	196	Val	8.73	122.17
145	Glu	7.44	123.41	171	Val	8.26	118.99	197	Glu	8.3	118.47
146	Cys	8.2	116.88	172	Leu	8.34	121.33	198	Gln	7.48	116.23
147	Glu	7.99	121.06	173	Ser	7.7	112.11	199	Ala	7.8	123.47
148	Thr	7.76	117.96	174	Gly	7.7	110.27	200	Lys	8.35	119.33
149	Leu	7.41	120.09	175	Lys	7.76	122.52	201	Ile	7.53	119.9
150	Ala	7.3	120.74	176	Leu	7.73	121.26	202	His	7.84	120.48
151	Ile	8.13	122.24	177	Ser	7.57	119.29	203	Ser	8.21	113.56
152	Trp	8.47	120.33	178	Phe	10.14	126.61	204	Gln	7.86	120.08
153	Gln	7.56	113.06	179	Arg	8.54	119.21	205	Lys	7.74	119.08

Table 8.3 *^{1}H and ^{15}N chemical shifts for the DnaD DDBH2 domain amide backbone assignment.*

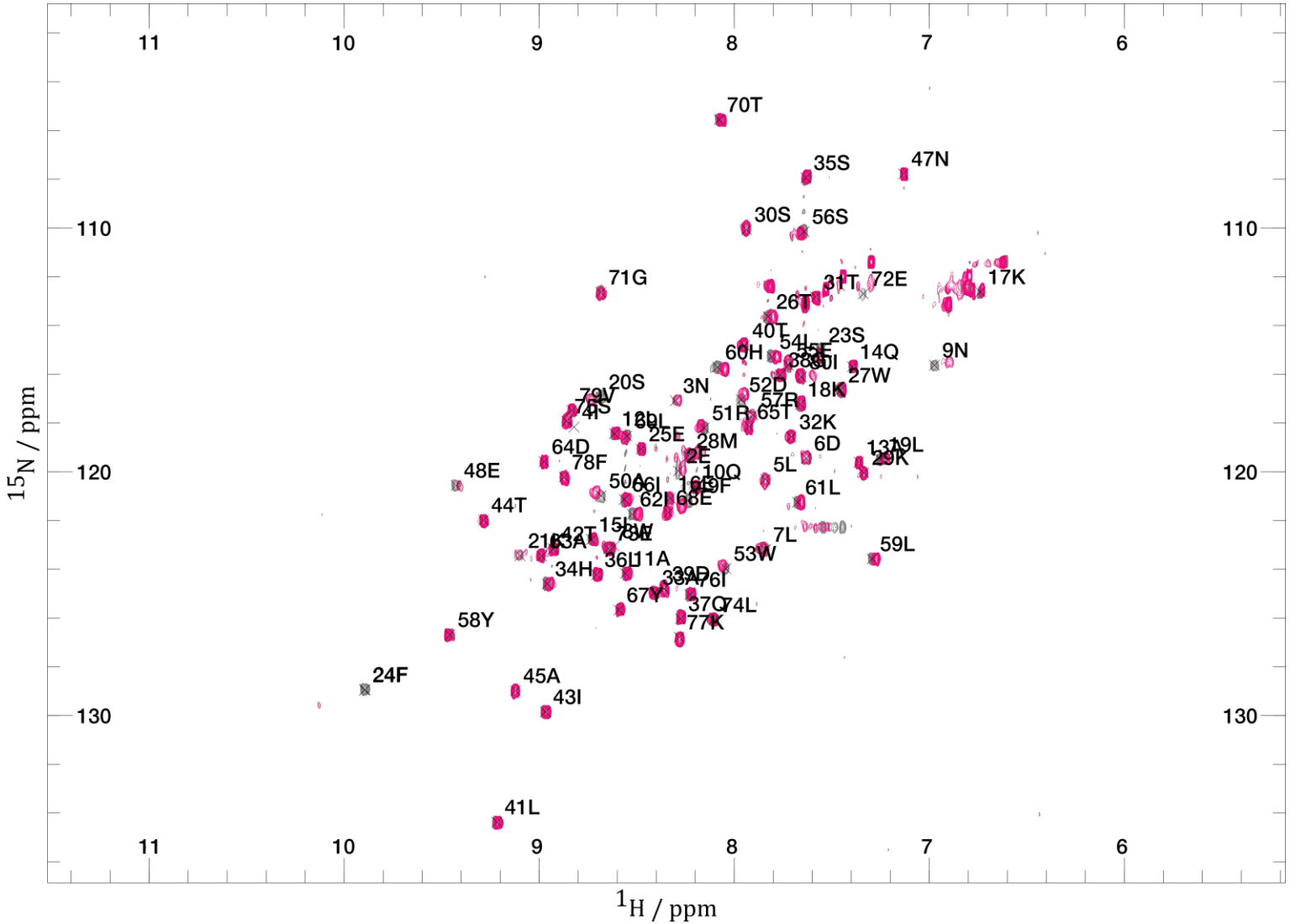


Figure 8.4 *Overlay of the ^{15}N HSQC spectrum of DnaA domain I titrated against DnaD DDBH2 domain. End points shown for DnaA domain I only sample in grey and 8:1 excess DnaD DDBH2 domain sample in pink.*

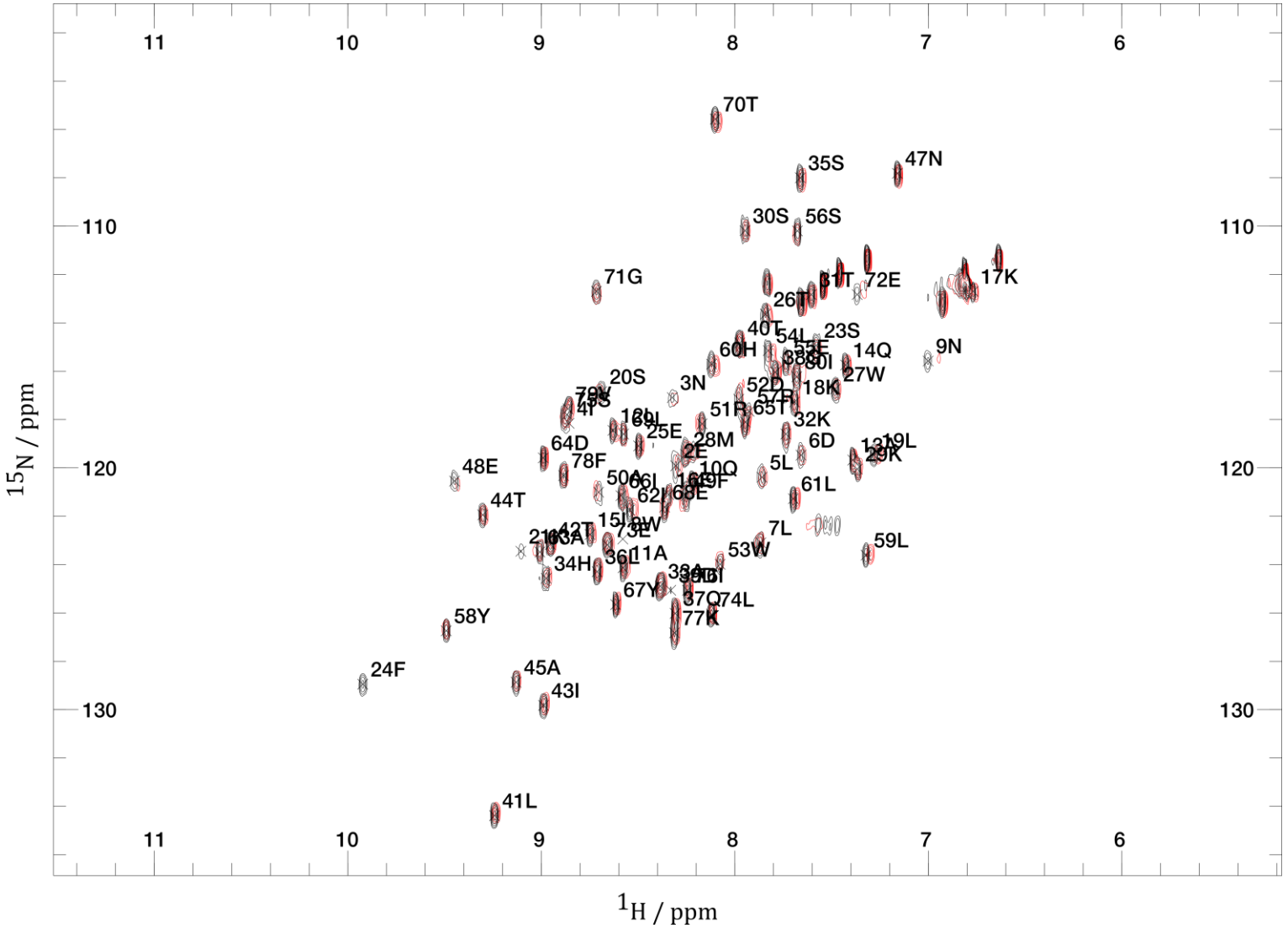


Figure 8.5 *Overlay of the ^{15}N HSQC spectrum of DnaA domain I titrated against DnaD DDBH2 domain in the presence of 800 µM ssDNA. End points shown for DnaA domain I only sample in grey and 8:1 excess DnaD DDBH2 domain sample in red.*

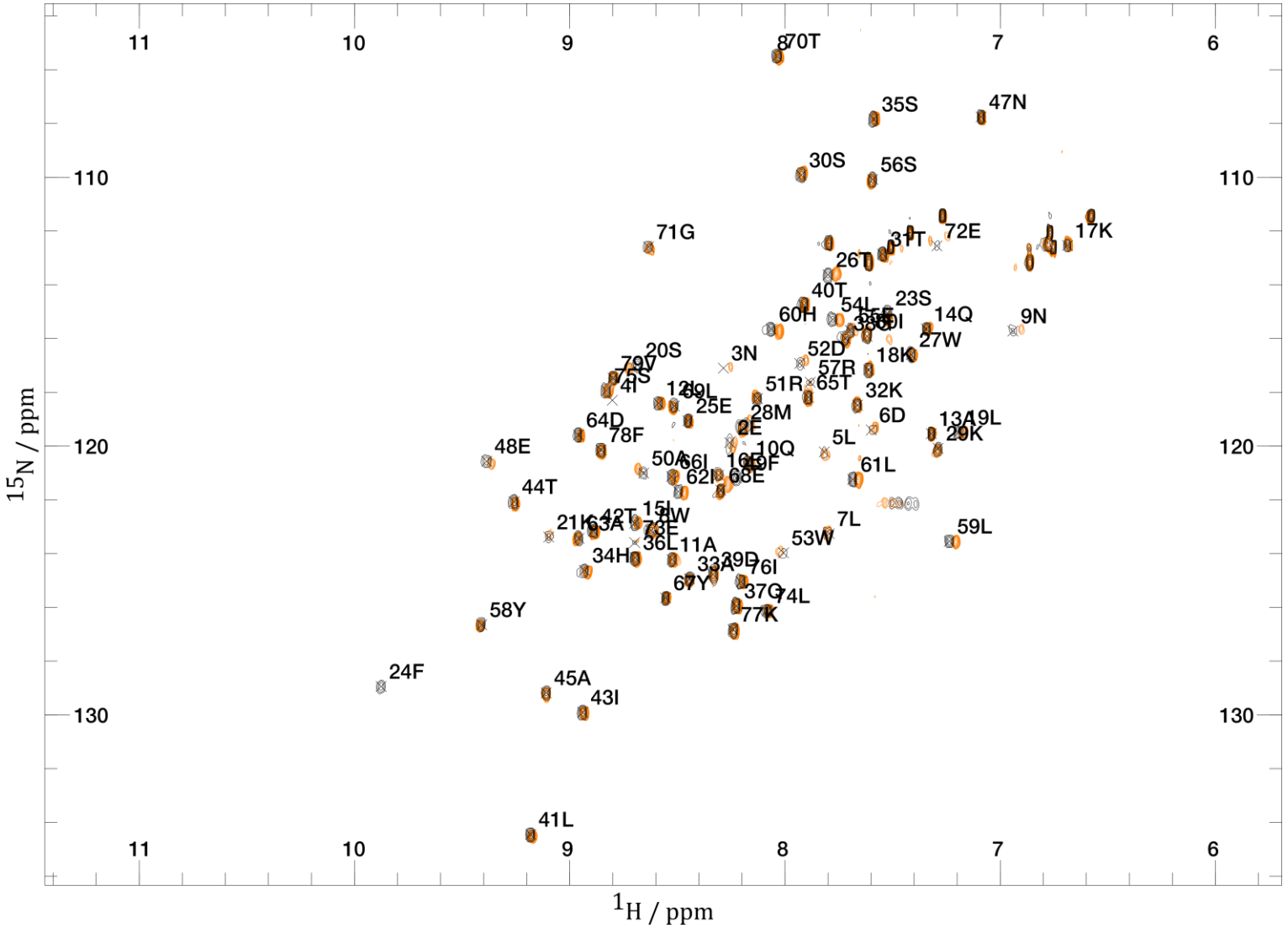


Figure 8.6 *Overlay of the ^{15}N HSQC spectrum of DnaA domain I titrated against truncated DnaD DDBH2 domain (residues 129-196). End points shown for DnaA domain I only sample in grey and 8:1 excess truncated DnaD DDBH2 domain sample in orange.*

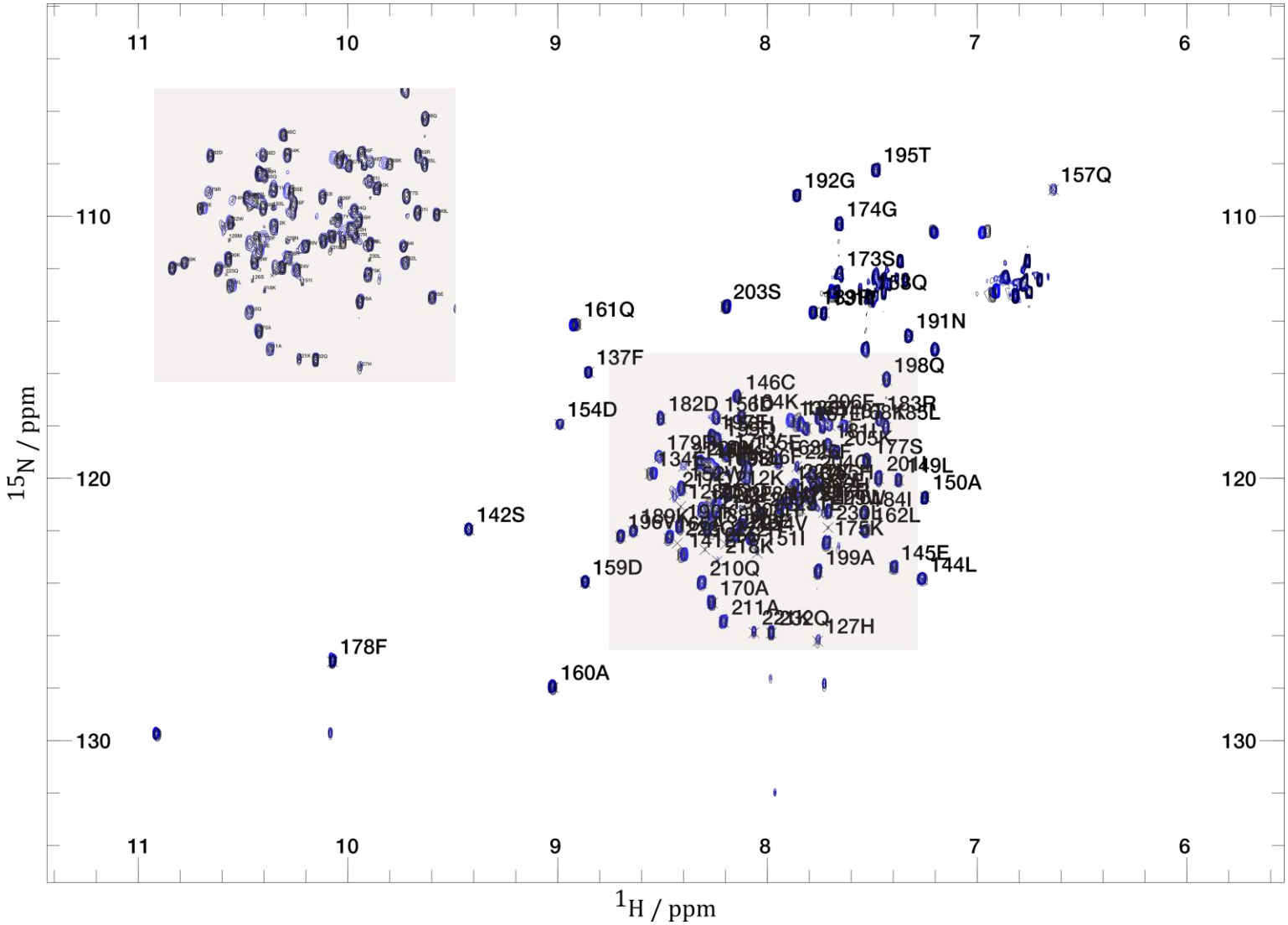


Figure 8.7 *Overlay of the ^{15}N HSQC spectrum of the DnaD DDBH2 domain titrated against DnaA domain I. End points shown for DnaD DDBH2 domain only sample in grey and 8:1 excess DnaA domain I sample in blue.*

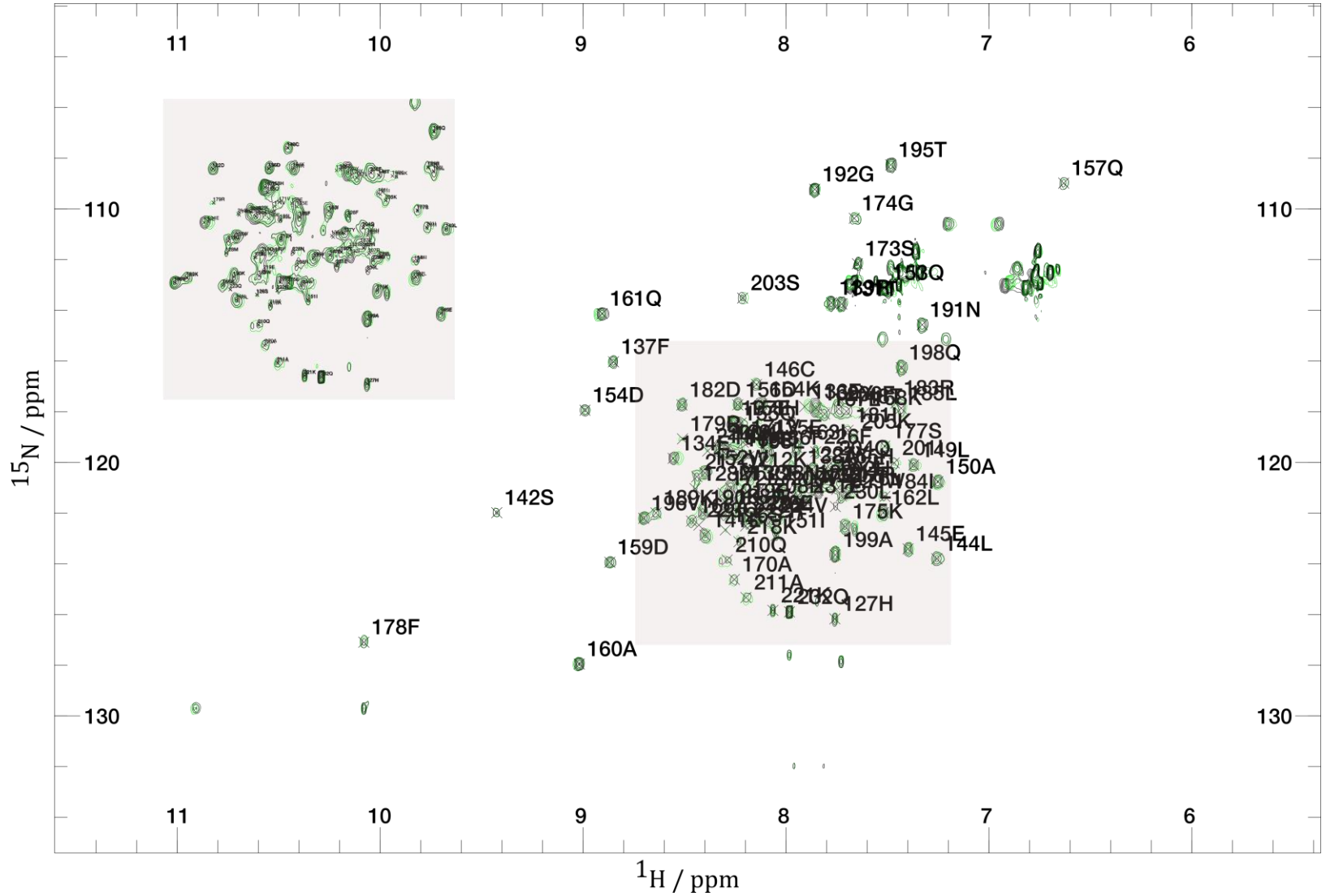


Figure 8.8 *Overlay of the ^{15}N HSQC spectrum of the DnaD DDBH2 domain titrated against DnaA domain I in the presence of 800 μM ssDNA. End points shown for DnaD DDBH2 domain only sample in grey and 8:1 excess DnaA domain I sample in green.*

8.3 SPR Data

Included in the following section are the SPR data, for the DnaA domain I concentration series with DnaD DDBH2 domain, fitted using the BIAevaluation software. As discussed in Chapter 5 (section 5.3) the data gave a poor fit to the models available in this software (Langmuir 1:1 binding model and the bivalent analyte model) limiting analysis using this method and consequently this data has been included in the appendices.

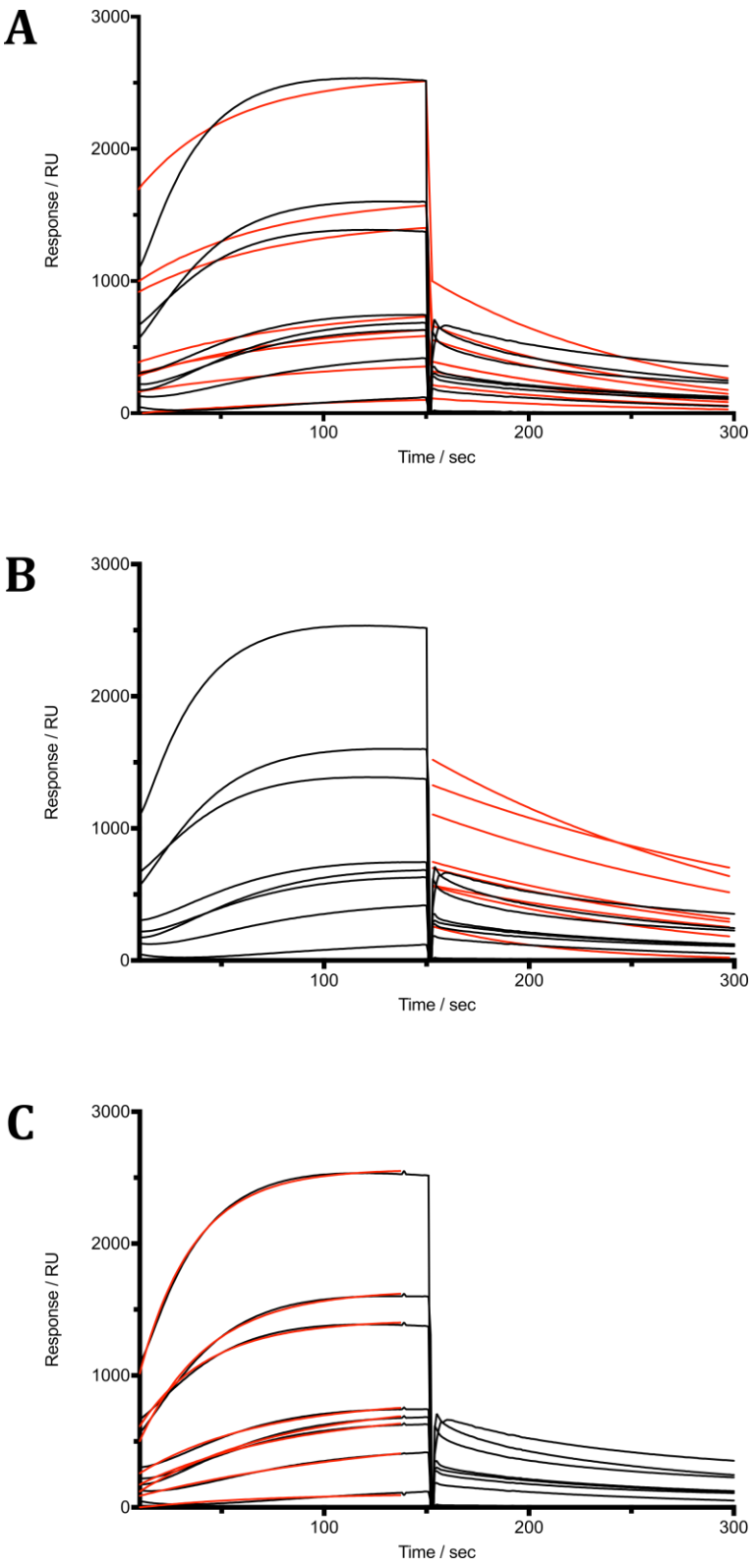


Figure 8.9 *The Langmuir 1:1 binding model was used to fit the sensograms for the DnaA domain I surface concentration series with DnaD DDBH2 domain. Experimental curves (48.125 – 770 μM) are shown in black and fitted curves are shown in red.* ***A.*** *The simultaneous fit of the association and dissociation phases* ***B.*** *The separate fit of the dissociation phase* ***C.*** *The separate fit of the association phase, using* k_d *values obtained from the dissociation phase fit. Kinetic values and* χ^2 *parameters obtained for each fit are displayed in Table 8.4.*

Fit	k_a / $M^{-1}s^{-1}$	k_d / s^{-1}	K_D / µM	R_{max}	χ^2
Langmuir Association	44 ±1.687	-	2140 ±0.003	8.98 x10^3	272
Langmuir Dissociation	-	6.771 x10^{-3} ±3.96 x10^{-3}	-	-	166
Langmuir Simultaneous	16.9 ±0.719	9.23 x10^{-3} ±2 x10^{-4}	547	1.82 x10^3	4.98 x10^3

Table 8.4 *Kinetic analysis of the interaction between DnaA domain I and the DnaD DDBH2 domain. The kinetic equilibrium binding constants were calculated using the Langmuir 1:1 binding model fitting both the simultaneous and separate association and dissociation phases.*

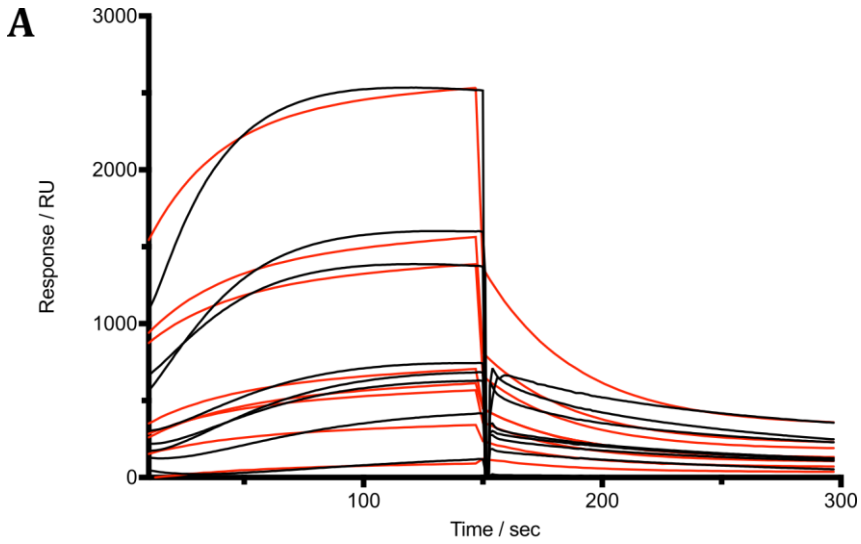


Figure 8.10 *The Bivalent Analyte binding model was used to fit the sensograms for the DnaA domain I surface concentration series with DnaD DDBH2 domain. Experimental curves (48.125 – 770 µM) are shown in black and fitted curves are shown in red. Kinetic values and χ^2 parameters obtained for each fit are displayed in Table 8.5.*

k_{a1} / $M^{-1}s^{-1}$	k_{d1} / s^{-1}	k_{a2} / $RU^{-1}s^{-1}$	k_{d2} / s^{-1}	R_{max}	χ^2
5.72 ±0.0683	0.0242 ±3.19 x10^{-4}	5.06 x10^{-7} ±2.26 x10^{-8}	4.93 x10^{-5} ±1.49 x10^{-4}	4.89 x10^{3}	4.21 x10^{3}

Table 8.5 *Kinetic analysis of the interaction between DnaA domain I and the DnaD DDBH2 domain. The kinetic equilibrium binding constants were calculated using the Bivalent Analyte model. The association constant for the second site binding (k_{a2}) is recorded in $RU^{-1}s^{-1}$ as the reaction involves the complex in RU binding to ligand in RU.*

References

1. Alberts, B., et al., *Molecular Biology of the Cell.* 2002.
2. Meselson, M. and F.W. Stahl, *The replication of DNA in Escherichia coli.* 1958.
3. Sakabe, K. and R. Okazaki, *A unique property of the replicating region of chromosomal DNA.* Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis, 1966. **129**(3): p. 651-654.
4. Briggs, G.S., W.K. Smits, and P. Soultanas, *Chromosomal replication initiation machinery of low-G+C-content Firmicutes.* J Bacteriol, 2012. **194**(19): p. 5162-70.
5. Hori, H. and S. Osawa, *Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences.* Mol Biol Evol, 1987. **4**(5): p. 445-72.
6. Battistuzzi, F.U., A. Feijao, and S.B. Hedges, *A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land*, in *BMC Evol Biol*. 2004. p. 44.
7. Robinson, A., R.J. Causer, and N.E. Dixon, *Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target.* Curr Drug Targets, 2012. **13**(3): p. 352-72.
8. Leonard, A.C. and J.E. Grimwade, *The orisome: structure and function.* Front Microbiol, 2015. **6**: p. 545.
9. Sanyal, G. and P. Doig, *Bacterial DNA replication enzymes as targets for antibacterial drug discovery.* Expert Opinion on Drug Discovery, 2012. **7**(4): p. 327-339.
10. Gille, H. and W. Messer, *Localized DNA melting and structural pertubations in the origin of replication, oriC, of Escherichia coli in vitro and in vivo.* EMBO J, 1991. **10**(6): p. 1579-84.
11. Wolański, M., et al., *oriC-encoded instructions for the initiation of bacterial chromosome replication.* Front Microbiol, 2014. **5**.
12. Ogasawara, N. and H. Yoshikawa, *Genes and their organization in the replication origin region of the bacterial chromosome.* Mol Microbiol, 1992. **6**(5): p. 629-34.
13. Krause, M., et al., *Complexes at the replication origin of Bacillus subtilis with homologous and heterologous DnaA protein.* J Mol Biol, 1997. **274**(3): p. 365-80.
14. Fukuoka, T., et al., *Purification and characterization of an initiation protein for chromosomal replication, DnaA, in Bacillus subtilis.* J Biochem, 1990. **107**(5): p. 732-9.
15. Rozgaja, T.A., et al., *Two oppositely-oriented arrays of low affinity recognition sites in oriC guide progressive binding of DnaA during E. coli pre-RC assembly.* Mol Microbiol, 2011. **82**(2): p. 475-88.

16. Woelker, B. and W. Messer, *The structure of the initiation complex at the replication origin, oriC, of Escherichia coli.* Nucleic Acids Res, 1993. **21**(22): p. 5025-33.
17. Zawilak-Pawlik, A., et al., *Architecture of bacterial replication initiation complexes: orisomes from four unrelated bacteria.* Biochem J, 2005. **389**(Pt 2): p. 471-81.
18. Zawilak, A., et al., *DNA binding specificity of the replication initiator protein, DnaA from Helicobacter pylori.* J Mol Biol, 2003. **334**(5): p. 933-47.
19. Ogasawara, N., et al., *Conservation of genes and their organization in the chromosomal replication origin region of Bacillus subtilis and Escherichia coli.* EMBO J, 1985. **4**(12): p. 3345-50.
20. Rajewska, M., K. Wegrzyn, and I. Konieczny, *AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons.* FEMS Microbiol Rev, 2012. **36**(2): p. 408-34.
21. Fuller, R.S., B.E. Funnell, and A. Kornberg, *The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites.* Cell, 1984. **38**(3): p. 889-900.
22. Zorman, S., et al., *Topological characterization of the DnaA–oriC complex using single-molecule nanomanipuation*, in *Nucleic Acids Res*. 2012. p. 7375-83.
23. Erzberger, J.P., M.L. Mott, and J.M. Berger, *Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling.* Nat Struct Mol Biol, 2006. **13**(8): p. 676-83.
24. Kowalski, D. and M.J. Eddy, *The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin.* EMBO J, 1989. **8**(13): p. 4335-44.
25. Bramhill, D. and A. Kornberg, *Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome.* Cell, 1988. **52**(5): p. 743-55.
26. Marszalek, J. and J.M. Kaguni, *DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli.* J Biol Chem, 1994. **269**(7): p. 4883-90.
27. Kobori, J.A. and A. Kornberg, *The Escherichia coli dnaC gene product. III. Properties of the dnaB-dnaC protein complex.* J Biol Chem, 1982. **257**(22): p. 13770-5.
28. Nishida, S., et al., *A nucleotide switch in the Escherichia coli DnaA protein initiates chromosomal replication: evidnece from a mutant DnaA protein defective in regulatory ATP hydrolysis in vitro and in vivo.* J Biol Chem, 2002. **277**(17): p. 14986-95.
29. Sekimizu, K., D. Bramhill, and A. Kornberg, *ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome.* Cell, 1987. **50**(2): p. 259-65.
30. Margulies, C. and J.M. Kaguni, *Ordered and sequential binding of DnaA protein to oriC, the chromosomal origin of Escherichia coli.* J Biol Chem, 1996. **271**(29): p. 17035-40.
31. Langer, U., et al., *A comprehensive set of DnaA-box mutations in the replication origin, oriC, of Escherichia coli.* Mol Microbiol, 1996. **21**(2): p. 301-11.
32. Miller, D.T., et al., *Bacterial origin recognition complexes direct assembly of higher-order DnaA oligomeric structures.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18479-84.
33. Kaur, G., et al., *Building the bacterial orisome: high-affinity DnaA recognition plays a role in setting the conformation of oriC DNA.* Mol Microbiol, 2014. **91**(6): p. 1148-63.

34. Zorman, S., et al., *Topological characterization of the DnaA–oriC complex using single-molecule nanomanipuation.* Nucleic Acids Research, 2012. **40**(15): p. 7375-7383.
35. Marszalek, J., et al., *Domains of DnaA protein involved in interaction with DnaB protein, and in unwinding the Escherichia coli chromosomal origin.* J Biol Chem, 1996. **271**(31): p. 18535-42.
36. Abe, Y., et al., *Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on oriC.* J Biol Chem, 2007. **282**(24): p. 17816-27.
37. Arias-Palomo, E., et al., *The bacterial DnaC helicase loader is a DnaB ring breaker.* Cell, 2013. **153**(2): p. 438-48.
38. Soultanas, P., *Loading mechanisms of ring helicases at replication origins.* Mol Microbiol, 2012. **84**(1): p. 6-16.
39. Bruck, I. and M. O'Donnell, *The DNA replication machine of a gram-positive organism.* J Biol Chem, 2000. **275**(37): p. 28971-83.
40. Velten, M., et al., *A two-protein strategy for the functional loading of a cellular replicative DNA helicase.* Mol Cell, 2003. **11**(4): p. 1009-20.
41. Beattie, T.R. and R. Reyes-Lamothe, *A Replisome's journey through the bacterial chromosome.* Front Microbiol, 2015. **6**: p. 562.
42. Sanders, G.M., H.G. Dallmann, and C.S. McHenry, *Reconstitution of the B. subtilis replisome with 13 proteins including two distinct replicases.* Mol Cell, 2010. **37**(2): p. 273-81.
43. Dervyn, E., et al., *Two essential DNA polymerases at the bacterial replication fork.* Science, 2001. **294**(5547): p. 1716-9.
44. Hill, T.M., J.M. Henson, and P.L. Kuempel, *The terminus region of the Escherichia coli chromosome contains two separate loci that exhibit polar inhibition of replication.* Proc Natl Acad Sci U S A, 1987. **84**(7): p. 1754-8.
45. Hidaka, M., et al., *Purification of a DNA replication terminus (ter) site-binding protein in Escherichia coli and identification of the structural gene.* J Biol Chem, 1989. **264**(35): p. 21031-7.
46. Neylon, C., et al., *Replication termination in Escherichia coli: structure and antihelicase activity of the Tus-Ter complex.* Microbiol Mol Biol Rev, 2005. **69**(3): p. 501-26.
47. Kamada, K., et al., *Structure of a replication-terminator protein complexed with DNA.* Nature, 1996. **383**(6601): p. 598-603.
48. Mulcair, M.D., et al., *A molecular mousetrap determines polarity of termination of DNA replication in E. coli.* Cell, 2006. **125**(7): p. 1309-19.
49. Berghuis, B.A., et al., *Strand separation establishes a sustained lock at the Tus-Ter replication fork barrier.* Nat Chem Biol, 2015. **11**(8): p. 579-85.
50. Hill, T.M., *Arrest of bacterial DNA replication.* Annu Rev Microbiol, 1992. **46**: p. 603-33.
51. Lewis, P.J., et al., *Identification of the replication terminator protein binding sites in the terminus region of the Bacillus subtilis chromosome and stoichiometry of the binding.* J Mol Biol, 1990. **214**(1): p. 73-84.
52. Vivian, J.P., et al., *An asymmetric structure of the Bacillus subtilis replication terminator protein in complex with DNA.* J Mol Biol, 2007. **370**(3): p. 481-91.
53. Langley, D.B., et al., *Protein-nucleoside contacts in the interaction between the replication terminator protein of Bacillus subtilis and the DNA terminator.* Mol Microbiol, 1993. **10**(4): p. 771-9.
54. Helmstetter, C.E. and O. Pierucci, *DNA synthesis during the division cycle of three substrains of Escherichia coli B/r.* J Mol Biol, 1976. **102**(3): p. 477-86.

55. Atlung, T., A. Lobner-Olesen, and F.G. Hansen, *Overproduction of DnaA protein stimulates initiation of chromosome and minichromosome replication in Escherichia coli.* Mol Gen Genet, 1987. **206**(1): p. 51-9.
56. Flatten, I., et al., *The DnaA Protein Is Not the Limiting Factor for Initiation of Replication in Escherichia coli.* PLoS Genet, 2015. **11**(6): p. e1005276.
57. Murray, H. and A. Koh, *Multiple regulatory systems coordinate DNA replication with cell growth in Bacillus subtilis.* PLoS Genet, 2014. **10**(10): p. e1004731.
58. Goranov, A.I., et al., *A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12932-7.
59. Braun, R.E., K. O'Day, and A. Wright, *Autoregulation of the DNA replication gene dnaA in E. coli K-12.* Cell, 1985. **40**(1): p. 159-69.
60. Collier, J., S.R. Murray, and L. Shapiro, *DnaA couples DNA replication and the expression of two cell cycle master regulators*, in *EMBO J*. 2006. p. 346-56.
61. Fernandez-Fernandez, C., D. Gonzalez, and J. Collier, *Regulation of the activity of the dual-function DnaA protein in Caulobacter crescentus.* PLoS One, 2011. **6**(10): p. e26028.
62. Kaguni, J.M., *DnaA: controlling the initiation of bacterial DNA replication and more.* Annu Rev Microbiol, 2006. **60**: p. 351-75.
63. Costa, A., I.V. Hood, and J.M. Berger, *Mechanisms for initiating cellular DNA replication.* Annu Rev Biochem, 2013. **82**: p. 25-54.
64. Katayama, T., et al., *Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC*, in *Nat Rev Microbiol*. 2010: England. p. 163-70.
65. Bonilla, C.Y. and A.D. Grossman, *The primosomal protein DnaD inhibits cooperative DNA binding by the replication initiator DnaA in Bacillus subtilis.* J Bacteriol, 2012. **194**(18): p. 5110-7.
66. Scholefield, G. and H. Murray, *YabA and DnaD inhibit helix assembly of the DNA replication initiation protein DnaA.* Mol Microbiol, 2013. **90**(1): p. 147-59.
67. Messer, W., et al., *Functional domains of DnaA proteins.* Biochimie, 1999. **81**(8-9): p. 819-25.
68. Scholefield, G., J. Errington, and H. Murray, *Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA.* EMBO J, 2012. **31**(6): p. 1542-55.
69. Natrajan, G., et al., *The structure of a DnaA/HobA complex from Helicobacter pylori provides insight into regulation of DNA replication in bacteria.* Proc Natl Acad Sci U S A, 2009. **106**(50): p. 21115-20.
70. Lowery, T.J., et al., *NMR structure of the N-terminal domain of the replication initiator protein DnaA.* J Struct Funct Genomics, 2007. **8**(1): p. 11-7.
71. Nozaki, S. and T. Ogawa, *Determination of the minimum domain II size of Escherichia coli DnaA protein essential for cell viability.* Microbiology, 2008. **154**(Pt 11): p. 3379-84.
72. Ogura, T. and A.J. Wilkinson, *AAA+ superfamily ATPases: common structure--diverse function.* Genes Cells, 2001. **6**(7): p. 575-97.
73. Katayama, T., *Roles for the AAA+ motifs of DnaA in the initiation of DNA replication.* Biochem Soc Trans, 2008. **36**(Pt 1): p. 78-82.
74. Speck, C. and W. Messer, *Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA*, in *EMBO J*. 2001. p. 1469-76.
75. Duderstadt, K.E., K. Chuang, and J.M. Berger, *DNA stretching by bacterial initiators promotes replication origin opening.* Nature, 2011. **478**(7368): p. 209-13.
76. Fujikawa, N., et al., *Structural basis of replication origin recognition by the DnaA protein.* Nucleic Acids Res, 2003. **31**(8): p. 2077-86.

77. Katayama, T., et al., *The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase.* Cell, 1998. **94**(1): p. 61-71.
78. Kato, J. and T. Katayama, *Hda, a novel DnaA-related protein, regulates the replication cycle in Escherichia coli*, in *EMBO J*. 2001. p. 4253-62.
79. Su'etsugu, M., et al., *Molecular mechanism of DNA replication-coupled inactivation of the initiator protein in Escherichia coli: interaction of DnaA with the sliding clamp-loaded DNA and the sliding clamp-Hda complex.* Genes Cells, 2004. **9**(6): p. 509-22.
80. Su'etsugu, M., et al., *Protein Associations in DnaA-ATP Hydrolysis Mediated by the Hda-Replicase Clamp Complex.* 2005.
81. Nakamura, K. and T. Katayama, *Novel essential residues of Hda for interaction with DnaA in the regulatory inactivation of DnaA: unique roles for Hda AAA Box VI and VII motifs.* Mol Microbiol, 2010. **76**(2): p. 302-17.
82. Duderstadt, K.E., et al., *Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator.* J Biol Chem, 2010. **285**(36): p. 28229-39.
83. Hayashi, M., et al., *Bacillus subtilis YabA is involved in determining the timing and synchrony of replication initiation.* FEMS Microbiol Lett, 2005. **247**(1): p. 73-9.
84. Noirot-Gros, M.F., et al., *An expanded view of bacterial DNA replication.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8342-7.
85. Noirot-Gros, M.F., et al., *Functional dissection of YabA, a negative regulator of DNA replication initiation in Bacillus subtilis.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2368-73.
86. Iyer, L.M., et al., *Evolutionary history and higher order classification of AAA+ ATPases.* J Struct Biol, 2004. **146**(1-2): p. 11-31.
87. Zhang, W., et al., *The Bacillus subtilis primosomal protein DnaD untwists supercoiled DNA.* J Bacteriol, 2006. **188**(15): p. 5487-93.
88. Smits, W.K., A.I. Goranov, and A.D. Grossman, *Ordered association of helicase loader proteins with the Bacillus subtilis origin of replication in vivo.* Mol Microbiol, 2010. **75**(2): p. 452-61.
89. Bruand, C., et al., *DnaB, DnaD and DnaI proteins are components of the Bacillus subtilis replication restart primosome.* Mol Microbiol, 2001. **42**(1): p. 245-55.
90. Marston, F.Y., et al., *When simple sequence comparison fails: the cryptic case of the shared domains of the bacterial replication initiation proteins DnaB and DnaD.* Nucleic Acids Res, 2010. **38**(20): p. 6930-42.
91. Nunez-Ramirez, R., et al., *Loading a ring: structure of the Bacillus subtilis DnaB protein, a co-loader of the replicative helicase.* J Mol Biol, 2007. **367**(3): p. 764-9.
92. Carneiro, M.J., et al., *The DNA-remodelling activity of DnaD is the sum of oligomerization and DNA-binding activities on separate domains.* Mol Microbiol, 2006. **60**(4): p. 917-24.
93. Grainger, W.H., et al., *DnaB proteolysis in vivo regulates oligomerization and its localization at oriC in Bacillus subtilis.* Nucleic Acids Res, 2010. **38**(9): p. 2851-64.
94. Ioannou, C., et al., *Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in Bacillus subtilis.* Nucleic Acids Res, 2006. **34**(18): p. 5247-58.
95. Rokop, M.E., J.M. Auchtung, and A.D. Grossman, *Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of Bacillus subtilis.* Mol Microbiol, 2004. **52**(6): p. 1757-67.
96. Hoshino, T., et al., *Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for DNA replication initiation and membrane attachment.* Proc Natl Acad Sci U S A, 1987. **84**(3): p. 653-7.
97. Marsin, S., et al., *Early steps of Bacillus subtilis primosome assembly.* J Biol Chem, 2001. **276**(49): p. 45818-25.

98. Turner, I.J., et al., *The Bacillus subtilis DnaD protein: a putative link between DNA remodeling and initiation of DNA replication.* FEBS Lett, 2004. **577**(3): p. 460-4.
99. Zhang, W., et al., *The Bacillus subtilis DnaD and DnaB proteins exhibit different DNA remodelling activities.* J Mol Biol, 2005. **351**(1): p. 66-75.
100. Schneider, S., et al., *Structure of the N-terminal oligomerization domain of DnaD reveals a unique tetramerization motif and provides insights into scaffold formation.* J Mol Biol, 2008. **376**(5): p. 1237-50.
101. Schneider, S., et al., *Crystallization and X-ray diffraction analysis of the DNA-remodelling protein DnaD from Bacillus subtilis.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2007. **63**(Pt 2): p. 110-3.
102. Zhang, W., et al., *Single-molecule atomic force spectroscopy reveals that DnaD forms scaffolds and enhances duplex melting.* J Mol Biol, 2008. **377**(3): p. 706-14.
103. Bruand, C., et al., *Functional interplay between the Bacillus subtilis DnaD and DnaB proteins essential for initiation and re-initiation of DNA replication.* Mol Microbiol, 2005. **55**(4): p. 1138-50.
104. Collier, C., et al., *Untwisting of the DNA helix stimulates the endonuclease activity of Bacillus subtilis Nth at AP sites.* Nucleic Acids Res, 2012. **40**(2): p. 739-50.
105. Chang, W., et al., *Global transcriptome analysis of Staphylococcus aureus response to hydrogen peroxide.* J Bacteriol, 2006. **188**(4): p. 1648-59.
106. Ishigo-Oka, D., N. Ogasawara, and S. Moriya, *DnaD protein of Bacillus subtilis interacts with DnaA, the initiator protein of replication.* J Bacteriol, 2001. **183**(6): p. 2148-50.
107. Jameson, K.H., et al., *Structure and interactions of the Bacillus subtilis sporulation inhibitor of DNA replication, SirA, with domain I of DnaA.* Mol Microbiol, 2014. **93**(5): p. 975-91.
108. Hulme, E.C. and M.A. Trevethick, *Ligand binding assays at equilibrium: validation and interpretation.* Br J Pharmacol, 2010. **161**(6): p. 1219-37.
109. Hochuli, E., et al., *Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent.* Nature Biotechnology, 1988. **6**(11): p. 1321.
110. Xiong, S., L. Zhang, and Q.Y. He, *Fractionation of proteins by heparin chromatography.* Methods Mol Biol, 2008. **424**: p. 213-21.
111. Sheehan, D. and R. FitzGerald, *Ion-exchange chromatography.* Methods Mol Biol, 1996. **59**: p. 145-50.
112. PORATH, J. and P. FLODIN, *Gel Filtration: A Method for Desalting and Group Separation.* Nature, 1959. **183**(4676): p. 1657.
113. Kay, L.E., *Protein dynamics from NMR.* Nat Struct Biol, 1998. **5 Suppl**: p. 513-7.
114. Dobson, C.M. and P.J. Hore, *Kinetic studies of protein folding using NMR spectroscopy.* Nat Struct Biol, 1998. **5 Suppl**: p. 504-7.
115. Ernst, R.R. and W.A. Anderson, *Application of Fourier Transform Spectroscopy to Magnetic Resonance.* http://oasc12039.247realmedia.com/RealMedia/ads/click_lx.ads/www.aip.org/pt/adcenter/pdfcover_test/L-37/1778822819/x01/AIP-PT/RSI_ArticleDL_1218/scilight717-1640x440.gif/434f71374e315a556e61414141774c75?x, 1966.
116. Wüthrich, K., K. Nagayama, and R.R. Ernst, *Emerging techniques Two-dimensional NMR spectroscopy.* Trends in Biochemical Sciences, 1979. **4**(8): p. N178-N181.
117. Wider, G., et al., *Homonuclear two-dimensional 1H NMR of proteins. Experimental procedures.* Journal of Magnetic Resonance (1969), 1984. **56**(2): p. 207-234.
118. Montelione, G.T. and G. Wagner, *Conformation-independent sequential NMR connections in isotope-enriched polypeptides by 1H・13C・15N triple-resonance experiments.* Journal of Magnetic Resonance (1969), 1990. **87**(1): p. 183-188.

119. Bax, A. and S. Grzesiek, *Methodological Advances in Protein NMR*, in *NMR of Proteins*, G.M. Clore and A.M. Gronenborn, Editors. 1993, Macmillan Education UK: London. p. 33-52.
120. Clore, G.M. and A.M. Gronenborn, *Applications of three- and four-dimensional heteronuclear NMR spectroscopy to protein structure determination.* Progress in Nuclear Magnetic Resonance Spectroscopy, 1991. **23**(1): p. 43-92.
121. Wuthrich, K., *The second decade--into the third millenium.* Nat Struct Biol, 1998. **5 Suppl**: p. 492-5.
122. Page, R., et al., *NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline.* Proc Natl Acad Sci U S A, 2005. **102**(6): p. 1901-5.
123. Wishart, D.S., et al., *1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J Biomol NMR, 1995. **5**(1): p. 67-81.
124. Cavanagh, J., et al., *CHAPTER 4 - MULTIDIMENSIONAL NMR SPECTROSCOPY*, in *Protein NMR Spectroscopy (Second Edition)*. 2007, Academic Press: Burlington. p. 271-332.
125. Grzesiek, S. and A. Bax, *An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins.* Journal of Magnetic Resonance (1969), 1992. **99**(1): p. 201-207.
126. Grzesiek, S. and A. Bax, *Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins.* J Biomol NMR, 1993. **3**(2): p. 185-204.
127. Muhandiram, D.R. and L.E. Kay, *Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity.* Journal of Magnetic Resonance, Series B, 1994. **103**(3): p. 203-216.
128. Grzesiek, S. and A. Bax, *Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein.* Journal of Magnetic Resonance (1969), 1992. **96**(2): p. 432-440.
129. Sattler, M. and C. Griesinger, *Coherence Selection by Gradients without Signal Attenuation: Application to the Three-Dimensional HNCO Experiment*. Vol. 32. 1993. 1489-1491.
130. Kay, L.E., G.Y. Xu, and T. Yamazaki, *Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H2O Saturation.* Journal of Magnetic Resonance, Series A, 1994. **109**(1): p. 129-133.
131. Clubb, R.T., V. Thanabal, and G. Wagner, *A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C' chemical shifts in 15N・13C-labelled proteins.* Journal of Magnetic Resonance (1969), 1992. **97**(1): p. 213-217.
132. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.
133. Goto, N.K. and L.E. Kay, *New developments in isotope labeling strategies for protein solution NMR spectroscopy.* Curr Opin Struct Biol, 2000. **10**(5): p. 585-92.
134. Muchmore, D.C., et al., *Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance.* Methods Enzymol, 1989. **177**: p. 44-73.
135. Krishnarjuna, B., et al., *Amino acid selective unlabeling for sequence specific resonance assignments in proteins.* J Biomol NMR, 2011. **49**(1): p. 39-51.
136. Kelly, M.J., et al., *Application of amino acid type-specific 1H- and 14N-labeling in a 2H-, 15N-labeled background to a 47 kDa homodimer: potential for NMR structure determination of large proteins.* J Biomol NMR, 1999. **14**(1): p. 79-83.
137. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2013. **73**: p. 1-16.

138. Williamson, R.A., et al., *Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation.* Biochemistry, 1997. **36**(45): p. 13882-9.
139. Jonsson, U., et al., *Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology.* Biotechniques, 1991. **11**(5): p. 620-7.
140. Jonsson, U., et al., *Introducing a biosensor based technology for real-time biospecific interaction analysis.* Ann Biol Clin (Paris), 1993. **51**(1): p. 19-26.
141. Karlsson, R. and A. Fält, *Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors.* Journal of Immunological Methods, 1997. **200**(1): p. 121-133.
142. Nelson, B.P., et al., *Surface plasmon resonance imaging measurements of DNA and RNA hybridization adsorption onto DNA microarrays.* Anal Chem, 2001. **73**(1): p. 1-7.
143. Jennifer M. Brockman, a. Anthony G. Frutos, and R.M. Corn*, *A Multistep Chemical Modification Procedure To Create DNA Arrays on Gold Surfaces for the Study of Protein−DNA Interactions with Surface Plasmon Resonance Imaging.* 1999.
144. Fagerstam, L.G., et al., *Detection of antigen-antibody interactions by surface plasmon resonance. Application to epitope mapping.* J Mol Recognit, 1990. **3**(5-6): p. 208-14.
145. Karlsson, R., A. Michaelsson, and L. Mattsson, *Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system.* Journal of Immunological Methods, 1991. **145**(1): p. 229-240.
146. Huber, W., *A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions.* J Mol Recognit, 2005. **18**(4): p. 273-81.
147. Johnsson, B., S. Löfås, and G. Lindquist, *Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors.* Analytical Biochemistry, 1991. **198**(2): p. 268-277.
148. Nieba, L., et al., *BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip.* Anal Biochem, 1997. **252**(2): p. 217-28.
149. Karlsson, R., *Affinity analysis of non-steady-state data obtained under mass transport limited conditions using BIAcore technology.* J Mol Recognit, 1999. **12**(5): p. 285-92.
150. Gershon, P.D. and S. Khilko, *Stable chelating linkage for reversible immobilization of oligohistidine tagged proteins in the BIAcore surface plasmon resonance detector.* J Immunol Methods, 1995. **183**(1): p. 65-76.
151. O'Shannessy, D.J., et al., *Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods.* Anal Biochem, 1993. **212**(2): p. 457-68.
152. Rahn-Lee, L., et al., *The Conserved Sporulation Protein YneE Inhibits DNA Replication in Bacillus subtilis.* 2009.
153. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
154. Tovchigrechko, A., C.A. Wells, and I.A. Vakser, *Docking of protein models.* Protein Sci, 2002. **11**(8): p. 1888-96.
155. Soni, N. and M.S. Madhusudhan, *Computational modeling of protein assemblies.* Curr Opin Struct Biol, 2017. **44**: p. 179-189.
156. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic Acids Res, 2018.
157. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit-5.6.

158. Contreras-Moreira, B., et al., *Novel use of a genetic algorithm for protein structure prediction: searching template and sequence alignment space.* Proteins, 2003. **53 Suppl 6**: p. 424-9.
159. Yu, J., et al., *InterEvDock: a docking server to predict the structure of protein-protein interactions using evolutionary information.* Nucleic Acids Res, 2016. **44**(W1): p. W542-9.
160. Ramirez-Aportela, E., J.R. Lopez-Blanco, and P. Chacon, *FRODOCK 2.0: fast protein-protein docking server.* Bioinformatics, 2016. **32**(15): p. 2386-8.
161. Andreani, J., G. Faure, and R. Guerois, *InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution.* Bioinformatics, 2013. **29**(14): p. 1742-9.
162. Dong, G.Q., et al., *Optimized atomic statistical potentials: assessment of protein interfaces and loops.* Bioinformatics, 2013. **29**(24): p. 3158-66.
163. Rodrigues, J.P., et al., *Clustering biomolecular complexes by residue contacts similarity.* Proteins, 2012. **80**(7): p. 1810-7.
164. Viswanath, S., D.V. Ravikant, and R. Elber, *DOCK/PIERR: web server for structure prediction of protein-protein complexes.* Methods Mol Biol, 2014. **1137**: p. 199-207.
165. Baspinar, A., et al., *PRISM: a web server and repository for prediction of protein-protein interactions and modeling their 3D complexes.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W285-9.
166. Roberts, V.A., et al., *DOT2: Macromolecular docking with improved biophysical models.* J Comput Chem, 2013. **34**(20): p. 1743-58.
167. Pierce, B.G., et al., *ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers.* Bioinformatics, 2014. **30**(12): p. 1771-3.
168. Vakser, I.A., *Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex.* Proteins, 1997. **Suppl 1**: p. 226-30.
169. Kozakov, D., et al., *The ClusPro web server for protein-protein docking.* Nat Protoc, 2017. **12**(2): p. 255-278.
170. Kozakov, D., et al., *PIPER: an FFT-based protein docking program with pairwise potentials.* Proteins, 2006. **65**(2): p. 392-406.
171. van Zundert, G.C.P., et al., *The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes.* J Mol Biol, 2016. **428**(4): p. 720-725.
172. Lensink, M.F. and S.J. Wodak, *Docking and scoring protein interactions: CAPRI 2009.* Proteins, 2010. **78**(15): p. 3073-84.
173. Wassenaar, T.A., et al., *WeNMR: Structural Biology on the Grid.* Journal of Grid Computing, 2012. **10**(4): p. 743-767.
174. *InterfaceResidues - PyMOLWiki*. 2018; Available from: https://pymolwiki.org/index.php/InterfaceResidues.
175. *The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.* 2018; Available from: https://www.schrodinger.com/suites/pymol.
176. Higgins, D. and J. Dworkin, *Recent progress in Bacillus subtilis sporulation.* FEMS Microbiol Rev, 2012. **36**(1): p. 131-48.
177. Castilla-Llorente, V., et al., *Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication*, in *EMBO J*. 2006. p. 3890-9.
178. Wagner, J.K., K.A. Marquis, and D.Z. Rudner, *SirA enforces diploidy by inhibiting the replication initiator DnaA during spore formation in Bacillus subtilis.* Mol Microbiol, 2009. **73**(5): p. 963-74.
179. Zawilak-Pawlik, A., et al., *DiaA/HobA and DnaA: a pair of proteins co-evolved to cooperate during bacterial orisome assembly.* J Mol Biol, 2011. **408**(2): p. 238-51.