Role of the bacterial nucleoid associated protein “HU” in cell-cell and cell-DNA interactions through the binding of HU with eDNA and Lipopolysaccharide

Bhishem

A Thesis submitted for the partial fulfilment

of the degree of

Doctor of Philosophy

Indian Institute of Science Education and Research Mohali

Department of Biological Sciences

Mohali-140306

September, 2019

Dedicated to my family and supervisor...

Declaration

The work presented in this thesis has been carried out by me under the guidance of Professor Purnananda Guptasarma at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

(Bhishem)

In my capacity as supervisor of the candidate's thesis work we certify that above statements made by candidate are true of our knowledge.

(Professor Purnananda Guptasarma)

Supervisor

ACKNOWLEDGEMENTS

"The influence of teachers extends beyond the classroom, well into the future"— F. Sionil Jose

As stated, firstly I would like to express my sincere gratitude to my Supervisor Prof. Purnananda Guptasarma for his constant support, positive attitude, providing free space for exploring new ideas, sharing scientific and spiritual wisdom with me. His immense scientific knowledge and out of the box thinking ability inculcated the ability to imagining scientific problems and results interpretation. I felt free to explore science and he always encouraged me in my exploration. His ability of troubleshooting problem is evident by this incidence when I was stuck with my fluorescence anisotropy experiments and after repeating many times failed to get the results, he simply calculated molecules and asked me to replicate the same. Any scientific discussion with him was a pleasurable learning experience. I was a nice time to discuss science with him. He was the person who taught me the real meaning of science. He enabled me with qualities to make my future professionally and personally better.

Secondly, I would like to thank my wife Shipra Chandel for being a constant driving force in my Ph.D. I would like to thank my parents for their love and support.

I would like to thank my doctoral committee members Dr. Sharavan Kumar Mishra and Dr. Rachna Chabba for sparing time from their busy schedule for conducting annual presentations and giving scientific suggestions.

Apart from sir and family members, many people were associated with to make my Ph.D Journey joyful. Firstly I would like to thank Dr. Kanika Arora, for teaching me handling and experiment designing, who was my senior after joining the lab. I would like to thank Angel D Samual and Archit Gupta for helping me with my experiments. Dr.Prince, Dr.Pallavi, Dr.Nitin, and Dr.Sukhdep were amazing seniors and It was nice to discuss science with them. Apart from past lab members, I would like to thank amazing present lab members, Archit Gupta, Arpita Mrigwani, Dr.Neeraj Dhaunta, Arpita Sarkar for providing positive and joyful lab environment. Their helping nature, whether it be correcting presentation, thesis, helping with experiments and consoling in difficult time was quite commendable. It was nice to spend time with new Ph.D Harman Kaur and Snehal Waghmare.

Apart ftom lab, I had many friends especially K. Abhijeet "Sultaan" IISER Mohali BS-MS student, who was a constant source of inspiration. I would remember road trips with him on weekends. I would like to thank Nagesh Kadam, Mohinder Pal, Bhupinder Singh, Mayank

Sarasvat, Dr.Vadla Rajkumar, Dr.Dominic Narang for their caring nature and for guiding me throughout my Ph.D.

I would like to thank Dr. Shantanu Pal and Ipshita Pani for providing their facility of Liquid crystal for interaction studies and Namrata Ramasakha from Dr.Samarjit Bhatacharya lab for teaching me confocal.

Above all, I owe eveerything to almighty for providing me wisdom, good teacher, family and friends.

List of Publications

Dhamija S, **Thakur B**, Guptasarma P, De AK. *Probing the excited state dynamics of Venus: origin of dual-emission in fluorescent proteins.* Faraday Discuss, 2018. **207**(0): p. 39-54.

Synopsis for whole thesis

Introduction

DNA within the nucleoid of the *E.coli* bacterium is associated with 12 nucleoid associated proteins (NAPs). These help in compacting 4.6 MB worth of DNA inside the small volume of the bacterium's cytoplasm. Of these 12 NAPs, 2 NAPs are classified under the DNABII family of DNA binding proteins, and they are known as HU and IHF. HU shares a common structural fold with IHF (integration host factor) and also with another non-NAP DNA-binding protein known as TF-1 (an HU like protein from bacillus subtilis bacteriophage SPO1).

HU was first identified in the *Escherichia coli* strain U93 in 1975 [1]. It is a small dimeric protein. In *E.coli,* HU is composed of two polypeptides of molecular weights 9535 Da (HU-A) and 9225 Da (HU-B), sometimes abbreviated as HU-2 and HU-1, respectively, and encoded by genes hupA, and hupB [2]. Western blot analysis conducted with antibodies against a mixture of the two HU subunits indicates that about 30,000 to 55,000 dimers of HU can exist inside a single *E. coli* cell during the bacterium's exponential growth phase, causing HU to be its major NAP and also one of its most abundant intracellular proteins. Besides *E.coli* and other members of the family *Enterobacteriaceae, Vibrionaceae* also have two homologous HU polypeptides, whereas most other bacteria contain only a single gene encoding HU, and a single HU polypeptide [3].

The expression of the hupB gene is regulated by three promoters, whereas the hupA gene is under the control of a single promoter, and the hupB and hupA genes are located at positions of 90 minutes and 10 minutes, respectively, on the *E.coli* chromosome [4, 5]. The two subunits of HU share about 70 % overall amino acid identity [6]. Expression of HU-B from its different promoters varies during the growth phase of the bacterial cell cycle, and this is associated with the accumulation of different homo and heterodimeric HU populations in a bacterial cell, i.e., HU-A2, HU-B2 and HU-AB [4]. The homodimer, HU-A2, is present during the lag phase and early exponential phase, and this is followed by a transient increase in the presence of HU-B2 homodimers in the mid exponential phase, with the heterodimer, HU-AB replacing HU-A2 and

HU-B2 homodimers in later stages of the cell cycle [7]. At the level of transcription, the formation of HU is regulated by proteins known as FIS, and CRP NAPs [7]. Double mutants of *E.coli* lacking both HU-A and HU-B show growth defects, but remain viable, apparently on account of the presence of other NAPs[8].

Till date, the crystal structures of HU from various organisms have been solved (e.g., 1P51, from *Anabaena*, in complex with a stretch of synthetic DNA, and 1MUL from *E.coli* in the absence of DNA). In all of these organisms, HU shares a common fold. As shown in Figure 1, the N-terminal region is made up of two helices joined by a turn, and the C-terminal region is made up of three anti-parallel beta strands (in red), with a beta finger extension (in magenta), with this last sub-structure being mainly involved in the binding of DNA. The protein ends with a C-terminal alpha helix. The beta hairpin extension is unstructured in the absence of DNA [9]; however, in the presence of DNA it is seen that the two hairpins from the two subunits of a homodimer wrap around the DNA in different directions, clasping it from all around the minor groove, with some nominal contact with the major groove. A conserved proline residue at position 63 (Pro63) on the tip of each of the beta hairpin loops inserts itself into the DNA. The beta hairpin is rich in positively-charged amino acids and these stabilize its interactions with the negatively-charged phosphate groups on the DNA backbone.

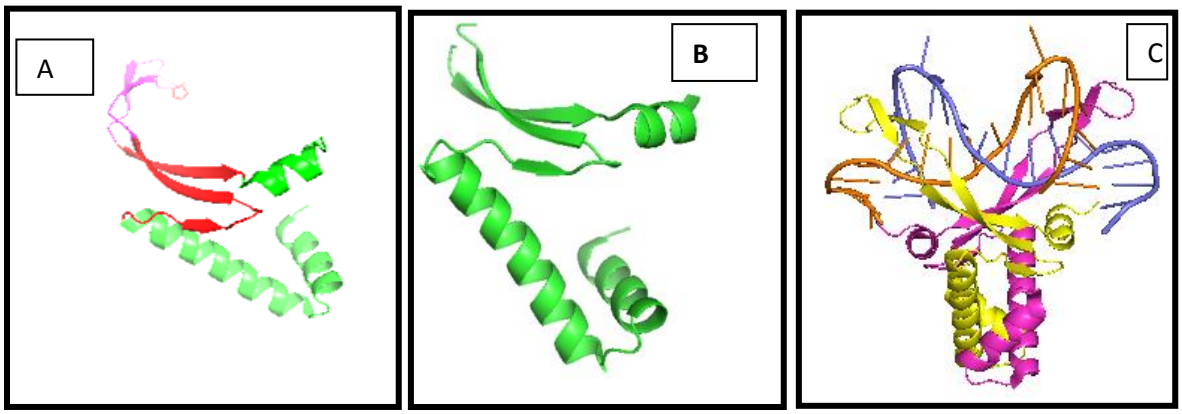


Figure 1: Panel A: Anabaene HU monomer (PDB 1P51) showing a alpha helical N-terminal core and a network of three antiparallel beta strands along with a beta hair pin loops shown as pink. ***PanelB:*** The structure of *E.coli* HU-A with invisible beta hair pin loop in absence of DNA. ***PanelC:*** The structure of Anabaena HU in complex with DNA.

Apart from this binding mode of HU with DNA, recently it has been shown that HU also binds DNA through certain conserved lysine residues K3, K83, present on the side of each HU dimers, which bind to DNA in cooperation with a phenylalanine residue, Phe47 [10]. The binding of HU through its sides has been shown in the Figure 2. The binding of DNA by HU through the beta hairpin will be called binding involving the 'canonical' binding site, and the binding through the lysine and phenylalanine residues will be called binding through the 'non-canonical' site.

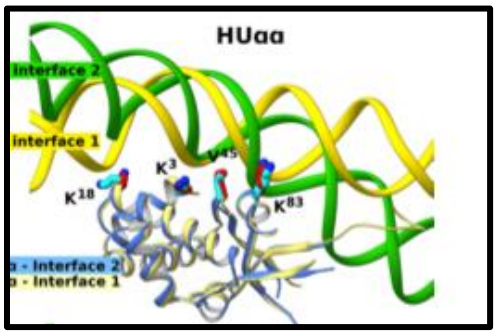


Figure 2: Shows the bonding of HU-A homodimers with DNA through lysine mediated electrostatic interactions.

Apart from their association with bacterial nucleoids, DNABII family proteins have also been shown to be a part of the extracellular matrix, in association with eDNA (extracellular DNA) which is a major component of the bacterial extracellular matrix in biofilms, along with polysaccharides and proteins. The importance of DNA and DNABII proteins in the extracellular matrix is seen in recent papers, in which DNAase treatment of bacterial biofilms and the use of antibodies against DNA binding tips of the DNABII family proteins (IHF and HU) have been shown to disrupt biofilms and release bacteria [11, 12]. Also, the exogenous addition of DNABII proteins promotes biofilm formation [13]. The source of this eDNA and the protein associated with it is either explosive cell lysis or release of DNABII proteins through secretion, e.g., involving the type IV secretion system [14, 15]. Cells with deletion mutants of type IV secretion components, traGC and ComE, fail to secrete DNA extracellularly [15].

The cytoplasm of Gram-negative bacteria is surrounded by a phospholipid inner membrane and a lipopolysaccharide outer membrane. The outer membrane is composed of phospholipids in its

inner leaflet and predominantly lipopolysaccharide (or LPS) in its outer leaflet. The outer membrane is traversed by many beta barrel integral proteins known as porins and ion channels [16]. LPS is crucial for the structural integrity and function of gram negative bacteria [17]. In *E.coli* and *Salmonella* serovar *typhimurium* the overall ratio of LPS to phospholipid is 0.15:1.00 with ~$2x10^6$ LPS molecules present on the surface of each cell, imparting it an overall negative charge.

The embedment of the LPS-displaying and, hence, negatively-charged, bacterial cells in the extracellular matrix made up of eDNA in biofilms raised the question of how a negatively-charged bacterial surface is compatible with the negative charges on the surface of DNA, The overall neutralization of the negative charge in the extracellular matrix is a prerequisite for biofilm stability [18]. Since DNABII proteins and, especially HU, have been shown to associate with eDNA in the extracellular matrix, we decided to examine whether HU can play an important role as a charge neutralizer on eDNA. For the entrapment of bacterial cells in eDNA associated with DNABII protein, the DNABII proteins needs to be interact with the bacterial cell surface as well. As shown earlier, HU a member of DNABII class of proteins is present in bacterial biofilms and can potentially interact with DNA through its multiple (canonical and non-canonical) surfaces, as shown by the studies with HU from Adhya group.

The availability of more than one surface for binding to negatively-charged DNA makes HU a potential candidate for additionally acting as a charge neutralized and a glue for holding bacterium-to-bacterium and bacterium-to-DNA. In this thesis, I have explored this possibility fully. I present extensive biophysical, biochemical and cell-biological evidence involving studies with both molecules, and cells, to show that HU binds to LPS, that HU can coagulate bacterial cells, and also that HU can act as a charge-neutralizer and glue binding bacteria to DNA in biofilms. I also present evidence that both the canonical and non-canonical sites on HU engage in binding of both (or either) of DNA and/or LPS. I support the experimental evidence with the argument that LPS contains a sugar-phosphate arrangement involving 6-carbon sugars which is similar in some ways to that of DNA, causing HU to recognize either DNA or LPS using either of its sites. I argue that this, together with the sheer abundance of HU, distinguishesit from other proteins carrying positive charge as a glue working to stabilize biofilms.

Further, HU has been show to form different oligomeric states depending upon the growth phase of bacterial cell, and HU-A as well as HU-B are reported to form not just dimers but also tetramers, and HU-B is even known to form homomeric octamers. These subunit-subunit associations of HU are worthy of exploration, not just because they might facilitate biofilm formation, but also because there is the question of whether subunit exchange occurs between the subunits or whether subunits associated once remain associated until they undergo turnover. The genes encoding HU-A and HU-B are located at 90 and 10 minutes, respectively, on the bacterial chromosome, and their differential expression during bacterial life cycles and this makes it difficult to understand how HU-A and HU-B subunits can associate into heterodimers through subunit exchange, in a protein (the HU homodimer) in which the subunits interact so intimately that there appears to be no scope for the existence of independently folded monomeric subunits, i.e., any subunit exchange must involve substantial unfolding of subunits along with dissociation, prior to reassociation during subunit exchange. It has been shown that the mixing of HU-A and HU-B homodimers in 1:1 proportion for 30 minutes at 27 ^{o}C temperature results in 100 % formation of a population of heterodimers [4]; although it is not clear how this can come to be, unless the dissociation constants of the two are compatible. An alternative possibility is that two homodimers or HU-A and HU-B first associate into a hetero-tetramer (consisting of HU-A and HU-B homodimers) and that this is followed by dissociation and reassociation of subunits in a partially folded, but still associated state, obviating the need for subunits to fully unfold during dissociation.

To probe these issues regarding the scope for subunit exchange, I have examined the stability of the dimeric interface of both HU-A and HU-B homodimers, using chemical (urea) denaturation combined with glutaraldehyde cross-linking. The urea treated samples of HU were also passed through a Superdex-75 column to check the effect of urea on the protein's hydrodynamic volume. These studies showed that the interface of both the HUs is stable enough to survive substantial unfolding of the HU subunits, with dimer dissociation being preceded by substantial unfolding of HU subunits. The heterodimer formation *in vivo* was studied by making co-trasformant cells containg one tagged and one untagged HU and pulling-down untagged HU through affinity purification by tagged HU. The *in-vivo* studies showed the presence of heteromeric populations; however, the question remains regarding whether these are hetermeric associations of homodimers, or heterodimers.

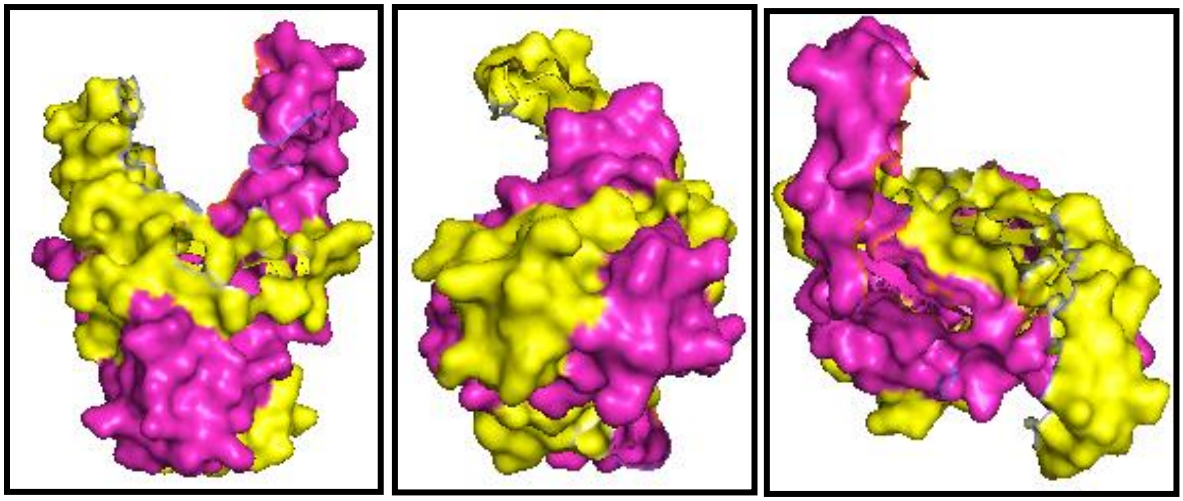


Figure 3: The surface representation of Anabaena HU showing extensive interaction between two subunits.

The third section of the thesis deals with the creation of a novel monomeric DNA binding protein derived from HU, with specificity towards dsDNA. HU has been shown to unfold via two unfolding intermediate states where the unfolding of the alpha helical core precedes the breaking of the hydrophobic interaction, leading to dissociation of partially unfolded HU monomers. The antiparallel beta strands are stabilized by an extensive network of hydrophobic interactions amongst phenylalanine residues as shown in the Figure 4. The interaction of HU with DNA is canonically mediated by two unstructured beta hair pin loops and four antiparallel beta strands. As HU unfolds via two state unfolding, we thought of making an artificial DNA binding protein containing only the beta hair pin loops and four antiparallel beta strands, lacking any alpha helical core. The protein was made. It was found to be extremely thermostable and showed preference for dsDNA over ssDNA.

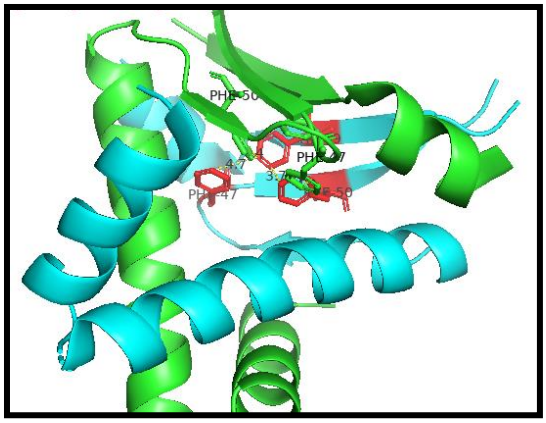


Figure 4: Showing extensive network of phenylalanine residues in antiparallel beta strands.

References:

1. Rouviere-Yaniv, J. and F. Gros, *Characterization of a novel, low-molecular-weight DNA-binding protein from Escherichia coli.* Proc Natl Acad Sci U S A, 1975. **72**(9): p. 3428-32.
2. Pettijohn, D.E., *Histone-like proteins and bacterial chromosome structure.* J Biol Chem, 1988. **263**(26): p. 12793-6.
3. Oberto, J., K. Drlica, and J. Rouviere-Yaniv, *Histones, HMG, HU, IHF: Meme combat.* Biochimie, 1994. **76**(10-11): p. 901-8.
4. Claret, L. and J. Rouviere-Yaniv, *Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival.* J Mol Biol, 1997. **273**(1): p. 93-104.
5. Kano, Y., et al., *Molecular cloning and nucleotide sequence of the HU-1 gene of Escherichia coli.* Mol Gen Genet, 1985. **201**(2): p. 360-2.
6. Dillon, S.C. and C.J. Dorman, *Bacterial nucleoid-associated proteins, nucleoid structure and gene expression.* Nat Rev Microbiol, 2010. **8**(3): p. 185-95.
7. Claret, L. and J. Rouviere-Yaniv, *Regulation of HU alpha and HU beta by CRP and FIS in Escherichia coli.* J Mol Biol, 1996. **263**(2): p. 126-39.
8. Huisman, O., et al., *Multiple defects in Escherichia coli mutants lacking HU protein.* J Bacteriol, 1989. **171**(7): p. 3704-12.
9. Boelens, R., et al., *Structure and dynamics of the DNA binding protein HU from Bacillus stearothermophilus by NMR spectroscopy.* Biopolymers, 1996. **40**(5): p. 553-9.
10. Hammel, M., et al., *HU multimerization shift controls nucleoid compaction.* Sci Adv, 2016. **2**(7): p. e1600650.
11. Novotny, L.A., et al., *Monoclonal antibodies against DNA-binding tips of DNABII proteins disrupt biofilms in vitro and induce bacterial clearance in vivo.* EBioMedicine, 2016. **10**: p. 33-44.
12. Tetz, G.V., N.K. Artemenko, and V.V. Tetz, *Effect of DNase and antibiotics on biofilm characteristics.* Antimicrob Agents Chemother, 2009. **53**(3): p. 1204-9.
13. Devaraj, A., et al., *DNABII proteins play a central role in UPEC biofilm structure.* Mol Microbiol, 2015. **96**(6): p. 1119-35.
14. Turnbull, L., et al., *Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms.* Nat Commun, 2016. **7**: p. 11220.

15. Jurcisek, J.A., et al., *Nontypeable Haemophilus influenzae releases DNA and DNABII proteins via a T4SS-like complex and ComE of the type IV pilus machinery.* Proc Natl Acad Sci U S A, 2017. **114**(32): p. E6632-E6641.
16. Jiang, X., et al., *Ligand-specific opening of a gated-porin channel in the outer membrane of living bacteria.* Science, 1997. **276**(5316): p. 1261-4.
17. Erridge, C., E. Bennett-Guerrero, and I.R. Poxton, *Structure and function of lipopolysaccharides.* Microbes Infect, 2002. **4**(8): p. 837-51.
18. Dengler, V., et al., *An Electrostatic Net Model for the Role of Extracellular DNA in Biofilm Formation by Staphylococcus aureus.* J Bacteriol, 2015. **197**(24): p. 3779-87.

Abbreviations

%	Percent
mg	Milligram
M	Molar
ml	Milliliter
min	Minute
nM	nanomolar
mM	millimolar
µM	micromolar
KDa	Kilodalton
CD	Circular dichroism
PCR	Polymerase chain reaction
4WJ-DNA	Four way junction DNA
EMSA	Electrophorectic mobility shift assay

Contents

CHAPTER 1

Introduction

Contents

Introduction

1. DNABII proteins

The genetic information is stored in DNA in all life forms like Eukarya, Bacteria and Archaea, regardless of the differences in their cellular organization and other features. In all organisms, DNA is required to be stored and organized in such a way that it is compatible with processes like DNA replication, transcription and segregation into daughter chromosomes. The DNA of eukaryotes is organised by proteins called histones. A similar function is served in prokaryotes by a group of proteins categorized under the term, NAPs (nucleoid-associated proteins). In *E.coli*, till date, 12 NAPs have been discovered, and these help DNA to become compact so that 4.6MB of the bacterium's chromosomal DNA, which is over 1 mm in length if it is stretched out, can be fitted inside a bacterial cell whose longest dimension is only about 1µm. Of these 12 NAPs, a few are grouped under a class of proteins called the DNABII family of proteins, because they share common protein folds and similar DNA binding modes.

The DNABII family of proteins includes HU (Histone-like protein HU), IHF (integration host factor), and TF1, an HU-like protein from bacillus subtilis bacteriophage SPO1 [1]. The bacteriophage SPO1 genome encodes a basic DNA-binding small transcription factor 1, named 'TF1'. TF1 takes advantage of the substitution of all thymine residues with 5-hydroxymethyluracil (hmU), which is a characteristic of all SPO1 DNA, by binding specifically and with high affinity only to DNA containing this modification [2]. TF1 selectively inhibits the transcription of DNA containing 5-hydroxymethyluracil by RNA polymerase) [3]. Unlike HU, the other two DNABII family proteins, IHF and TF1, bind to specific DNA sequences, and they differ mainly in the length of DNA used by them for binding, which is 9 base pairs for HU and 34 base pairs for IHF and TF1 [4,5].

The members of DNABII proteins have a core which is stabilized by hydrophobic interaction(s) between two helices from each monomer and a set of three anti-parallel beta strands. All the members of this family share the existence of a common motif, a 'beta hairpin' extension from each monomer, which has been implicated in DNA binding. IHF and HU have been shown to facilitate DNA replication at the chromosomal origin of replication, and also play roles in recombination and transposition of genes [6]. DNABII family proteins are dimeric, and in *E. coli*

HU can be a homodimer of HU-A or HU-B, or a heterodimer of HU-A and HU-B, whereas in bacteria containing a single HU, of course, HU is always a homodimer; in contrast, IHF is always a heterodimer, and TF-I is always a homodimer [7,8]. In the *E. coli* HU-A and HU-B proteins, which are made from separate genes, there are lysine residue at positions 3 (Lys3) and 86 (Lys86) and these are conserved among all DNABII family members. It has been shown that Lys3 in the case of IHF and TF-1 interacts with the negatively charged phosphate backbone of DNA at a site distal to the site of 9 base pairs in which HU binds to DNA and causes kinking, and it is this difference (i.e., in HU, Lys3 binds to DNA within the 9 base pair site site) which results in the formation of a longer DNA-binding site of 37 base pairs (because Lys3 binds DNA at sites separated by over twenty base-pairs.

The other members of the family of DNABII proteins differ from the 90 amino acids-long HU in that they have a C-terminal extension which is used for DNA binding, whereas this extension is absent in HU [9,10]. In HU, the positive charge present on Lys3 forms a salt bridge with the aspartic acid residue at position 26 (Asp26), resulting in a shorter binding-site of 9 base pairs. Most of the NAPs affect DNA compaction by altering the trajectory of the DNA backbone, by wrapping, bridging and bending DNA. The expression of nucleoid-associated proteins varies with the growth phase of bacterial cells [2,24]. The list of all nucleoid associated proteins is shown in the Table shown below as Figure 1.

Name of the nucleoid associated protein	**Role and the mode of DNA binding**
HU (Histone like DNA binding protein)	HU can exist as homodimers and heterodimers depending upon bacterial growth phase [11]. HU been shown to bend DNA by wrapping it around itself.
IHF (Integration host factor)	IHF is related to HU in amino acid sequence but shows preferences for AT sequences [1].
Dps (Protein associated with bacterial starvation)	This protein is expressed during bacterial starvation and its expression is regulated by IHF, Fis and H-NS [12].
Lrp (Leucine responsive regulatory protein)	It can exist in diverse oligomeric states as

	dimer, octamer and hexadecamer. It binds to DNA on degenerate consensus sequences and alters DNA trajectory [13] .
DnaA (DNA binding protein A)	This protein is associated with DNA replication at the bacterial chromosomal origin of replication.
CpbA/B (curved DNA binding protein)	Affects bacterial growth at lower and high temperature. Its transcription is under the control of Lrp and is regulated by Leucine and RpoS [14].
Dan (DNA binding protein under anaerobic condition)	As the name suggests, the protein binds to DNA under anaerobic conditions.
StpA (suppressor of td2 phenotype A)	This protein is similar to H-NS but has been found to show preferences for RNA. It can form heterodimers with H-NS [15].
Fis (Factor for inversion stimulation)	It exists as a homodimer and binds to a 17 base pair AT rich DNA sequence, except at positions 2 and 16, where G and C residues are commonly present. This protein has been implicated in many DNA based cellular activities like transcription, replication and recombination [16].
H-NS	It exists as homodimers and heterodimers, It forms DNA–protein-DNA bridges and helps in constraining negative supercoils.
MukB	Associated with bacterial nucleoid partitioning during cell divison.It has role in the formation of separate topological domains in nucleoids, in association with DNA gyrase [17].
EbfC	It is a homodimeric DNA binding protein

	which binds and bends DNA at specific sequences [18].
Crp	Dimeric DNA binding protein which acts as activator or suppressor of transcription by binding to conserved nucleotide sequences. Cylic AMP (C-AMP) acts as co-factor for this protein [19].
SPO1 TF-1	DNA binding protein isolated from bacillus subtilis lytic SPO1 bacteriophage. This protein is classified under NAP for sharing common folds with HU and IHF.

Figure 1. Table with the list of all DNA binding proteins in bacteria.

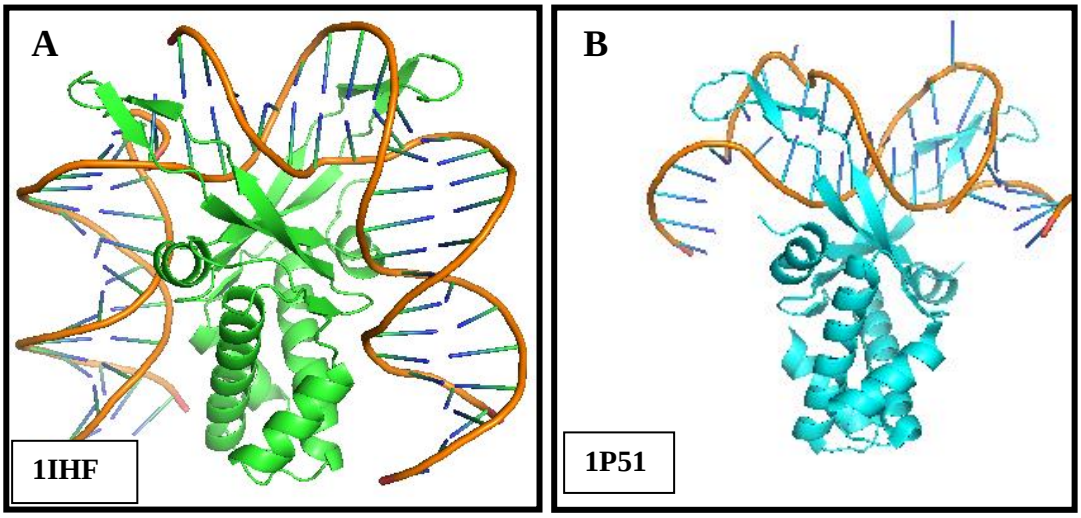


Figure 2: Panel A, B: Showing the tertiary structure and DNA binding mode of two DNA binding proteins IHF (integration host factor with a PDB ID of 1IHF) and HU (Histone like protein HU with a PDB ID 1P51). The conservation of overall three dimensional structure and the similarity of geometry of association with DNA is evident from the image above.

The protein sequences of the members of DNABII family

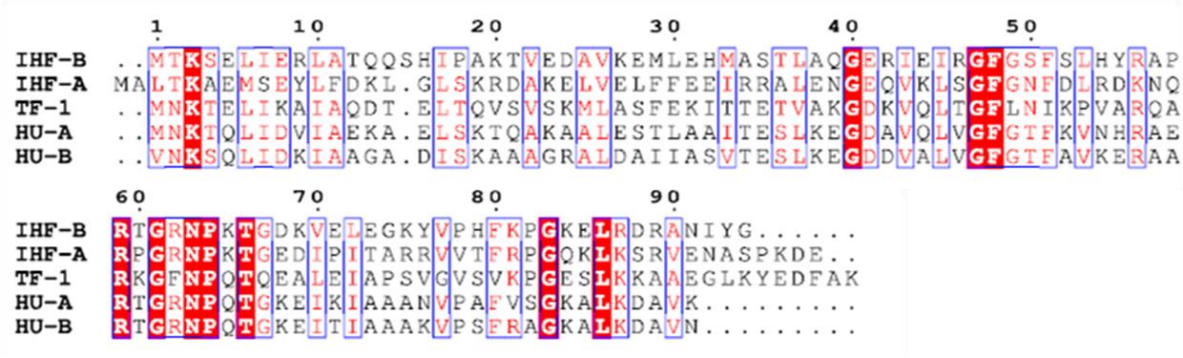
HUA:MNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHRAERTG
RNPQTGKEIKIAAANVPAFVSGKALKDAVK

HUB:VNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGFGTFAVKERAARTGR
NPQTGKEITIAAAKVPSFRAGKALKDAVN

IHFA:MALTKAEMSEYLFDKLGLSKRDAKELVELFFEEIRRALENGEQVKLSGFGNFDLRDKNQR
PGRNPKTGEDIPITARRVVTFRPGQKLKSRVENASPKDE

IHFB:MTKSELIERLATQQSHIPAKTVEDAVKEMLEHMASTLAQGERIEIRGFGSFSLHYRAPRTG
RNPKTGDKVELEGKYVPHFKPGKELRDRANIYG

TF1:MNKTELIKAIAQDTELTQVSVSKMLASFEKITTETVAKGDKVQLTGFLNIKPVARQARKGF
NPQTQEALEIAPSVGVSVKPGESLKKAAEGLKYEDFAK



1.1. HU protein

HU was first identified in the Escherichia coli strain U93 in 1975[20]. HU is a small dimeric protein. In *E.coli,* it is composed of subunits of molecular weight 9535 Da and 9225 Da, named HU-A and HU-B, respectively (sometimes abbreviated as HU-2 and HU-1, respectively, or HUP-A and HUP-B) which are encoded by genes hupA and hupB, respectively[21]. Western blot analysis conducted with antibodies against a mixture of the two HU subunits indicates that about 30,000 to 55,000 dimers of HU can exist inside an *E. coli* cell during the exponential growth phase of *E. coli*. It is due to this that HU is called the major nucleoid associated protein, and it is possible that it is the most abundant protein in *E. coli*. Besides *E.coli,* two homologs of HU exist also in other members of *Enterobaceriaceae* and *Vibrionaceae* families. In most other bacteria, however, HU is encoded by single gene [22].The expression of the hupB gene is regulated by three promoters, whereas the hupA gene is under the control of a single promoter,

and the hupB and hupA genes are located at positions of 90 minutes and 10 minutes, respectively, on the *E.coli* chromosome [11,23]. The two subunits of HU share about 70% amino acid similarity [24]. Expression of HUB from different promoters varies with the growth phase of the bacterial cell cycle, and this leads to accumulation of different homo and heterodimeric HU populations in a bacterial cell [11]. The homodimer, HUA2, is present during the lag phase and early exponential phase, and this is followed by a transient increase in the presence of HUB2 homodimers in the mid exponential phase, with the heterodimer, HUAB, replacing HU homodimers in later stages of the cell cycle [25]. At the level of transcription, the formation of HU is regulated by proteins known as FIS, and CRP NAPs [25]. Double mutants of *E.coli* lacking both HU-A and HU-B show growth defects, but remain viable, apparently on account of the presence of other NAPs[26].

Till date, the crystal structures of HU from various organisms have been solved (e.g., 1P51, from *Anabaena*, in complex with a stretch of synthetic DNA, and1MUL from *E.coli* in the absence of DNA). In all of these organisms, HU shares a common fold. As shown in Figure 3, the N-terminal region is made up of two helices joined by a turn, and the C-terminal region is made up of three anti-parallel beta strands (in red), with a beta finger extension (in magenta), with this last sub-structure being mainly involved in the binding of DNA. The protein ends with a C-terminal alpha helix. The beta hairpin extension is unstructured in the absence of DNA [27]; however, in the presence of DNA it is seen that the two hairpins from the two subunits of a homodimer wrap around the DNA in different directions, clasping it from all around the minor groove, with some nominal contact with the major groove. A conserved proline residue at position 63 (Pro63), located on the tips of each of the beta hairpin loops, inserts itself into the DNA. The beta hairpin is rich in positively-charged amino acids and these stabilize its interactions with the negatively-charged phosphate groups on the DNA backbone.

HU has been shown to form higher oligomers of different geometry which can wrap DNA around itself [28]. The effect of HU upon DNA topology is concentration-dependent. At lower concentrations, it bends linear DNA by binding to it and forming a kink, but at higher concentrations it keeps DNA in an extended form by binding to it at very high density [29]. The minimum binding length of HU for linear DNA is 9 base pairs in the case of *E.coli* HU whereas for *Thermotoga maritime* HU, the DNA binding site is more than 35 base pairs in length [30].

The binding of HU with DNA introduces a kink whose angle varies from the 109-135^{o} as has been reported from studies of DNA bending involving HU sourced from different organisms. The minimum binding length and affinity of HU for different DNA substrates apparently depends upon the number of basic amino acid on HU's surface which stabilize HU-DNA interactions through salt bridge formation. HU can bind with high nanomolar (nM) affinity to naturally-kinked DNA possessing mismatches and single strand breaks, nicked DNA, negatively super coiled DNA, cruciform DNA, as well as RNA folded into secondary structures [31-34].

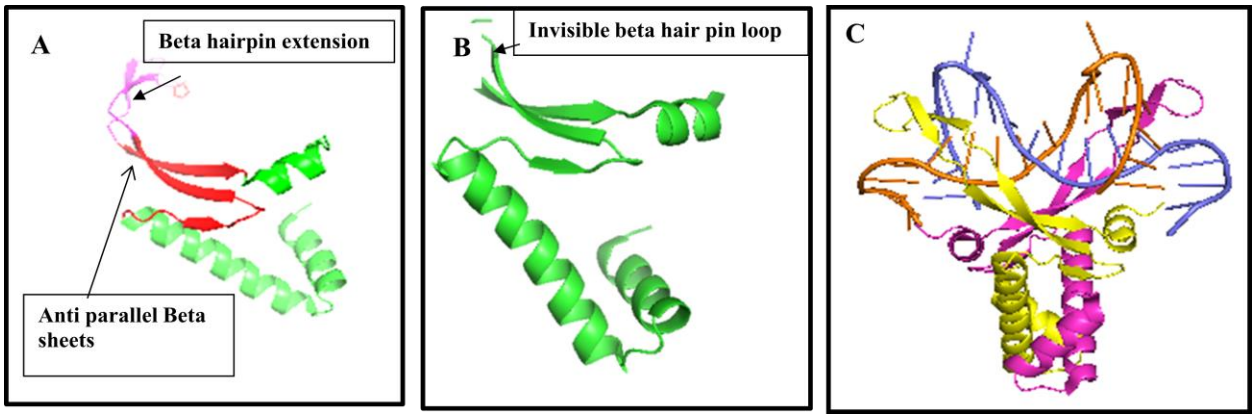


Figure 3: Panel A: Showing the structure of Anabaena HU monomer (1P51), after removing the other subunit and associated DNA using protein visualization software PYMOL along with highlighting various structural elements. ***Panel B***: Showing the structure of monomeric *E.coli* HU-A with unstructured loops in absence of DNA (the unstructured loops are commonly present in all DNABII family proteins and they constitute the canonical DNA binding site in HU). ***Panel C:*** Showing the bend introduced in the DNA backbone by the insertion of beta hair loops in minor groove in *Anabaena* HU dimer.

Besides the task of compacting DNA, HU in *E.coli* has been implicated in other cellular processes, like replication-initiation, through some role in the recruitment of DnaA proteins to the chromosomal origin of replication, and also to cell divison, DNA repair, homologous recombination and transcription of numerous genes [35-37]. *E.coli* cells lacking both types of subunits of HU have been shown to be extremely sensitive toward gamma radiation, and the over expression of either HU-A or HU-B protects bacteria from radiation damage[38]. In

Mycobacterium tuberculosis, unlike in *E.coli*, there is only one HU, homologous to *E. coli* HU-B, which has two domains: an N-terminal domain that resembles bacterial HU (~40% identity to *E. coli* HU) and a long distinctive C-terminal domain (CTD) [39]. The long C-terminal domain extension contains several PAKK/KAAK motifs, which are characteristic of the members of the eukaryotic histone H1/H5 protein family. The overall structure of HU is shown in the Figure 3, representing the archetype.

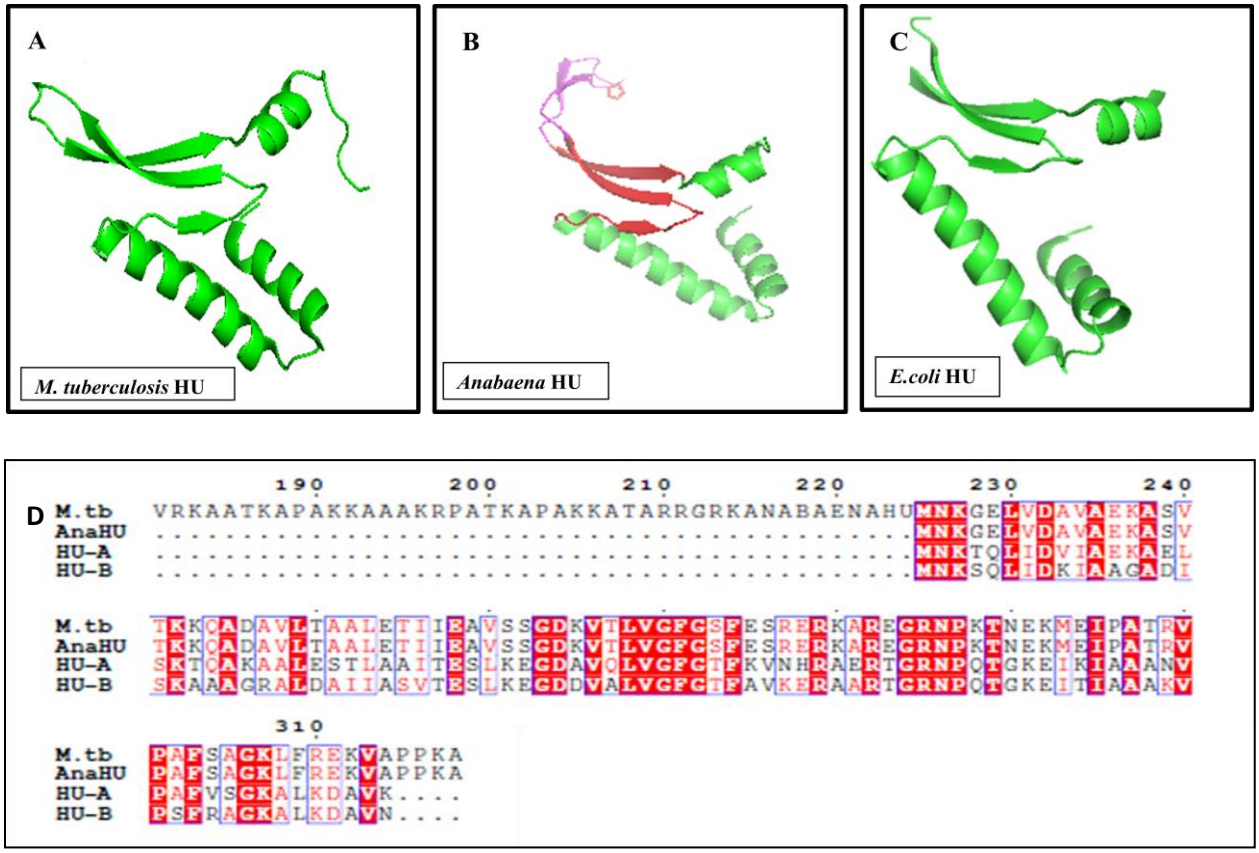


Figure 4: Panel A, B, C: Showing the comparison of HU's tertiary structure in HU from *Mycobacterium tuberculosis*, *Anabaena* and *E.coli*. Apart from the extended C-terminal helix in the case of *M.tuberculosis* HU, the overall structural fold is similar in all HUs. Panel D : Multiple sequence alignment of various HUs, where M.tb is Mycobacterium tuberculosis, HU-A and B are from *E.coli*, and AnaHU is *Anabaena* HU.

HU up-regulates gene expression by the constraining negative supercoils formed by topoisomerase and RNA polymerase diffusing upstream and positive supercoils downstream of transcribing polymerases [40-42]. There is some correlation of HU function in *E.coli* with genes associated with the survival of the bacterium under alkaline conditions, and in *Salmonella enteric* serovar *Typhimurium*, there is some association with the Type 3 secretion system (T3SS) [43]. In contrast to the clustered distribution of another major NAP, H-NS, the distribution of HU is largely scattered throughout the DNA of the nucleoid comprising the *E. coli* chromosome [44].

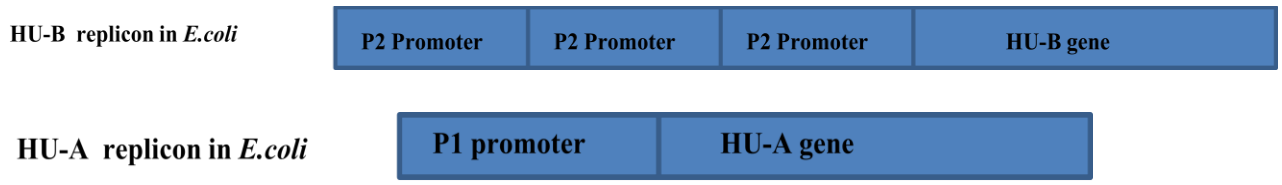


Figure 5: The schematic of replicon for HU-A and B, showing that the latter is made up of three promoters, which are expressed under different growth phases of the bacterial cell cycle.

1.1.1. The structure of HU

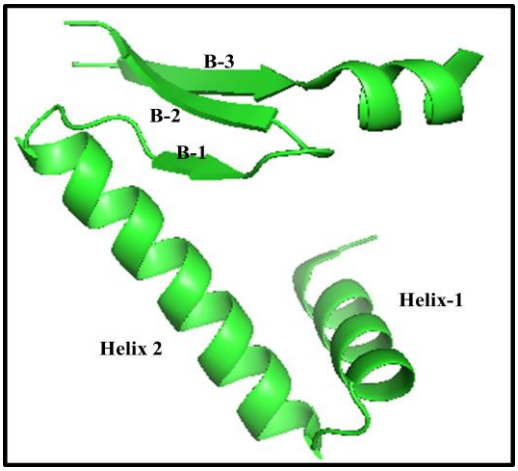


Figure 6: Showing the constituent structural elements of an HU monomer. These structural elements are common amongst all HUs.

The N terminal region of the HU protein is made up of two alpha helices (residues 3-13 and 17-37), shown in Figure 6, as helix-1 and helix-2. These two helices form a big V shaped structure, the base of which is connected by conserved alanine residues at positions 11 (Ala11) and 21 (Ala21). The C terminal of the protein is made up of three anti-parallel beta strands which cover the V formed by the alpha helix, from the DNA-binding side of the protein. The chain terminates with another short alpha helix (helix-3). The three anti-parallel beta strands labelled B1, B2 and B3, cover residue positions 42-45, 48-53 and 76-81. Between strandsB2 and B3 there is a long anti-parallel beta hairpin loop which binds to DNA, in the minor groove, non-specifically with regard to DNA sequence. The strands 1 and 2 are connected by conserved Gly-Phe-Gly residues. The folding of the monomer is governed by core hydrophobic residues Leu6, IIe7, Ile10, Ala11 , Ala21 and Leu25 which intercalate with each other to form a 'V' shape between helix-1 and helix-2. Along with their role in maintaining structure, they interact with the residues of the other subunit (through some of their atoms) thus making the folding and structure of the monomer, and the dimer, interdependent.

It is thus difficult to conceive of HU as a monomer. The residues, Ile32 or Val32, Leu36, Val42, Phe50, Val52, and Pro77, hold helix-2 and the three strands in a given conformation, in both subunits. Most of these residues also participate in dimerization (see below), making the formation of monomers and dimers interdependent [45].

HU has a highly hydrophobic core and hydrophilic surface. The hydrophobic residues at the positions 29,47,50,79,6,32,36, and 44, are conserved amongst HU from different bacteria, and have been shown to form a stable hydrophobic dimeric interface with the other subunit [28]. The interior of the dimer is formed by a number of aromatic rings and aliphatic side chains, originating from Leu6,Ile7, Leu10, Leu16,Leu25, Leu29, Ile32, and Leu36 in HU-A, and Leu6,IIe7,IIe10, Ile16, Leu25,Ile28,Ile29, and Leu36 in HU-B. The residues Ile32, Val42, Phe50,Val52, and Pro77 hold the three beta strands against helix-2 [28]. In HU-A and HU-B there are three conserved phenylalanine residues (Phe47, Phe49 and Phe79), some of which pack against each other as well as against the phenylalanine residues from the other subunit, to form a strong hydrophobic core under the anti-parallel beta strands (beta hairpins) as shown in the figure. The clustering of hydrophobic cores differentially under beta strands and in the region of

the helix motifs of the two subunits, is proposed to result in a three-states unfolding transition of HU [46].

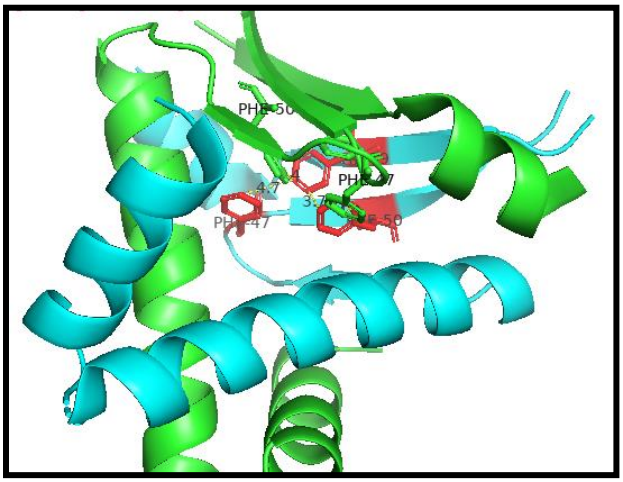


Figure 7: Showing the extensive hydrophobic interactions amongst phenylalanine residues from adjacent anti-parallel beta strands coming from each subunits.

Apart from the conserved phenylalanine in anti-parallel beta strands, the core also contains an intensive network of hydrophobic leucine and isoleucine amino acids, which is shown in the figure below.

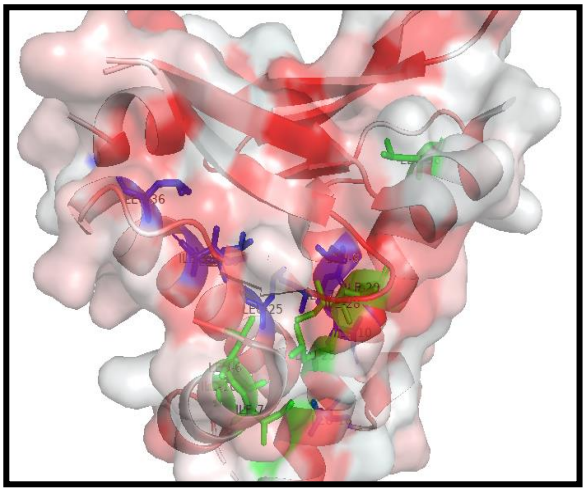


Figure 8: Showing the stabilization of the protein core by hydrophobic amino acids. The core hydrophobic amino acids form intra and inter subunit hydrophobic interaction making folding and dimer stability interdependent.

1.1.2. HU's surface

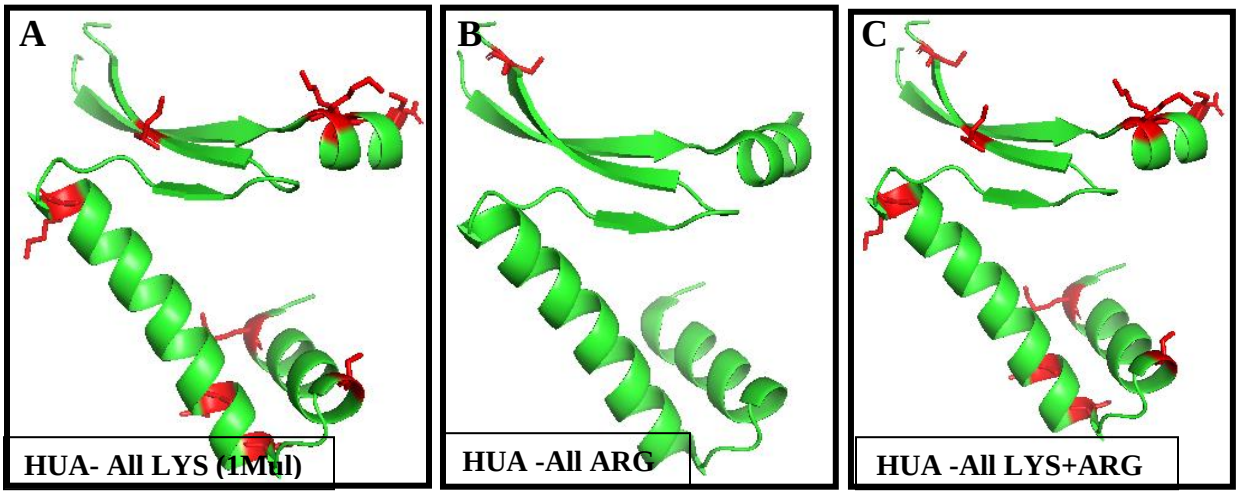


Figure 9: Panel A: Showing in red the position of all lysines in HU-A (1MUL) except the ones present in the beta hairpin loop region (as these loops are not visible in the crystal structure when they are not in complex with DNA). ***Panel B:*** The positions of all arginines and ***Panel C:*** HU with all charged residues.

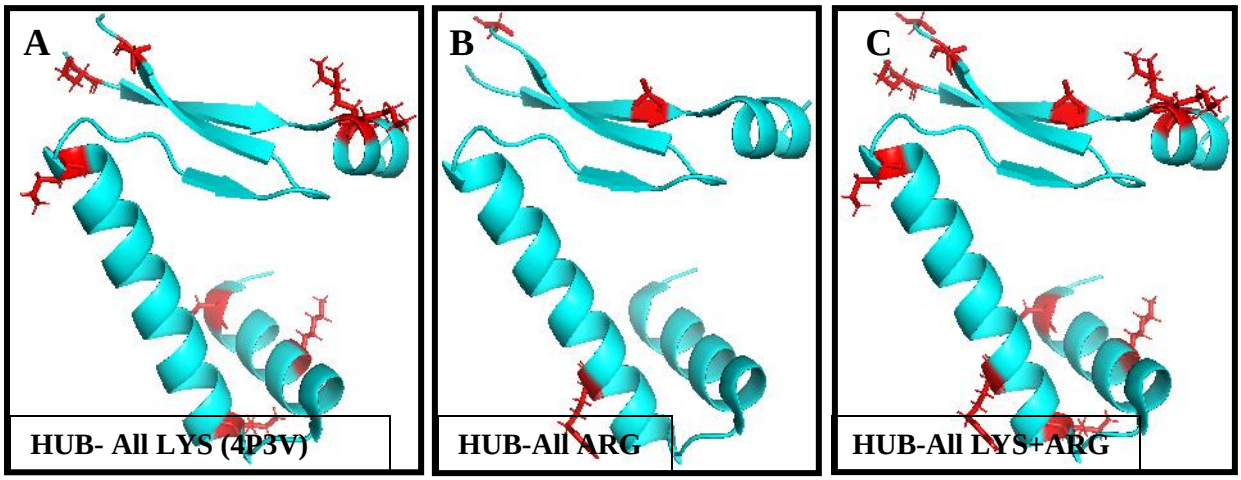


Figure 10: Panel A: Showing the position of all lysines in HU-B (4P3V), other than those in the beta hairpin loops (as in Figure 9). ***Panel B:*** The positions of all arginines. ***Panel C:*** HU with all charged residues.

HU-A Lysines with amino acid positions	HU-B Lysines with amino acid positions
K-3	K-3
K-13	K-9
K-18	K-18
K-22	
K-37	K-37
K-51	K-53
K67	K67
K70	K75
K83	K83
K-86	K86
K90	
HU-A Arginines with amino acid positions	HU-B Arginines with amino acid positions
	R-23
R-55	R-55
R-58	R-58
R-61	R-61
	R-80

Apart from positively charged residues the surface is also rich in Glutamate and Aspartate.

HU-A Aspartic acid residues with amino acid positions	HU-B Aspartic acid residues with amino acid positions
D-8	D-8
	D-15
	D-26
	D-40
D-41	D-41
D-87	D-87

HU-A Glutamic acid residues with amino acid positions	HU-B Glutamic acid residues with amino acid positions
E-12	
E-15	
E-26	
E-34	E-34

E-38	E-38
E-57	E-54
E-68	E-68

Figure 11: Table showing the number and the location of all charged residues from *E.coli* HU-A and B.

The total number of charged residues in HU-A and HU-B are 24 and 24 respectively, in chain lengths of 90 amino acids each. The fewer numbers of one charged type of residue are compensated by the higher numbers of another residue of the same charge, as shown in the table of Figure 10. The beta hairpin loops, which are not visible in any of the crystal structures shown immediately above (because the electron density for them is only visible when they are stabilized in a particular conformation, in complex with DNA) contain a proline residue which intercalates between bases in the minor groove of DNA. This proline residue is conserved amongst the members of the DNABII family. The amino acid sequences of the beta hairpin loops are also conserved between HU-A and HU-B as shown in the sequence marked in red. HU has been shown to bind DNA through beta hairpin loops and anti-parallel beta strands as shown in the Figure 3 Panel C. The role of lysine residues in HU which lie away from the DNA binding region was recently revealed by Adhya and coworkers in a paper in which it was shown that K83, K18 and K3 can interact with linear DNA in a sequence independent manner, through electrostatic interactions. The beta hairpin loops of both HU-A and HU-B are identical, with only one amino acid different (threonine is present in HU-B in place of lysine in HU-A, at position 70, as shown as the red labeled amino acids in both HU-A and HU-B in the sequences below. HU from thermophilic organisms contains even more positively-charged residues, as compared to HU from mesophilic organisms, as is illustrated by considering that HU from *Thermotoga maritima* contains 18 lysine residues and 5 arginine residues, whereas HU-A from *E. coli* contains 11 lysine residues and 3 arginine residues. The DNA intercalating beta hairpin loop also differs, between mesophilic bacteria, in composition and length.

HU-A

MNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHRAERT
GRNPQTGKEIKIAAANVPAFVSGKALKDAVK

HU-B

VNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGFGTFAVKERAART
GRNPQTGKEITIAAAKVPSFRAGKALKDAVN

HU *from Thermotoga maritima*

MTKKELIDRVAKKAGAKKKDVKLILDTILETITEALAKGEKVQIVGFGSFEVRKAAARK
GVNPQTRKPITIPERKVPKFKPGKALKEKVK.

1.2. Subunit exchange in bacterial nucleoid associated protein HU

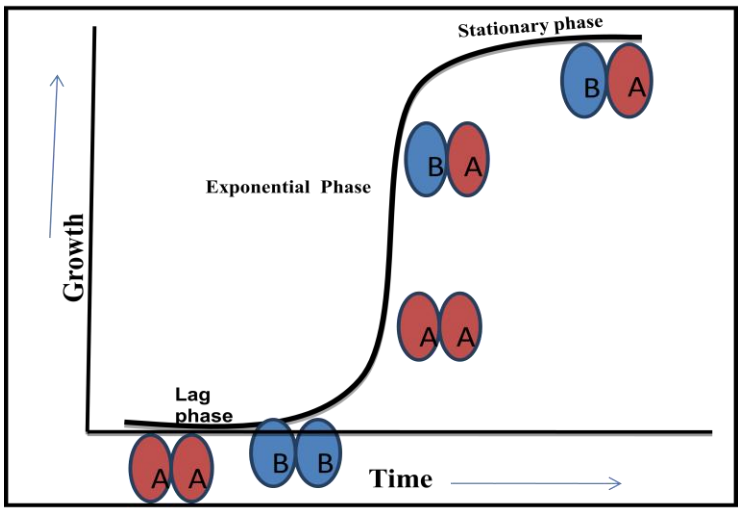


Figure 12 : Showing the distribution of different HU dimers at different times during bacterial growth where the early lag phase is dominated by HU-A and B homodimers, respectively, with the switching of dimers between A and B followed by presence of only heterodimers in stationary phase.

The fact that the genes encoding HU-A and HU-B are located far apart, at positions 10 and 90 minutes on the bacterial chromosome, becomes interesting when one considers that the polypeptides synthesized by them must come into contact with each other during folding, when they are in a partially-folded monomeric form, in order that they are able to form a heterodimer. It may be noted that with homodimer formation, this is not an issue because a ribosome translating the mRNA for one either HU-A or HU-B is likely to produce high local concentrations of the encoded polypeptide in the vicinity of the gene, where transcription and translation take place simultaneously. However, this issue needs some thought and consideration because the native purified form of HU from *E.coli* is thought to be mostly heterodimeric. When purified protein is separated using phosphocellulose chromatography and eluted using NaCl, all HU populations tend to be present, including the two homodimers and the heterodimer; however,the HU heterodimer apparently dominates the population [48], especially in the late exponential and early stationary phase. Polypeptide compositions of different HU dimers (i.e., heterodimers *versus* homodimers) can only be discerned through separations performed on Urea Triton-X 100 gels [48]. The formation of HU-A and HU-B dimers is apparently regulated by the protein 'Fis' which positively regulates HU-A expression. The concentration of Fis is at its highest during the early exponential phase, resulting in maximal expression of the hupA gene with the resultant appearance of HU-A homodimers in the cell, during this phase. The formation of HU-B homodimers, on the other hand, is not highly favoured, as these homodimers seem to be somehow deleterious to the bacterial cell. HU-B2 has been shown to be degraded by Lon protease in HU-A null mutants [49], or converted to HU-AB after association with HU-A homodimers.

Thermal denaturation experiments on HU homodimers have been shown to follow an apparently three-step unfolding process in which the alpha helical core is first unfolded followed by the unfolding of the remaining sections of both polypeptides in the dimers [45]. The stability of the two helices and the turn depicts the stability of individual homodimers as these regions are least conserved in HU dimers. It has been shown that the native conformations of HU homodimers involves a partially unfolded state, with 85% of an intermediate unfolded state present for HU-B and 60% of the comparable unfolded state present for HU-A [45]. The formation of heterodimers

depends upon subunit exchange between homodimers and it has been shown that the mixing of HU-A and HU-B homodimers in 1:1 for 30 minutes at 27°C temperature results in 100% population of heterodimers (although it is not clear how this can come to be, unless the dissociation constants of the two are compatible) [11].The formation of heterodimeric HU inside the bacterial cell is quite controversial. The thermal stability of HU-B is more than that of HU-A (Kanika Arora, personal communication). Both the forms of HU have been shown to refold after denaturation and retain their ability to bind to '4-way junction' DNA (Kanika Arora, personal communication). HU-B has been described as inefficient in DNA binding and it has been observed that it binds to DNA with higher affinity than HU-A. In the structure of HU, the two monomers are tightly packed against each other, making formation of the heterodimer quite controversial and poorly understood.

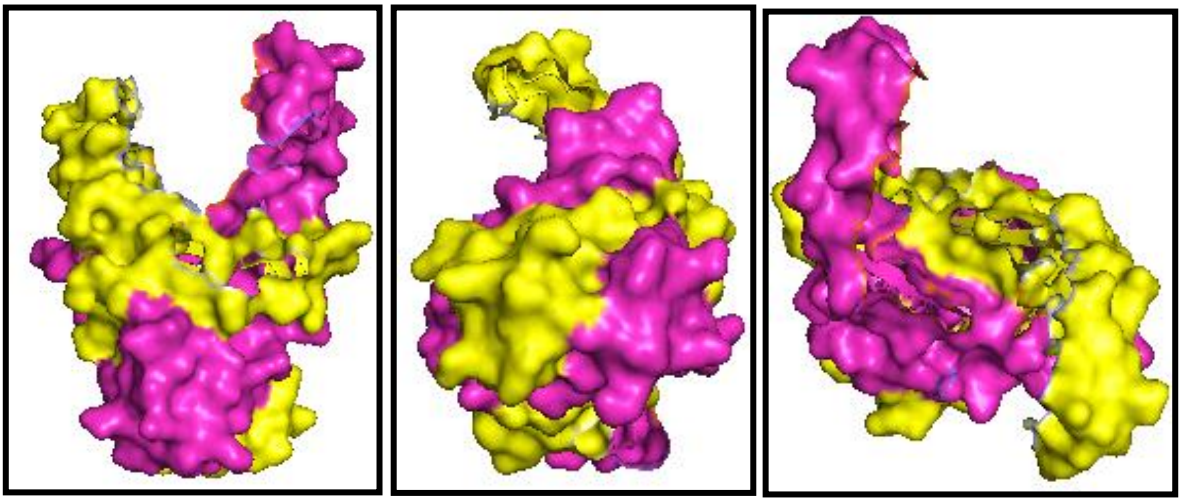
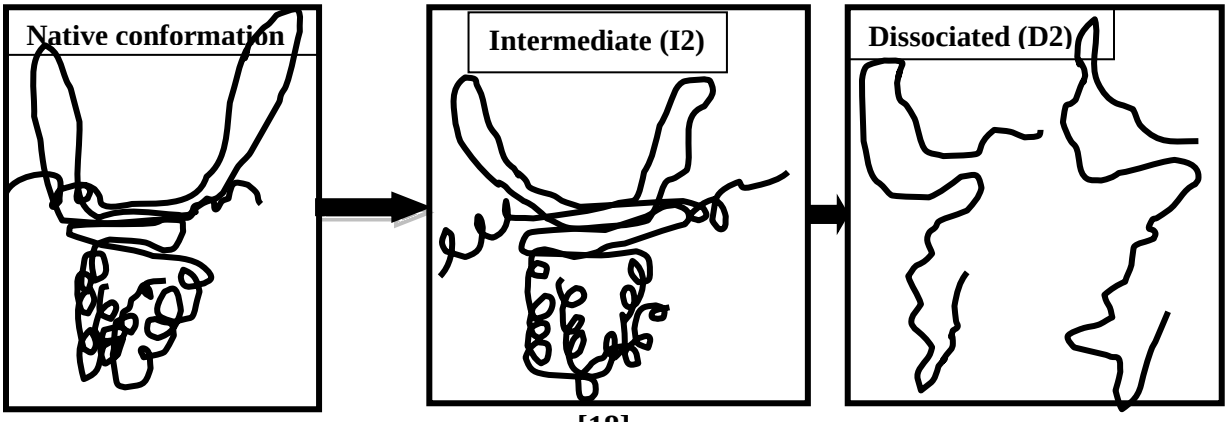


Figure 13: Showing the surface representation of Anabaena HU (1P51) in different orientations showing a close association between monomer interfaces.



[18]

Figure 14: Showing schematically the unfolding of HU dimers via a partially-unfolded dimeric intermediate, where the two dimers remain associated with each other through inter-antiparallel beta strands interactions, lying below the canonical DNA binding sites, before dissociating into completely unfolded monomers. ***Panel A:*** Shows the completely folded dimer followed by partially- unfolded dimer in ***Panel B***. ***Panel C:*** Shows the dissociation of partially unfolded dimers into unfolded monomers.

1.3. HU, biofilms and the extracellular matrix

Biofilms are aggregates of bacterial cells embedded in a self-secreted matrix of extracellular polymeric substance(s), or EPS. In biofilms, bacteria are adhered to each other to form aggregates, as well as associated with the extracellular matrix.The EPS plays several functional roles, including adhesion to biotic and/or abiotic surfaces, serving as a source of nutrients, retention of water in the environs of bacteria, and providing a protective barrier to bacteria [50]. The behaviour of bacteria in biofilms is completely different from the behaviour of free-living planktonic bacteria. The extracellular matrix in biofilms is made up of polysaccharide, extracellular DNA (eDNA) and proteins. The eDNA is of bacterial (self) origin, either secreted by bacteria or released after bacterial lysis. The composition of exopolysaccharide in the biofilm matrix varies with bacterial species [51]. The exopolysachharide of *Pseudomonas aeruginosa* is a mixture of psl, pel and alginate [52]. The secretion of these polysachharides helps bacteria to adapt to changing environmental conditions and withstand dessication, withstand attack by the immune response of an animal host, and resist the effect of oxidizing free radicals [53, 54]. Isolates of *P. aeruginosa*from the lungs of cystic fibrosis patients express huge amounts of alginates along with some small colony variants which have pel and psl polysachharides in place of LPS, as their EPS components [55]. The extracellular matrix of NTHI (Non typable *Haemophilus influenzae*) is lipopolysaccharide (LPS) which is modified by the addition of phosphatidylcholine [56].eDNA is an important component of the biofilm matrix. eDNA is released by lysing bacterial cells and sometimes bacteria are also thought to exude eDNA through secretion systems. eDNA can be a mixture of DNA released by the biofilm-forming organisms, or DNA released by phagocytic neutrophills at a site of bacterial infection [57].In salmonella sp. and *Pseudomonas aeuginosa*, eDNA binds pel polysaccharide, and amyloid fibers, and helps in organising the biofilm matrix [58]. The matrix proteins present in the ECM are proteins secreted by bacterial cells and proteins which are the components of the bacterial motility system, such as the protein components of the pilus. In *Pseudomonas aeruginosa*, LecA,

and LecB are the matrix-associated carbohydrate binding lectins, and CdrA is a large adhesion protein associated with providing integrity to ECM. Proteins associated with NTHI and *Salmonella typhimurium* typeIV pili and a protein called 'curli' are associated with promoting bacterial cell-cell and cell-ECM contacts.

Another important class of proteins associated with biofilms are DNABII family nucleoid associated proteins. DNABII proteins are reported to be associated with biofilms formed by NTHIand *Pseudomonas aeruginosa*, and they bind to and stabilize the interwoven mesh-like structure of eDNA in biofilms formed in vivo by NTHI [59].The presence (and association) of DNABII proteins makesthe eDNA resistant to DNase treatment in full-grown bacterial biofilms [59]. In *in vitro* experiments, anti-IHF antibodies used against non typable haemophilus influenza completely abolish biofilm formation and the newly released bacteria tend to be 4-8 times more sensitive to antibiotic action [60]. To determine the *in vivo* effect of anti-IHF antibodies, members of a model organism, Chichilas, were immunized with IHF. Disruption of NTHI biofilms was observed in the middle ear. After mapping for the epitopes on DNABII proteins responsible for disruption of biofilms by antibodies, it was found that antibodies against DNA binding sections of the DNABII proteins are potent in preventing biofilm formation [59, 60].

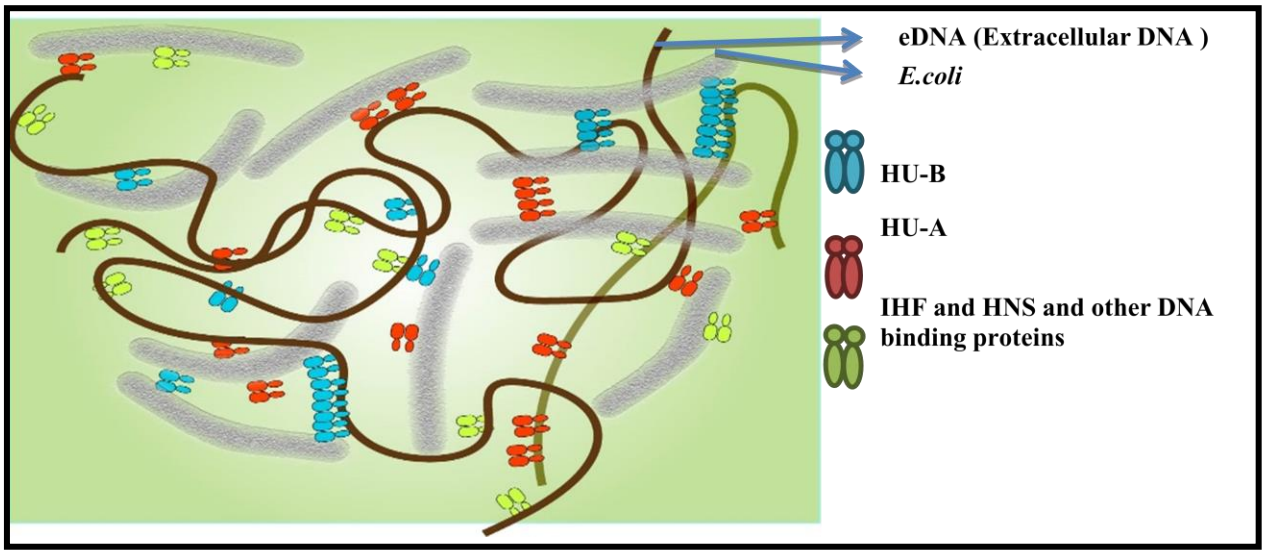


Figure 15: Showing the association of DNABII proteins with eDNA and bacterial cells in biofilm matrix.

1.3.1. Bacterial nucleoid associated protein HU is secreted by type IV secretion system

Bacteria employ different types of secretion system to secrete substances which help them in killing other bacteria, help them in attaching to host cells and acting as a sink for nutrients outside the bacterial cell. The secretion systems of Gram negative bacteria have to transfer proteins through double membranes, i.e., through the innermost plasma membrane and the lipopolysaccharide (LPS)-displaying outermost membrane. The secretion can occur in two steps, where the protein to be secreted is first released into the periplasmic space in unfolded form by Sec or tat secretion system and then translocated across outer membrane through second transporter [61]. The other secretion systems, named as Sec/Tat-independent secretion systems, are made up of protein complexes which span both the inner and outer membrane .These secretion systems are of number types I to VI in *E.coli*. These secretion systems are differentiated on the basis of types of cargo translocated through them [61].

Type I secretion system (T1SS)

This secretion system resembles ATP binding cassette (ABC) transporters and is associated with the transport of antibiotics and small toxins out of the cell [62]. Bacteria can have multiple T1SS with each transporter associated with different unfolded substrates [63]. These substrates can be unfolded digestive enzymes such as protease and lipases. T1SS is divided into three parts: an inner membrane ABC transporter protein, MFP membrane fusion protein that passes through inner membrane and binds to OMF(outer membrane factor) in the outer membrane [64].ABC proteins provide energy by hydrolyzing ATP for the transport of substrate [65]. N- terminal of MFP is associated with the recognition of substrate which is then translocated through the pore formed in outer membrane by OMF in unfolded state [66,67].These transporters recognize a C-terminal recognition sequence on the target substrate. TISS is associated with the release of virulence factors in *V.cholera*, *Serratia marcescens* and uropathogenic *E.coli* [68-70].

Type II secretion system (T2SS)

This secretion system transports only folded proteins and is assisted by Sec or Tat transporters because TIISS channel is present in outer membrane. Protein destined to secrete by T2SS must

have N-terminal Sec/Tat type cleavable signal [71]. The proteins are folded into the periplasm before secretion.T2SS is divided into four parts: outer membrane complex, inner membrane complex, pseudopilus, and the secretion ATPase [71]. The outer membrane channel is made up of a multimeric protein complex called secretin. Secretin has a long N-terminal region which extends upto the periplasm to make contact with inner membrane complex [72].This secretion system is employed by various bacteria to release their virulence factors. Cholera toxin of *V. cholera* and exotoxin A of *P. aeruginosa* is released by this system [73].

Type III secretion system (T3SS)

Type III secretion system is known as "injectosome" because of their injection-needle-like structure. Secretion of T3SS substrates is generally a one step process except in *Yersinia pestis*. The protein to be secreted through this system carries an N-terminal secretory signal which is not cleaved after secretion. Many of the T3SSs have associated chaperones which guide the substrate to the base of T3SS, where they are secreted in an ATP-dependent, unfolded state.

The T3SS is divided into three components: basal body, needle like component and translocon [74]. Basal body contains cytoplasmic complex and spans inner and outer membrane, forming a ring like structure with a central rod [75]. Emanating from this structure is a needle which extends into extracellular space [75]. This needle has a hollow core which permits unfolded substrate to transverse [76].

Type IV secretion systems (T4SS)

This secretion system evolutionarily resembles bacterial DNA conjugation and is associated with the transport of variety of substrates including proteins, protein-DNA complexes and protein-protein complexes. They can transfer their substrates into bacterial (same and different) and eukaryotic cells. Despite substrate variability all T4SSs are evolutionary related and share the same components and operons.The most well studied T4SS is of *Agrobcterium tumifecans*. *A. tumifecans* uses this system to transport oncogenic T-DNA into plant cells [77]. T4SS is composed of 12 VirB/D proteins [78]. VirB6-10 is found in inner membrane, periplasm and outer membrane, interacting with each other to form a secretion channel. VirB4, VirB11 and

VirD4 serve as ATPs in the inner membrane. This system also include one pilus composed of VirB2 and VirB5 subunits which helps in making contacts with target cells to deliver substrate directly into recipient cytoplasm [78].

The mechanism of secretion through this system is still an active area of research. However, it is believed that DNA and proteins first make contact with VirD479. The substrate is then transferred to VirB11, which finally transfers it into the inner membrane channel passing through periplasm and outer membrane. The role of the pilus is still an active area of research. Some believe it is used to make contact with recipient cell [78]. Still, others have shown the pilus can serve as a conduit for release substrate directly into recipient cell [79].T4SS is associated with the pathogenesis of bacterial pathogens like *Neisseria gonorrhoeae, Brucella suis* and *Helicobacter pylori*[80, 81].

Type 5 secretion system (T5SS)

The secretion through this system is coupled with the Sec apparatus in the inner membrane. These proteins contain a beta barrel domain which forms a channel in the outer membrane [86].The unfolded periplasmic proteins are folded with the help of chaperones in the periplasm before secretion. These secretion systems secrete proteins in the unfolded state. Chaperone-mediated systems are used for the assembly of pilins on the surface of uropathogenic *E.coli* [82].

Type 6 secretion system (T6SS)

T6SS is the most recently discovered secretion system and is conserved in a fairly wide range of gram negative bacteria. T6SS is capable of transporting effectors molecules from one cell to another in contact-dependent manner [83].

HU is secreted by T4SS

Bacterial nucleoid associated proteins of DNABII family have been found to stabilize the extracellular matrix made up of eDNA during biofilm formation. The possible explanation for the release of eDNA and the protein associated with them is bacterial cell lysis and extrusion through OMVs (outer membrane vesicles). The release of DNA through explosive cell lysis through rounded giant cell formation was shown in *Pseudomonas aeruginosa* recently [84].

The release of ssDNA through T4SS was firstly shown for *Neisseria.* The secretion system in this organism is made up of an inner membrane complex composed of TraC, D, and B which spans both membranes and TraK and V forming outer membrane complex along with many cytoplasmic chaperones [78,85]. Recently NTHI, a causative agent of many upper and lower respiratory tract disorders was shown to release DNA and DNABII proteins by a similar mechanism. The release of DNA and proteins was found to be independent of cellular lysis and startingas early as 3 hours after inoculation in the log phase of growth. The release was independent of cellular lysis although there was formation of rounded giant cells as found in pseudomonas (1 in 1000) but their formation started very late after inoculation [86].Cells with deletion mutants of traGC and ComE failed to secrete DNA extracellularly [86].

1.3.2. Lipopolysaccharide outer membrane of gram negative bacteria

The cytoplasm of Gram-negative bacterium is surrounded by an inner phospholipid membrane and an outer lipopolysaccharide membrane. The region between the inner and outer membranes harbors a peptidoglycan cell wall and this region is known as the periplasm, which serves as the site for disulphide bond formation in proteins, because of its oxidizing environment [87, 88].The outer membrane is composed of phospholipids in the inner leaflet and lipopolysaccharide in the outer leaflet. The outer membrane is transversed by many beta barrel integral proteins known as porins and ion channels [89]. Porins helps in non specific uptake of small compounds by passive diffusion and are major components of OMPs. LPS is crucial for the structural integrity and function of gram negative bacteria [90]. In *E.coli* and *Salmonella* serovar *typhimurium* the LPS to phospholipid ratio is 0.15:1 with $2x10^6$LPS molecules per cell surface.

Structurally LPS can be divided into three parts: innermost lipid-A, middle core oligosaccharide and outer O-antigen of varying oligosaccharide chain length [91]. Lipid-A which is known as bacterial endotoxin causes septic shock and fever. It anchors the whole LPS on the outer membrane and acts as site for the nucleation of core and O antigen. Lipid-A is made up of glucosamine dimers associated covalently with lipid acyl chains of varying length, depending upon bacterial strains and environmental conditions [92].The core of LPS is divided into outer and inner cores. The inner core is covalently attached to lipid-A. The inner core is made up of highly conserved sugars such as L-glycerol-D-mannose heptose (Hep) and 2-keto-3-

deoxyoctulosonate (Kdo). The outer core has repeating units of sugars like hexose and hexosamines [90].The core and lipid A are negatively charged because of PO^{4-}(phosphates) associated with sugars. The negatively charged core and lipid-A are always associated with divalent cations, which provide them structural integrity [93]. The O-antigen contains repeating units of monosaccharides with a mixture of branched and linear chains. In bacterial membrane, negative charge potential of LPS is associated with the Min-D oscillation and partitioning of cells during division [94].

Synthesis of LPS in Gram negative bacteria

The synthesis of LPS in most of the gram negative bacteria occurs via the Raetz pathway using Lpx enzymes at the interface of the cytosol and inner membrane. First step of the reaction is addition of acyl chain to the 3-OH end of uridine diphosphate N- acetylglucosamine (UDP-GlcNAc) catalysed by LpxA enzymes. In *E.coli* LpxA is specific for beta hydroxymyristate (3-OH-C14:0) [95].The difference in LpxA active site translates into the structural difference of lipid A amongst various bacterial species [96].The second step of Kdo2-lipid A synthesis is catalysed by the LpxC enzyme which is a deacetylase. Deacetylation provides a free amino group for the addition of a second acyl chain by LpxD [97]. In *E.coli*, LpxD is specific for beta hydroxymyristate [98]. Part of the LpxD product is cleaved by LpXH to form UMP(uridinemonophosphate) and lipid X [99].

The fully synthesized LPS assembles on the inner membrane towards the periplasm and the assembled LPS is transported by lpt (LPS transport proteins) bridgingthe inner and outer membranes [100] driven by ATP hydrolysis [101].

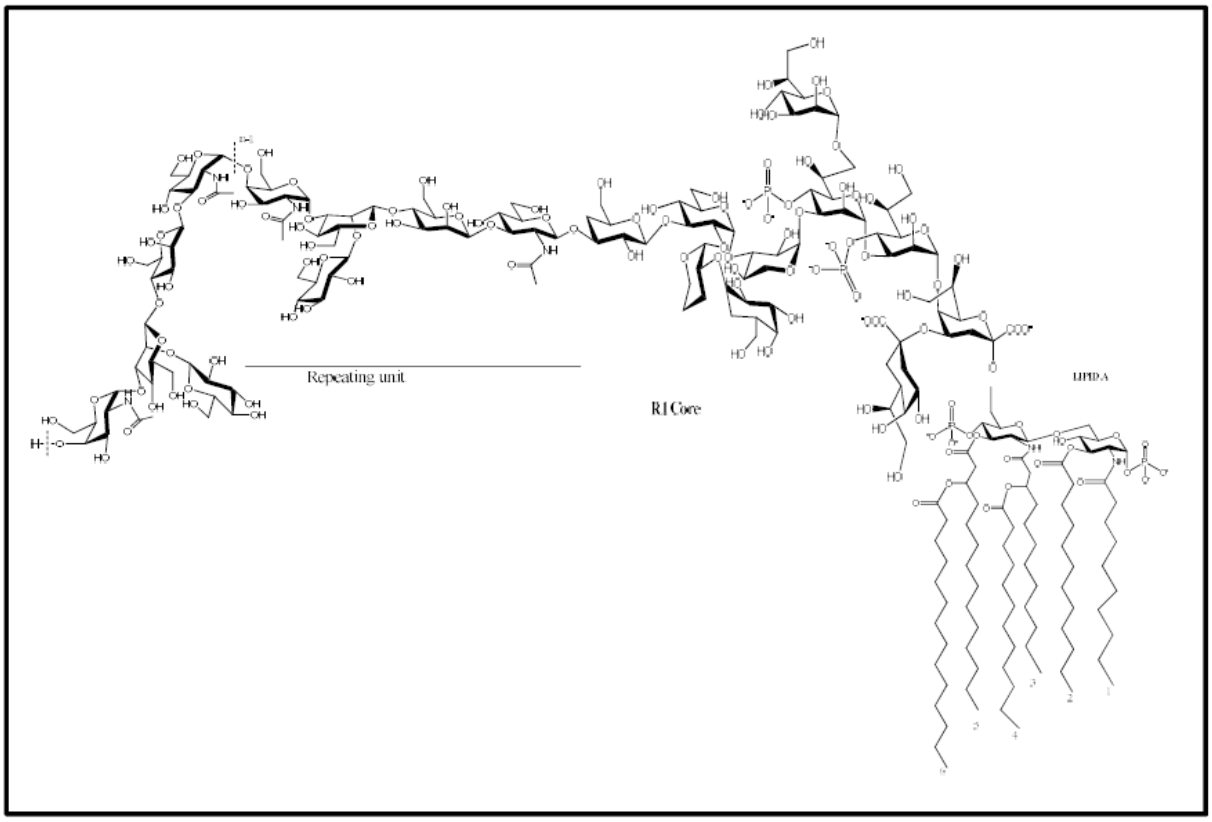


Figure 16: Showing the structure of LPS with Lipid A made up of made up of repeating units of D-glucosamine joined by beta (1-6) linkage, two monophospho ester group at O1and O40,and six amide/ester linked fatty acids. The core is made up of Kdo residues and three Hep residues,two of which have monophosphoesterester group at O4' position in the inner core. The outer core consists of five hexopyranoses, three D-glucoses, and two D-galactoses, all of which are alpha-linked, except for the beta-linkage between the last two glucose residues. The O6 antigen polysaccharide is attached to the O3 position of the terminal glucosyl residue in the outer core and its linkage has the b-configuration. This is in contrast to the corresponding alpha-(1/4)-linkage between the repeating units 'O' antigen containing repeating units of sugars with a 3-substituted N-acetyl-D-glucosamine residue at the reducing end. The additional sugars in the repeating unit are two beta-D-mannoses, one N-acetyl-alpha-D-galactosamine, leading to four sugars in the backbone of the polymer, and a beta-D-glucose residue forming a branched structure via its (1/2)-linkage to the second mannose residue.

1.3.3. Physical/Chemical conundrum arising from embedding of negatively charged bacteria in negatively charged eDNA matrix.

Polysaccharide, proteins and *eDNA* are major components of the biofilm extracellular matrix in all biofilm forming bacteria (Gram positive and negative) and the interaction between them is imperative for biofilm stability [102].The interaction between exopolysaccharide psl and *eDNA* has been shown to stabilize biofilms in the opportunistic pathogen *Pseudomonas aeruginosa [103]*. The binding of the bacterial cells to eDNA was shown with a food borne pathogen *Listeria monocytogenes,* where removing eDNA with *DNAse* drastically reduced bacterial adhesion on a glass surface in a static chamber attachment assay [104]. Immobilized bacteria were detached after treatment with *DNAse* and there was no effect of RNAse or proteinase-K suggesting that eDNA has a role in bacterial immobilization [104].The biofilm formation starts with embedding bacteria in *eDNA*, which subsequently secretes exopolysaccharides to form a protective covering around bacterial cells as shown by the *DNAse* experiments in *Pseudomonas aeruginosa*, where biofilms were liable to *DNAse* treatment only for short duration after inoculation [105]. The core and lipid-A portion of LPS contains sugar phosphates which impart negative charge to gram negative bacterial cell surface. The charge of bacterial surface is measured by zeta potential studies and it has been shown to be negative for both gram positive and gram negative bacteria [106]. Gram positive bacterial surface is also negatively charged because of techoic acid, which resembles sugar phosphate present in the core LPS of gram negative bacteria. There are no reports about the initial association of bacteria with eDNA in literature, which raises a very fundamental question: how can similarly negatively-charged LPS and eDNA interact?

The DNA has shown to bind with core oligosaccharide portion of LPS only in presence of $CaCl_2$, where the charge on DNA backbone is neutralized by Ca^{2+} ions [107]. The binding of *E.coli* and other gram negative bacteria directly to *eDNA* is not possible because of charge similarity between LPS and DNA backbone. The entrapment of bacteria in *eDNA* matrix is only possible in presence of a charge neutralizer present between them, and this charge neutralizer must be associated with both negatively-charged bacteria, and negatively-charged *eDNA*.

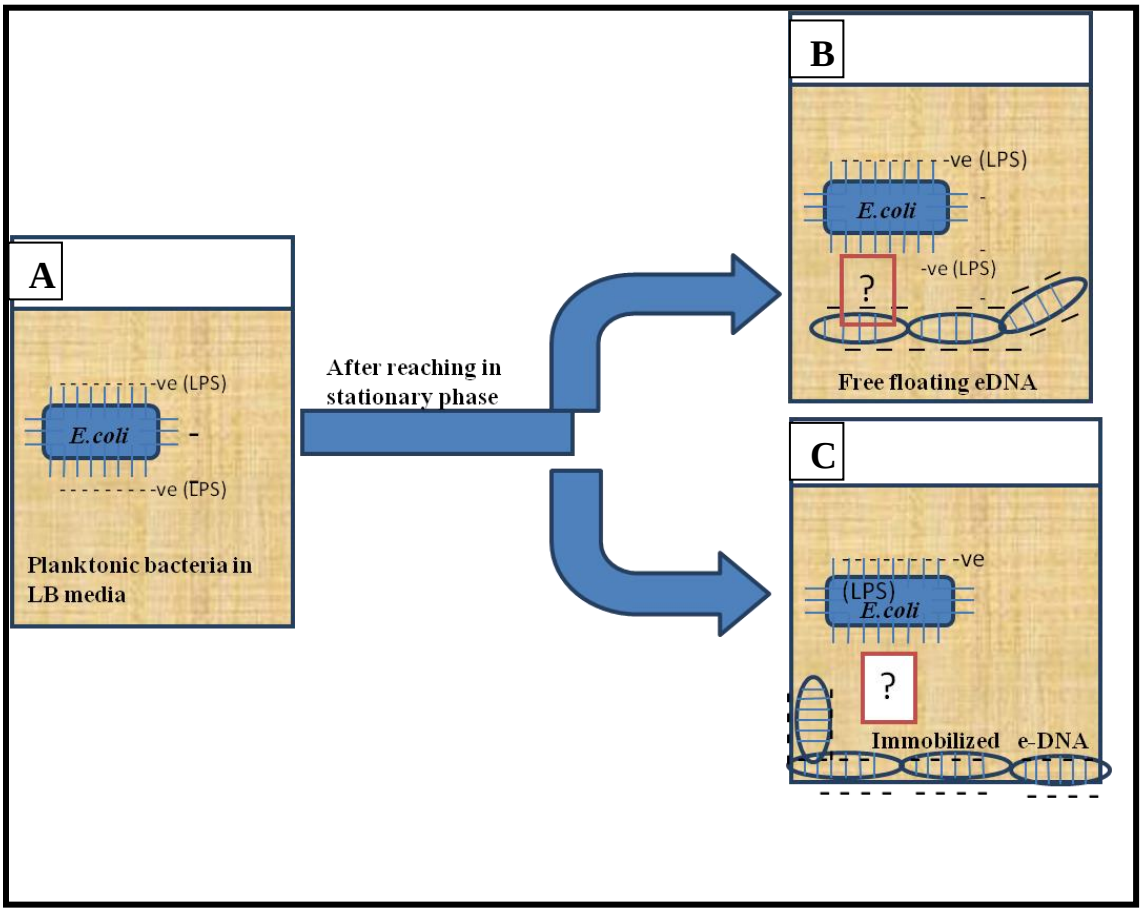


Figure 17: Schematic showing the release of eDNA by bacterial cells after reaching late exponential phase and early stationary phase and the appearance of released eDNA as free floating and surface adhered entities, in culture vials. ***Panel A :*** Shows the bacterial cells immediately after inoculation in fresh LB media. Growing bacterium releases DNA into the media which remains free floating and can adhere or immobilize on the bottom of culture vials as shown in the panel B and of the same figure. The adsorption of the negatively charged bacterial cells on negatively charged DNA (which is essential for the formation of biofilm assembly) necessitates the presence of positively charged molecules which can bind both bacterial cells as well as DNA. The location of that molecule is represented by the question mark in ***Panel B*** and ***C*** of the same figure.

1.3.4. Structural similarities between DNA and LPS

The core region of LPS is made up of two repeating units of sugar phosphates,where the sugars arehexoses. The DNA backbone similarly contains sugar phosphates where phosphates projects outwards and imparts to the backbone an overall negative charge. The molecules which can recognise LPS through the phospohate backbone can also bind with DNA, as has been shown in the recent paper where antibodies against lipid-A region of LPS have been shown to cross react with ssDNA of the patients and to cause the autoimmune disorderknown as SLE (Systemic lupus erythromatosus) and rheumatic autoimmune disorder (reference from anti LPS antibodies). From these studies, we can conclude that LPS can behave like ssDNA (or dsDNA without helicity) and can thus probably bind to positively-charged DNA binding proteins.

1.3.5. The need for a neutralizer: can HU play the role?

Biofilm formation proceeds with the bacteria making contact with negatively charged eDNA, which is released through lysis, active secretion via T4SS or through OMVs (Outer membrane vesicles. The initial temporary binding of bacterial cell with eDNA is regulated by electrostatic interaction as bacterial surface is negatively charged and tends to repel eDNA [108].Bacteria can bind to eDNA, if the backbone is neutralized by DNA binding proteins. DNABII proteins are present in abundance in biofilm matrix in association with eDNA. The release of eDNA along with associated DNA-binding proteins has been shown to be mediated by quorum sensing molecules in a density-dependent manner [109]. Presence of DNA binding proteins prevents DNAse-mediated digestion of eDNA and the antibodies against DNA binding tip of HU/IHF have been shown to disrupt biofilm assembly in *in vitro* and *in vivo* experiments. HU, which is one of the 13 nucleoid associated proteins, is present at 45000-50000 copies of dimers in exponentially growing bacteria, and is replaced by Dps (protein associated with bacterial starvation) in stationary phase after being secreted out by T4SS [110, 111]. HU has been shown to bind DNA through multiple surfaces (as its surface is rich in positively charged amino acids) and its oligomeric nature adds more surface to bind with opposite charges [112]. DNABII proteins which are secreted out can mask charges on LPS and DNA and can act as a “glue” by bridging them via dimer oligomerization. The binding of charged surface with HU is entropically favoured as it results in dissociation of salt bridges on protein surface and the release of ions into

the solution. Along with acting as a glue between LPS and DNA, HU can adhere to bacterial cells and transform them in sessile form. This form is necessary for initial stages of biofilm formation [113]. Bacteria have been shown to form clumps or aggregates in planktonic forms prior to entrapment into eDNA and biofilm formation by secreting extracellular polymeric substances(EPS) [113].The bacteria inside biofilms have very less access to nutrients which reduces their metabolism hence making them less susceptible to antibiotics [114]. The schematic for the entrapment of bacterial in eDNA by HU is shown in the Figure 19.

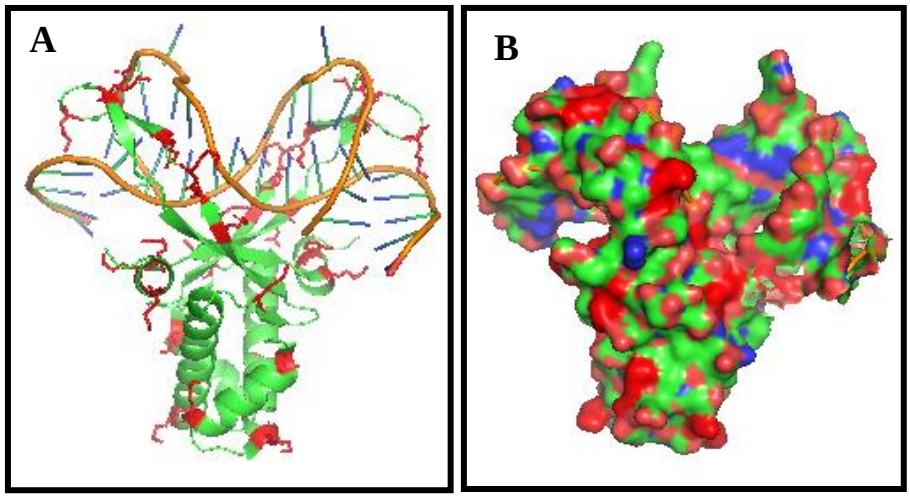


Figure 18: Showing the distribution of positive charge all over the surface of HU (shown as red spots). ***Panel A :*** Shows the binding of DNA through canonical DNA binding site on HU (1P51). ***Panel B :*** Shows the avalibilty of postively charged sufaces away from the canonical DNA binding site, which can be used for the non specfic electroctrostatic intercactions with oppositely charged molecules, including DNA.

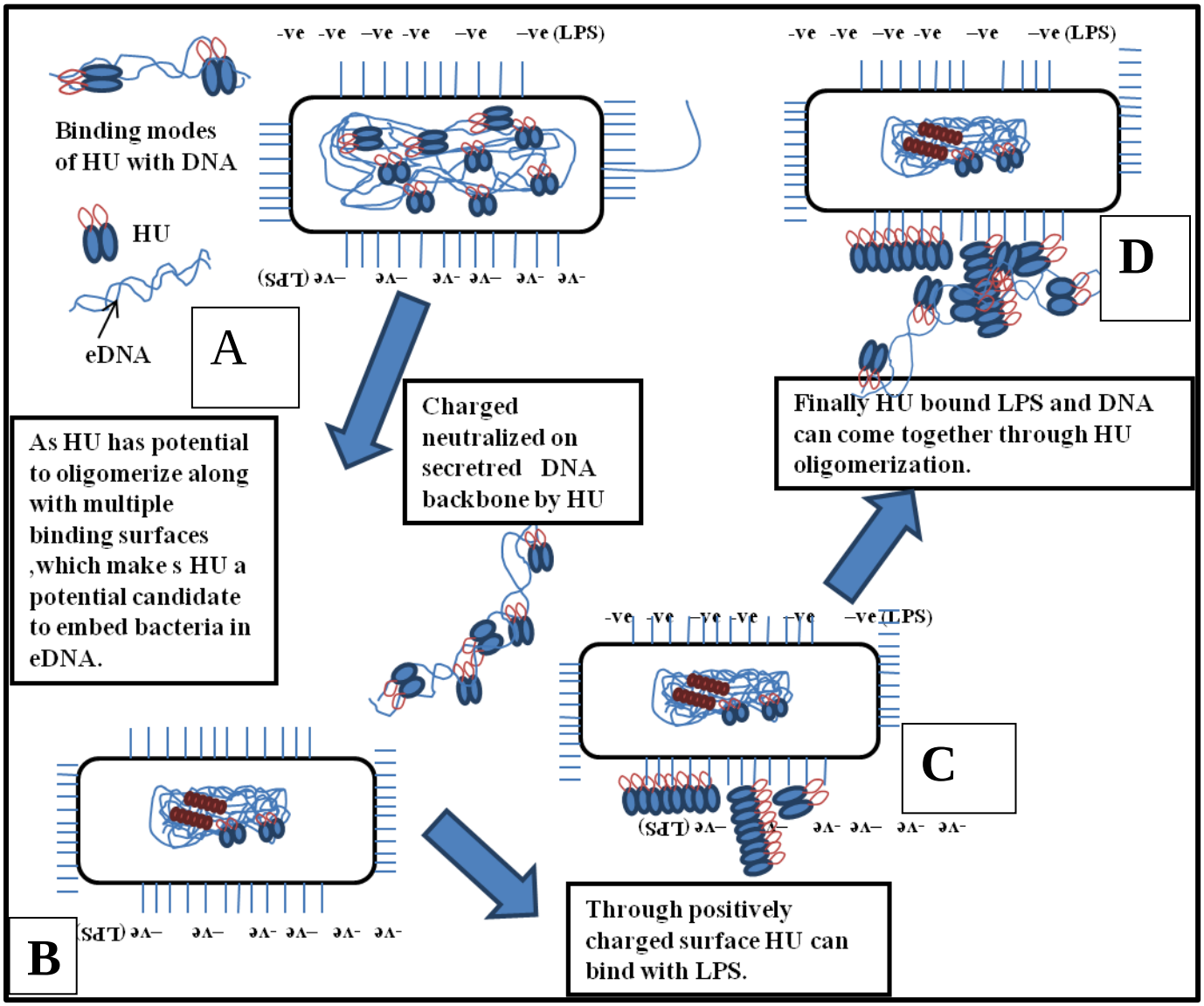


Figure 19: Showing the bacterial growth phase transition and nucleoid remodelling in context of HU protein only. The bacterial exponential phase nucleoid is less compacted and associated with a large numbers of HU dimers and oligomers shown in the ***Panel A:*** Along with possible modes with which HU can interact with DNA. ***Panel B:*** In late exponential or early stationary phase bacteria will start secreting DNA along with HU and other DNA binding proteins bound to it. ***Panel C:*** The released HU can bind with LPS molecules on bacterial cell surface through electrostatic interaction as LPS is negatively charged due to sugar-linked phosphates, in a manner similar to DNA. ***Panel D:*** The LPS bound HU can interact with DNA though its multiple DNA

binding sites, hence helping bacterial cell to interact with similarly charged EPS (extracellular polymeric substance) mainly made up of eDNA.

1.3.6 Can HU bind to LPS?

The binding of HU with dsDNA through beta hair pin loop and introducing a kink in helix backbone is a widely accepted binding model. Recently a paper showed that HU can also associate with DNA through conserved lysine residues, away from the canonical DNA binding beta hair pin loop [7,115]. The same group showed that lysine mutations E38K, and V42L of HU turned non-invasive *E.coli* into invasive *E.coli* (116). Mutation of other residues into lysine residues away from the DNA binding loops on HU has been shown to change the transcription pattern of the whole *E.coli* genome. The mutant HU-A E38K-V42L showed altered morphology on LB plates as compared to control cells. The cells were more opaque, glossy and round along with showing defect in growth on different sugar substrates [116]. The quiescent virulence genes like haemolysin and genes encoding for curli got activated along with the change in nucleoid condensation suggesting the region away from the DNA binding loops can interact with DNA in conformation-independent manner. The extended arms (beta hair pin loops) of HU can interact with DNA independent of lateral interaction trough K83, K18 and K3 residues which can sustain weak interaction with linear DNA at lower salt concentration without bending of DNA [115].

In the absence of DNA, the nearby opposite charges on a protein's surface form salt bridges and compete for subsequent interaction with DNA. The negatively charged surface causes accumulation of salt cations, which are released into the solution upon protein binding to favor the binding reaction entropically. The Lipid–A region and RI core of LPS contains repeating units of sugar phosphates which imparts LPS an overall negative charge and the positive surface of HU with large number of lysine and arginines favors the chances of electrostatic interactions between them.

The canonical DNA binding site is specific to helical DNA and the RNA containing secondary structure without any sequence specificity but the lysines and arginines on non canonical DNA binding sites can form electrostatic interaction with LPS molecules along with involvement of beta hair pin loops positively charged amino acids.

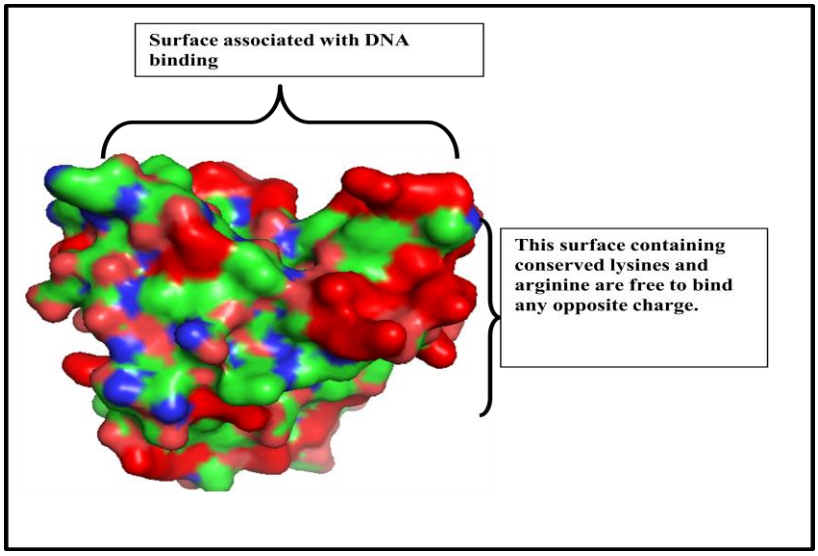


Figure 20: Showing the surface view of protein HU, PDB ID 2O97. Positively charged surface is marked as red. The dense cluster of positively charged residues away from the canonical DNA binding site can possibly bind non- specifically with oppositely charged molecules, including DNA.

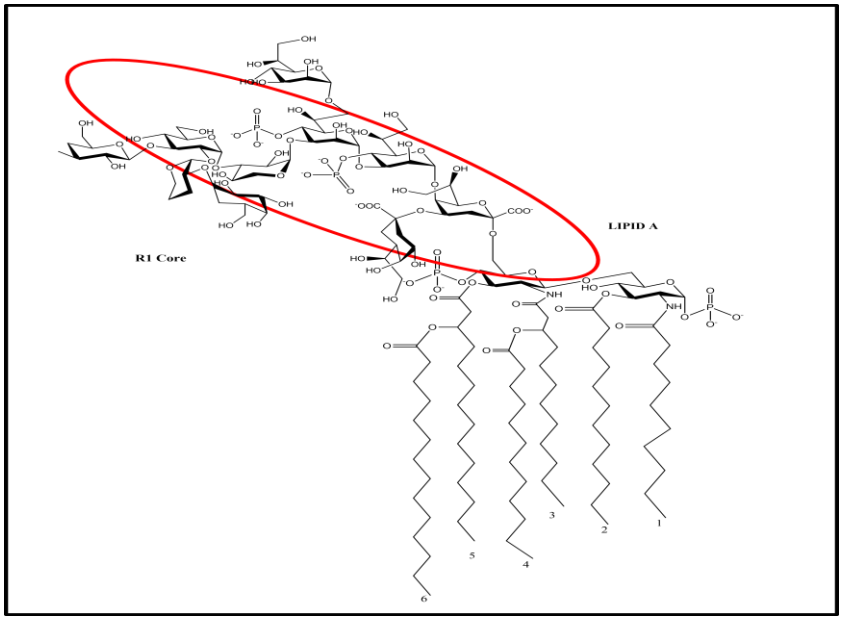


Figure 21: Showing the core lipid A made up of sugar phosphates which impart to LPS an overall negative charge.

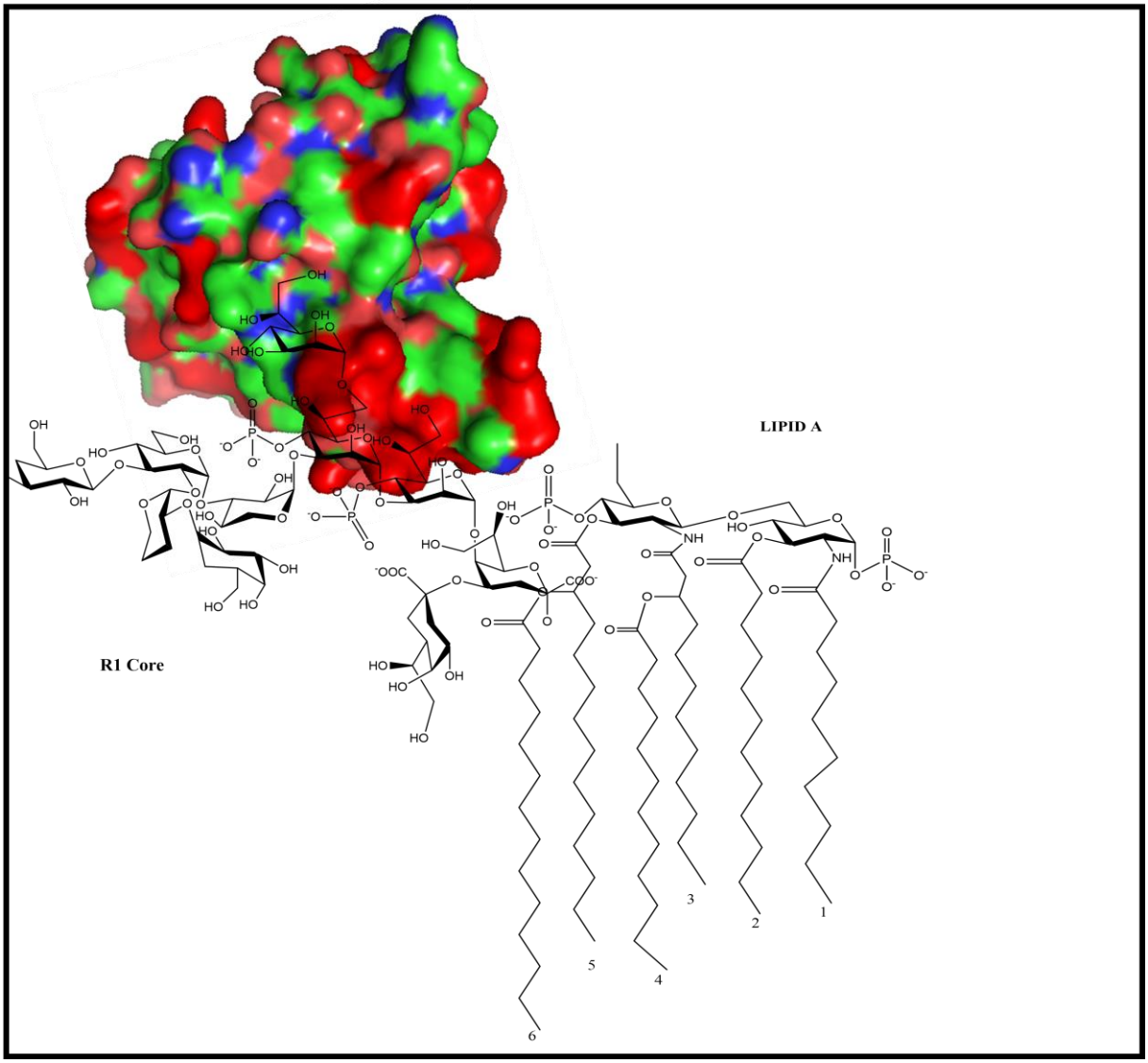


Figure 22: Showing the posssible binding site for LPS on HU molecules throug non canonical DNA binding sites.

References :

1. Swinger, K. K., and Rice, P. A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol***14**, 28-35
2. Grove, A., Galeone, A., Mayol, L., and Geiduschek, E. P. (1996) On the connection between inherent DNA flexure and preferred binding of hydroxymethyluracil-containing DNA by the type II DNA-binding protein TF1. *J Mol Biol***260**, 196-206
3. Greene, J. R., Brennan, S. M., Andrew, D. J., Thompson, C. C., Richards, S. H., Heinrikson, R. L., and Geiduschek, E. P. (1984) Sequence of the bacteriophage SP01 gene coding for transcription factor 1, a viral homologue of the bacterial type II DNA-binding proteins. *Proc Natl Acad Sci U S A***81**, 7031-7035
4. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell***87**, 1295-1306
5. Greene, J. R., and Geiduschek, E. P. (1985) Site-specific DNA binding by the bacteriophage SP01-encoded type II DNA-binding protein. *EMBO J***4**, 1345-1349
6. Haniford, D. B. (2006) Transpososome dynamics and regulation in Tn10 transposition. *Crit Rev Biochem Mol Biol***41**, 407-424
7. Grove, A. (2011) Functional evolution of bacterial histone-like HU proteins. *Curr Issues Mol Biol***13**, 1-12
8. Jia, X., Grove, A., Ivancic, M., Hsu, V. L., Geiduscheck, E. P., and Kearns, D. R. (1996) Structure of the Bacillus subtilis phage SPO1-encoded type II DNA-binding protein TF1 in solution. *J Mol Biol***263**, 259-268
9. Granston, A. E., and Nash, H. A. (1993) Characterization of a set of integration host factor mutants deficient for DNA binding. *J Mol Biol***234**, 45-59
10. Sayre, M. H., and Geiduschek, E. P. (1990) Effects of mutations at amino acid 61 in the arm of TF1 on its DNA-binding properties. *J Mol Biol***216**, 819-833
11. Claret, L., and Rouviere-Yaniv, J. (1997) Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival. *J Mol Biol***273**, 93-104
12. Almiron, M., Link, A. J., Furlong, D., and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev***6**, 2646-2654
13. Wang, Q., and Calvo, J. M. (1993) Lrp, a major regulatory protein in Escherichia coli, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. *EMBO J***12**, 2495-2501
14. Chenoweth, M. R., and Wickner, S. (2008) Complex regulation of the DnaJ homolog CbpA by the global regulators sigmaS and Lrp, by the specific inhibitor CbpM, and by the proteolytic degradation of CbpM. *J Bacteriol***190**, 5153-5161
15. Dorman, C. J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol***2**, 391-400
16. Dorman, C. J. (2009) Nucleoid-associated proteins and bacterial physiology. *Adv Appl Microbiol***67**, 47-64
17. Hsu, Y. H., Chung, M. W., and Li, T. K. (2006) Distribution of gyrase and topoisomerase IV on bacterial nucleoid: implications for nucleoid organization. *Nucleic Acids Res***34**, 3128-3138
18. Riley, S. P., Bykowski, T., Cooley, A. E., Burns, L. H., Babb, K., Brissette, C. A., Bowman, A., Rotondi, M., Miller, M. C., DeMoll, E., Lim, K., Fried, M. G., and Stevenson, B. (2009) Borrelia burgdorferi EbfC defines a newly-identified, widespread family of bacterial DNA-binding proteins. *Nucleic Acids Res***37**, 1973-1983

19. Browning, D. F., and Busby, S. J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol***2**, 57-65
20. Rouviere-Yaniv, J., and Gros, F. (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from Escherichia coli. *Proceedings of the National Academy of Sciences***72**, 3428-3432
21. Pettijohn, D. E. (1988) Histone-like proteins and bacterial chromosome structure. *J Biol Chem***263**, 12793-12796
22. Oberto, J., Drlica, K., and Rouviere-Yaniv, J. (1994) Histones, HMG, HU, IHF: Meme combat. *Biochimie***76**, 901-908
23. Kano, Y., Yoshino, S., Wada, M., Yokoyama, K., Nobuhara, M., and Imamoto, F. (1985) Molecular cloning and nucleotide sequence of the HU-1 gene of Escherichia coli. *Mol Gen Genet***201**, 360-362
24. Dillon, S. C., and Dorman, C. J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol***8**, 185-195
25. Claret, L., and Rouviere-Yaniv, J. (1996) Regulation of HU alpha and HU beta by CRP and FIS in Escherichia coli. *J Mol Biol***263**, 126-139
26. Huisman, O., Faelen, M., Girard, D., Jaffe, A., Toussaint, A., and Rouviere-Yaniv, J. (1989) Multiple defects in Escherichia coli mutants lacking HU protein. *J Bacteriol***171**, 3704-3712
27. Boelens, R., Vis, H., Vorgias, C. E., Wilson, K. S., and Kaptein, R. (1996) Structure and dynamics of the DNA binding protein HU from Bacillus stearothermophilus by NMR spectroscopy. *Biopolymers***40**, 553-559
28. Guo, F., and Adhya, S. (2007) Spiral structure of Escherichia coli HUalphabeta provides foundation for DNA supercoiling. *Proc Natl Acad Sci U S A***104**, 4309-4314
29. Luijsterburg, M. S., White, M. F., van Driel, R., and Dame, R. T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol***43**, 393-418
30. Bonnefoy, E., and Rouviere-Yaniv, J. (1991) HU and IHF, two homologous histone-like proteins of Escherichia coli, form different protein-DNA complexes with short DNA fragments. *EMBO J***10**, 687-696
31. Kamashev, D., Balandina, A., and Rouviere-Yaniv, J. (1999) The binding motif recognized by HU on both nicked and cruciform DNA. *EMBO J***18**, 5434-5444
32. Shindo, H., Furubayashi, A., Shimizu, M., Miyake, M., and Imamoto, F. (1992) Preferential binding of E.coli histone-like protein HU alpha to negatively supercoiled DNA. *Nucleic Acids Res***20**, 1553-1558
33. Bonnefoy, E., Takahashi, M., and Yaniv, J. R. (1994) DNA-binding parameters of the HU protein of Escherichia coli to cruciform DNA. *J Mol Biol***242**, 116-129
34. Balandina, A., Kamashev, D., and Rouviere-Yaniv, J. (2002) The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids. *J Biol Chem***277**, 27622-27628
35. Chodavarapu, S., Felczak, M. M., Yaniv, J. R., and Kaguni, J. M. (2008) Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin. *Mol Microbiol***67**, 781-792
36. Kamashev, D., and Rouviere-Yaniv, J. (2000) The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *The EMBO journal***19**, 6527-6535
37. Berger, M., Farcas, A., Geertz, M., Zhelyazkova, P., Brix, K., Travers, A., and Muskhelishvili, G. (2010) Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU. *EMBO reports***11**, 59-64

38. Boubrik, F., and Rouviere-Yaniv, J. (1995) Increased sensitivity to gamma irradiation in bacteria lacking protein HU. *Proc Natl Acad Sci U S A***92**, 3958-3962
39. Sharadamma, N., Khan, K., Kumar, S., Patil, K. N., Hasnain, S. E., and Muniyappa, K. (2011) Synergy between the N-terminal and C-terminal domains of Mycobacterium tuberculosis HupB is essential for high-affinity binding, DNA supercoiling and inhibition of RecA-promoted strand exchange. *FEBS J***278**, 3447-3462
40. Berger, M., Farcas, A., Geertz, M., Zhelyazkova, P., Brix, K., Travers, A., and Muskhelishvili, G. (2010) Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU. *EMBO Rep***11**, 59-64
41. Pruss, G. J., and Drlica, K. (1989) DNA supercoiling and prokaryotic transcription. *Cell***56**, 521-523
42. Liu, L. F., and Wang, J. C. (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A***84**, 7024-7027
43. Mangan, M. W., Lucchini, S., Cróinín, T. Ó., Fitzgerald, S., Hinton, J. C., and Dorman, C. J. (2011) Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in Salmonella enterica serovar Typhimurium. *Microbiology***157**, 1075-1087
44. Wang, W., Li, G. W., Chen, C., Xie, X. S., and Zhuang, X. (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science***333**, 1445-1449
45. Ramstein, J., Hervouet, N., Coste, F., Zelwer, C., Oberto, J., and Castaing, B. (2003) Evidence of a thermal unfolding dimeric intermediate for the Escherichia coli histone-like HU proteins: thermodynamics and structure. *J Mol Biol***331**, 101-121
46. Mukherjee, A., Sokunbi, A. O., and Grove, A. (2008) DNA protection by histone-like protein HU from the hyperthermophilic eubacterium Thermotoga maritima. *Nucleic Acids Res***36**, 3956-3968
47. Grove, A., and Saavedra, T. C. (2002) The role of surface-exposed lysines in wrapping DNA about the bacterial histone-like protein HU. *Biochemistry***41**, 7597-7603
48. Rouviere-Yaniv, J., and Kjeldgaard, N. O. (1979) Native Escherichia coli HU protein is a heterotypic dimer. *FEBS Lett***106**, 297-300
49. Bonnefoy, E., Almeida, A., and Rouviere-Yaniv, J. (1989) Lon-dependent regulation of the DNA binding protein HU in Escherichia coli. *Proc Natl Acad Sci U S A***86**, 7691-7695
50. Flemming, H. C., and Wingender, J. (2010) The biofilm matrix. *Nat Rev Microbiol***8**, 623-633
51. Karatan, E., and Watnick, P. (2009) Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol Mol Biol Rev***73**, 310-347
52. Franklin, M. J., Nivens, D. E., Weadge, J. T., and Howell, P. L. (2011) Biosynthesis of the Pseudomonas aeruginosa Extracellular Polysaccharides, Alginate, Pel, and Psl. *Front Microbiol***2**, 167
53. Berry, A., DeVault, J. D., and Chakrabarty, A. M. (1989) High osmolarity is a signal for enhanced algD transcription in mucoid and nonmucoid Pseudomonas aeruginosa strains. *J Bacteriol***171**, 2312-2317
54. Govan, J. R., and Deretic, V. (1996) Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. *Microbiol Rev***60**, 539-574
55. Starkey, M., Hickman, J. H., Ma, L., Zhang, N., De Long, S., Hinz, A., Palacios, S., Manoil, C., Kirisits, M. J., Starner, T. D., Wozniak, D. J., Harwood, C. S., and Parsek, M. R. (2009) Pseudomonas aeruginosa rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. *J Bacteriol***191**, 3492-3503
56. Hong, W., Mason, K., Jurcisek, J., Novotny, L., Bakaletz, L. O., and Swords, W. E. (2007) Phosphorylcholine decreases early inflammation and promotes the establishment of stable

biofilm communities of nontypeable Haemophilus influenzae strain 86-028NP in a chinchilla model of otitis media. *Infect Immun***75**, 958-965

57. Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D. S., Weinrauch, Y., and Zychlinsky, A. (2004) Neutrophil extracellular traps kill bacteria. *Science***303**, 1532-1535
58. Jennings, L. K., Storek, K. M., Ledvina, H. E., Coulon, C., Marmont, L. S., Sadovskaya, I., Secor, P. R., Tseng, B. S., Scian, M., Filloux, A., Wozniak, D. J., Howell, P. L., and Parsek, M. R. (2015) Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the Pseudomonas aeruginosa biofilm matrix. *Proc Natl Acad Sci U S A***112**, 11353-11358
59. Goodman, S. D., Obergfell, K. P., Jurcisek, J. A., Novotny, L. A., Downey, J. S., Ayala, E. A., Tjokro, N., Li, B., Justice, S. S., and Bakaletz, L. O. (2011) Biofilms can be dispersed by focusing the immune system on a common family of bacterial nucleoid-associated proteins. *Mucosal Immunol***4**, 625-637
60. Brockson, M. E., Novotny, L. A., Mokrzan, E. M., Malhotra, S., Jurcisek, J. A., Akbar, R., Devaraj, A., Goodman, S. D., and Bakaletz, L. O. (2014) Evaluation of the kinetics and mechanism of action of anti-integration host factor-mediated disruption of bacterial biofilms. *Mol Microbiol***93**, 1246-1258
61. Green, E. R., and Mecsas, J. (2016) Bacterial Secretion Systems: An Overview. *Microbiol Spectr***4**
62. Songer, J. G. (1997) Bacterial phospholipases and their role in virulence. *Trends Microbiol***5**, 156-161
63. Delepelaire, P. (2004) Type I secretion in gram-negative bacteria. *Biochim Biophys Acta***1694**, 149-161
64. Thomas, S., Holland, I. B., and Schmitt, L. (2014) The Type 1 secretion pathway - the hemolysin system and beyond. *Biochim Biophys Acta***1843**, 1629-1641
65. Kanonenberg, K., Schwarz, C. K., and Schmitt, L. (2013) Type I secretion systems - a story of appendices. *Res Microbiol***164**, 596-604
66. Pimenta, A. L., Young, J., Holland, I. B., and Blight, M. A. (1999) Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the E. coli haemolysin translocator. *Mol Gen Genet***261**, 122-132
67. Balakrishnan, L., Hughes, C., and Koronakis, V. (2001) Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by Escherichia coli. *J Mol Biol***313**, 501-510
68. Dolores, J. S., Agarwal, S., Egerer, M., and Satchell, K. J. (2015) Vibrio cholerae MARTX toxin heterologous translocation of beta-lactamase and roles of individual effector domains on cytoskeleton dynamics. *Mol Microbiol***95**, 590-604
69. Welch, R. A., Dellinger, E. P., Minshew, B., and Falkow, S. (1981) Haemolysin contributes to virulence of extra-intestinal E. coli infections. *Nature***294**, 665-667
70. Mackman, N., and Holland, I. B. (1984) Functional characterization of a cloned haemolysin determinant from E. coli of human origin, encoding information for the secretion of a 107K polypeptide. *Mol Gen Genet***196**, 129-134
71. Korotkov, K. V., Sandkvist, M., and Hol, W. G. (2012) The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol***10**, 336-351
72. Korotkov, K. V., Gonen, T., and Hol, W. G. (2011) Secretins: dynamic channels for protein transport across membranes. *Trends Biochem Sci***36**, 433-443
73. Lu, H. M., and Lory, S. (1996) A specific targeting domain in mature exotoxin A is required for its extracellular secretion from Pseudomonas aeruginosa. *EMBO J***15**, 429-436
74. Abrusci, P., McDowell, M. A., Lea, S. M., and Johnson, S. (2014) Building a secreting nanomachine: a structural overview of the T3SS. *Curr Opin Struct Biol***25**, 111-117

75. Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E., and Aizawa, S. I. (1998) Supramolecular structure of the Salmonella typhimurium type III protein secretion system. *Science***280**, 602-605
76. Demers, J. P., Habenstein, B., Loquet, A., Kumar Vasa, S., Giller, K., Becker, S., Baker, D., Lange, A., and Sgourakis, N. G. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nat Commun***5**, 4976
77. Bundock, P., den Dulk-Ras, A., Beijersbergen, A., and Hooykaas, P. J. (1995) Trans-kingdom T-DNA transfer from Agrobacterium tumefaciens to Saccharomyces cerevisiae. *EMBO J***14**, 3206-3214
78. Fronzes, R., Christie, P. J., and Waksman, G. (2009) The structural biology of type IV secretion systems. *Nat Rev Microbiol***7**, 703-714
79. Babic, A., Lindner, A. B., Vulic, M., Stewart, E. J., and Radman, M. (2008) Direct visualization of horizontal gene transfer. *Science***319**, 1533-1536
80. Hamilton, H. L., and Dillard, J. P. (2006) Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Mol Microbiol***59**, 376-385
81. Backert, S., and Meyer, T. F. (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol***9**, 207-217
82. Waksman, G., and Hultgren, S. J. (2009) Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol***7**, 765-774
83. Russell, A. B., Peterson, S. B., and Mougous, J. D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol***12**, 137-148
84. Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., Osvath, S. R., Carcamo-Oyarce, G., Gloag, E. S., Shimoni, R., Omasits, U., Ito, S., Yap, X., Monahan, L. G., Cavaliere, R., Ahrens, C. H., Charles, I. G., Nomura, N., Eberl, L., and Whitchurch, C. B. (2016) Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat Commun***7**, 11220
85. Hamilton, H. L., Dominguez, N. M., Schwartz, K. J., Hackett, K. T., and Dillard, J. P. (2005) Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol***55**, 1704-1721
86. Jurcisek, J. A., Brockman, K. L., Novotny, L. A., Goodman, S. D., and Bakaletz, L. O. (2017) Nontypeable Haemophilus influenzae releases DNA and DNABII proteins via a T4SS-like complex and ComE of the type IV pilus machinery. *Proc Natl Acad Sci U S A***114**, E6632-E6641
87. Ruiz, N., Kahne, D., and Silhavy, T. J. (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol***4**, 57-66
88. Van Wielink, J. E., and Duine, J. A. (1990) How big is the periplasmic space? *Trends Biochem Sci***15**, 136-137
89. Jiang, X., Payne, M. A., Cao, Z., Foster, S. B., Feix, J. B., Newton, S. M., and Klebba, P. E. (1997) Ligand-specific opening of a gated-porin channel in the outer membrane of living bacteria. *Science***276**, 1261-1264
90. Erridge, C., Bennett-Guerrero, E., and Poxton, I. R. (2002) Structure and function of lipopolysaccharides. *Microbes Infect***4**, 837-851
91. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem***71**, 635-700
92. Pontes, F. J., Rusu, V. H., Soares, T. A., and Lins, R. D. (2012) The Effect of Temperature, Cations, and Number of Acyl Chains on the Lamellar to Non-Lamellar Transition in Lipid-A Membranes: A Microscopic View. *J Chem Theory Comput***8**, 3830-3838

93. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zahringer, U., Seydel, U., Di Padova, F., and et al. (1994) Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J***8**, 217-225
94. Colville, K., Tompkins, N., Rutenberg, A. D., and Jericho, M. H. (2010) Effects of poly(L-lysine) substrates on attached Escherichia coli bacteria. *Langmuir***26**, 2639-2644
95. Stenutz, R., Weintraub, A., and Widmalm, G. (2006) The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol Rev***30**, 382-403
96. Dotson, G. D., Kaltashov, I. A., Cotter, R. J., and Raetz, C. R. (1998) Expression cloning of a Pseudomonas gene encoding a hydroxydecanoyl-acyl carrier protein-dependent UDP-GlcNAc acyltransferase. *J Bacteriol***180**, 330-337
97. Kelly, T. M., Stachula, S. A., Raetz, C. R., and Anderson, M. S. (1993) The firA gene of Escherichia coli encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase. The third step of endotoxin biosynthesis. *J Biol Chem***268**, 19866-19874
98. Bartling, C. M., and Raetz, C. R. (2009) Crystal structure and acyl chain selectivity of Escherichia coli LpxD, the N-acyltransferase of lipid A biosynthesis. *Biochemistry***48**, 8672-8683
99. Babinski, K. J., Ribeiro, A. A., and Raetz, C. R. (2002) The Escherichia coli gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *J Biol Chem***277**, 25937-25946
100. Chng, S. S., Gronenberg, L. S., and Kahne, D. (2010) Proteins required for lipopolysaccharide assembly in Escherichia coli form a transenvelope complex. *Biochemistry***49**, 4565-4567
101. Okuda, S., Freinkman, E., and Kahne, D. (2012) Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. *Science***338**, 1214-1217
102. Chiba, A., Sugimoto, S., Sato, F., Hori, S., and Mizunoe, Y. (2015) A refined technique for extraction of extracellular matrices from bacterial biofilms and its applicability. *Microb Biotechnol***8**, 392-403
103. Wang, S., Liu, X., Liu, H., Zhang, L., Guo, Y., Yu, S., Wozniak, D. J., and Ma, L. Z. (2015) The exopolysaccharide Psl-eDNA interaction enables the formation of a biofilm skeleton in Pseudomonas aeruginosa. *Environ Microbiol Rep***7**, 330-340
104. Harmsen, M., Lappann, M., Knochel, S., and Molin, S. (2010) Role of extracellular DNA during biofilm formation by Listeria monocytogenes. *Appl Environ Microbiol***76**, 2271-2279
105. Whitchurch, C. B., Tolker-Nielsen, T., Ragas, P. C., and Mattick, J. S. (2002) Extracellular DNA required for bacterial biofilm formation. *Science***295**, 1487
106. Halder, S., Yadav, K. K., Sarkar, R., Mukherjee, S., Saha, P., Haldar, S., Karmakar, S., and Sen, T. (2015) Alteration of Zeta potential and membrane permeability in bacteria: a study with cationic agents. *Springerplus***4**, 672
107. Panja, S., Aich, P., Jana, B., and Basu, T. (2008) Plasmid DNA binds to the core oligosaccharide domain of LPS molecules of E. coli cell surface in the CaCl2-mediated transformation process. *Biomacromolecules***9**, 2501-2509
108. Bullitt, E., and Makowski, L. (1995) Structural polymorphism of bacterial adhesion pili. *Nature***373**, 164-167
109. Ibanez de Aldecoa, A. L., Zafra, O., and Gonzalez-Pastor, J. E. (2017) Mechanisms and Regulation of Extracellular DNA Release and Its Biological Roles in Microbial Communities. *Front Microbiol***8**, 1390
110. Azam, T. A., and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J Biol Chem***274**, 33105-33113

111. Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol***181**, 6361-6370
112. Suryanarayana, T., and Subramanian, A. R. (1978) Specific association of two homologous DNA-binding proteins to the native 30-S ribosomal subunits of Escherichia coli. *Biochim Biophys Acta***520**, 342-357
113. Melaugh, G., Hutchison, J., Kragh, K. N., Irie, Y., Roberts, A., Bjarnsholt, T., Diggle, S. P., Gordon, V. D., and Allen, R. J. (2016) Shaping the Growth Behaviour of Biofilms Initiated from Bacterial Aggregates. *PLoS One***11**, e0149683

114. Anderson, G. G., and O'Toole, G. A. (2008) Innate and induced resistance mechanisms of bacterial biofilms. *Curr Top Microbiol Immunol***322**, 85-105
115. Hammel, M., Amlanjyoti, D., Reyes, F. E., Chen, J. H., Parpana, R., Tang, H. Y., Larabell, C. A., Tainer, J. A., and Adhya, S. (2016) HU multimerization shift controls nucleoid compaction. *Sci Adv***2**, e1600650
116. Koli, P., Sudan, S., Fitzgerald, D., Adhya, S., and Kar, S. (2011) Conversion of commensal Escherichia coli K-12 to an invasive form via expression of a mutant histone-like protein. *MBio***2**

CHAPTER 2

General Materials and Methods and specific Methods for result section

Contents

General materials and methods

2.1. Materials

2.1.1. Bacterial strains and plasmids used for cloning of genes and expression of proteins.

Strains	Genotype	Used for
1. *E.coli* XL1 Blue	recA1 endA1 gyr96(nal^R) thi-1 hsdR17(rk-mk+) supE44 relA1 lac [F'proABlacIqΔ(lacZ)M15 Tn10(Tetr)] • Derived from E.coli K12 strain • Nalidixic acid resistant • Tetracycline resistant (from F plasmid)	• Used as cloning host for all these constructs. • As expression host for some genes cloned in pQE-30 vector.
2. *E. coli* BL21star(DE3)pLyS	F^-ompT[lon]gal dcmhsdS_B($r_B^-m_B^-$)λ(DE3)pLysS(cm^R)rne131 • An E.coli B strain • T7 RNA polymerase gene carrying λprophage DE3 • IPTG inducible lac UV5 promoter • pLysS plasmid encodeschloramphenicol resistant gene. • T7 phage lysozyme (inhibitor for T7 polymerase) which reduces expression from transformed T7 promoter containing plasmid when not induced. • Mutation in RNaseE gene	Used as expression host for all gene cloned in pET vectors.

	(involved in mRNA degradation) offering enhanced mRNA stability for protein expression.	
3.Rosetta (DE3)	These cells are derived from *E.coli* Bl21 cells and are used for enhanced production of eukaryotic genes. The genotype of these cells is F^- *ompT* *hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R). • Contains DE3 which means it is lysogenic for chromosomal copy of λ DE3 which encodes for T7 polymerase under the control of UV5 promoter. • These strains supply tRNA forAGG,AGA,AUA,CUA,CCC, GGA onchloramphenicol plasmids. • These plasmids are suitable for protein expression under IPTG induction although they are deficient in pLyS.	Used for the expression of Archaeal genes.

2.1.2. Chemicals and Kits

Reagents used in this study were of analytical grade, and obtained from commercial sources. Restriction modification enzymes were purchased from NEBand Fermentas. Protein molecular weight markers were from Pierce. Plasmidminiprep kit, gel extraction and PCR clean up, Ni-NTA agarose beads were purchased form Qiagen,USA. All other fine chemicals were obtained from Sigma,USA.

2.1.3. Media

Luria Bertani Broth(LB)

Components	Amount of each components per 1L
Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Agar(For LB agar plates)	2 %
pH	7.4
Total volume	1L

The media was sterilized by autoclaving(15 psi and 121^{o}C for 15 minutes).

2.1.4. Antibiotics

Kanamycin,Chloramphenicol, Tetracycline and Ampicillin used in this study were procured from sigma USA. The 1000X stock of these antibiotics was prepared as follows

Antibiotic	1000X stock concentration
Ampicillin	100 mg/ml in water
Chloramphenicol	35 mg/ml in methanol
Tetracycline	12.5 mg/ml in 70% ethanol
Kanamycin	25 mg/ml in water

Stock of each antibiotics were sterilized by filter sterilization using 0.22 µM Millipore filters. Stocks were stored at -20^{o}C.

2.1.5. Buffers for molecular biology work

2.1.5.1. Buffer for making chemical competent cells

Calcium chloride	60 mM
Glycerol	15% v/v
PIPES	10 mM
pH	7.0

The solution was sterilized by passing through 0.22 µM filter followed by autoclaving. The solution was stored at 4^{o}C.

2.1.5.2. 6X DNA gel loading buffer(In deionized water)

Bromophenol blue	0.25%
Glycerol	30%

2.1.5.3. 50X TAE

Components	Amount per 1L
Tris.Cl	242 g
Glacial acetic acid	57.1ml
0.5 M EDTA (pH 8.0)	100 ml
pH	8

2.1.5.4. Ethidium bromide stock solution (1% w/v)

Ethidium bromide	0.1g
Deionized water	10 ml

The stock solution was stored in amber colored MCT and stored at 4^{o}C.

2.1.5.5. TE buffer(Tris EDTA buffer in deionized water)

Tris.Cl(pH 8.0)	10 ml
EDTA	1 mM

2.1.6. Buffers and solutions for SDS-PAGE

2.1.6.1. Acrylamide

Acrylamide	30 g
N.N'-Methylene bisacrylamide	0.8 g
Total Volume	100 ml

2.1.6.2. Upper Tris (4X), pH 6.8

Tris	6.06 g
10%SDS	4 ml
pH	6.8
Total volume	100 ml

2.1.6.3. Lower Tris (4X), pH 8.8

Tris	18.17 g
10% SDS	4 ml
pH	8.8
Total volume	100 ml

2.1.6.3. 5X sample loading buffer

Tris.Cl(pH6.8)	0.15 M
SDS	5 %
Glycerol	25 %
Beta-mercaptoethanol	12.5 %
Bromophenol blue	0.06 %
Total volume	100 ml

2.1.6.4. Laemmli buffer (Laemmli, 1970)

Tris buffer	3.00 g
Glycine	14.4 g
SDS	1 g
Total volume	1000 ml

2.1.6.4. Gel staining solution

Methanol	40 %
Glacial acetic acid	10 %
Coomassie brilliant Blue R-250	0.1 %
Deionized water	50 %
Total volume	100 ml

2.1.6.5. Gel destaining solution

Methanol	40 %
Glacial acetic acid	10 %
Deionized water	50 %
Total volume	100 ml

2.2. Methods

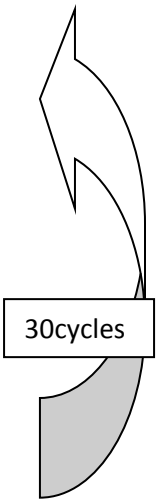
2.2.1. Polymerase chain reaction

Standard PCR reaction components

		Supplier	Stock concentration	Final concentration
1	Template		Variable	Depends upon type of DNA For plasmid DNA 1ng-1pg And for genomic DNA 1ug
2	Forward primer	IDT,Sigma	10 mM	0.5 mM
3	Reverse primer	IDT,Sigma	10 mM	0.5 mM
4	dNTPs	NEB	100 mM	250 µM
5	Polymerase(Taq,Vent,Deep Vent,Phusion	NEB	2units/µl	0.02 units/µl
6	Buffer	NEB	10X and 5X for Phusion	1 X
7	$MgSO_4$(optional)	NEB	10 mM	2-14 mM

All the PCR reactions were carried out in Eppendorf PCR machine with the following PCR reaction.

	Steps	Temperature	
1.	Initial denaturation	95 °C	5 minutes for colony PCR and 1-2 minutes for genomic and plasmid DNA
2.	Denaturation	95 °C	30 seconds
3.	Annealing	50-65 °C depends upon primers GC content	30-45 seconds for all polymerases other than Phusion polymerase which requires only 20 seconds
4.	Extension	72 °C	1kb/30seconds for Phusion polymerase and 1kb/minute for other polymerases
5.	Final extension	72 °C	10 minutes



2.2.2. Agarose gel electrophoresis

The separation of DNA fragments was done on agarose gel. The polysaccharide gel was formed by taking different concentrations (0.8 to 1.5%) of agarose for longer and shorter DNA. Gel was prepared by dissolving required amount of Agarose in 1XTEA buffer. Ethidium bromide (0.5 µg/ml) was supplemented for visualizing under UV trans-illuminator. 6X loading was added to the sample with the final concentration of 1X. Electrophoresis was carried out in 1XTEA buffer

at 8V/cm or 90 volts. 50, 100 or 1000 base pairs ladder wereused for calculating sizes of DNA fragments.

2.2.3. Purification of desired DNA bands from agarose gel

After electrophoresis, the DNA fragments were visualized under trans-illuminator and the desired band was excised out. Qiagen gel extraction kit was used to extract DNA from excised pieces of agarose gel. The steps for the protocol are as following:

1. **Solubilization:** Excised gel was weighed and dissolved in Qiagen's solubilization and binding buffer, which is named as QG buffer (300µl/100mg of gel piece). Incubation was done at 55°C till the gel is completely dissolved with intermittent mild mixing.
2. **Binding:** The dissolved agarose solution contacting DNA was then poured onto the QIAquick spin column (provided with kits by the manufacturer) to allow the adsorption of DNA onto the silica gel matrix.
3. **Washing:** The columns were washed with wash buffer PE which contains ethanol to remove impurities.
4. **Elution:** Finally the adsorbed DNA was eluted either using elution buffer which is a Tris buffer or autoclaved distilled water.

2.2.4. Direct PCR purification

The PCR product giving neat band without contaminants on DNA was directly purified from PCR reaction mixture using Qiagen's PCR cleanup kit. This procedure includes the following steps:

1. **Binding:** 3 volumes of column bindingbuffer was added to PCR reaction and loaded onto the DNA binding columns.
2. **Washing:** washing was done with wash PE wash buffer containing ethanol to remove impurities.
3. **Elution:** The elution was done either with Tris buffer or autoclaved distilled water.

Quantitation of DNA

The eluted DNA after gel extraction or plasmid purification or PCR cleanup was quantitated visually after running on agarose gel and then visually comparing band intensities with marker or by using Nano drop spectrophotometer. The Nano drop was done by putting 1µl of DNA on probe after doing 100% transmittance using solvent in which the DNA was eluted by measuring absorbance at 260nm.An absorbance of 1 at 260nm is considered equivalent to 50ng/ul of double stranded DNA(2). The purity of eluted DNA was ascertained by looking at OD_{260}/OD_{280}. Pure DNA has this ratio around 1.8-2.0.

Restriction digestion

After quantitation, the PCR and the plasmid DNA were digested using specific restriction enzymes purchased from Fermentas. The following reaction was set was restriction digestion:

1.	Template	200 ng of PCR product and 1000ng of plasmid DNA
2.	Fast digest buffer(10X)	To the final concentration of 1X
3.	Restriction enzymes	1µl each for the given concentration of DNA
4.	Water	To make up final volume

The digestion reaction was incubated at 37°C for not more than 30 minutes. The digested product was either PCR clean up or run on agarose gel for purification.

2.2.5. Ligation

The digested vector and insert were ligated by taking vector to insert ration of 1:3. The following reaction was set up:

1.	Vector	50ng

2.	Insert	xng(Depend upon insert size)
3.	Buffer(10X in case of T4 DNA ligase and 5X in Quick ligase	1X
4.	Ligase	1µl
5.	Water	For volume make up

The amount of insert was calculated using this formula:

$$\text{Amount of insert x (ng)} = \frac{\text{Amt. of digested vector(ng)} \times \text{molar ratio (insert: vector,3:1)} \times \text{insert size(bp)}}{\text{Plasmid size (bp)}}$$

The reaction was incubated at 24°C or 16°C for 2 and 12 hours, respectively. the reaction time for Quick ligase was 25-30 minutes.

2.2.6. Preparation of *E.coli* competent cells

1. A single colony of *E.coli* cells from LB agar plate was inoculated into 5ml of LB and grown overnight.
2. On the next day, the secondary culture was set up by re-inoculating 200ml of LB to the final concentration of 1% and was grown till the early log phase (O.D. at 600nm of 0.3-0.4)
3. The cells were chilled on ice for 15 minutes and 50ml of these cells centrifuged at 1600x g for 7 minutes at 4°C in pre-chilled centrifuge bottles. Cells were kept on ice during all subsequence competent cells preparation steps.
4. The cell pellet was re-suspended in 20ml of pre-chilled 60mM calcium chloride solution, after which they were again centrifuged at 1100xg for 5 minutes at 4^{o}C.

5. Supernatant was again discarded and step 4 was again repeated.
6. Again, the pellet was dissolved in 20ml of calcium chloride solution but this time cells were left on ice for 30 minutes.
7. The cells were again centrifuged at 1100x g for 5 minutes at 4^{o}C and supernatant was discarded after which the cells were re-suspended in 4ml of ice- cold calcium chloride solution.
8. Finally the aliquots of 100µl were made and stored at -80^{o}C for further use.

2.2.7. Transformation

1. Competent cells were thawed on ice for 15 minutes.
2. Ligation mixture was added and left for adsorption on calcium chloride coated competent cells for 15-20 minutes on ice.
3. Heat shock was given by heating cells at 42^{o}C for 90 seconds in water bath.
4. After giving heat shock, 1ml of sterile LB was added to these cells and left for growing at 37^{o}C for 45-50 minutes in incubator shaker set at 220rpm.
5. After incubation, cells were pelleted down by centrifuging at 5000rpm for 3 minutes.
6. Pelleted cells were re-suspended in 100µl of fresh LB media and spread on LB agar plates containing desired antibiotics selection.
7. Plates were incubated at 37°C overnight and observed for colonies on the next day.

2.2.8. Plasmid DNA purification

For plasmid DNA purification, Qiagen miniprep kit was used. Standard protocol given by the manufacturer was used. The steps are as following:

1. Cell growth: Few of the transformants from the plates were picked and inoculated into LB media overnight with appropriate antibiotics.
2. Pelleting and re-suspension: cells were pelleted in the same MCT by taking 5ml of grown bacterial culture and finally re-suspending in 250µl of P1 buffer. P1 buffer is kept at 4^{o}C as it contains RNase.

3. Lysis: Cells were lysed by adding 250 µl of P2 buffer (lysis buffer) and gently mixing by inverting up and down.
4. Neutralization: 350µl of N3 buffer (neutralization buffer) was added and mixed gently. The precipitated solution wascentrifuged at 13000rpm for 15-20 minutes.
5. Loading: Clear supernatant containing plasmid DNA is loaded over DNA binding column and spun at 13000 rpm for 1 minute.
6. Washing: Column was washed with 750µl of PE wash buffer by spinning at 13,000 rpm for1 minute. An empty spin of 2 minutes was given to dry off residual ethanol of PE buffer.
7. Elution: Finally the DNA was eluted with 10mM Tris buffer, pH 8.5 or autoclaved distilled water.

2.2.9. Screening of the transformants

Two step screening of transformants before sending for sequencing was done.

1. **Colony PCR.** Individual colonies were picked and suspended into the PCR reaction mixture to a desired final concentration of each PCR components in PCR tubes. For colony PCR, Taq polymerase (from NEB) was used. Vector specific primers (T5/T7 promoter forward and T5/T7 terminator reverse) for pQE and pET vectors were used for amplification. The amplified PCR product was run on agarose gel and the size was measured by comparing with DNA ladder. The positive clones were further verified by restriction digestion.

2. **Restriction digestion:** Following the plasmid isolation from the culture of selected clones, the plasmid was checked for the excisionof the amplified DNAby digesting with the restriction enzymes. The digested plasmid was run on agarose gel and the fallout was compared with DNA ladder and the size of the gene inserted between those restriction sites.

2.2.10. Expression of recombinant protein in *E.coli*

For the expression of the gene from the pET series vectors the plasmid was transformed into the BL21star (DE3) plysS cells and the gene cloned in pQE vectors were transformed into the *E. coli* XL1-blue or M-15 cells depending upon the toxicity of the protein. Protein expression check includes following steps:

1. **Primary culture:** cells were inoculated in 5ml of primary culture supplemented with required antibiotics and incubated at 37^{o}C and 220rpm.
2. **Secondary culture:** 1% of the overnight grown primary culture was inoculated into the fresh LB media. Cells were allowed to grow till mid log phase (O.D. of 0.6 at 600nm) and were induced with 1mM IPTG depending upon the type of expression system used. The cells were left for 4-6 hours after induction with IPTG.
3. **SDS-PAGE:** Harvested cells were boiled for 5 minutes in 50µl of SDS-PAGE sample loading buffer. Then the samples were analyzed by running on the SDS-PAGE. The over expressed band among thousands of proteins of *E.coli* with required size represents the protein of interest.

To check whether theprotein is coming into the soluble fraction, 1 ml cells were pelleted and lysed in non-denaturing buffer. After sonication, cell debris was settled down by centrifuging at high speed.10µl supernatant, after mixing in sample loading dye and boiling at 99^{o}C for 2 minutes, was run on SDS-PAGE and analyzed for desired band.

2.2.11. Protein purification by Ni-NTA agarose beads purchased from Qiagen

All theproteins cloned contain either N or C terminal 6X histidine tag and were purified using Ni-NTA affinity chromatography with the steps described below:

1. **Lysis:** Secondary culture of the protein expressing cells were pelleted by spinning at 8000rpm for 10 minutes at 4^{o}C.

2. Pellet was lysed in lysis buffer which depends upon type of protein to be purified.
3. **Loading:** Cell debris was removed after centrifugation at 11500 rpm for 45 minutes at 4^{o}C and the clear supernatant was loading on Ni-NTA column pre-equilibrated with lysis buffer.
4. **Washing:** In order to remove the non-specific protein bound to the column,washing was done with a buffer containing very less concentration of imidazole.
5. **Elution:** Immobilized protein was eluted with elution buffer containing an imidazole concentration more than that of wash buffer.
6. SDS-PAGE- Eluted fractions were analyzed for purity by running on SDS-PAGE.

2.2.12. Glycerol stock preparation

For glycerol stocks, 1500 µl of overnight grown bacterial culture was mixed with 500µl of 60% glycerol and stored at -80^{o}C for further use.

2.3. Difference absorption spectroscopy

The interaction between LPS and HU-A RFP was detected by performing difference absorption spectroscopy. The experiment was performed on Cary 50 UV-Visible spectrophotometer, using a tandem absorption quartz cuvette containing two compartments of path length approximately 0.45cm separated by a thin wall of 0 .1cm so that the total path lengths of the whole cuvette becomes 1cm. 10 µM of Tag-HU-A-RFP was taken in one compartment, and 2mg/ml LPS (Lipopolysaccharide) in PBS in the other compartment.All the reactants were kept in PBS(pH7.4).After collecting the baseline the cuvette was inverted to let both the components mix with each other and inverting again to allow mixed solution to fall back in roughly equal volume in the each compartments. The absorption scan was again performed against the collected baseline of 200 to 650 nm.

2.4. Fluorescence spectroscopy

Fluorescence spectroscopy was carried out on Cary Varian Eclipse Fluorimeter. For the fluorescence quenching experiments of intrinsic Tryptophan fluorescence the excitation was performed at 295nm and the emission spectra was collected between 305-400nm with setting excitation and emission slits at 5/5nm.The quenching experiments were performed on 0.7mg/ml HU-B mut-1 and Mut-3 containing

Tryptophan in place of phenyalnine and 47 and 79 amino acid positions with increasing concentration of LPS starting form 0.14 mg/ml till 0.56mg/ml. All the reactions were performed in PBS (pH7.4).

2.5 Fluorescence anisotropy

The binding of HU to bacterial cell surface LPS was studied by using fluorescence anisotropy of HU fused with fluorescence protein after mixing with bacterial cells. For this experiment to the 50 µl of 10nM HU-A-RFP and HUB-Venus,10 µl of bacterial cells with O.D 1.43 were added successively. Fluorescence anisotropy was measured on Horibba fluoremeter. The excitation wavelength was set at 515nm and the emission spectra were scanned for 540 to 600nm wavelength by setting excitation and emission slit at 5 nm.

2.6. Flow cytometery

Bacterial cells clumping experiment and the fluorescence associated with clumped cells was done on BD acuri C6 instrument. The cells for these experiments were prepared by growing XL-1Blue bacterial cells overnight in LB media. After overnight incubation secondary culture was setup by inoculating fresh LB media with 1% of primary culture. The cells were grown to the OD of 0.6 at 600nm. From this culture 1ml was taken in 1.5ml MCT and centrifuged at 13000 rpm for 1 minute. After discarding supernatant the pellet was re-suspended in 1ml PBS. This step was repeated for 3 times to wash off all media components bound to bacterial cells. Finally the pellet was dissolved in 500 µl PBS (pH7.4).For bacterial cells clumping experiments 10 µl of bacterial cells from 500 µl re-suspended pellet was incubated with 50ul of 30uM HU-A and HU-B proteins dialysed against the same PBS buffer and kept for incubation at room temperature for half an hour. After incubation cells were passed through the flow cytometry machine. The extent of clumping was analysed by looking at forward (FSC) and side scatter (SSC) detector. For looking at the fluorescence associated with bacterial clumps the same experimental procedure was followed apart from HU used were fused with RFP and Venus proteins. After incubation the unbound fluorescent protein was removed by pelleting and re-suspending final pellet in PBS. RFP fluorescence was measured by looking at FL-2 detector and for Venus FL-1 detector was used.

2.7. Fluorescence microscopy

The fluorescence associated with bacterial cell surface upon exogenous addition of HUA-RFP and HUB-Venus was viewed under Olympus tabletop confocal microscope used under fluorescence conditions without sectioning. The sample preparation for microscopy experiments was same as that of flow cytometry experiments. The slide preparation was done by dropping 5ul of bacterial cells on glass slide.

After spreading with slide the cells were air dried and finally covered with cover slip. The cells were visualized under 100X and for RFP with an excitation maxima of 560nmnm() laser line and for Venus 520nmlaser line was used.

2.8. Gel filtration chromatography and protein subunits dissociation experiments.

Gel filtration chromatograpy was performed on Akta Purifier 10 workstation (GE healthcare), connected to 500ul loop using superdex 75 [10/300GL(Catalog No. 17-5174-01, GE)] or superdex 200 [10/300GL(Catalog No. 17-5175-01, GE)] equilibrated with the buffer against which the protein was dialysed.The sample loading was doen using 500ul injection(in inject mode). Auto zero UV was done before loading the sample.After sample injection, the remaining run was continued with the load mode. The fractions of 1ml were collected . The flow rate was set at 5ml/min. To calculate the molecular weight of eluted protein its elution volume ml was compared with the standard curve generated for each columns after running known molwecular weight markers. The calibration curves for superdex 75 and 200 are shown in the figure below. .5mg/ml protein was incubated overnight in different molar concentration of Urea and Guanididum hydrochloride ranging from .5 to 6 M in PBS(pH.-7.4). On the following day the incubated protein samples were run on the Superdex 75 column pre-equilibrated with denaturant concentration buffer.

2.9. Circular Dichrosim(CD) spectroscopy

Circular dichroism measurement was done on MOS -500 CD spectrophotometer (Biologic France) using a quartz cuvette of 1mm path lenth(Sterna scientific, UK). The slit width was 2nm for most of the experiment depending upon the values of HV. Data was collected in the form of raw ellipticity in millidegree (mdeg) and was converted into mean residual ellipticity (MRE), using following formula. The protein concentration used for CD experiments was 0.2mg/ml. The thermal study experiments with E.coli-Thermus thermophilus DNA binding domain were done on Applied photophysics Chirascan CD instrument with same cuvelle of same path length.

$[\Theta]^{MRE} = [\Theta_{obs} \text{ X MRW X } 100]/ [\text{concentration (mg/ml) X path length (cm)}]$

MRW = mean residual weight (Total molecular weight of the protein/ Total number of amino acids)

$[\Theta]^{MRE}$= Mean residual ellipticity (deg cm^2dmol^{-1})

Θ_{obs} = Raw Ellipticity (deg).

2.10. Microscale thermophoresis (MST)

Microscale thermophoresis was used to study interaction between HUA-RFP and LPS. The interaction between the molecules is measured by the measuring alteration in migration in solution upon application of IR laser. This system uses Infrard laser for heating capillaries locally and LASER used for the excitation of the chromophore. The experiments were performed on Nanotemper monolith NT. 115. The results were analyzed with MO affinity analysis software provided with the instrument. The instrumental setup was made up of 16 quartz capillaries in which the concentration of analyte was kept fixed and the ligant was serially diluted. The analyte in this experiment was HUA-RFP. The concentration of analyte was10nM.The experiment was performed in PBS (pH.7.4).The first step was the addition of 10ul buffer in 16 PCR tubes except first one. Then add 10ul of 5mg/ml stock concentration LPS was added in first two PCR tubes and make a serial dilution by taking 10ul from PCR tune 2 into PCR tube 3. Repeat this step for remaining dilution process. Mix the solution gently. Finally add 10ul of 10nM HUA-RFP in all 16 PCR tubes. After mixing them properly by pipetting the tune were left at room temperature for half an hour to promote interation between LPA and HUA-RFP and to reach binding equilibrium. Then 16 MST capillaries were filled by dipping them in 16 PCR tubes contain solution for analysis. After loading capillaries in MST device the system was left to reach temperature at 25^{o}C. LED power was set to gain fluorescence signal between 500-1000 units by using red LED setting for RFP. The first step before analyzing sample is to perform a capillary scan to ensure the quality of sample in each capillary. The fluoresce yield should be same in each capillary The fluorescence scan was performed under standard conditions with measuring fluoresce for initial 5 seconds followed by heating sample and recording thermophoresis for 30 seconds and again measuring recovered fluoresce for 5 seconds. The laser power was set between 20-40%. High laser will lead to the fluorescence quenching.

2.11. Biolayer interferometry

The intercation of LPS with both wild type and loop replaced HU was done on Forte bio Biolayer interferrometry. The proteins with which the the interaction has to be studied was immobelized on Ni-NTA biosensor . The concentration of the proteins and LPS used was 6 µM and 1mg/ml respectively. The steps for performing the experiment are written in results and discussion section. For the interaction studies of *E.coli-T.thermophilus* DNA binding domain the 600 nMconcentration of the protein was loaded on Ni-NTA probe followed by interaction studies with 10 and 20 µM double stranded and single stranded DNA respectively.

2.12. Dynamic light scattering

The oligomeric status and the population size homogeneity of the protein were checked by Dynamic light scattering on Wyatt QUELS dynamic light scattering instrument. The concentration of the protein used was decided on the requirement of experiment conditions. The samples before analysis were passed through a 0.02µm.

2.13. Isothermal titration calorimetry (ITC)

ITC was used to study the intercation of loops replaced HU with 20 mer DNA. The experiments were performed on the MicroCAl ITC 200 from GE . The amount of DNA and protein used in this experiment are written in the result and discussion section. The ligand was injectd into the sample cell as pulse of 1 to 2 µl injections untill the saturation reached.

2.14. Liquid crystal

The effect on the birefringence of liquid crystal was employed for studying interaction between HU and LPS. The liquid crystals were coated with LPS, whose orderedness got affected upon the interaction with HU molecules. The extent of disorderedness was visualized through optical polarized microscope.

2.15. Electrophoresis mobility shift assay (EMSA)

The activity in terms of binding ability with double and single stranded DNA was checked by incubating a fixed concentration of 20-mer DNA with an increasing concentration of protein and running the samples on 2% agarose gel.The construction of double stranded 20 mer DNA was done by mixing equal volume from a 100µM stock solution followed by heating at 40^{o}C followed by cooling down the mixture. A DNA concentration of 4µM in case of dsDNA and 8µM for ssDNA was used for the reaction. As the fusion protein lacks tryptophan, the protein concentration was estimated by Bradford assay. In this experiment, the protein was taken randomly by increasing its concentration by a µleach time. The protein and DNA after mixing were kept at room temperature for 30 minutes to maximize the interaction between them. The binding of protein with DNA was established by the appearance of smeared band on agarose gel. The sequence of the 20mer oligomer used for DNA binding studies is shown in the following table.

Strand 1	GTTCAATTGTTGTTAACTTG
Strand 2	CAAGTTAACAACAATTGAAC

Electrophoretic mobility shift assay was done by incubating proteins with single stranded and double stranded DNA for 30 minutes at room temperature followed by separation on 0.6% agarose gel.

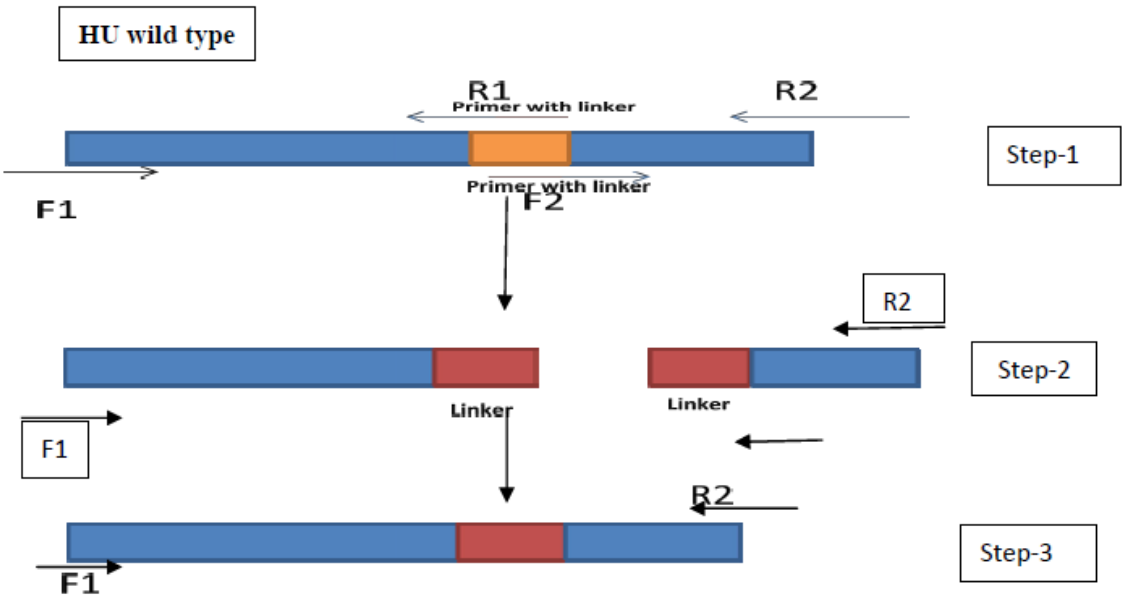
2.16. Primers used in this theis for making clones

Loop replaced HU-A	
Forward Primer	5-ATATATGGATCCATGAACAAGACTCAACTGATTGATGTAATTGC-3
Linker forward	5-TCAGGTGGAGGAGGTTCAGGCGGAGGTGGCAGCAACGTACCGGCATTTGTTTC-3

Linker Reverse	5-GCTGCCACCTCCGCCTGAACCTCCTCCACCTGACACTTTGAAGGTACCGAAACC-3
Reverse Primer	5-TGATATAAGCTTTTACTTAACTGCGTCTTTCAGTGCCTTGCC-3
Loop replaced HU-B	
Forward primer	5-ATATATGGATCCGTGAATAAATCTCAATTGATC-3
Linker Forward	5-TCAGGTGGAGGAGGTTCAGGCGGAGGTGGCAGCAAAGTACCGAGCTTCCG-3
Linker Reverse	5-GCTGCCACCTCCGCCTGAACCTCCTCCACCTGAAACGGCAAAAGTACCAAAACC-3
Reverse Primer	5-TATGATAAGCTTTTAGTTTACCGCGTCTTTCAGTGCTTTACC-3
HU lysine mutants	
HU-A, K83A	
Forward Primer	5-ATATATGGATCCATGAACAAGACTCAACTGATTGATGTAATTGC-3
Reverse Primer	5-ATATATAAGCTTTTACTTAACTGCGTCTGCCAGTGCCTTGCCAGAAAC-3
HU-A, K86A	
Forward Primer	5-ATATATGGATCCATGAACAAGACTCAACTGATTGATGTAATTGC-3
Reverse Primer	5-ATATATAAGCTTTTACTTAACTGCGTCTTTCAGTGCTGCGCCAGAAACAAATGCCGG-3
HU-B, K83A	
Forward Primer	5-ATATATGGATCCGTGAATAAATCTCAATTGATC-3
Reverse Primer	5-ATATATAAGCTTTTAGTTTACCGCGTCTTTCAGTGCTGCACCTGCACGGAAGCTCGG-3
HU-B, K86A	
Forward Primer	5-ATATATGGATCCGTGAATAAATCTCAATTGATC-3
Reverse Primer	5-ATATATAAGCTTTTAGTTTACCGCGTCTGCCAGTGCTTTACCTGC-3
DNA binding finger construction (Ecoli-Thermus thermophilus HU)	
E.coli Forward(D1 Froward)	5-ATATATCATATGGTTGGATTCGGAAGCTTC-3
T.T Reverse	
E.coli-TT Linker Forward	5-GCCCTCAAAGAGAAGGTCAAGAAGGTCCAGCTCACGGGCTTC-3
E.coli-TT Linker reverse	5-GAAGCCCGTGAGCTGGACCTTCTTGACCTTCTCTTTGAGGGC-3

T.T HU Reverse	5-ATTAATCTCGAGCTTGACCTTCTCTTTGAGGGCTTTTCC-3

2.17 Creation of Loop deleted HU variants by SOE PCR



A schematic showing the steps used for the construction of loop-replaced HU, made through SOE-PCR. The orange part of the gene in the first step represents the removable part, i.e., the original loop. The second step represents fragments of the genes containing the linker sequences (in red). The third step represents construction of the gene containing the spliced Glycine-Serine loop replacement (marked in red).

2.18. The sequence of 4-way junction DNA used

Strand 1	CCCTATAACCCCTGCATTGAATTCCTGTCTGATAA
Strand 2	GTAGTCGTGATAGGTGCAGGGGTTATAGGG
Strand 3	AACAGTAGCTCTTAATTCGAGCTCGCGCCCTATCACGACTA
Strand 4	TTTATCAGACTGGAATTCAAGCGCGAGCTCGAATAAGAGCTACTGT

2.19 Creation and purification of tandem-HU-DßF derived from stitching protein sequences forming DNA binding region together from E.coli and Thermus thermophilus

E.coli HU-A sequence

MNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHRAERT GRNPQTGKEIKIAAANVPAFVSGKALKDAVK

Strand 1GTFKVNHR

Strand 2 ANVPAFV

T. thermophilus HU sequence

AAKKTVTKADLVDQVAQATGLKKKDVKAMVDALLAKVEEALANGSKVQLTGFGTFE VRKR KARTGVKPGTKEKIKIPATQYPAFKPGKALKDKVKK

Strand 1GTFEVRKR

Strand 2 TQYPAFKP

Sequence of HU from both the organisms along with the sequences of beta strands. The sequences marked as yellow and green were fused to make a DNA binding chimera. The order of placement was *E.coli* followed by *T. thermophillus.* The sequence were fused together suing a spacer peptide linker of sequence TRKSPQREGQEGPA using splicing by overlap extension PCR methods. The primer sequences used are given in the previous sections.The sequence of the final protein was as following :

MVGFGSFKANHRAERTGRNPQTGKEIKIAAANVPAFVSGKALKDEVQTRKSPQRE GQEGPATGFGTFEVRKRKARTGVKPGTKEKIKIPATQYPAFKPGKALKDKVKKLE HHHHHH

Number of amino acids: 106

Molecular weight: 11996.31 Daltons, pI: 10.93

The assembled gene was cloned in pET-23a vector between Nde-1 and Xho-1 restriction sites and the protein was purified by Ni-NTA affinity purification.

CHAPTER 3

Results and Discussions

Content

3.0 Results and discussion

3.1 *SECTION 1: HU and its involvement in biofilm formation*

3.1.1A hypothesis relating HU to biofilms

Bacteria within biofilms are embedded in a matrix of extracellular polymeric substance (EPS), primarily consisting of extracellular DNA (e-DNA) derived from cellular secretions, cell death and lysis. A practical issue with such embedment is that negative charges decorate the surfaces of both DNA and bacteria. These negative charges create scope for charge-charge repulsions between bacteria and DNA in the biofilm, and raise the question of how bacteria coexist with DNA inside biofilms.

It is possible, of course, that there is a 'mediator' substance or'glue' which is present within biofilms (and which attenuates the repulsions between DNA and bacteria); a glue capable of binding to DNA as well as to the surfaces of bacterial cells; preferably a highly positively-charged substance capable of neutralizing the negative charges on both DNA and bacteria through the making of suitable electrostatic contacts; and lastly, a glue whose binding through electrostatic contacts happens to be facilitated by specific conformations of DNA and/or chemical configurations involving the glue itself.

In this thesis, we have begun by working to examine whether the hypothesized glue could be a protein molecule and, in particular, a DNA-binding protein molecule; even more particularly, whether it could be the nucleoid-associated protein, HU.

Therefore, in the rest of Section 3.1 of this chapter below, in sequential order, are presented the following:

(i) a hypothesis that the protein glue that binds biofilms together is the abundant DNA-binding protein, HU, which forms a mediating link between DNA, on the one hand, and the molecule called lipopolysaccharide (LPS) present on the surfaces of Gram negative bacteria, on the other hand;

(ii) results from experiments in biophysics and physical chemistry, using many different methods, which show that HU does indeed bind to LPS, in solution;

(iii) results from microscopy, cytometry and other studies to show that HU does indeed bind to the surfaces of bacterial cells, ostensibly by binding to LPS present on the bacterial cell surface;

(iv) results indicating that HU can coagulate and hold together bacterial cells, ostensibly by forming multimers that can bind both to LPS and also to DNA;

(v) results from biophysical studies providing insight into why HU has a strong subunit interface, and also into how and why HU forms stable multimers that aid in its role as a biofilm glue;

(vi) results from mutational studies involving HU, which point towards specific surface regions with binding preferences for DNA or LPS; and

(vii) results demonstrating that an engineered (redesigned) minimalistic DNA-binding domain, based on HU's DNA-binding loop,actually binds to DNA and/or to LPS.

Below, we begin with arguments supporting our hypothesis and building a case for the experiments that were performed to test whether indeed HU binds to LPS as well as to DNA.

The case for the existence of a 'molecular glue' capable of trapping bacteria in a matrix of e-DNA.DNA is negatively charged, owing to its backbone of pentose sugars linked with phosphate groups. Bacteria such as *Escherichia coli* are also negatively-charged, in that their surfaces display high densities of the molecule, lipopolysaccharide (LPS). It is not widely appreciated that LPS contains phosphate groups in association with hexose sugars in the lipid A component of LPS. However, when one considers this fact, at least theoretically, it could be argued that LPS and DNA might mimic each other in respect of each substance possessing some sugar-phosphate moieties. Since LPS imparts negative charge to bacterial surfaces, and DNA is also negatively-charged, it also stands to reason that any interaction of bacteria with e-DNA (including embedment of bacteria within e-DNA) must be facilitated by the presence of a positively-charged substance that is capable of acting as a mediator[1]. Such a mediator substance (or molecule) would need to be capable of neutralizing negative charges on both bacteria and DNA by using the positive charges on its own surface to form ionic interactions with DNA and LPS. Without such an arrangement, the physical-chemical issues inherent in embedding a negatively-charged material (such as a bacterium) within a matrix of another negatively-charged material (such as DNA) would become quite serious. Thus, through simple reasoning we can appreciate a

mediator (or mediators, in plural) with an overall preponderance of surface positive charges must exist within biofilms, and act like a glue, binding bacteria to e-DNA.

The required characteristics of a 'molecular glue' for trapping bacteria in e-DNA. Since DNA is an acid, a good candidate for a molecular glue capable of binding bacteria to e-DNA would be a basic protein displaying DNA-binding as well as the following properties : (i) the molecule would need to be capable of binding to DNA quite non-specifically, in order to locate itself everywhere upon e-DNA, rather than only at locations corresponding to specific DNA sequences; (ii) the molecule would need to be highly abundant, in order to be present in sufficient quantities to coat e-DNA at all possible locations; and (iii) the molecule would need to be capable of binding simultaneously to both LPS and to DNA, such that it could trap bacteria while in a DNA-bound state.

The protein, HU, could be an ideal molecular glue. The DNA-binding proteins, HU-A and HU-B, satisfy two of the above three criteria. Firstly, HU-A and HU-B constitute possibly the most abundant proteins in an *Escherichia coli* cell, with combined numbers being as high as 50,000 dimeric HU molecules per cell[2, 3]. Secondly, it is also well-known that HU binds nucleic acids non-specifically, i.e., with no sequence-specificity, mainly by using lysine-rich beta hairpins from each of the two subunits in each HU dimer to wrap around the minor groove of DNA[3]. A promising third possibility arises from the fact that HU does not just bind to DNA through a single type of surface region (the canonical DNA-binding beta hairpin regions). Rather, it is now known that HU also has an additional DNA-binding site, which is a non-canonical site consisting of a cluster of lysine residues on the side of each subunit in the HU dimer, away from the canonical DNA-binding beta hairpins[4]. Furthermore, HU also forms tetramers and octamers. Thus, with HU binding to itself and also to DNA through multiple sites, multiple copies of HU engages in the compaction of DNA.

The next question to consider is whether it is possible for spare DNA-binding sites on HU to additionally bind to LPS? *E.coli* contains two variants of HU, called HU-A and HU-B. In work from an earlier member of the research group (Kanika Arora), we have described engineered forms of these proteins consisting of genetic fusions of fluorescent proteins such as red fluorescent protein (tag-RFP), or Venus, with HU. We have also described engineered forms of

HU in which phenylalanine residues have been replaced by tryptophan to allow spectroscopic monitoring of HU's interactions with DNA and other molecules.

A fraction of over-expressed Venus-HU-B is found in a halo in recombinant bacterial: secretion or lysis of lysine decorated HU. Overexpression of HU fused to fluorescence protein Venus leads to appearance of excessive Venus-HU protein outside of the bacterial cells (personal communication from Kanika Arora). The secreted protein appears as a halo after immobilization around bacterial cells, which also display poorer separation after cell division, forming long filamentous assemblies. The appearance of Venus HU on the bacterial surface supports the finding of Goodman et al showing that bacteria can secrete HU through some secretion systems, irrespective of cellular lysis.

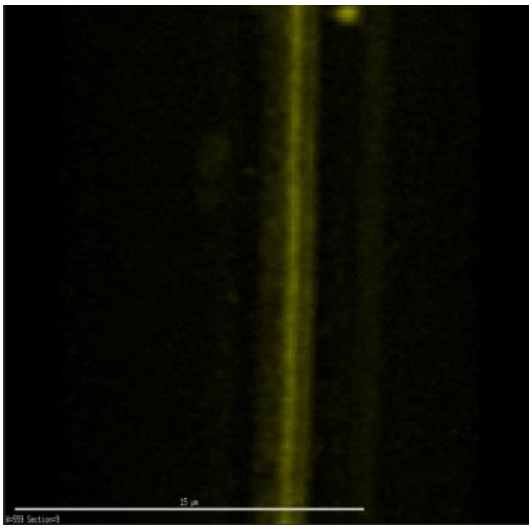


Figure 1: Showing XL-1 Blue cells overexpressing Venus-HU-B protein. The protein is mainly concentrated within the nucleoid region of the filamentous assembly of *E. coli* cells, but is also immobilized on the bacterial cell suface, giving rise to a halo of fluorescence around bacterial cells.

3.1.2. Binding of free LPS to free HU

In the experiments described below, we used both native HU and such engineered forms of HU (both HU-A and HU-B) to investigate whether LPS binds to HU, tryptophan-containing HU, RFP-HU, or Venus-HU.

3.1.21LPS binding to HU (Experiment I): A microscale thermophoresis (MST) experiment.

A description of thermophoresis :Thermophoresisis, a technique used to detect molecular interaction(s), is based on molecular 'translational' and 'diffusional' movements occurring after the heating of a solution by an IR laser for few seconds, followed by the bringing-back of the temperature to the ambient temperature, in heating and cooling cycles that last for 30, and 5, seconds, respectively. During heating, molecules diffuse away from the region of solution undergoing localized heating, and during cooling the molecules that have diffused away on account of heating happen to diffuse back into the same original region. When either the receptor or ligand in an intermolecular interaction happens to be fluorescently labelled, the return of molecules to the heated region during cooling can be followed simply by monitoring of the restoration of the fluorescence signal. Since molecular diffusion rates depend upon the hydrodynamic volume(s) of molecules, and since these are likely to be different when a fluorescently-labelled molecule interacts with another molecule, changes in diffusion rates can be used to interpret whether any interactions are occurring to slow down molecular diffusion. Furthermore, by keeping the concentration of one molecule (e.g., a fluorescently-labelled receptor) a constant and by varying the concentration of the non-fluorescently-labelled ligand, over a series of experiments, the ligand-receptor interaction can be characterized through the determination of the $k_d$from the different rates of restoration of the receptor's fluorescence, which are caused by different levels of binding of the receptor to the ligand.

In practical terms, the heating is done with the receptor and ligand present in a capillary and the behavioral differences arising due to the use of different concentrations of ligand are examined by using multiple capillaries which are simultaneously examined. Due to the use of capillaries and small volumes, the technique is called microscale thermophoresis. All capillaries have a window through which the IR laser and the fluorescence-exciting laser access the solution inside the capillary, and through which the fluorescence signal emerges and gets recorded.

***The LPS-RFP-HU thermophoresis experiment*:** In the experiment shown below in Figure 2, the binding of Tag-RFP-HU-A (receptor) to LPS (ligand) was analysed by microscale thermophoresis. The protein and LPS were incubated for 30 minutes at room temperature to

maximize the intercation between them.The protein concentration was kept a constant, at 10 nM, in all capillaries, and the LPS was serially diluted. There was 2.5mg/ml in the first capillary, 1.25 mg/ml in the second, 0.0625 mg/ml in the third capillary, and so on, using serial half-dilutions, over a total of 16 capillaries. The depletion and appearance of fluorescence in heating and cooling cycles in each capillary is dependent upon the rate of brownian motion of protein bound to the LPS. In the first few capillaries, the ligand (LPS) concentration was very high, leading to the formation of big aggregates. These big aggregates failed to migrate during the heating cycle, as is evident from the persistence of fluorescence in the first few capillaries for which data can be seen in the panel A of Figure 2. However, with decreasing ligand(LPS) concentration, the free molecules of Tag-RFP-HU-A migratefaster than the molecules of Tag-RFP-HU-A bound to LPS, for heating cycles of 30 seconds. In other words, what we see is a form of thermophoresis in which LPS-bound HU fails to migrate away upon heating, but when LPS concentrations are lower, there is some migration of HU (which, however, does not return over 5 seconds of cooling). The data plotted is integrated over multiple cycles of heating-and-coooling. The migrating free Tag-RFP-HU-A fails to return to the region from which it is depleted through heating, in the cooling cycle of 5 seconds, and this is clear from the lower levels of restored fluorescence obtained after cooling, in the capillaries with lower LPS concentrations. The migration behaviour of free Tag-RFP-HU-A, and of the RFP-HU-A-LPS complex, is evident from the panel B of Figure 2, as free RFP-HU-A fluorescenceis depleted at faster rates, as compared to the RFP-HU-A, in the presence of LPS, and this is suggestive of the formation of RFP-HU-A-LPS complexes. This differential behaviour of free and bound HU, with fluctuation in fuorescence signal, establishes that LPS binds to Tag-RFP-HUA. The $K_d$calculation of Tag-RFP-HUA interaction with LPS was difficult as LPS is a polymeric macromolecule, with no defined molecular mass.

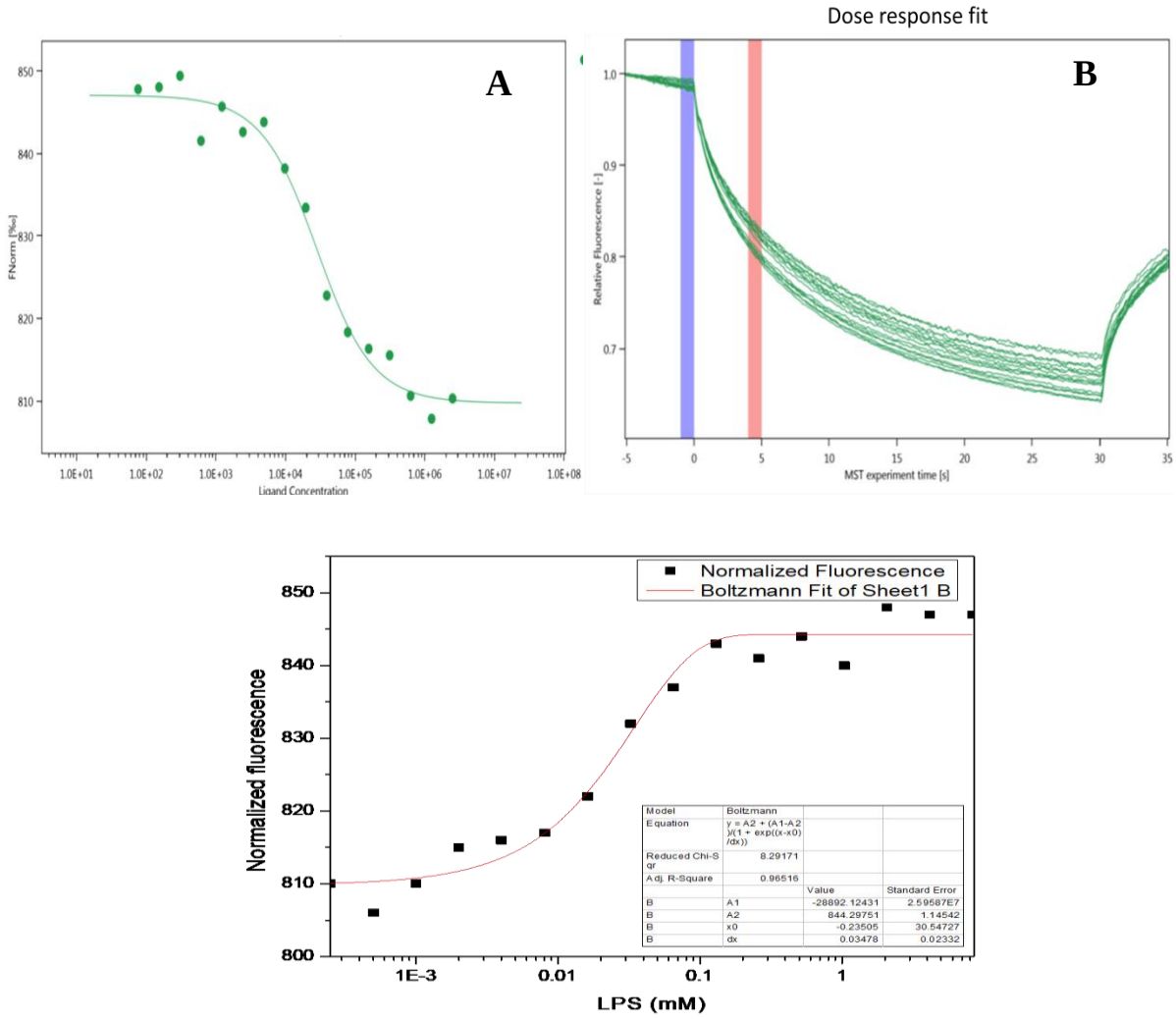


Figure 2: Panel A: Shows the values of normalized fluorescence from Tag-RFP-HU-A in the heated region, after cooling, for every capillary. ***Panel B:*** Shows the variation in fluorescence intensity coming from the brownian motion of the fluorophore Tag-RFP-HUA, after binding with LPS upon heating and cooling cycle, for all sixteen capillaries.The curves show the fluorescence signal from RFP present – at the location being observed - within each of 16 different capillaries (containing different concentrations of LPS and the same concentration of RFP-HU-A). The blue line shows the point at which the heating IRLASER is switched-on. The pink line shows the time point at which data was collected from each of the 16 capillaries (from the 16 different curves) to calculate the ratio of the fluorescence in the capillary during heating and the fluorescence prior to heating. This data was used in the other curve to calculate the diffusion of fluorescent protein as a function of presumed-interaction with varied concentrations of ligand LPS. ***PanelC:*** Fitted data to calculate a dissociation constant using origin pro by employing sigmoid, Boltzmann fitting.

The data was fitted using MO analysis software provided with MST instrument. The principle of MST (Microscale Thermophoresis) is temperature mediated depletion and recovery of fluorescence upon switching on and off of an IR (infrared Laser). The data is plotted by taking the ratio of fluorescence before switching on of the IR LASER, and after switching on of the same laser, for each 16 capillaries, and then by comparing that ratio across the 16 capillaries, which have different ligand (LPS) and fixed fluorophore (RFP-HU-A) concentrations, to calculate an apparent Kd. The final Kd was calculated by taking an apparent weight of 15000 Da for LPS. The fitting with the calculated kd has been incorporated in the final submission. The approximate kd is calculated to be about 34 µM.

3.1.2.2. LPS binding to HU (Experiment II): A difference absorption spectroscopyDAS) experiment.

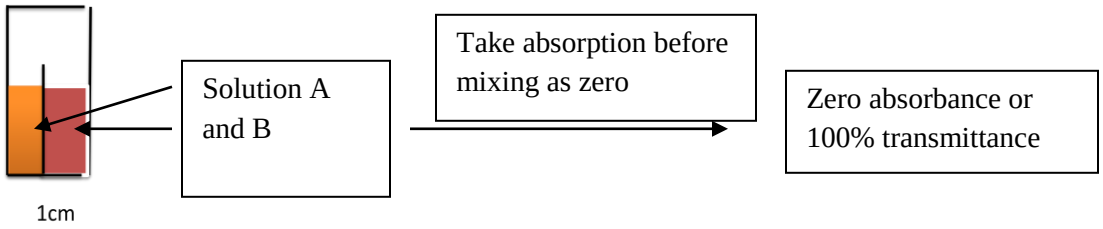
A description of difference absorption spectroscopy :Molecular interactions can affectthe absorptivity of a chromophore present on a molecule which is an interaction partner in a binary, ternary, quaternary or higher-order molecular complex. This is especially so if the chromophore happens to be a part of the region of the molecule which bears it, and which also happens to be the region involved in the intermolecular interaction. Essentially, what is seen is that the probability of occurrence of an electronic excitationtransition in an atom, or group of atoms, is affected either positively, or negatively, or not at all, by the involvement of the carrier molecule in an interaction. Thus, apositive or negative change in absorptivity of one of the binding partners, or both partners, is a reliable indicator of an interaction having occurred, whereas the absence of a change cannot be definitively interpreted.

The molecules whose interactionsare to be studied are taken in the two chambers of a tandem quartz cuvette, as shown in Figure 3. Each of the two chambers has a path length of 0.5cm. The wall between the two chambers stops short of the top of the cuvette, allowing for the contents of the two chambers to be mixed at will, through the physical inversion of the cuvette. During an experiment, when both chambers are filled to the same height (with the same volume), with one chamber filled with the receptor solution and the other with the ligand solution, a baseline spectrum for absorbance can be recorded. In this baseline spectrum, the absorptions due to the two substances, each over 0.5 cm path length of interaction with the light beam, become constituent features in the overall spectrum, both contributing to the overall absorption. Following this, the baseline spectrum is digitally ‘zeroed’ using the microprocessor and software

controls of the spectrophotometer. Then the cuvette is stoppered, inverted and the receptor and ligand solutions are allowed to mix. Since the volumes are identical, the concentration of each molecule (receptor/ligand) is effectively halved through the mixing. Thereafter the cuvette is set upright again and the mixed solution falls back into the two chambers. Now, the concentrations of both receptor and ligand have become halved and the path lengths of the solutions have become doubled, because both receptor and ligand are in both compartments in equal volumes before the mixing. According to the Beer-Lambert law, if the concentration is halved and the path length is doubled, there should be no effective change in absorbance. Recording of the spectrum against the zeroed spectrum should thus produce a flat line. However, if there is any interaction, and this affects the absorptivity of one or more chromophores, this is detected as deviations from the zero line. In a difference absorption spectroscopic experiment, the changes observed are changes brought about by mixing, resulting from changes in the electronic absorption spectra of one or both of the molecules being mixed. If there were light scattering, e.g., from aggregation caused by the mixing, this would show wavelength dependence, increasing with the fourth power of reducing wavelength. Instead, we see changes in two discrete regions corresponding to the absorption bands of aromatics (phenylalanine-centered at ~260 nm) and RFP's chromophore. The signal changes in the electronic absorption characteristics of the RFP-HU-A as a function of binding to LPS. The changes in the bands are seen in relation to the flat baseline obtained for the pre-mixing state of the tandem cuvette, before the mixing of the contents of the two compartments (LPS, and HU). The concentration of LPS and RFPHU-A used were 2 mg/ml and 10 µM respectively.

***The LPS-RFP-HU DAS experiment*:** In our experiment, we used LPS and Tag-RFP-HU-A as the ligand, and receptor, respectively, in the two compartments. The range of wavelengths used for recording of the spectrum and for 'zeroing' of the spectrum was selected to include both the peptide bond (220 nm) and aromatic (260 nm) absorptions of HU. HU lacks tryptophan and tyrosine, and only contains three phenylalanine residues, and also shows the absorption due to RFP (555 nm).The mixing was done by inverting the cuvette and taking absorbance readings repeatedly, with a brief incubation of 1.5 minutes in between scans, until the absorption spectrum changed no more, as shown in the Figure 4. As the path length was doubled after mixing, but the concentration was reduced to half, for both the Tag-RFP-HU-A and the LPS, the change of

absorptivity seen for regions of the spectrum corresponding to both Tag-RFP-HU-A (~555 nm) and HU (~260 nm and below) is indicative of changes brought by binding of LPS to Tag-RFP-HU-A. The changes in the region of~260nm correspond to altered absorptivities of the phenyalanine residues present in HU. The absorption at ~555nm corresponds to the changed environment of the RFP fluorophore. Presumably, the absorptivities of both were affected by binding of HU with LPS, suggesting that the two molecules interact.



Step2: Mixing both the solution by inverting cuvette gently

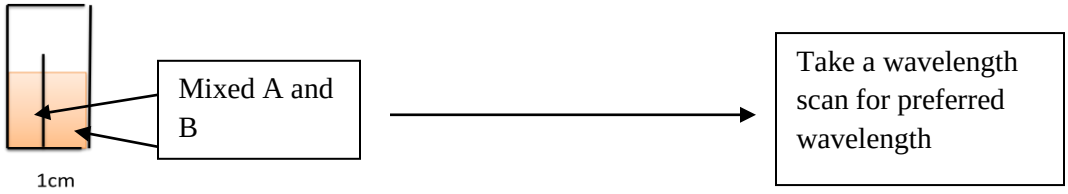


***Figure 3*:** Schematic illustrating the principle and the way in which the difference absorption spectroscopy experiment is performed.

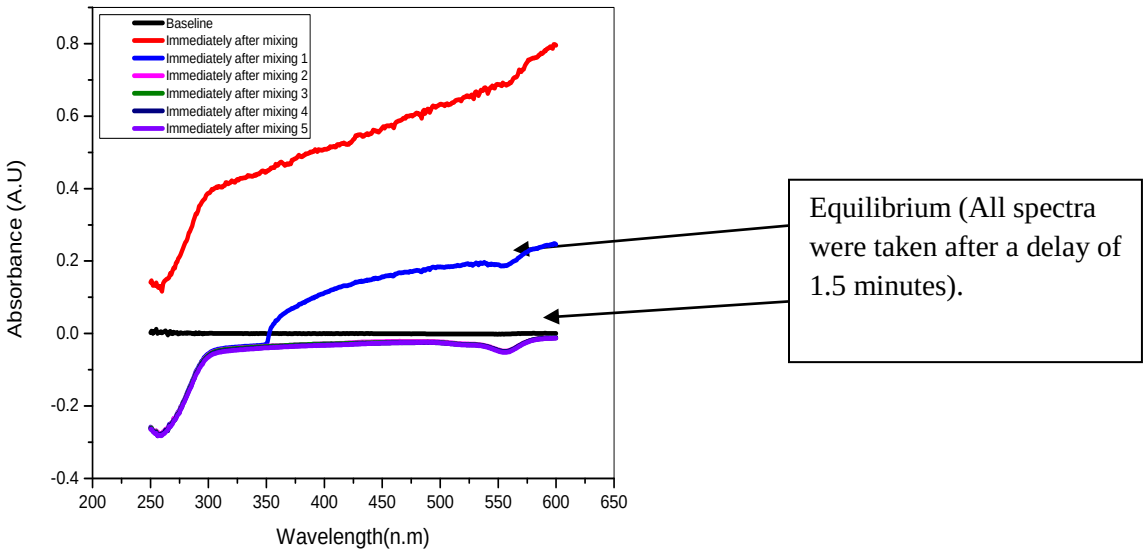


Figure 4: Absorption scans of RFP-HUA after mixing with LPS. The red curve was obtained immediately after mixing. The navy blue curve (with arrow pointing towards it) corresponds to a scan taken 1.5 minutes later. The other arrow points to a collection of overlapping scans collected thereafter at intervals of 1.5 minutes each, indicating attainment of an equilibrium state after mixing of LPS with RFP-HU. The black line is the baseline collected prior to mixing. The concentrations of RFP-HU (i.e., Tag-RFP-HU-A) and LPS used in this experiment were 10 µM, and 2 mg/ml, respectively. The negative peaks seen in the equilibrium spectra (lower arrow) are at the absorption maxima of HU protein (260 nm) and RFP fluorophore/chromophore (560 nm). The central black line is for the control (or blank) corresponding to the situation before the mixing of the samples, and the other lines correspond to a series of spectra taken post mixing of the LPS and HU solutions in the two compartments. The spectra were collected one after the other immediately after mixing, until the stabilization of the situation in the tandem cuvette suggestive of chemical and physical equilibrium.

3.1.2.3. LPS binding to HU (Experiment III):A Biolayer Interferometery(BLI) experiment

A description of Bio-layer interferometry :Biolayer interferometry is a label-free technique for the detection of molecular interactions in which one molecule is immobilized on the tip of a biosensor, while the other molecule remains in solution around the tip. Thus, opportunity exists for the binding of the molecule present in solution to the molecule present on the biosensor tip.

This leads to a thickening of the layer of molecules on the sensor tip, and thereby to a difference in the interference pattern of white light reflected from two surfaces (one the tip, and the other a reference). As molecules bind to the tip, or to whatever is already bound on the tip, this interference pattern changes, and this is monitored as a change in the apparent wavelength of the light reflected from the tip, in proportion with the attachment or detachment of molecules from the tip. The monitored changes are plotted in the form of a sensorgram.

***The LPS-HU BLI experiment*:**We studied the binding of HU to LPS using bio-layer interferometry. In this case, we used the protein, HU-B, without any fluorescent protein present in fusion. The protein, HU-B,was immobilized through its 6xHis N-terminal tag, which was allowed to bind to the Ni-NTA functionalized surface on the tip of the biosensor probe. The probe was dipped in LPS solution. The addition of layers of LPS molecules over the HU-B molecules already immobilized on the surface of the Ni-NTA probe resulted in an altered interference pattern and wavelength of reflected light. The initial binding of the HU-B protein to the Ni-NTA biosensor surface showed a response of 8 response units after a baseline collection was done with buffer as shown in Figure 5. The experiment was done in PBS buffer (pH7.4). After protein binding, the baseline was collected with HU bound biosensor, and the response curve was monitored upon addition of LPS to the solution surrounding the biosensor tip.

The bound LPS on HU becomes dissociated after dipping of the probe in the buffer solution suggesting weak associations between them. The variable molecular mass of the LPS macromolecule prevented us from calculating K_d as the exact molecular mass of ligand (HU-B) and the analyte (LPS) arerequired for this calculation, and LPS consists of molecules with variable lipid chain length.

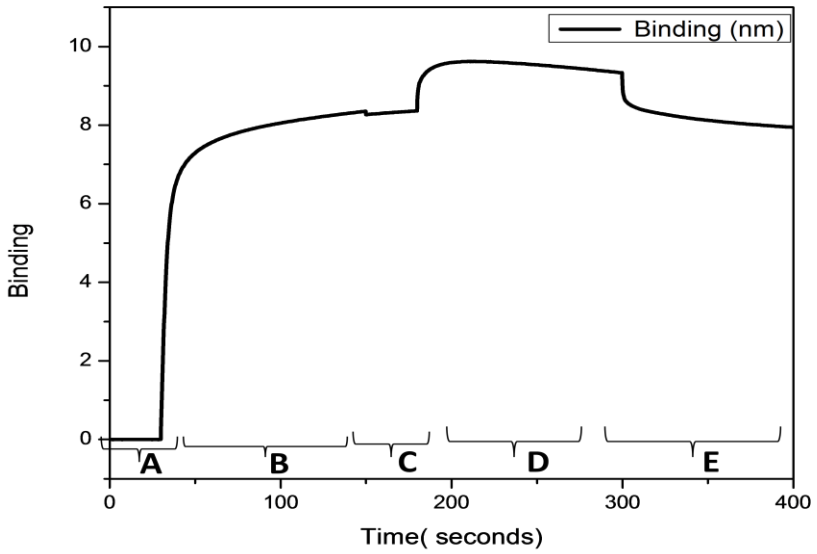


Figure 5: The sensorgram of LPS binding to HU-B determined by Bio-layer interferometery (BLI) along with the steps labelled as A to E. The Step A represents the baseline collected by dipping the Ni-NTA biosensor in PBS buffer. Step B represents association of HU-B (concentration of 6 µM) with the Ni-NTA biosensor, after dipping it in HU-B protein solution. Step C represents the baseline taken after dipping protein immobilized to the Ni-NTA probe in PBS. Step D represents estimation of LPS binding to HU-B by dipping protein immobilized upon the Ni-NTA biosensor in a solution of LPS (concentration of 1 mg/ml). Step E corresponds to dissociation of LPS from the HU-B protein after dipping of the NI-NTA containing LPS associated with immobilized protein in buffer PBS.

Typically, the association curve rises and becomes horizontal in any biolayer interferometry (BLI) or surface plasmon resonance (SPR) experiment. It does not show any sign of dissociation before the dissociation step has been deliberately initiated by ceasing the presence of the analyte. Ultimately, when the rate of association and the rate of dissociation are matched, the association curve becomes flat. However, in our observation, there is some dissociation during the association step (i.e., step D). There are only two possible explanations for this nominal dissociation occurring before the actual dissociation step is initiated : (i) One possibility is that during the association, the binding of LPS molecules is somehow affecting the interaction of the HU with the Ni-NTA (e.g., causing the bulk of the material in the complex to move away from the surface as more binding occurs, to physically accommodate the bound LPS, resulting in a change in the interferometry (note: if this were SPR, it would be a change in

refractive index); (ii) the other possibility is that the LPS binding is somehow causing the HU to dissociate from the Ni-NTA, or causing the Ni-NTA to fall away from the functionalized surface of the tip, i.e., a leaching of the Ni-NTA from the tip's surface is occurring (note: in BLI, the Ni-NTA-based HU-bound tip is inserted into a solution of LPS). We do not know which of these is occurring, but we know from an exploration of available questions on sites such as researchgate that others have also seen such behavior.

3.1.2.4. LPS binding to HU (Experiment IV): A Dynamic Light Scattering (DLS) experiment

A description of dynamic light scattering: The binding of LPS with HU was also analysed using the technique of dynamic light scattering. The core idea of this experiment was that if LPS and HU do not interact, the sizes and diameters of light scattering entities observed separately for the two solutions of molecules would continue to be separately observed when the molecules are added to each other, with no emergent, or new, entity resulting from interaction(s). On the other hand, if there were interactions between LPS and HU, and especially if there were interactions between multiple LPS molecules and multimers of HU, then additional entities with different average diameters would be generated, and also observed.

***The LPS-HU DLS experiment*:** The DLS scattering experimentwas performed on a Wyatt Dawn MALS Helios-8(multi angle light scattering) instrument, and results were analysed using ASTRA software for calculations of size distributions. The form of HU used was HU-B (purified by affinity chromatography on Ni-NTA resin), at a concentration of 0.5mg/ml, after being passed through a 0.02 μm PVDF membrane filter to remove any aggregates of larger diameter. The solution of protein and the LPS were made in phosphate buffered saline (PBS) of pH7.4. The LPS was used at a final concentration of 1mg/ml. Before the experiment, LPS was sonicated to break down any micelles. The solution of protein and LPS, after mixing, was kept at room temperature for 30 minutes. At the start of the experiment, a baseline was collected using the PBS used subsequently to dissolve the protein and LPS. After measuring the sizes of free protein and LPS,the sizes of scattering entities in the mixture of LPS with HU-B was measured. The hydrodynamic radius of the free HU-B was found to be ~3nm, with few aggregates of ~ $1x10^5$ nm size range as shown in the panel A of Figure 6. The size is plotted according to the weight fraction, as in intensity fraction larger aggregates formed in trace amounts tend to dominate the

distribution disproportionately. The size of free LPS was around ~8nm, as shown in panel B of the Figure5. Addition of HU-B to LPS led to the formation of large species/entities with a size centred around $1x10^6$nm, with a concomitant reduction in the intensity of the LPS peak from 4 (in the control) to 1.5(in the LPS-HU-B mixture), along with a complete disappearance of the peak corresponding to the free protein, suggesting its incorporation into aggregates formed with LPS, whether of a vesicular/micellar nature or otherwise, as shown in the panel C of the Figure 6.

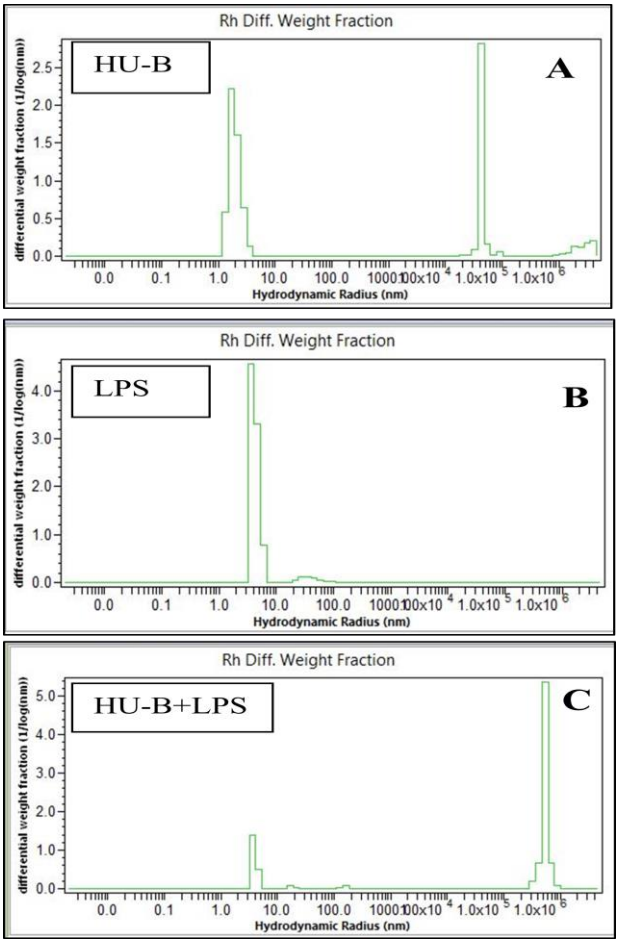


Figure 6: The dynamic light scattering based estimation of molecular sizes. ***Panel A:*** Shows the size distribution curve of free HU-B protein at a concentration of 0.5 mg/ml, plotted in term of weight fraction. ***Panel B:*** Shows the size distribution curve of free LPS at a concentration of 1 mg/ml, plotted in terms of weight fraction. ***Panel C:*** Shows the size distribution curve after incubating HU-B with LPS.

A question that arises is why there were no signs of increase in apparent absorption through scattering, if large aggregates are being formed through LPS-HU interactions. There are two possible answers: (i) The time that is allowed to pass. If the binding proceeds slowly, at initial time points (after initial physical equilibrium of mixing is achieved) the objects resulting from associations will not be as big as they would be afterwards. Therefore, since the DLS experiment was done after incubating HU and LPS together for 30 minutes at room temperature, whereas difference absorption spectroscopy experiments were done in a time span of a few tens of seconds (rapid and repetitive scanning involving a few seconds of time per spectrum, until most of the difference absorption spectrum is at the level of the baseline), it is possible that the interactions did not lead to very large objects in the difference absorption spectroscopy experiment, causing little scattering; (ii) the other possibility is that the objects, however, large, cause an amount of scattering that is underwhelming in relation to the change seen due to interactions and the effects of LPS binding upon the electronic excitation transitions of the phenylalanine residues and the RFP fluorophore.

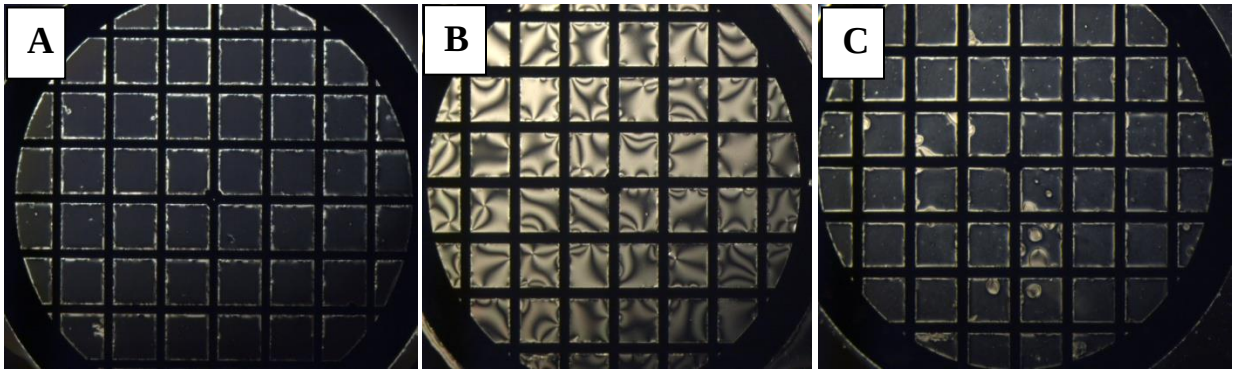
3.1.2.5. LPS binding to HU (Experiment V): A birefringence experiment involving LPS bound liquid crystals

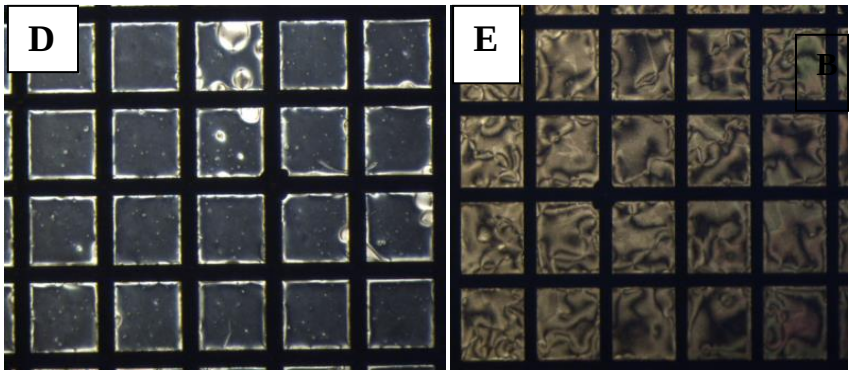
***A description of birefringence in respect of liquid crystals*:**Liquid crystals (LCs) are a state of matter lying between solids and liquids, with order in the arrangement of molecules like in a solid, but with potential for changes in the order, e.g., changes brought about by binding of ligands. The binding of the ligand affects the refractive index of the liquid crystal, leading to birefringence and to a change effected in the polarization of incident light (which is already polarized, to begin with). The changed polarization is visualized through a polarization-monitoring microscope, utilizing a pair of crossed polarizer plates in the light path. A change in refractive index upon binding of a ligand to a molecule pre-immobilized on the liquid crystal changes the appearance of the crystal, from dark to light, when visualized under a polarized microscope.

***The LPS-HU birefringence experiment* :** In the present experiment, the liquid crystal used for studying LPS-HU interactions was made by coating a hydrophobic chemical on a DMOAP-coated glass slide and aligning LPS on this coating, through the hydrophobic chains of LPS,

using a method already standardized by Dr.Shantanu Pal (IISER Mohali) and co-workers [5]. The binding of HU with LPS was established by monitoring a transition in the liquid crystal from a dark state to a light state, as follows.The aliphatic tails of the lipid in LPS molecules are aligned in parallel to the LC mesogens, as a result of hydrophobic interactions leading to homeotropic orientation (dark appearance) of the LCs. The specific binding of proteins to the core regions of the sugar phosphate groups of LPS trigger an orientation transition in the liquid crystal.When plane polarized light is incident on to the liquid crystal, it appears dark due to parallel arrangements of LPS hydrophobic chains aligned with the hydrophobic liquid crystal. After any interaction of the LPS polar head groups with another substance (e.g., HU molecules), or upon the binding of sugar phosphates in the core region with a protein, these arrangements are altered, leading to change in the polarization of the plane polarized light, visualized through a polarized microscope with a detector polarizer kept perpendicular to the incident light.

After the addition of LPS over the liquid crystals, their ordered state remained intact even after addition of water, as shown in the panels C and D of Figure 7. The addition of HU-B to this LPS-coated arrangement, however, could be seen to lead to a drastic change in ordered-ness, suggesting binding of HU-B to LPS, as shown in panel E of the Figure 7.





Liquid crystal with LPS after washing with water	Liquid crystal with LPS after addition of HU

Figure 7: Images of liquid crystals before and after addition of LPS in water. The liquid crystal remains intact in orientation after the addition of LPS even after washing with water. The orientation of LPS is affected after incubation with 0.5 mg/ml HU-B in PBS (pH7.4). ***Panel A:*** Shows the liquid crystal in air. ***Panel B:*** Shows the liquid crystal after washing with water, leading to the disruption in the liquid crystal's molecular arrangement***. Panel C:*** Shows coating of the liquid crystal with LPS, leading to the restoration of the original molecular arrangement, seen in air. ***Panel D:*** shows the stability of liquid crystal in water evident by its dark appearance. ***Panel E:*** Shows the LPS-coated liquid crystal after incubation with 0.5 mg/ml HU-B, indicating disruption of the liquid crystal, akin to that seen with water in panel B.

3.1.3. Binding of bacterial cell-surface LPS to free HU

In the preceding sections, numerous experiments have been described, the results of all of which point towards the discovery of the protein HU's being able to bind to a bacterial cell surface molecule, lipopolysaccharide (LPS). However, all of the above experiments establish that HU binds to free LPS, which is not the same thing as examining whether HU binds to LPS displayed upon the surfaces of bacterial cells, as a component of the outer membrane of *E.coli*.

Therefore, we performed experiments designed to examine whether HU (in either its HU-A or HU-B forms) also binds to the LPS present on bacterial cell surfaces. For this, we used fluorescent forms of HU and performed either spectroscopic or microscopic experiments to examine binding of HU to bacteria. For this purpose, we used the fluorescent proteins, Tag-RFP

or Venus, in genetic fusion with HU at the N-termini of the HU molecule(s). Such fusions were already available in the laboratory from the work of Dr.Kanika Arora.

3.1.3.1 LPS binding to HU (Experiment VI): Fluorescence anisotropy of Tag-RFP-HU-A increases in presence of bacterial cells displaying LPS

A description of anisotropy of fluorescence: Measurements related to rotation freedom of fluorophores, under steady state illumination conditions, display an increase in both fluorescence polarization and fluorescence anisotropy, when a fluorophore becomes associated with a larger object characterized by a larger mass and a slower rotational or 'tumbling' time in solution. Essentially, anisotropy results from restriction of rotational motion, causing more emission of fluorescence in certain directions than in other directions, whereas isotropic fluorescence involves equal intensities of emissions in all directions.

The E.coli-RFP-HU anisotropy experiment:The binding of Tag-RFP-HU-A to bacterial cells was studied by addition of HU to non-HU-over expressing XL-1 Blue *E.coli* cells, i.e., by monitoring whether the fluorescence anisotropy associated with the RFP fused to HU-A undergoes any increase upon addition of Tag-RFP-HU-A to bacteria, as a result of the binding of the protein to the cell surface and the associated changes in the rotational freedom of the RFP fluorophore. Addition of bacterial cells was observed to result in an enhancement in the fluorescence anisotropy of the signal associated with the RFP fluorophore in Tag-RFP-HU-A. This suggests that there is immobilization of Tag-RFP-HU-A on bacterial cell surfaces, as shown in the panel A of the Figure 8. The anisotropy values crossed the value of 0.4 ostensibly on account of scattering caused by bacterial cells, with the addition of larger numbers of bacterial cells. Light of a wavelength of 550 nm was used to excite the RFP fluorophore with 5nm excitation and emission slit widths set on the fluorimeter. Multiple experiments performed (panel A) were averaged and the mean anisotropy values (along with the variations around the mean) are plotted in panel B. The tightness of the bars is indicative of the reproducibility of measurements from three independent sets of data. The data suggests that Tag-RFP-HU-A binds to the *E.coli* and that these results in the increase in the anisotropy of fluorescence associated with Tag-RFP.

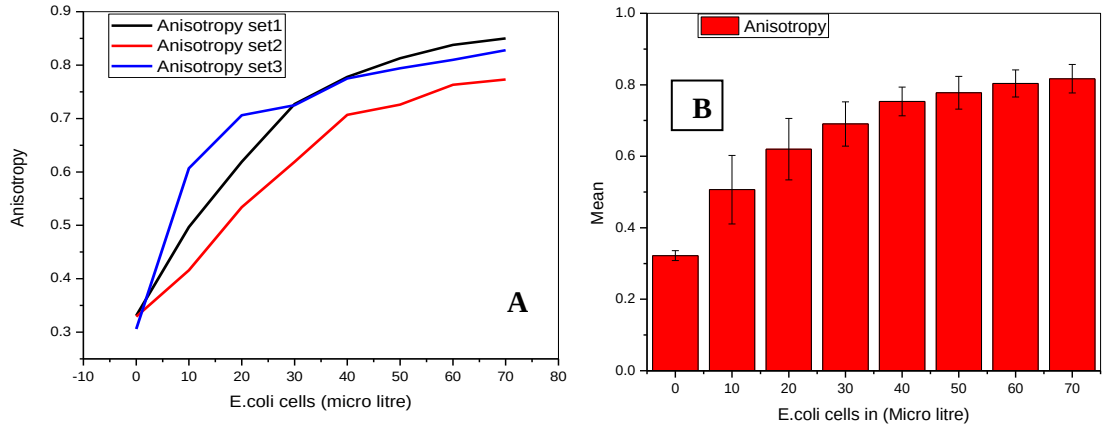


Figure 8: Outcomes assimilating results from three independent anisotropy experiments performed by taking 50µl HU-A RFP (10 nM) amounting to ~3×10^{11} molecules and adding to it 10 µl *E.coli* XL-1 Blue cells (1.14×10^{6} cells/µl) successively. With an estimated ~ 2million molecules of LPS on the surface of each bacterial cell, the total HU binding sites are likely to be ~2.28×10^{12} HU molecules/µl of bacterial culture**.** ***Panel A:*** Shows the results obtained after plotting results of three independent anisotropy experiments. ***Panel B:*** *Shows* results plotted in terms of mean fluorescence and standard deviation, along with error bars, where the tightness of the error bars represents reproducibility of anisotropy values across data sets collected from independent experiments.

When using cells, it is recognized that anisotropy is large due to back scattering of light from bacterial cells, the concentration of which, in this experiment, happened to be increased in every step.

3.1.3.2 LPS binding to HU (Experiment VII): Visual evidence for the titration of Tag-RFP-HU-A to the*E.coli*cell surface

To examine further whether HU binds to the *E.coli* cell surface (i.e., to the LPS present on the cell surface), we added Tag-RFP-HU-A to non-overexpressing XL-1 Blue *E.coli*cells in a manner similar to what was done for the anisotropy experiment, the difference being merely this: instead of carrying out spectroscopic measurements, as in the previous section, here we carried out microscopic observations. The purpose was to examine whether the Tag-RFP-HU-A moves from a diffused state of fluorescence in the solution surrounding cells, to a pin-pointed state of intense fluorescence associated with *E. coli* cells, through binding of Tag-RFP-HU-A to the bacteria. As a control, Tag-RFP was used in equal (molar) concentration to Tag-RFP-HU-A, in a

separate experiment, to examine whether Tag-RFP itself binds to *E.coli* cells. The experiment was done by adding the protein to overnight-grown XL-1-Blue *E.coli* cells pre-washed thrice with PBS (pH7.4). The mixture of cells and fluorescent protein was then kept at room temperature for 30 minutes, to allow adequate interaction between proteins and the cell surface (with the decision to increase the time of incubation, if necessary, in subsequent experiments; however, this proved not to be necessary). After incubation, cells were pelleted down and washed with PBS to remove all unbound protein, i.e., Tag-RFP in the control experiment, and Tag-RFP-HU-A in the main experiment.

The slides for viewing on the microscope were made by taking a drop of suspended bacterial cells and placing them on a glass slide and then putting a coverslip over the same. Figure8shows the results of fluorescence images collected under different settings, with Tag-RFP-HU-A and the Tag-RFP control. Fluorescence images show the binding and titration of Tag-RFP-HU-A to bacterial cell surfaces, as can be seen in Panel A of the Figure 9, whereas no fluorescently labelled cells were found upon addition of Tag-RFP (panel A1), showing that Tag-RFP itself does not bind to *E.coli*. The positions and outlines of cells were visualized by collecting phase contrast images of both the samples (Panels B and B1). The pictures were recorded on a Olympus laser scanning FV3000 confocal microscope without optical sectioning with 100X oil immersion objective lens with fluorophore excitation done using the red laser. Panels C and C1 show merged images of the panels in A and B, and A1 and B1, respectively.

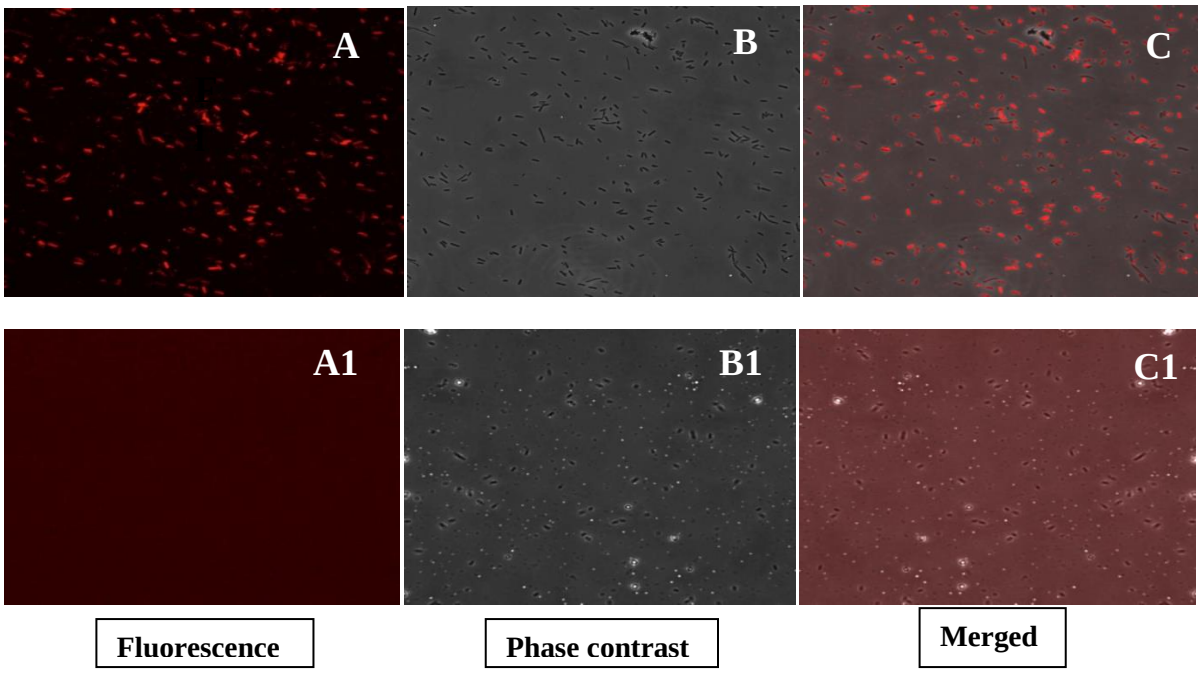


Figure 9: Microscopic images taken under various setting for *E.coli* XL-1 Blue cells treated with Tag-RFP-HU-A and Tag-RFP. The cells with treated with 10 µM of each proteins**. *Panel A, A1:*** Fluorescence images of the cells treated with Tag-RFP-HU-A and Tag-RFP. ***Panel B, B1:*** Phase contrast images and the merges of both the images are shown in ***Panel C,C1***.

3.1.3.3. LPS binding to HU (Experiment VIII): Visual evidence for the titration of Venus-HU-B to the *E.coli* cell surface

The binding of HU-B to XL1-Blue cells was also examined, as done above for HU-A. For this, the washed cells were incubated with HU-B displaying the fluorescent protein, Venus, fused at the N-terminal of HU-B. The incubation was done at room temperature for 30 minutes. After washing with PBS, cells were visualized under the confocal microscope with optical sectioning under 100X oil immersion objective lens, as shown in Figure 10. Panel A shows the fluorescence image, with localization of Venus-HU-B on bacterial cells, and especially onto a cluster of cells for which the image looks bleached. Panel B shows the phase contrast image of bacterial cells. Panel C shows the merged fluorescence and phase contrast images, indicative of co-localization of fluorescence with bacterial cells. The intensity of fluorescence is non-uniform because of the irregularity in sample height and localization of fluorescence along different planes.

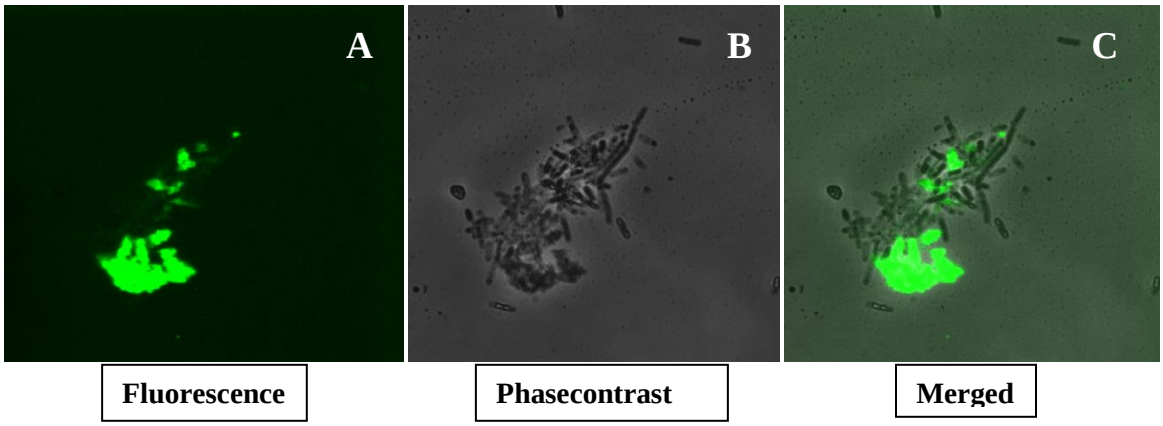


Figure 10: Shows the images of the *E.coli* XL1-Blue cells incubated with Venus-HU-B under various visualization settings**.** ***Panel A:*** Image collected under fluorescence settings followed by phase contrast and merged image for the same slide section shown in ***Panel B and C.***

3.1.3.4. LPS binding to HU (Experiment IX):Tag-RFP-HU-A is associated with *E. coli* coagulates/clumps/aggregates

RFP-HU-A also showed clumping (and associated fluorescence, with clumps) as was shown by Venus-HU-B. Figure 11 shows, (panel A), fluorescence images of clumps associated with Tag-RFP-HU-A. Clumps were visible in phase contrast images, with co-localization of fluorescence**.**

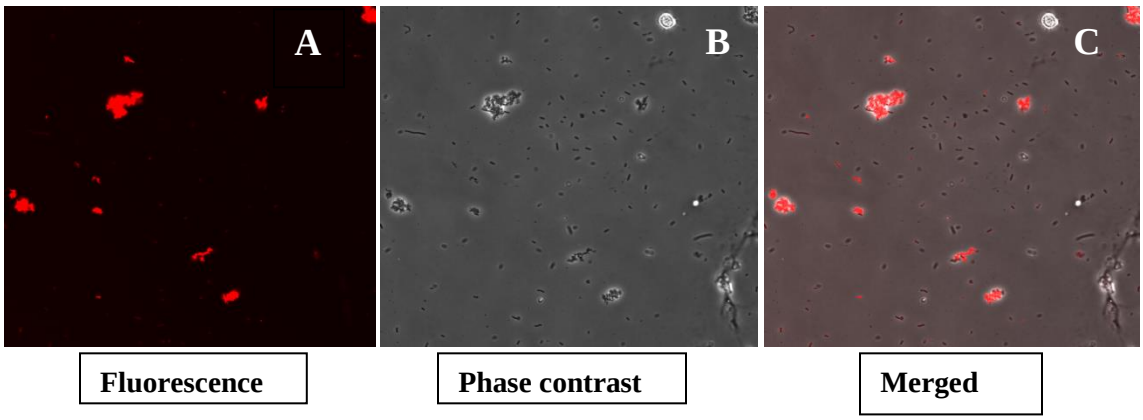


Figure 11: Images of the *E.coli* XL1-Blue cells incubated with Tag-RFP_HU-A under various visualization settings**.** ***Panel A:*** Image collected under fluorescence settings followed by phase contrast and merged image for the same slide section shown in ***Panel B and C.*** The presence of HU mediated clumping is apparent in all three images.

3.1.4. Clumping of bacterial cells by free HUmediated by HU-LPS interactions: Concepts

The protein, HU, is rich in charged residues on its surface. It has also been shown to bind to DNA through multiple regions of its surface. The canonical DNA-binding surface in HU is thought to have specificity for double-helical nucleic acids (DNA or RNA). The regions of the protein which are involved in binding to the minor groove of DNA at the canonical DNA-binding site are believed to be ordinarily unstructured (because they do not show up in crystal structures of the protein in the absence of bound DNA). These regions, primarily consisting of a loop made up of two beta strands (one from each of the two HU subunits in a homodimer or heterodimer), adopt stable structure(s) in complex with DNA, in interactions that are mediated through electrostatic contacts. The non-canonical DNA-binding surface on HU also uses electrostatic contacts, but there is no evidence that it has any preference for double-helical nucleic acids. Thus, it is possible that it binds to DNA, RNA (including single-stranded nucleic acids) or any negatively charged molecule, through electrostatic interactions. At the non-canonical DNA-binding site, there are two conserved lysine residues, Lys 83 and Lys 3, which interact with bound DNA. Given that HU has two sites for binding to negatively-charged substances like DNA, and that we have evidence that HU can bind to the surfaces of *E.coli* cells, it stands to reason that if HU (HU-A and/or HU-B) can bind to *E.coli* cells, it should also be able to aggregate, or coagulate, such cells, since every dimer, tetramer or octamer of HU would possess two or more LPS binding surfaces.

3.1.4.1. LPS binding to HU (Experiment X):HU-B addition leads to clumping/aggregation of *E.coli* cells.

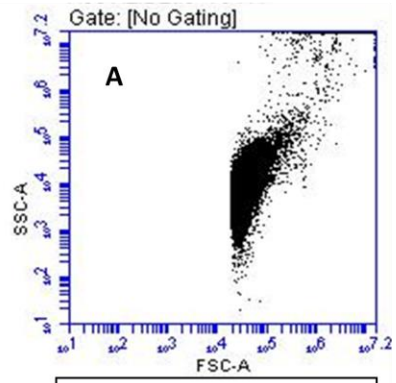
A description of the fluorescence activated cell sorting (FACS) cytometry experiment :The sizes and the granularity of cells cross linked non-covalently by surface-bound HU could, in principle, be analysed by passing cells through a flow cytometer (or FACS instrument) and by monitoring whether there is an elevation in the values of the forward scatter, or FSC (plotted on the x-axis), as well as side scatter, or SSC(plotted on the y-axis), leading to the formation of diagonal streaks at the top right corners of cell populations, measured as a function of cell or cell-cluster size.

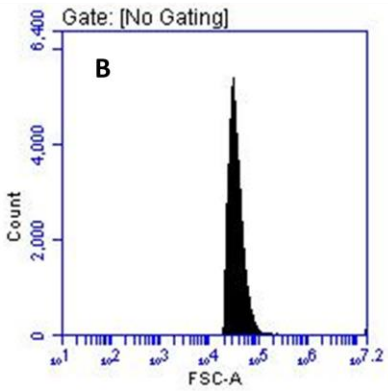
The cytometry experiment involving E. coli and added HU :The scatter plots of FSC-*vs*-SSC in panels A, A1 and A2 of Figure 12,show that whereas XL-1 Blue cells ordinarily display only a very poor natural streak indicative of a baseline level of cell cluster formation (panel A), when HU-B (panel A1) or HU-A (panel A2) are added to cells, there is the formation of a strong and distinct streak, indicative of the formation of cell clusters, which increase light scatter both when viewed from the front and from the side.

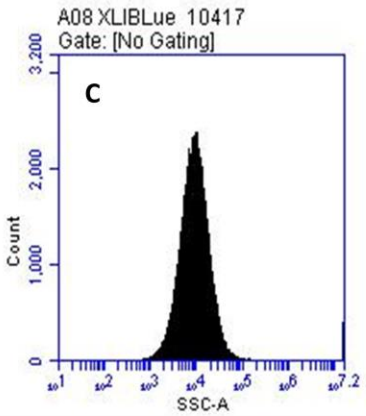
Thus, it may be said that addition of HU to cells results in substantial clumping of cells. The same data is reflected in the remaining panels of the Figure12, in which FSC or SSC are independently plotted against counts of bacterial particles flowing through the cytometer, and it is evident that there is an increase in the net forward as well as side scatter values.

Notably, the streaking is higher and more intense in case of HU-B, suggesting that its ability to coagulate, or clump, bacteria is higher. This,in turn, can probably be attributed to HU-B tending to be a multimer (tetramer, or octamer) to a much greater extent than HU-A (which is mostly a dimer), as will be demonstrated later in this thesis. The data is assimilated and overlayed in panels D and E, which drives home the same conclusions stated above, namely that both HU-A and HU-B cause clumping of *E. coli,* with HU-B performing better than HU-A in this respect. It may be noted that HU has been shown to be secreted by bacteria through a type IV secretion system (T4SS). Thus, it is possible that the clumping of cells seen during stationary phase is mediated at least partly by secreted HU which appears on the cell surface.

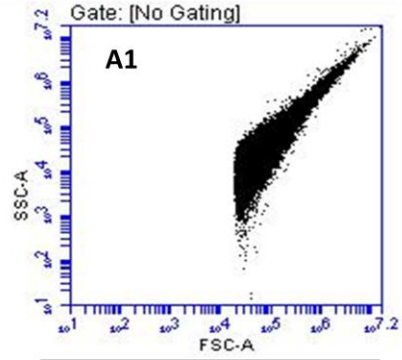
XL1-Blue cells

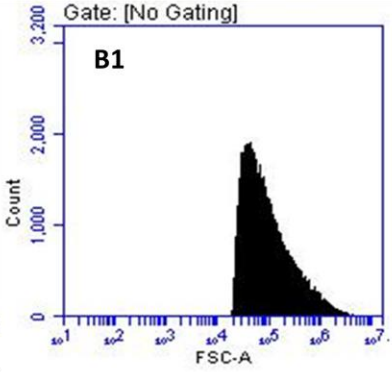


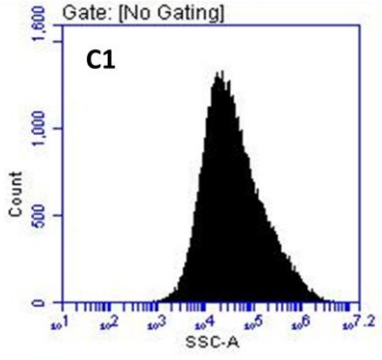




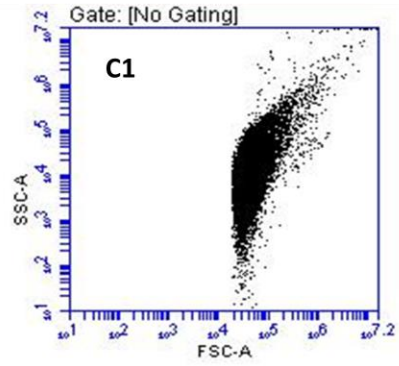
XL1-Blue cell +HU-B

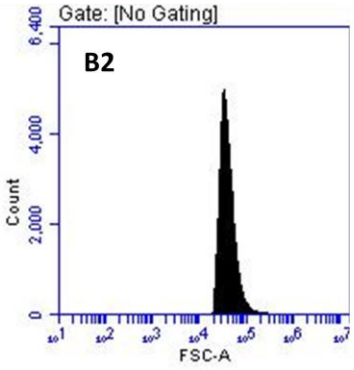


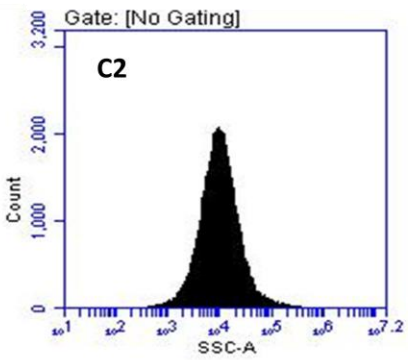




XL1-Blue cells +HU-A







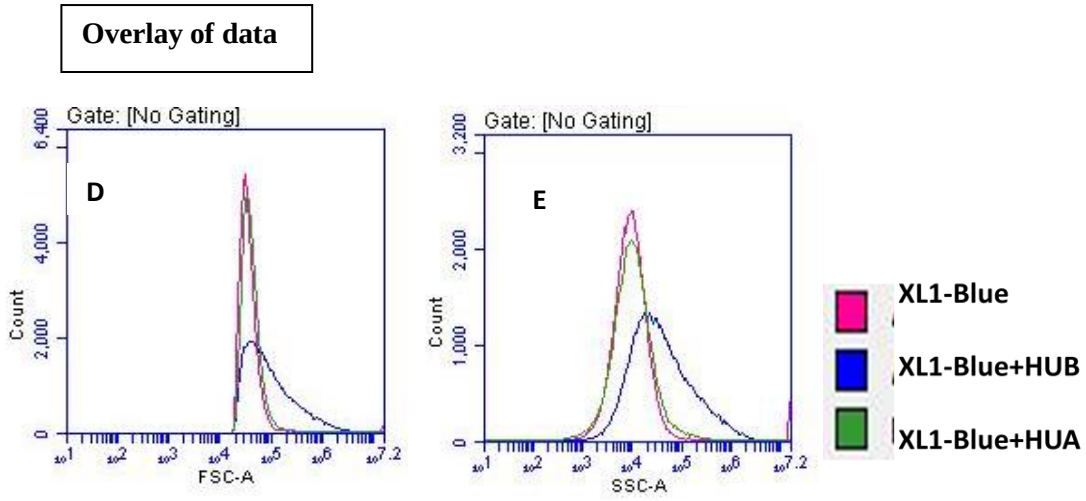


Figure 12: Flow cytometery images showing HU mediated clumping of bacterial cells**.** ***Panels A, A1, A2:*** Scatter plots of control and cells incubated with HU-B and HU-A**.** ***Panels B, B1, B2:*** Histograms of bacterial counts plotted against forward scatter for control and HU-B and A treated XL-1 Blue *E.coli* cells. ***Panels C, C1, C2:*** Histograms for control, HU-B and HU-A-treated XL1-Blue *E.coli* cells plotted against side (SSC-A) scatter. ***Panels D, E:*** Overlapping histograms of forward scatter (FSC-A) with bacterial cell count and side scatter (SSC-A) with bacterial count for all three sets.

3.1.4.2. LPS binding to HU (Experiment XI): DNA competes with HU driven clumping

If indeed, as has been proposed, HU has two DNA-binding sites, one canonical and the other non-canonical and, furthermore, if both of these sites can potentially also each bind to LPS individually (through the sugar-phosphate moieties in the structure of Lipid A in LPS), it stands to reason that one must examine whether DNA and LPS (both negatively-charged substances) can compete each other out, in terms of their binding to HU (a positively-charged substance). This issue was examined by incubating HU with increasing amounts of DNA, prior to the addition of these mixtures to *E. coli* cells, to see whether the streaks seen during flow cytometry which are indicative of the cell-clumping/coagulating property of HU are somehow reduced in the presence of DNA. The assumption here is that if DNA is pre-bound to HU, it will prevent

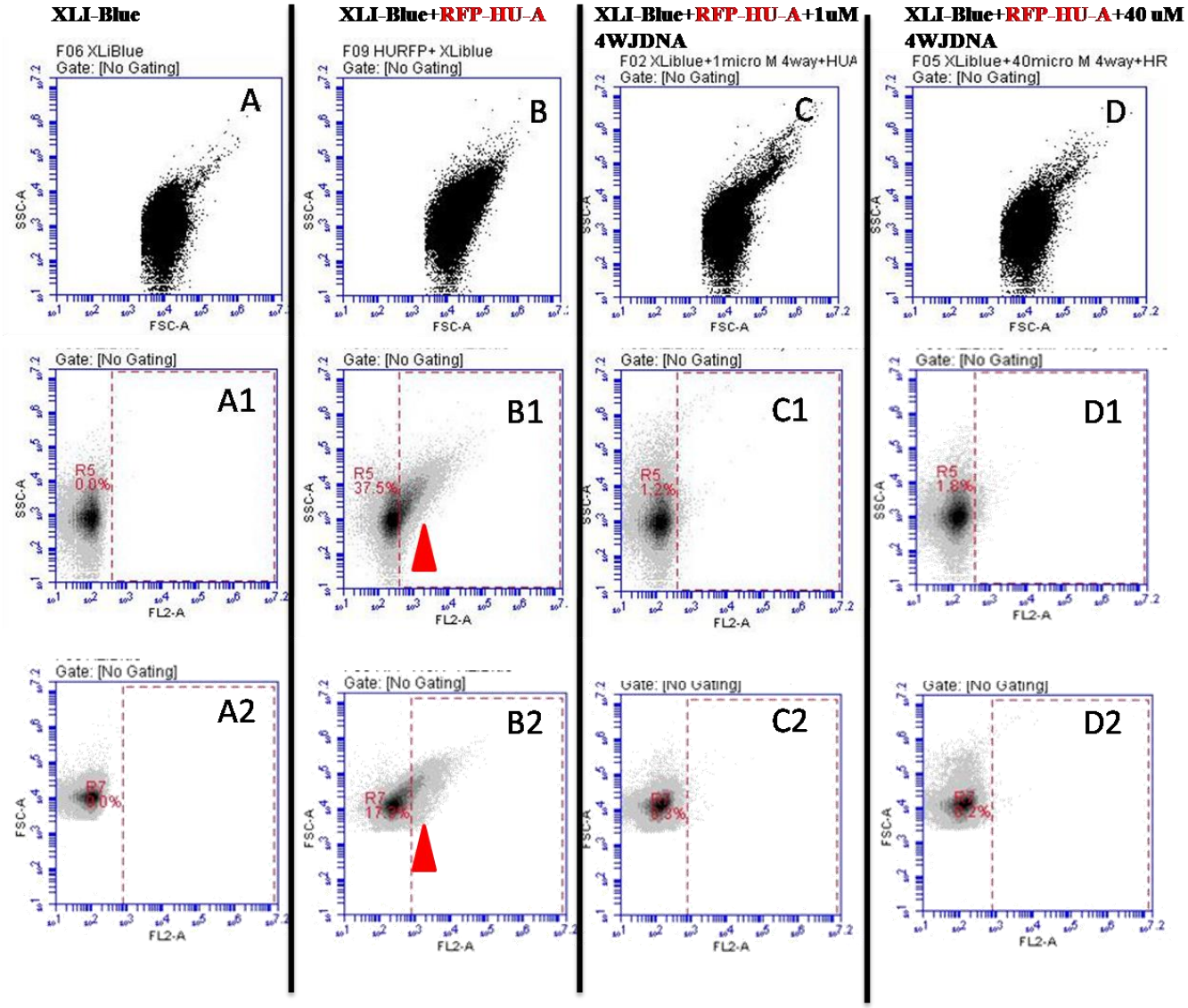
LPS from binding to HU and thus HU will not be able to effectively cause any clumping, as it will no longer be able to bind to cells. The DNA used for this experiment was a well-defined 4-way junction already known to bind to HU (see materials and methods). Further, in addition to collecting the scatter data displaying the streak, we collected fluorescence data relating to the fluorescence associated with the cell-clusters of clumps, by using Tag-RFP-HU-A instead of just HU-A as an additive to DNA prior to the addition of the mixture to bacterial cells.

The binding of Tag-RFP-HU-A to cells was observed to reduce significantly after prior incubation with the 4–way junction DNA (4WJ), in a very dose-dependent manner, as shown in Figure 13. The streak intensity was the highest when no DNA was added to Tag-RFP-HU-A and decreased progressively with increasing concentrations of 4WJ DNA (the sequence is shown in section 2.18 of materials and methods), as can be seen in (panels B1 to D1 and B2 to D2) of Figure13. In addition to the decrease in the streak intensity and distribution, we were also able to see that there is fluorescence associated with bacterial cells, and that this fluorescence increases or decreases in proportion with the intensity and distribution of the streak. The binding of Tag-RFP-HU-A to bacterial cells makes them fluorescent and the fluorescence associated with such cells is measured simultaneously with the FSC and the SSC, through the FL-2 detector, upon excitation with the red laser of the cytometer.

Panels A1 to D1 and A2 to D2 show dot plots of the cells registering on both the FL2-A (fluorescence) and side scatter detector(s). It is clear from the graph that the cells bound to Tag-RFP-HU-A show increased side scatter values. The change in the SSC-A values was not very significant with the use of HU-B, but there were increased FL2-A values associated with cells incubated with Tag-RFP-HU-A. The fluorescence intensity decreased progressively upon addition of more and more 4WJ DNA to cells, in a dose dependent manner, suggesting binding of HU to bacterial cell surfaces through the same sites used by HU for DNA binding. Clearly, binding of DNA made HU unable to bind to cells, by making its DNA-binding sites unavailable for binding to cell surface LPS. Similar results were obtained for FSC-A values when plotted against FL2-A.

Panels E,, show the overlay of all FSC-A, SSC-A and FL2-A data plotted against bacterial cell number and it is clear from this data that HU-A is less efficient in bacterial cell clumping. This is

consistent with our previous result. It is clear from our results that HU binds to the bacterial cell surface by using some of the very same sites that are used by DNA to bind to HU.



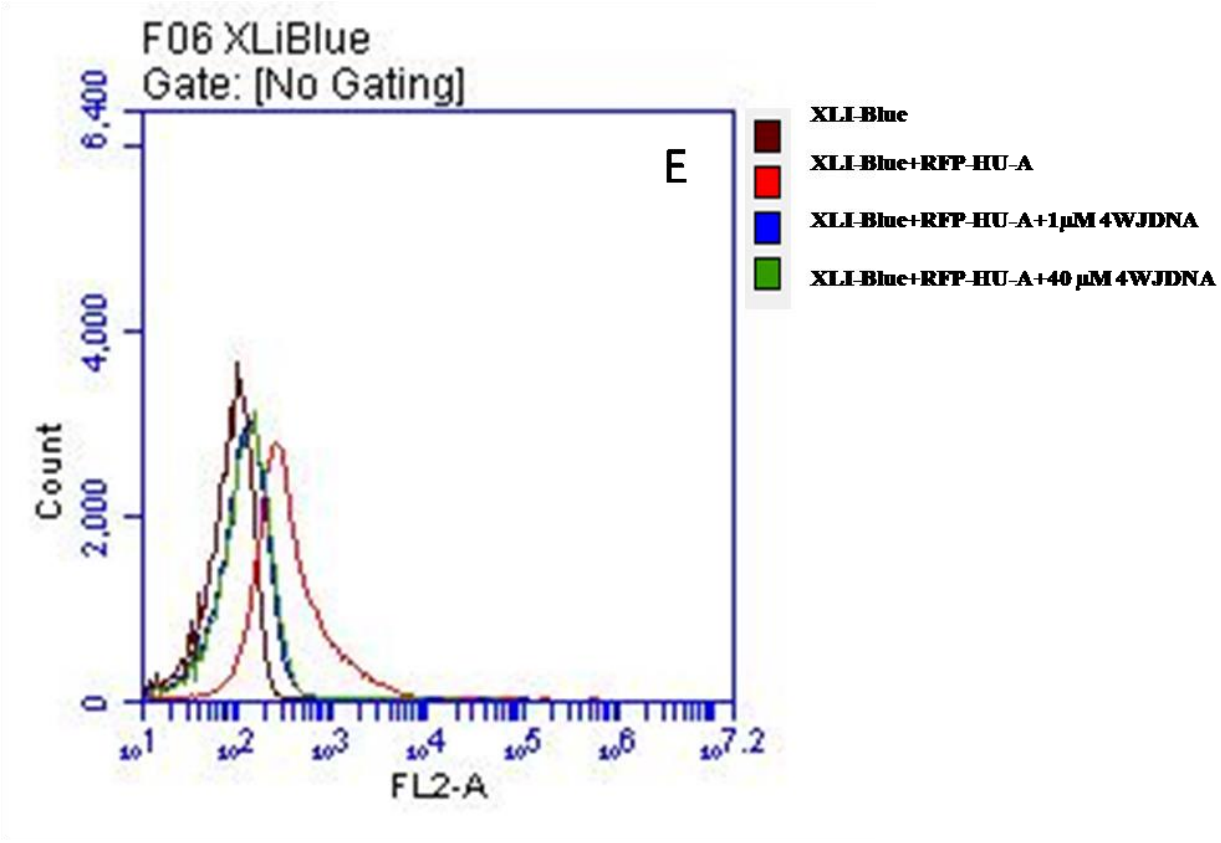
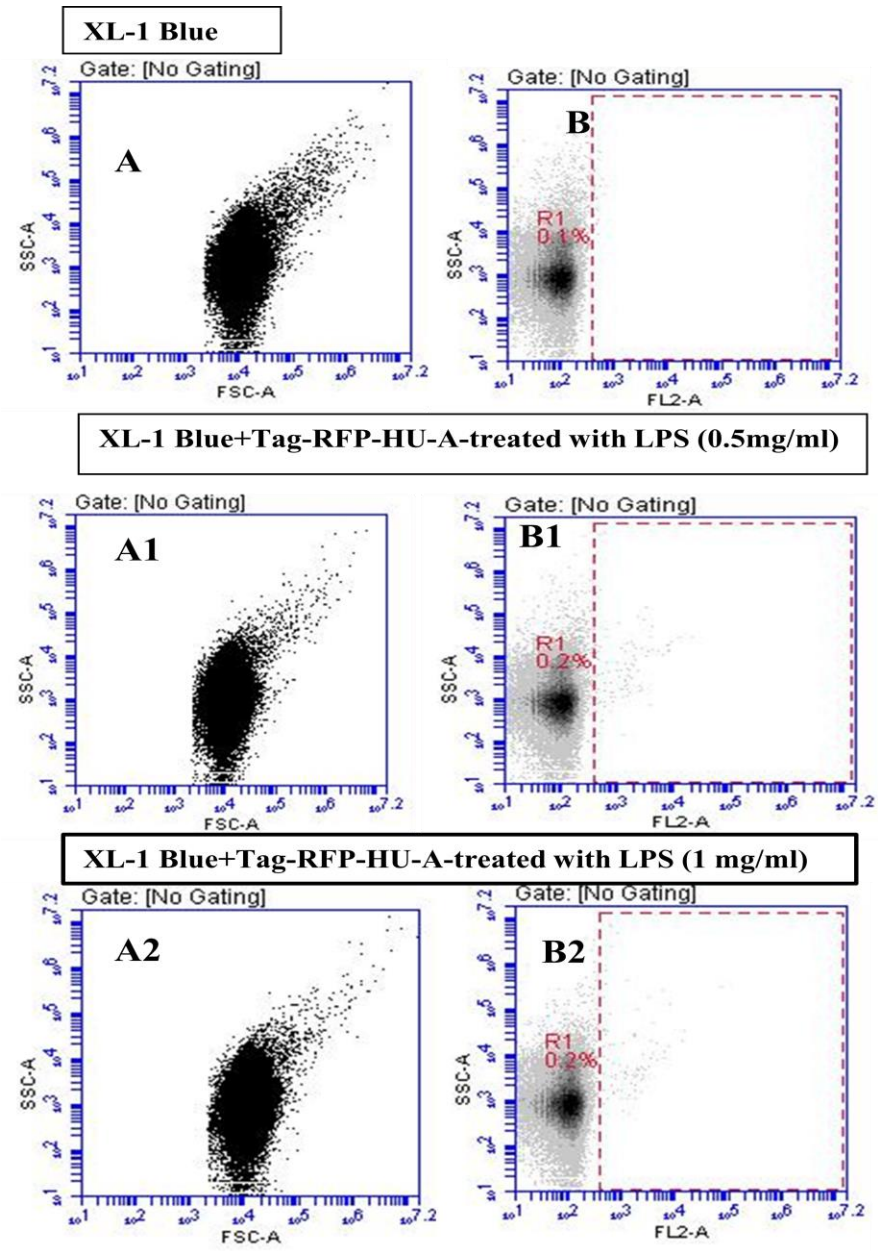


Figure 13: Flow cytometery images showing the effect of DNA on HU mediated clumping of bacterial cells. ***Panels A, B ,C ,D***: Scatter plots for the untrated and threated cells. The intensity of the streak can be seen to reduce after pre- incubating proteins with 4WJDNA. ***Panel A1, B1, C1, D1:*** scatter plot of fluorescence associated with bacterial cells with side scatter (SSC-A). FL-2 is the detector used for measuring fluorescence coming from RFP fluorophore emission. ***Panels A2, B2 ,C2, D2:*** Fluorescence asociated with bacterial cells against forrward scatter (FSC-A). Arrow heads in ***Panel B1*** and ***B2*** represents the cells associted with Tag-RFP-HU-A ***Panel E***: Overlay graphs showing the comparsion of all detected variables with each other. The Tag-RFP-HU-A concnetration used in this experiment was 10 µM.

3.1.4.3. LPS binding to HU (Experiment XI): LPS-prebound HU shows reduced clumping of *E.coli*

As 4WJDNA could be seen to compete with the clumping of bacterial cells (presumed to occur through the binding of bacterial cells to HU) through competition of DNA for HU which also binds to the cell surface, the next question to be asked was whether LPS pre-incubated with HU could similarly compete to destroy bacterial clumping or clumping, by pre-binding of LPS to HU preventing binding of HU to bacterial cell surfaces. The pre-incubation with LPS showed a similar effect with Tag-RFP-HU-A as was seen with incubation with the 4WJ DNA, as shown in (panels C and D) of Figure 14. Streaks resulting from bacterial clumping (through creation of populations displaying increased forward scatter, as well as increased side scatter), seen upon addition of HU to bacterial cells, become progressively abolished when the HU has been pre-incubated with LPS. Furthermore, since the HU used here was Tag-RFP-HU-A, the non-binding of HU-A to bacterial cells could be also established from the monitoring of the fluorescence

signal associated with RFP, which is no longer associated with bacteria upon pre-incubation of Tag-RFP-HU-A with LPS.


XL-1 Blue
Gate: [No Gating]
A
SSC-A
FSC-A
Gate: [No Gating]
B
R1
0.1%
SSC-A
FL2-A
XL-1 Blue+Tag-RFP-HU-A-treated with LPS (0.5mg/ml)
Gate: [No Gating]
A1
SSC-A
FSC-A
Gate: [No Gating]
B1
R1
0.2%
SSC-A
FL2-A
XL-1 Blue+Tag-RFP-HU-A-treated with LPS (1 mg/ml)
Gate: [No Gating]
A2
SSC-A
FSC-A
Gate: [No Gating]
B2
R1
0.2%
SSC-A
FL2-A

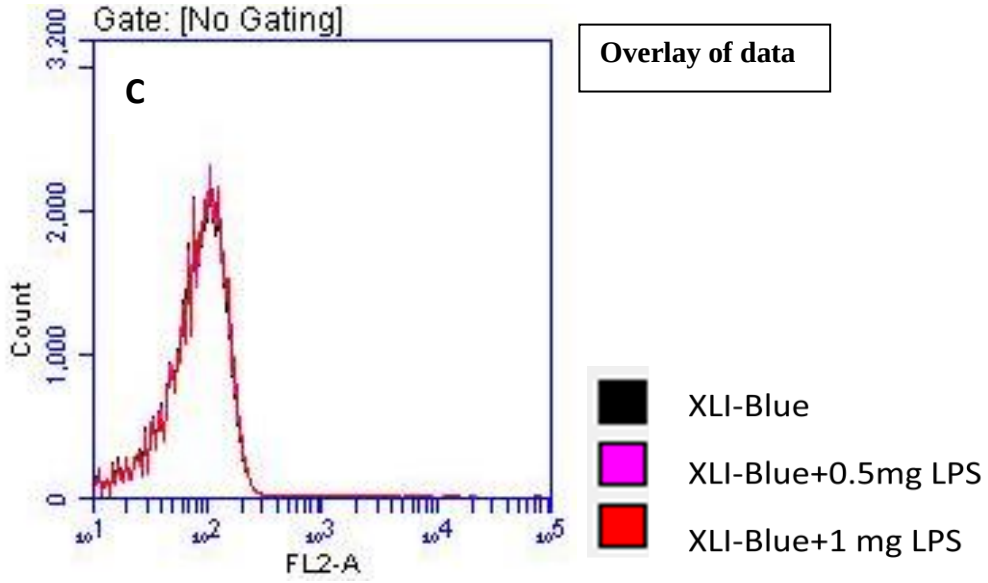


Figure 14: Showing the flow cytometry plots demonstrating the effect of exogenous LPS on Tag-RFP-HU-A mediated bacterial cells clumping by pre-incubating 10 µM Tag-RFP-HU-A with 0.5, 1 mg/ml LPS and adding this mixture over bacterial cells***. Panel A to A2:*** Showing scatter plot of bacterial cells plotted between FSC-A and SSC-A detectors. ***Panel B to B2***: Showing the fluorescence associated with bacterial cells where after incubation with LPS, HU treated cells behave like control XL-1 Blue cells. ***Panel C:*** Overlay of FL2-A with bacterial cells number for all three sets of experiments .

3.1.4.4. LPS binding to HU (Experiment XIII) : A primarily electrostatic effect, since Poly-D-Lysine also causes clumping of *E.coli*

Thus far, we have hypothesized that the binding of HU to both DNA and LPS is promoted primarily by electrostatic effects, and also by such electrostatic effects being supported in the context of specific conformations of macromolecules (protein/DNA), and configurations of smaller molecules (lipid A in LPS). To further assess the importance of electrostatics, we examined whether another positively-charged proteinaceous substance, such as poly-D-lysine, can behave like HU inasmuch as it displays any ability to clump or coagulate *E.coli* cells. We already know that poly-L-lysine is used to coat plastic ware to create a surface for the binding of eukaryotic cells, which also bear a high density of negative charges on the surface (due to the negatively-charged lipid head groups). Here, we examine whether poly-D-lysine is able to bind

to *E. coli* and cause it to clump. The inversion of amino acid chirality from L-to-D is also useful because naturally occurring proteins are not made of D-amino acids, causing poly-D-lysine to be an even more appropriate control for the effect of positively-charged lysine residues in mediating clumping.

An overnight grown culture of XL1-Blue*E.coli* cells was incubated with 0.5mg/ml poly-D-lysine after washing of cells with PBS as explained in flow cytometery experiments reported in earlier sections. After incubation, the cells were passed through the flow cytometer without washing. There was an enhanced FSC and SSC associated with a streak as can be seen in panels B and A of Figure 15, exactly as is seen with HU. This sugests that the 'charges on a compatible polymer' can cause clumping of bacteria as long as the charges are of the opposite nature to the charges on the bacterial cell surface.

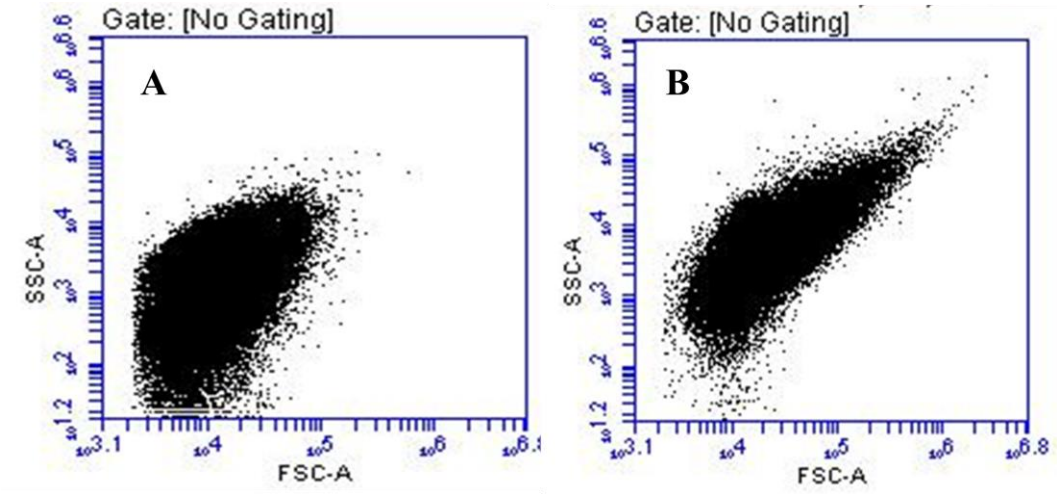


Figure 15: Panel A: Scatter plots for *XL1-Blue* cells. ***Panel B:*** Scatter plots for *XL1-Blue* cells treated with *poly D-Lysine*.

In the case of poly-D-lysine, the experiment with this polymer was itself a sort of positive control used, based on experiments previously performed by other groups, in which poly-D-lysine was shown to bind to bacterial cells (whereas other homopolymers of amino acids were not found to do so). Crosslinking of cells by any reagent (through non-covalent interactions) could only be considered conceivable as a secondary property, if the reagent molecule were to display binding to cells as its primary property, with scope available in the molecule for further interactions with other cells after binding to one cell, since only such scope would facilitate

cellcell crosslinking by the reagent. We believe that any polymer which possesses a surfeit of positive charges would display the capability of binding to bacterial cells, and potentially also for crosslinking such cells, if there were enough charges available for the polymer to bind to multiple cells. Thus, poly-D-lysine works, as does HU. The possibility of bacterial cells clumping with polyphosphate would be low due to the presence of the same charge on both bacterial outer membrane (i.e., LPS) and the proposed cross-linker (poly-phosphate). Admittedly, however, additional controls could have been used, including the negative control of polyphosphate. We have tried using a monomeric positively-charged protein (cutinase encoded by *Thermobifida fusca* produced in *E. coli*) and found that it does not cause clumping of E. coli cells, presumably because it is monomeric and folded into a globular shape. Positively-charged proteins are, of course, expected to binding to negatively-charged bacterial cell surfaces; however, we do not expect every protein to either bind to LPS efficiently or to be able to crosslink cells, unless the protein has a surfeit of positive charges and some multimeric character.

3.1.4.5. LPS binding to HU (Experiment XIV) : Primarily electrostatic effect, since salt reduces clumping of *E.coli*

The HU-LPS interaction seems to be electrostatic in nature, with involvement of surface lysine residues on HU, and LPS which bears negative charge on its phosphate groups. If the interactions are primarily electrostatic in nature, they can usually be disrupted by providing counter ions in the form of Na^+ and Cl^- by incubating samples in the presence of NaCl. Interestingly, the use of salt is advised for the removal of biofilms of bacteria in patients suffering from throat infections, which could probably be explained by the dissociation of bacteria from whatever it is that they are bound to through electrostatic interactions between negatively-charged bacterial surfaces and other positively-charged surfaces such as those of HU molecules. Therefore, we tested the effect of addition of salt, in assays examining clumping of bacterial cells by HU. The results of the incubation of HU-A and HU-B with bacteria in the presence of salt are shown in Figure 16, from which it is clear that there is abolishing of streaks (i.e., populations with high FSC and SSC) comprising clumped bacteria. Salt thus prevents clumping of bacteria by HU.

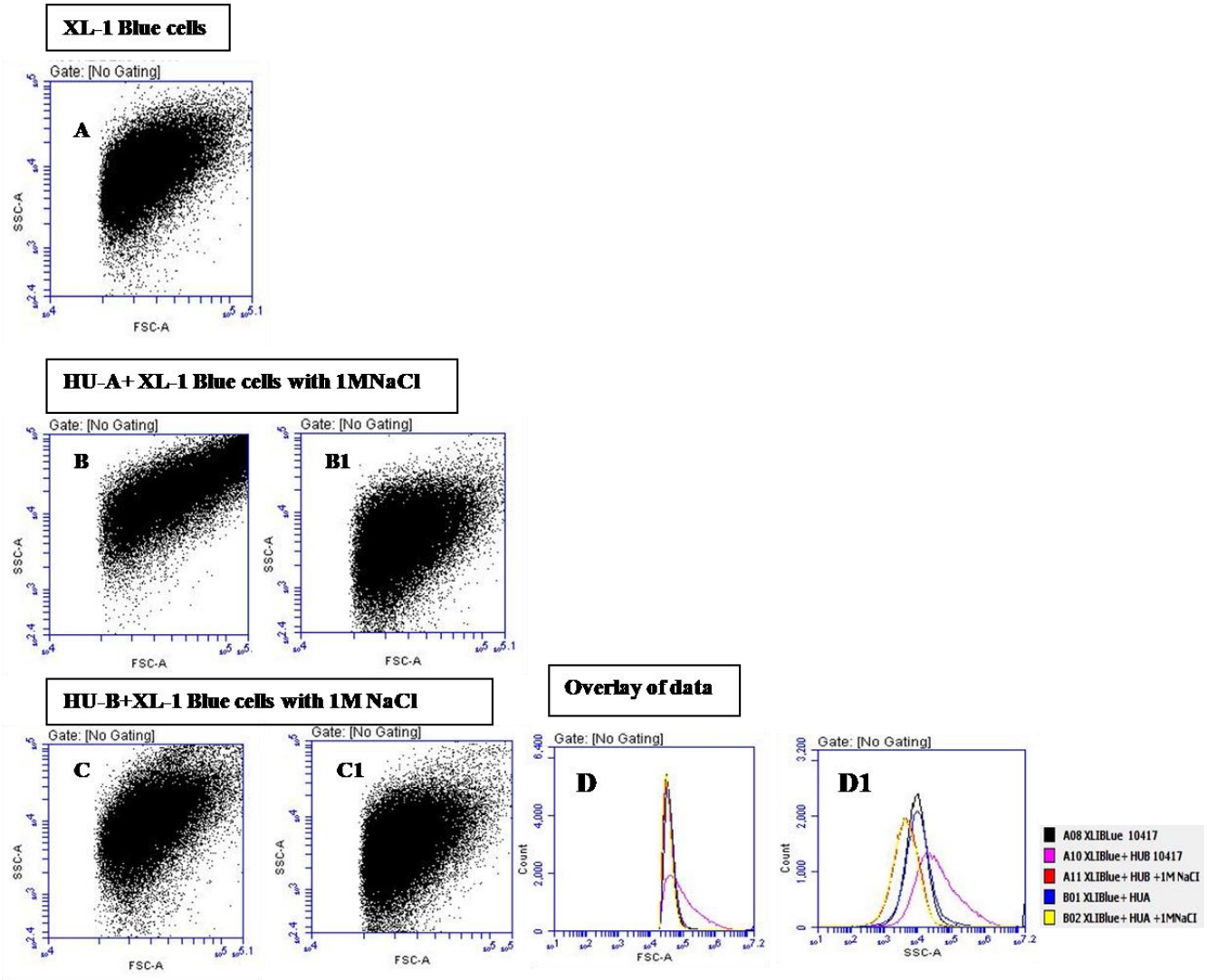


Figure 16: Flow cytometric analyses showing the effect of salt on HU-mediated bacterial cells clumping. ***Panel A:*** Showing the forward and side scatter plot of control XL-1 Blue cell size without any treatment. ***Panels B, B1*****:** Scatter plot of the cells treated with HU-B N-Tag in absence and presence of 1M NaCl. ***Panels C, C1:*** Scatter plot of the cell size treated with HU-A N-Tag in presence and absence of 1M NaCl. ***Panel D, D1:*** Showing the overlay of all samples plotted as histogram between FSC-A,SSC-A vs cell number.

It is also clear from the scatter plot (in the high SSC-A and FSC-A region) that the HU-B-mediated streaking of bacterial cells, in Panel B of Figure 16, is more than the HU-A-mediated

streaking of bacterial cells, in Panel C of Figure 16, suggesting that HU-A is less efficient in clumping bacterial cells than HU-B. A possible reason for this is that HU-B is known to form multimers to a greater degree than HU-A. In Panel B1 of Figure 16, and in Panel C1 of the same figure, comparisons with panels B and C, respectively, show clearly that salt 1M) abolishes clumping mediated by both HU-B and HU-C. Interestingly, the FSC-A and SSC-A levels of NaCl-treated cells are lower on the whole than even the control populations (see panel C1 *versus* panel C, or panel B1 *versus* panel B). A possible explanation for this behaviour is that there is a certain amount of clumping behaviour even in the control cells, based on electrostatic interactions with a positively-charged substance (perhaps even HU from within cells which appears on the cell surface through the lysis of some bacteria) and that these become neutralised by salt.

With the experiments involving inclusion of salt, the increases in FSC and SSC are not matched. The clumping was basically established by the occurrence of major changes in SSC as compared to FSC. Along with major changes in SSC upon de-clumping of bacterial cells, salt could have changed the granularity of bacterial cells resulting in an additive effect on change in SSC values. We recognize that this is not a satisfactory explanation. We don't really understand why the changes in the FSC are less than the changes in the SSC. One possibility is that the cells themselves are longer than they are wide, and the cells probably align with each other during clumping, causing the size in one direction to be disproportionately large compared to the clumping in the other direction. If so, this would affect how cellular clumps pass through pores during sorting on a cytometer (i.e., there is a greater probability of a clump passing through an orifice if it is aligned in a certain manner with the orifice, than in all other orientations). This could have also caused the changes in the SSC to be larger than the changes in the FSC at the exit from the orifice.

3.1.5. Concepts underlying creation of HU mutants to compare binding of sites on HU to LPS and DNA

We created HU mutants as follows, to ask questions about LPS binding to HU (HU-A or HU-B).

1) Loop-replacement mutants (LRM): The beta hairpin loop, consisting of a pair of antiparallel beta strands on each HU monomer, is the canonical DNA-binding motif in that the two beta hairpin loops in an HU dimer (one from each monomer) wrap around the minor groove of DNA and grip the DNA double helix. We created a deletion mutant of each HU (i.e., both HU-A and HU-B) lacking the beta hairpin loop. The intention was to see the effect of this upon DNA binding as well as upon LPS binding, by using one or two of the multiple methods for which results are described above.
2) Lysine-replacement mutants (KRM): We replaced Lys83 and Lys86 by alanine in both HU-A and HU-B, both singly and in combination. The intention was to see the effect of these mutations which are likely to affect binding of HU to DNA, or LPS, through the non-canonical (side) surface, by using one or two of the above methods.

3.1.5.1. Creation and basic characterization of loop-replacement mutant of HU

The binding of LPS with HU can happen through canonical as well as non-canonical DNA binding sites on HU. The canonical DNA binding site on HU has substrate specificity towards double stranded DNA and RNA (i.e., with nucleic acids possessing secondary structure), which makes the binding of this site with LPS somewhat less likely. The canonical DNA binding site in HU recognises sugar-phosphate moieties. This is done by two unstructured loops, one derived from each subunit of the HU dimer, which embrace the major groove of DNA (minimum size 9 base-pairs) through formation of anti-parallel beta sheets by each loop. The fact that the arrangement of the sugar-phosphate in LPS is not helical makes conformation-specific contacts between LPS and the DNA-binding loops of HU unlikely, from an intuitive viewpoint. So, (1) we probed the possible involvement of the lysine residues located away from the canonical DNA binding region (the region of the loop) by removing the DNA-clasping beta-hairpin loops and

replacing them with a neutral 11 amino acids-long linker made of glycine and serine. In other experiments, (2) we also mutated the lysine residues at the putative non-canonical DNA-binding site and replaced these with alanine, both individually and combinedly. The schematic of the construct(s) along with the sequence of the glycine-serine-rich replacement loop are shown in Figure 17.

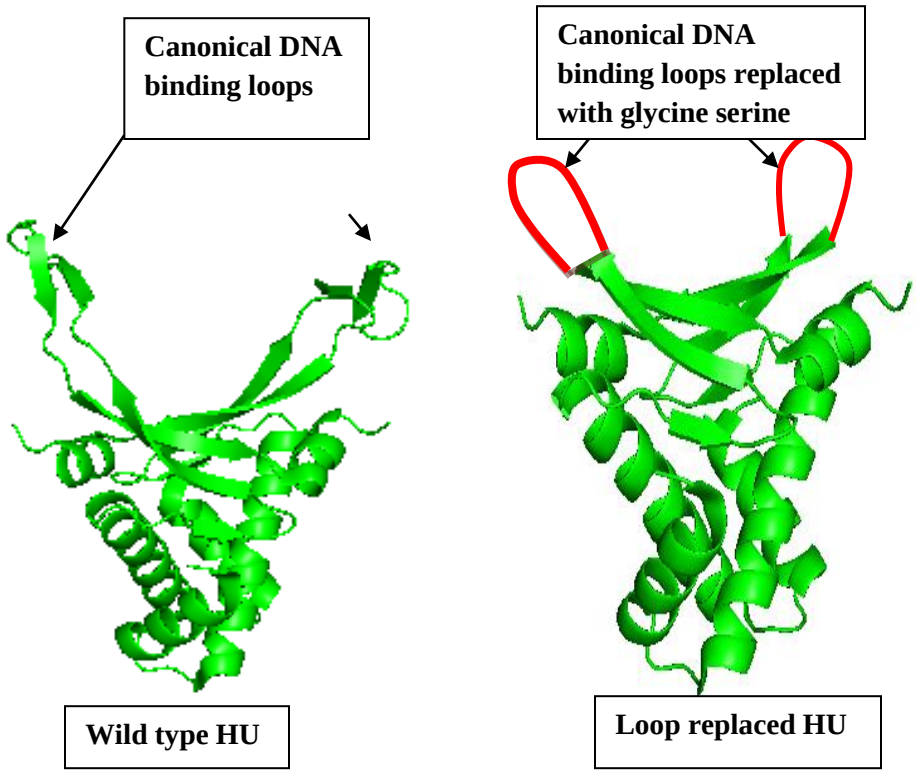
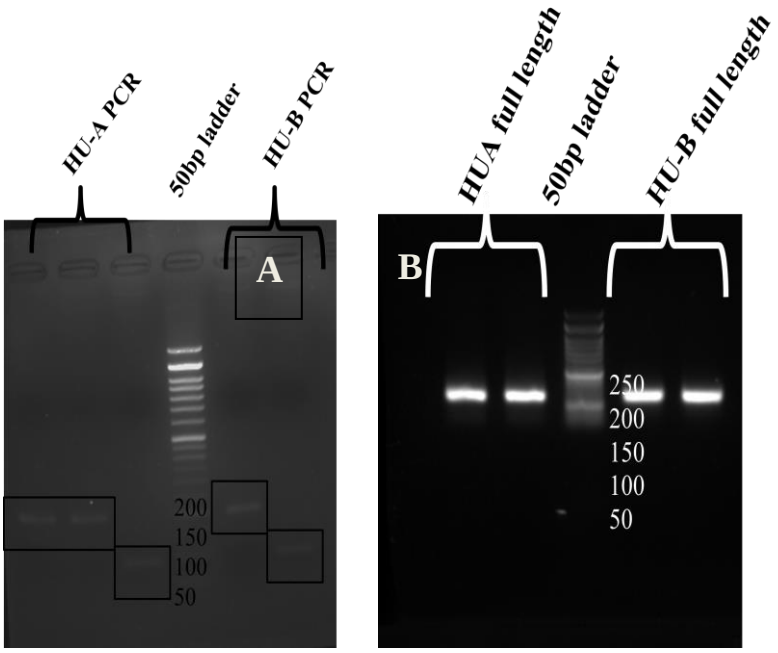


Figure17: Pymol image showing the replacement of beta hair pin loops with glycine serine linker. ***Panel A:*** Shows the full structure of wild type Anabaena HU, where the loops are presents and are marked as DNA binding canonical site. ***Panel B:*** Shows the same crystal; structure after the deletion of DNA binding loops and their replacement with glycine-serine linker. These cartoon are only representative images to explain the concept of loop replacement. The mutations were done on *E.coli* HU-A and B.

The extent of the DNA-binding loop region to be replaced was determined from the crystal structures of HU-A and HU-B, marking the starting and ending positions of the loop in each HU. As the length of the loops is identical in both HUs, loop replacement with a glycine-serine-rich loop was effected between residues 52 and 75 (excluding both of these residues). In other words, 22 residues comprising the DNA-binding loop were replaced by a 11 residues-long glycine-serine-rich stretch of residues, which may be called a linker (shown in red in the sequences depicted in Figure 17). The construction of the genetic fragments is shown in Figure 18.



HU-A

MRGSHHHHHHGSMNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVG
FGTFKVSGGGGSGGGGSNVPAFVSGKALKDAVK

Number of amino acids: 91;**Molecular weight:** 9250.44:**Theoretical pI:** 9.10

HU-B

MRGSHHHHHHGSVNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGF
GTFAVSGGGGSGGGGSKVPSFRAGKALKDAVN

Number of amino acids: 91:**Molecular weight:** 8988.03;**Theoretical pI:** 9.22

Figure 18: Agarose gels showing the individual fragments of genes encoding both the proteins, ***Panel A,*** along with the full length product obtained after SOE-PCR, shown in the ***Panel B.*** The Figure also shows the full amino acid sequences of loop-replaced HUs after cloning in pQE-30 vector along with their molecular masses and isoelectric points.

All experiments shown below were performed with loop-replaced HU-B, in which the 22 residues-long DNA-binding loop was replaced by the residues SGGGGSGGGGS. The loop-replaced protein was purified using the same protocols and buffers already used for purification of wild type HU-B. After purification, the oligomeric status and secondary structure of wild-type and loop-replaced HU-B were compared. The CD spectra were collected on an MOS-500 CD spectrometer (Biologic)using a cuvette of path length 1mm.

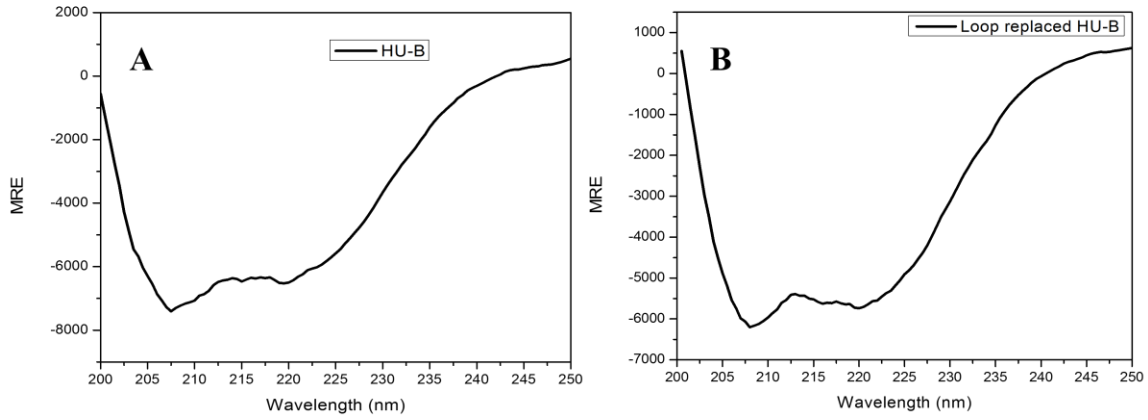


Figure 19: A comparison of the secondary structural contents of wild type HU-B in ***Panel A*** with loop-replaced HU-B in ***Panel B***.

It is clear from the circular dichroic spectra shown in Figure19 that the loop-replaced HU-B protein folds similarly to the wild-type protein, except for the fact that the contribution from randomly coiled structure is higher in wild-type HU-B, whereas loop-replaced HU-B appears to have a lower content of random coil, because it shows a zero-crossover above 200 nm, unlike wild-type HU-B. This can be explained by the fact that the 22 amino acids-long loop in wild-type HU-B is randomly coiled until it binds to DNA, and this loop is replaced by an 11 amino acids-long glycine-serine-rich peptide in loop-replaced HU-B (thus reducing the number of

residues in random coil conformation). We next compared the oligomeric status of loop-replaced HU-B with wild type HU-B, by performing gel filtration chromatography on a Superdex-75 10/300GL column (GE healthcare).The gel filtration chromatograms for both proteins are shown in Figure20.

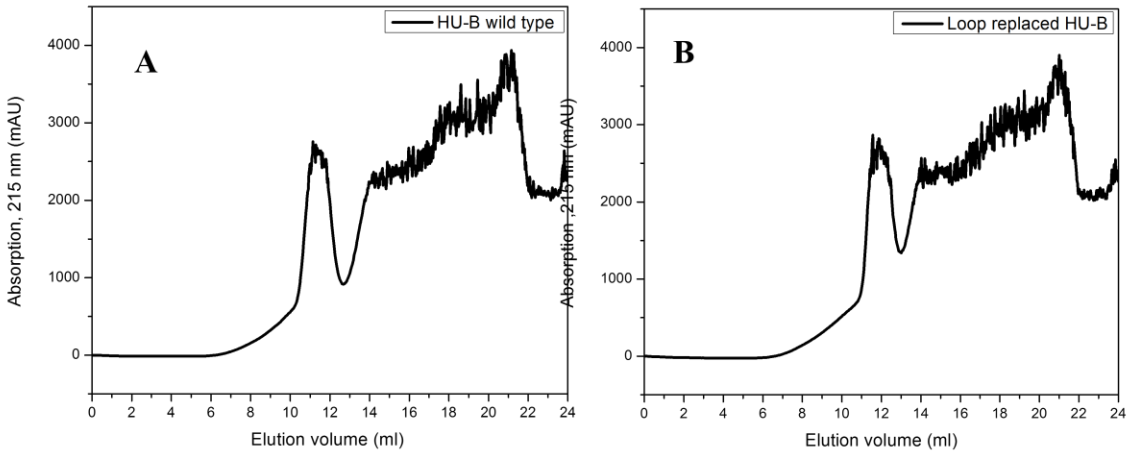


Figure 20: Size exclusion chromatograms of wild type and loop-replaced HU-B shown in the ***Panel A*** and ***B***. The absorption was monitored at 215 nm as HU lacks both tyrosine and tryptophan aromatic amino acid residues.

Gel filtration chromatogram peaked at ~11 ml elution volume for wild type HU-B and at ~11.7ml for loop-replaced HU-B. The chromatograms for both forms of HU were collected by monitoring absorbance at 215 nm by peptide bonds, and not at 280 nm as is usually done for most proteins, due of the absence of any tryptophan or tyrosine residues in the sequence of HU-B. The rising 215 nm elution peak towards the end of the column run is likely to be imidazole which is present in the samples.

3.1.5.2. Binding of HU loop-replacement mutant to DNA:Non-canonical site binding probed by EMSA and ITC

After performing initial structural characterization, the DNA binding abilities of both forms of HU were checked, using the electrophoretic mobility shift assay (EMSA) and isothermal titration calorimetry ITC). Decreasing concentration of both the proteins incubated with fixed amount of 4-way junction (a fixed volume of 2.7 µM 4-way junction) was electrophoresed on a 0.8%

agarose gel. It is clear from the gel shown in Figure 21 that both proteins can bind with DNA, as is evident from the appearance of smeared DNA band showing electrophoretic mobility shift in the lanes incubated with protein. With the same concentration of protein in loop-replaced HU-B as wild-type HU-B, more of the DNA was found remaining in unbound form even at the highest protein concentration used, as shown in the lanes following the control 4-way junction lane. The apparent higher net binding of wild type HU-B as compared to loop-replaced HU-B is obvious due to presence of DNA holding canonical site present in wild type protein. The binding of loop-replaced HU-B with 4-way junction favours the claims made by *Adhya et. al.*, who have shown that HU can bind with DNA through the sites other than the canonical DNA binding one

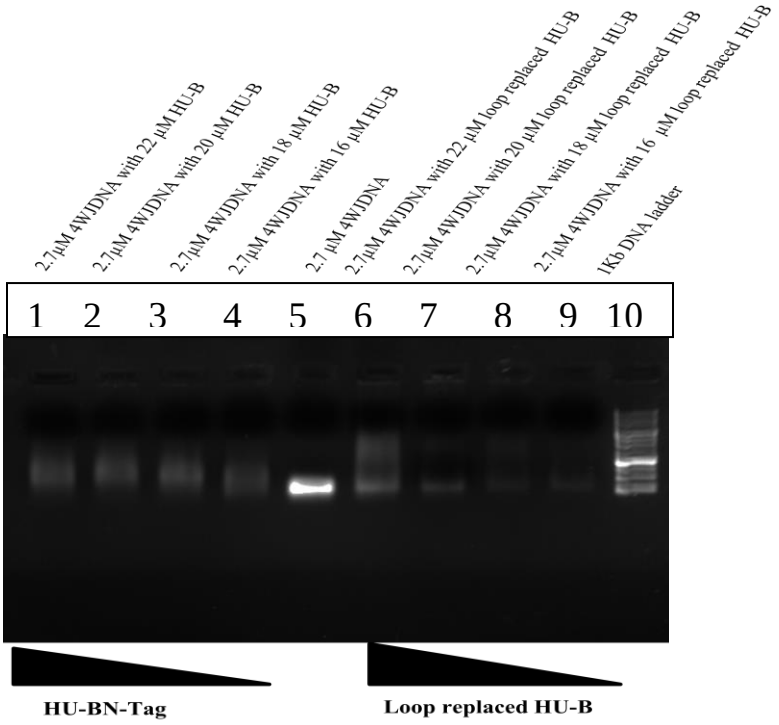


Figure 21: Showing the EMSA gel of wild-type and loop-replaced HU with 4-way junction. Lane 1 to lane 4, where the decreasing concentration of HU-B was incubated with a fixed concentration of 2.7 µM 4WJDNA, a high affinity binding substrate for HU protein. The protein concentration was decrease by a factor of 2 with 22 µM in first and 16 µM protein concentrations in the fourth lane. The fifth lane represents a DNA control followed by the electrophoretic mobility shift assay (EMSA) of the 2.7 µM 4WJDNA with the same concentration of loop-replaced HU-B protein as wild type HU-B and 1Kb DNA ladder loaded in the last lane. The binding of loop-replaced HU with 4WJDNA was also checked by isothermal titration calorimetry.

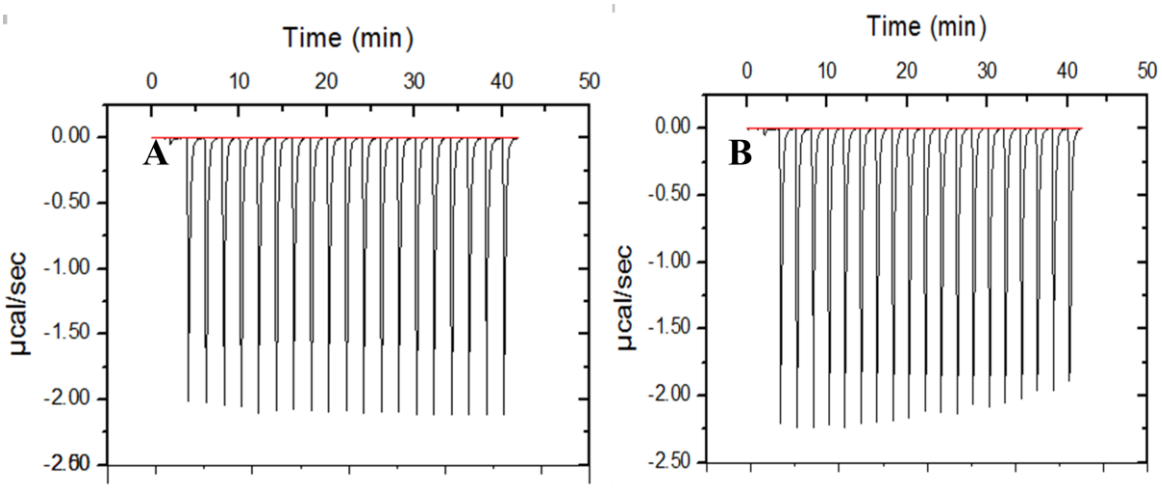


Figure22: Panel A: ITC graph represented as heat change of the system upon titration of 3µM 4WJ DNA with 20 µM loop-replaced HU-B. ***Panel B:*** The concentration of 4WJ DNA was reduced to 0.3 µM and was titrated with the same concentration of 20 µM loop-replaced HU-B.

The interaction between molecules can be studied by monitoring the heat change associated with the system in an ITC instrument. In ITC, the temperatures of chambers containing buffer and reacting molecules are monitored after every injection of ligand. The formation of bonds between interacting molecules releases heat into the system, triggering abstraction of heat from the system to keep the chamber temperature a constant. The release and abstraction of heat is seen as a negative spike in the ITC graph, indicative of association between molecules. The decreased negative peak with each injection represents a tendency towards some saturation of binding sites on the substrate. No saturation of heat exchange dips was observed with 3 µM 4-way junction (4WJ) concentration as shown in the Panel of the Figure with some decreasing trend observed after reducing concentration to 0.3 µM Panel B of Figure22. The intensity of heat exchange with each injection was found significant for establishing any interaction between loop-replaced HU-B with 4WJ. The non-saturating nature of heat exchange can be explained by having multiple DNA binding sites on loop-replaced HU for DNA and fast association dissociation of bound HU molecules from DNA molecules.As the magnitude of the changes seen was high, and also because upon reducing the concentration of the 4WJ DNA by 10-fold (from 3 micromolar to 0.3 micromolar) a trend of reduction in the heat changes observed with progressive injections was seen in the data, we do not think that these just represent dilution

heats. Rather, the affinity could be low, and in the range of tens or even hundreds of micromolar, since the protein has had its DNA-binding loops deleted and replaced with a glycineserine linker. So, what the data seems to suggest is a much more difficult approach to saturation, and one that we did not manage to reach.

3.1.5.3. Binding of HU loop-replacement mutant to LPS: Non-canonical site binding probed by liquid crystal birefringence

The binding of loop-replaced HU with LPS was also established by liquid crystal experiments, Dynamic light scattering and glutaradehyde cross-linking. In the liquid crystal experiment, loop-replaced HU-B from a 0.5mg/ml stock concentration was incubated with LPS liquid crystal and the changes in polarization of light upon interaction were monitored through observation on a polarized light microscope. The transition of the liquid crystal from a dark field to a light field is indicative of interaction between LPS and protein molecules, and the loop-replaced HU is demonstrated to show this interaction in Figure 23.

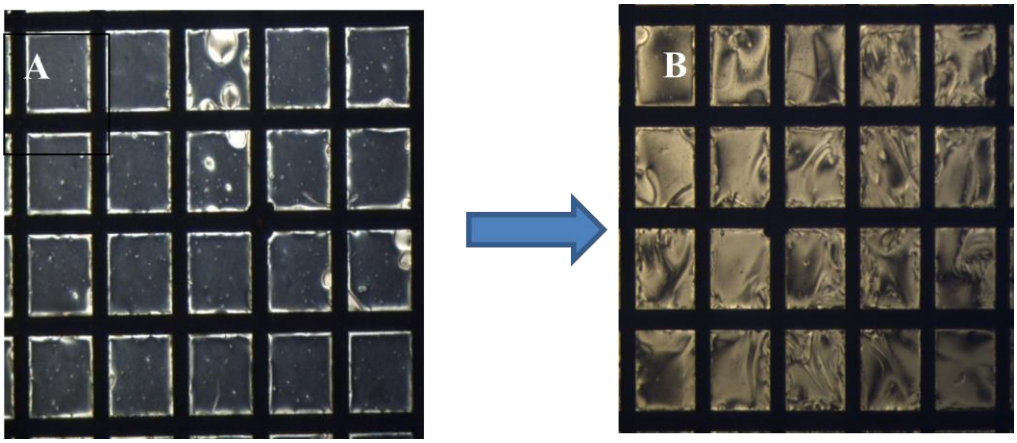


Figure 23: Panel A: showing the LPS liquid crystal in air. ***Panel B:*** showing the transition of the liquid crystal form dark to light state upon incubation with 0.5 mg/ml loop-replaced HU-B.

The loop-replaced HU-B was found to be as efficient as wild type HU-B in binding with LPS liquid crystal. The result from this experiment supports the likely role of sites other than canonical DNA binding sites in binding of HU with LPS.

3.1.5.4. Binding of HU loop-replacement mutant to LPS:Dynamic light scattering study of binding

The binding of loop-replaced HU-B with LPS was also further established by Dynamic Light Scattering (DLS) experiments. In dynamic light scattering, the sizes of control loop-replaced

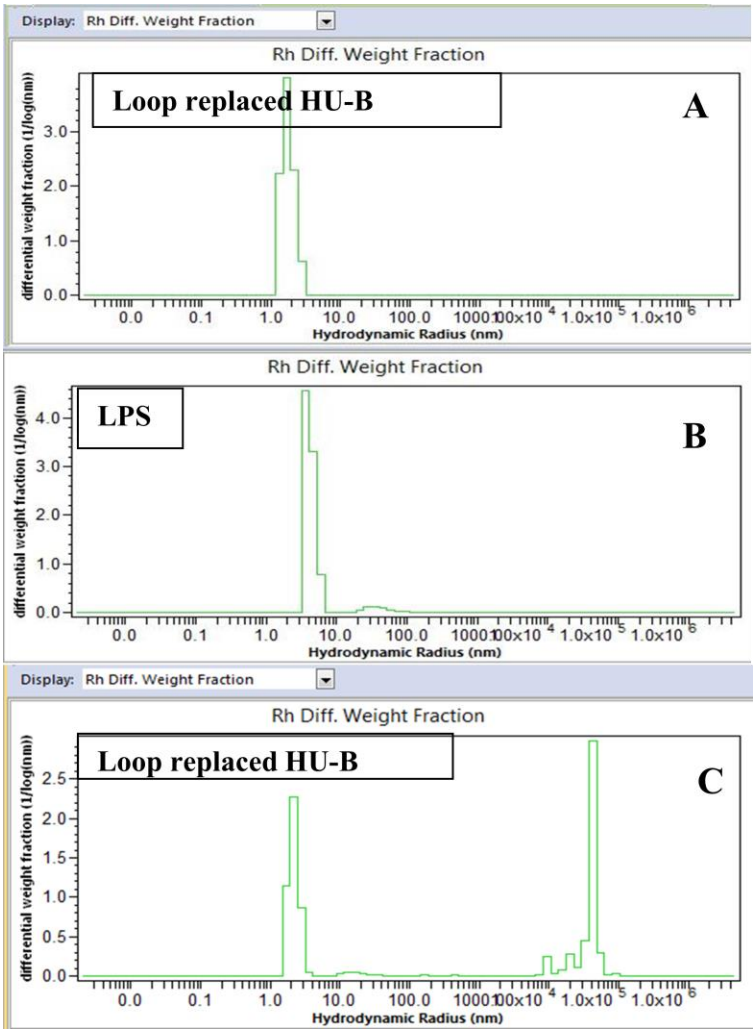


Figure 24: Panel A: Showing the dynamic light scattering images of only loop-replaced HU-B. ***Panel B:*** Showing the size distribution of LPS. ***Panel C:*** Showing the size distribution of combined HU and LPS after incubation.

HU-B was determined to be ~ 3 nm equivalent to the size of wild type HU. The sizes are plotted in term of weight-fraction with maximum population in 3 nm size range, as shown in the Panel A of the Figure 24. Only one peak is present in the whole DLS size distribution graph. In the case of LPS, more than 90 % of the molecules were found with sizes of ~ 8 nm, with some higher

species of bigger sizes as seen in Panel B of Figure24. The incubation of loop-replaced HU-B with LPS caused shifting of smaller aggregates into bigger aggregates which is evident by the diminishing of whole free HU population peak and conversion of LPS micelles from <10 nm size range to very big aggregates of size ~$1X10^6$ nm.

3.1.5.5. Binding of HU loop-replacement mutant to LPS :Glutaraldehyde cross-linking of LPS with loop-replaced HU-B compared to HU-B wild type

The same samples used for dynamic light scattering experiments were cross-linked with glutaraldehyde. Glutaraldehyde is a bi-functional cross-linker which cross-links two reacting groups in close proximity through the aldehyde group. The cross-linking reaction is stopped after 5 minutes through the addition of SDS-PAGE sample loading buffer, and boiling samples at 99^oC for 2 minutes. The boiling and SDS-containing sample loading dye denature proteins but leave covalent bonds formed by the cross-linker intact. The SDS-PAGE gel lane thus shows evidence of the crosslinking.

An SDS-PAGE gel is shown in Figure25. The wild type and loop replaced HU-B natively exists as dimers and tetramers which becomes evident after glutaraldehyde crosslinking shown in the second lanes of both of the gels. The binding of both the proteins with LPS can be seen by the presence of LPS-HU glutaraldehyde crosslinked aggregates which appears on the interface of stacking and separating gel. The LPS associated HU-B and Loop-replaced HU are pointed by with the red arrowhead in both the gels .These results indicate that the capacity of LPS binding is still retained in loop-replaced HU-B.

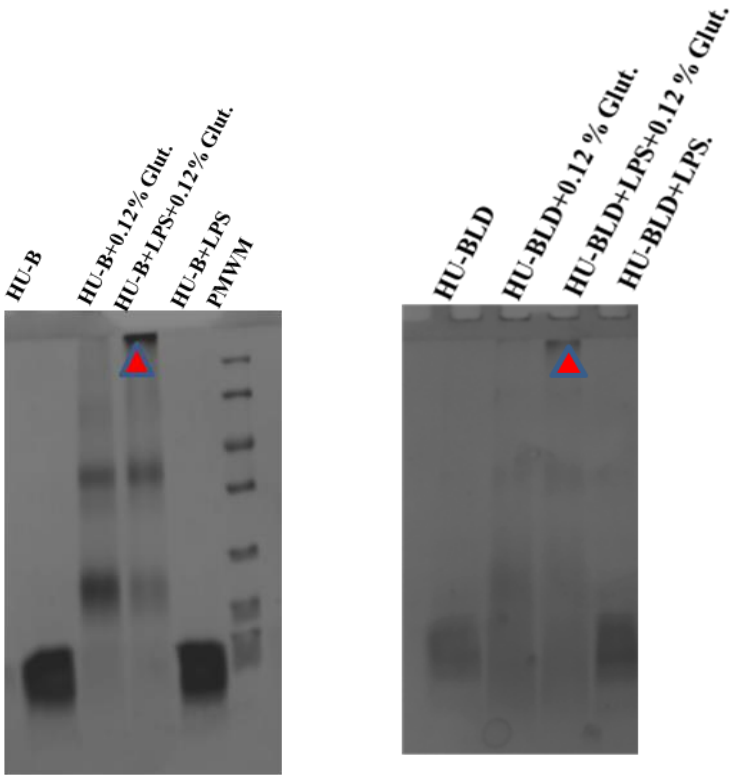


Figure 25: Showing the cross-linking of LPS with wild type and loop-replaced HU-B where the red arrow heads in 3rd lane of both the gels point towards the LPS HU-B and Loop-replaced HU-B cross-linked species with glutaraldehyde.

3.1.6. Contrasting HU binding to LPS through the canonical *versus* the non-canonical DNA-binding sites

As already discussed, HU binds to DNA principally through a canonical site involving a loop that becomes structured into a beta-sheet hairpin which grips the major groove of DNA (one loop from each unit of a dimer). We have already replaced this loop and shown that DNA-binding still survives in the loop-replaced mutant. We have also discussed the existence of a non-canonical DNA-binding site in HU, involving lysine mutants, K83 and K86, and discussed the possibility that these lysines bind to DNA as well as to LPS. The question being addressed in the subsequent sections is whether HU has a preference for binding to LPS through the canonical or the non-

canonical sites, assuming that both sites are capable of binding to DNA as well as to LPS. A schematic diagram is shown in Figure 26 to suggest that if the DNA-binding loops are involved in binding of extracellular DNA, the non-canonical DNA binding sites on the sides of the HU molecule could be more involved in the binding of LPS. The primary reason for this hypothesis is that this : The canonical site involves one DNA-binding site per dimer of HU. The non-canonical site involves two DNA-binding sites per dimer of HU, with one on the side of each monomer in the dimer. Obviously, the non-canonical site is more likely to be involved in clumping because of its having a greater 'valency', i.e., two instead of one, per HU dimer.

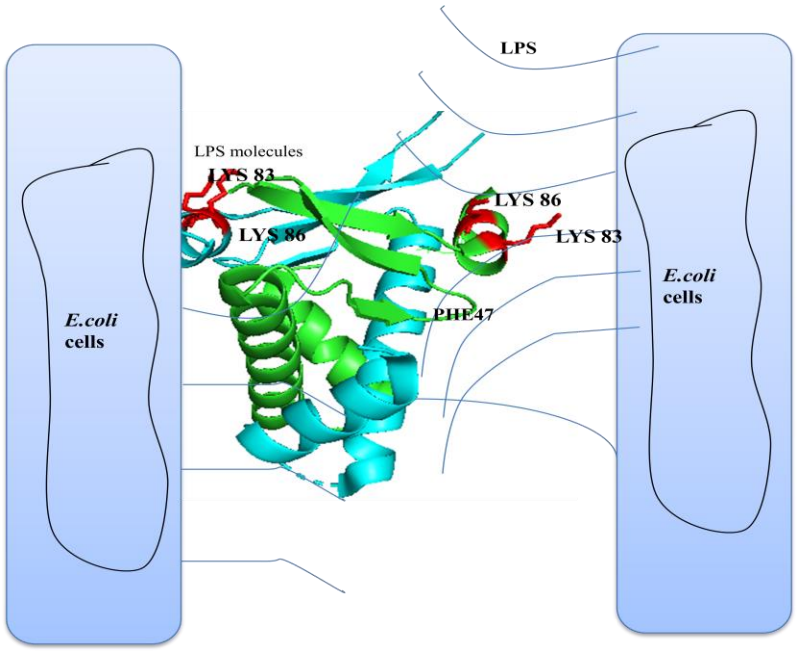


Figure 26: Showing a hypothetical schematic cartoon of HU mediated clumping of bacterial cells through the involvement of lysines 83, 86 (K83, K86) amino acids mediated electrostatic interactions with LPS molecules on bacterial cell surfaces. As LPS is very long molecule, after binding with lysines through polar head groups of Lipid-A, extended O-antigen sugar chain can wrap over whole HU surface leading to the quenching of fluorescence in P47W HU-B mutant-1.The involvement of beta hair-pin loops with LPS in this binding mode cannot be neglected as the loops are rich in positively charged amino acids. The binding of HU with LPS through this mode is independent of helix specificity.

Below, we show that the non-canonical site is better than the canonical site in terms of binding to LPS, by contrasting the fluorescence quenching of one 'probe' tryptophan each, introduced into HU, at the canonical DNA-binding site (F79W) and at the non-canonical DNA-binding site (F47W).

3.1.6.1. LPS binding to HU: Use of mutants (F47W and F79W) to distinguish between canonical and non-canonical site binding by fluorescence quenching of tryptophan

The role of the non-canonical DNA binding sites consisting of conserved lysine residues, K83 and K86, on the C-terminal terminating alpha helix in both the HUs was further probed by fluorescence quenching experiments with tryptophan mutants of wild-type HUs in whichF47 and F79 were replaced by tryptophan,to create F47W and F79W mutants. The selection of these mutants for quenching experiments was done on the basis of the locations of the F47 and F79 residues, respectively, known to be in the vicinity of the canonical site and the non-canonical site for DNA binding. The locations of each tryptophan are shown in the following Figure.

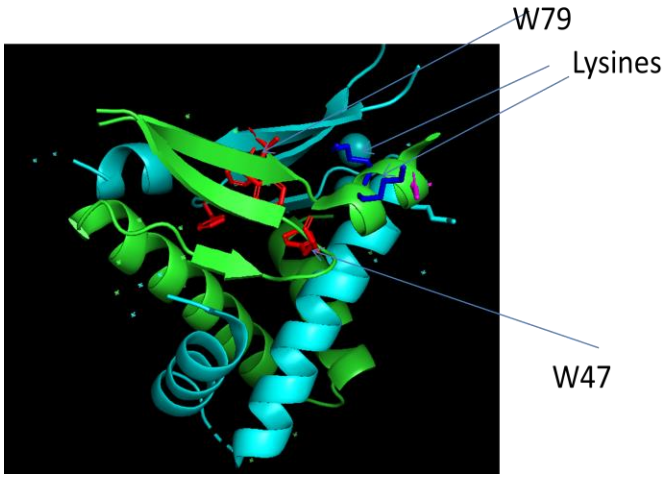


Figure 27: Cartoon image of HU heterodimer (PDB ID *2o97*) along with the positions of two conserved lysines forming non canonical DNA binding site labelled as blue sticks along with the positions of two phenylalanines 47, 79 (F47W, F79W) which are replaced with tryptophan amino acids shown as red sticks.

Each HU-B mutant containing F47W and F79W was incubated with increasing concentration of LPS. Fluorescence quenching of tryptophan was used as tool to establish interaction between HU

and LPS. The quenching studies were performed after incubating HU-B mutants and LPS for 30 minutes at room temperature to maximize interaction between them.

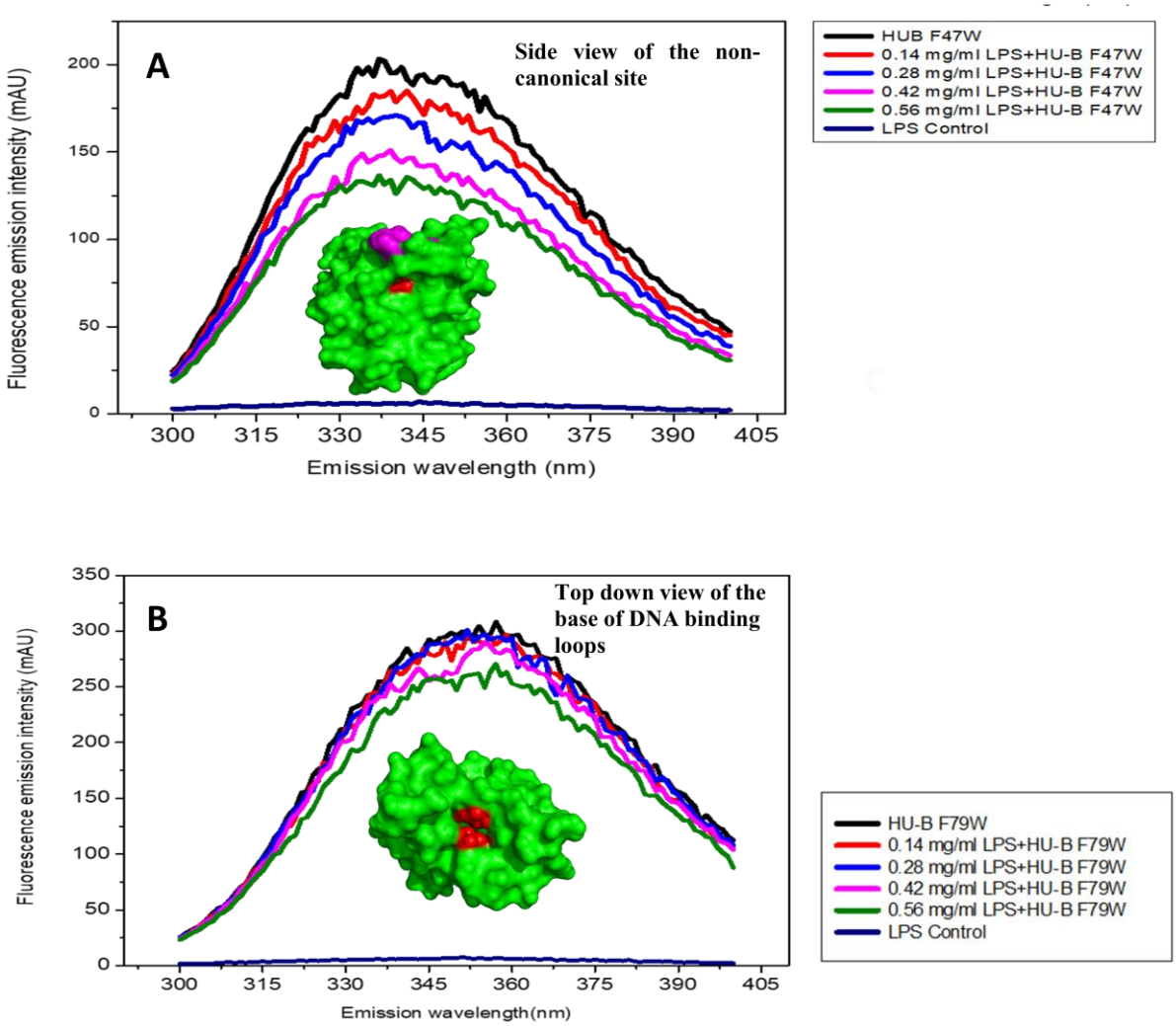


Figure 28: Panel A: Shows the intrinsic tryptophan fluorescence quenching of HU-B mutant, F47W, containing tryptophan in place of phenylalanine 47, with different concentrations of LPS. The concentration of reactants is written inside the individual panels. ***Panel B:*** Representing the fluorescence quenching of HU-B mutant 3, containing tryptophan in place of phenylalanine 79 (F79W) performed under same set of conditions as HU-B mutant 1. Both the mutants are abbreviated as mut-1 and mut-3.

The quenching was more pronounced in F47W than in F79W HU-B, where dose-dependent decrease in fluorescence quenching was observed in Figure 28(PanelA). The same effect seen in

F47W HU-B was observed at higher LPS concentration in F79W HU-B (Figure28PanelB). This suggests that both sites can bind to LPS, but that the F47 site represented by tryptophan in the F47W mutant is more affected (in terms of quenching) by lower concentrations of LPS, than the F79 site. This suggests that the non-canonical site is more often involved in LPS binding. The observation also makes sense in terms of a greater involvement of the non-canonical DNA binding site in LPS binding,since the canonical DNA binding sites could have evolved to be specific to helical DNA or RNA. The involvement of loops in LPS binding can't be deciphered from these studies.

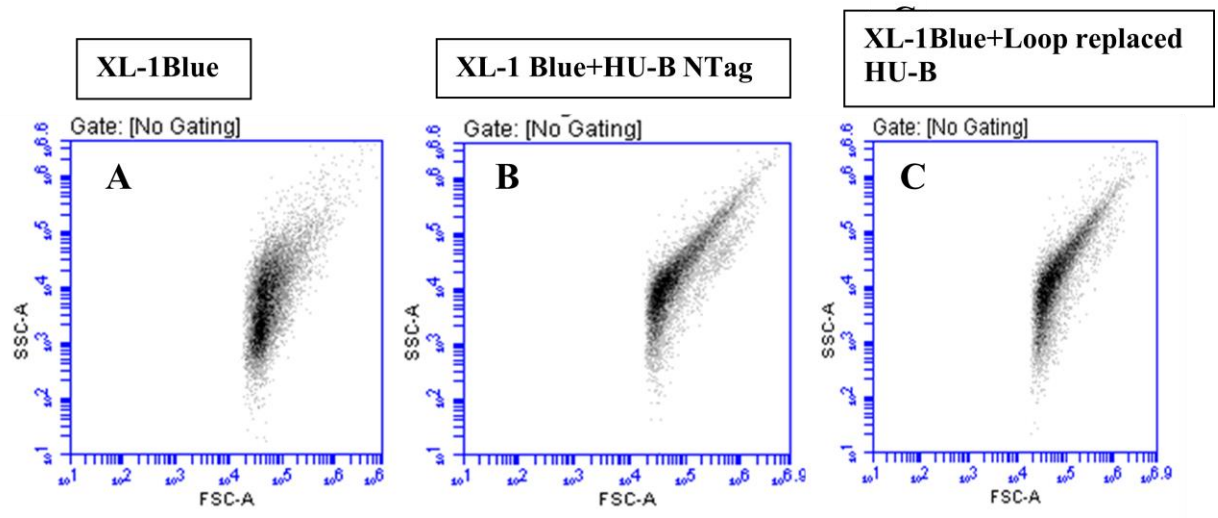
With tryptophan (W), it is known that its quantum yield, and its wavelength of maximal fluorescence emission can both be influenced by the polarity of its environment. Thus, for a tryptophan that is partially-exposed to the solvent, with atoms in contact with the solvent, it may be held that its quantum yield may be affected by the binding of a ligand in the vicinity since this would change the environment of the exposed tryptophan side-chain atoms. The question thus is whether W47 and W79 are exposed or buried. If they were buried, changes upon ligand binding would not be expected; however, if they were exposed, such changes could occur. In the figure below, the colour red is used to show the F residues at both the canonical and non-canonical sites. The F residues are shown with spheres representing side-chain atoms. The colour green is used to show the general surface of the protein, and the colour magenta is used to show the surfaces of lysine residue side-chains at the non-canonical site. The lysines of the canonical site are not seen because the loops carrying them in HU are not seen in the crystal structure (no electron density; mobile loops which only become fixed upon DNA binding; no DNA-bound HU structure exists in the database, but we know from Anabaena HU that its loops are only seen in complex with DNA). In the figure below, the left panel shows HU's structure as it would be seen if one were looking down upon the base of the canonical DNA binding site above. The two F79 residues are seen 'face-to-face' from the two monomeric HU chains. It can be seen (a) that they have atoms on the surface (which is why they are seen in this representation), (b) that if they were replaced by W which has two rings, a six-membered ring an a five-membered ring, there would be even more atoms exposed, and (c) that the lysines which are proposed to bind to LPS in the loop above this base of the canonical site would be distal to these F/W residues, but the hydrophobic tails of the LPS would create a hydrophobic environment which could change the

quantum yield of the partially-exposed tryptophan residues; enough to cause a quenching. Similarly, in the right panel (a) the F residue can be seen to exposed, and likely to be more exposed when replaced with W, because of the larger number of atoms in W, and (b) close to the pair of lysine residues and, therefore, likely to have the quantum yield of W changed by the presence of the hydrophobic tails of LPS. Thus, in both instances, W79F at the base of canonical site, and W47F at the non-canonical site have side-chain atoms exposed to the solvent and might display quantum yield changes upon binding of LPS to HU.

These experiments provided clue about involvement of conserved lysine along non-canonical DNA binding site alone or with the involvement of loops in binding with LPS along with supporting models of HU mediated clumping of bacterial cells shown in the following Figure.

3.1.6.2.LPS binding to HU : Similar clumping efficiencies of wild-type HU andloop-replaced HU, despite differences in DNA-binding efficiency

The binding of loop-replaced HU-B with XL-1 Blue *E.coli* cells was studied by flow cytometry experiments, by looking at increased forward and side scattering. The cells after incubation with equal amount of wild type and loop-replaced HU-B at room temperature were passed through BD AccuriC6 flow cytometer, to compare clumping efficiencies of two proteins. All the reactions were carried out in PBS (pH7.4 buffer).



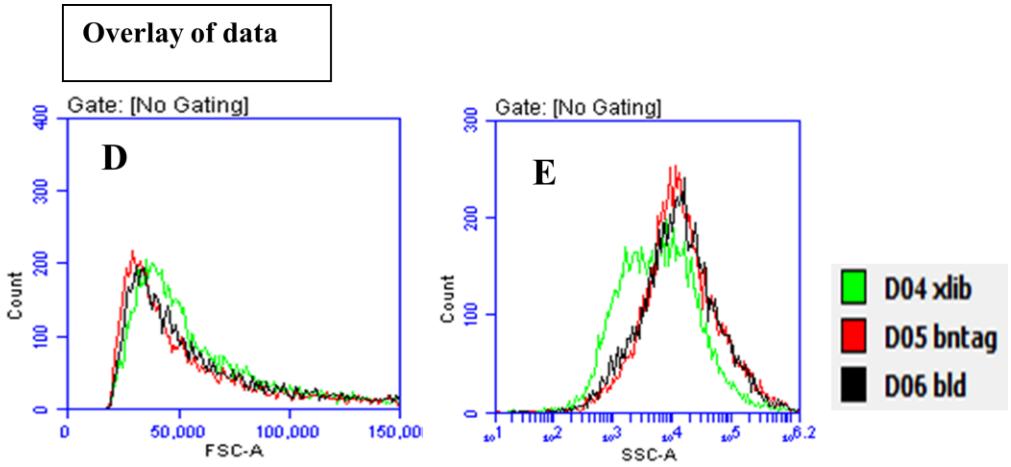


Figure 29: Showing the scatter plot of bacterial cells plotted between FSC-A and SSC-A detectors. ***Panel A:*** Scatter plot for XL-1 Blue cells which served as control. The number of the cells was same in all sets of flow cytometry experiments along with the concentration of protein at 0.4 mg/ml. ***Panel B****:* Scatter plot of XL-1 Blue bacterial cells treated with wild type HU-B***. Panel C:*** Scatter plot for the cells treated with loop-replaced HU-B. ***Panels D***, ***E:*** Showing the overlays of bacterial cell number plotted against the FSC-A and SSC-A detectors.

The change in side scatter was maximum for HU-B wild type followed by loop-replaced HU-Bs (Figure29,Panel E). The sizes of the cells was bigger in protein treated cells as compared to control XL-1 Blue cells in PBS only (Figure29,PanelsD,E) which can also be seen in the dot plot by the presence of streak which is absent in control XL-1 Blue cells. The chances of getting high side scattering in clumped cells will be more due to more light scattering from many bacterial surface planes associated with each other. The horizontally arranged bacterial clump after passing parallel through laser will behave like single cells in spite of clumping giving low intensity in forward scatter signal. The possibility of getting more side scattering in clumped bacterial cells is depicted in Figure 30.Loop replaced HU has only one LPS binding site (the non-canonical site) in contrast to wild-type HU. Thus, it could be expected to show less of an ability to cause clumping.

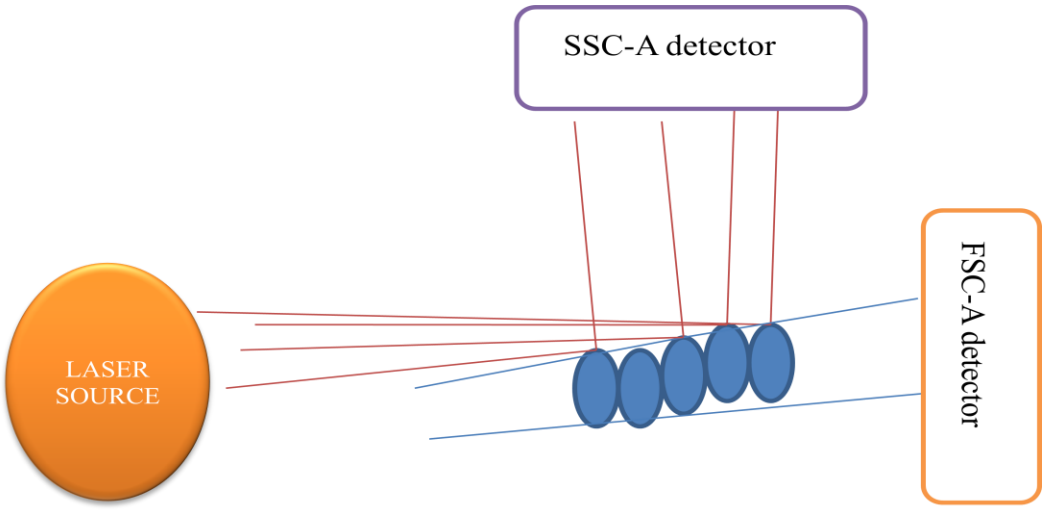


Figure 30: Schematic showing the forward and side scattering from bacterial clumps. The possibility of getting clumps of this arrangement is highest because of largest inter-bacterial surface area.

3.1.7. Non-canonical DNA-binding site lysine mutants

3.1.7.1. Binding of HU non-canonical site lysine-mutants to LPS: Fluorescence quenching study of binding

The fluorescence quenching of F47W HU-B relative to the poorer quenching of F79WHU-B and the clear evidence for binding of loop-replaced HU with bacterial cells(suggesting that the loop is expendable) led us to speculate that the role of the conserved lysines clustered above this tryptophan residue (F47W) is important in the binding of HU-B to LPS. The two lysines, K83 and K86, are conserved amongst HU proteins of mesophilic and archaeal origin, as shown in Figure 31, presenting the amino acid sequences of HU-A and HU-B with lysine residues marked in red, and in Figure 32, presenting the structures of HU-A and HU-B.

HUA:MNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHR AERTGRNPQTGKEIKIAAANVPAFVSGKALKDAVK

HUB:VNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGFGTFAVKERA ARTGRNPQTGKEITIAAAKVPSFRAGKALKDAVN

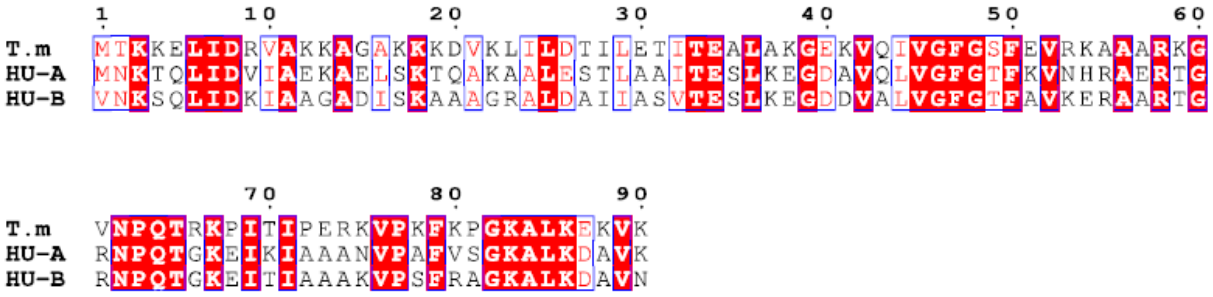


Figure 31: Sequences of both the HU from *E.coli*. The amino acid sequence marked as red in both HU represents a conserved C-terminal amino acid sequence. Apart from the bacteria of mesophilic origin this sequence is conserved among archaeal HU. Sequence alignment of Thermotoga and E.coli HU.

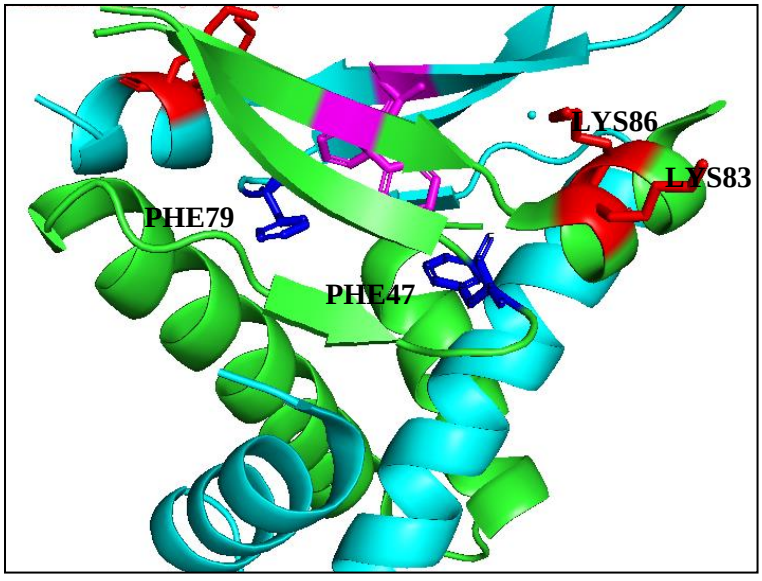


Figure 32: Showing the relative postions of conserved lysine clusture with respect to the phenyalanine residues (F47 and F79). The conserved lysines are represented as red sticks on both the subunits of HU heterodimer (PDB ID: 2O97)**.**

To decipher the role of these lysines in LPS binding, HU mutants containing single and double mutants of K83 and 86 were made by replacing them with the amino acid residue, arginine.The proximity of K83 and K86 to F47 led us to mutate these conserved residues in both the HUs singly as well as mutating both of them together. The mutations were done on HU-B containing W47 as template as quenching with LPS was observed in this mutant. Lysines were mutated to alanines. The genetic constructs are shown in Figure 33. The secondary structure of all mutants was checked by circular dichorism spectroscopy along with oligomeric structure analysis with size exclusion chromatography

The sequences of the gene after cloning and sequencing are as following

HU-B wild type

MRGSHHHHHHGSVNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGF
GTFAVKERAARTGRNPQTGKEITIAAAKVPSFRAGKALKDAVN

HU-B(K83)

MRGSHHHHHHGSVNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGF
GTFAVKERAARTGRNPQTGKEITIAAAKVPSFRAGAALKDAVN

Number of amino acids: 102: **Molecular weight:** 10534.91: **Theoretical pI:** 9.77

HU-B(K86)

MRGSHHHHHHGSVNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGF
GTFAVKERAARTGRNPQTGKEITIAAAKVPSFRAGKALADAVN

Number of amino acids: 102 : **Molecular weight:** 10534.91 :**Theoretical pI:** 9.77

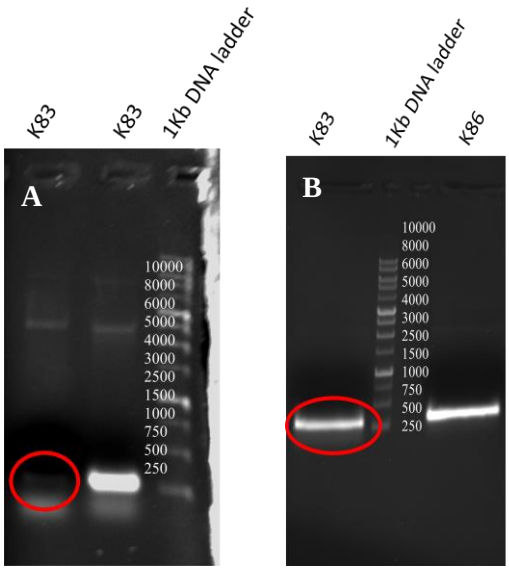


Figure 33: PCR product of the HU-B K83 and 86 mutants. ***Panel A:*** The low intensity of PCR product in the gel is marked in red circle. The intensity was low due to the formation of hairpin loop in the K83 reverse primer. The problem with hairpin loop was resolved by the addition of 10% DMSO to the reaction mixture. ***Panel B:*** Shows amplified products in huge quantity after the addition of 10% DMSO.

The genes were cloned in pQE-30 vector between *BamHI* and *Hind III* restriction sites. The plasmids were expressed in XL-1 Blue cells and protein was purified using N- terminal 6X-Histidine tag using same buffers used for the purification of wild type HU. Secondary structure comparison of wild type HU with HU mutants. The secondary structure comparison of wild and mutated HU-B is shown in Figure 34.

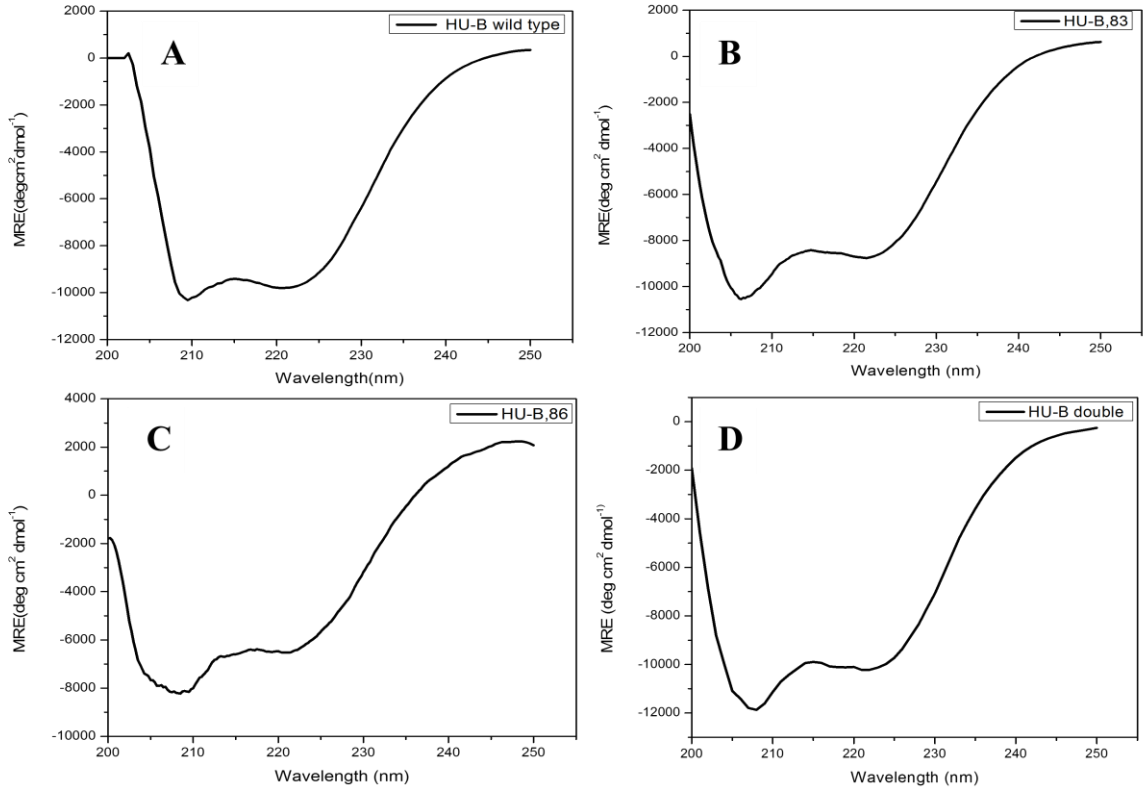


Figure 34: CD spectra based comparison of all three HU-B mutants with HU-B wild type. The CD spectrum for each mutant and wild type is shown in the ***Panel A*** to ***D***. The data is plotted in term of MRE (mean residual ellipticity) for all HU-B wild type and mutants containing Lysine 83 (K83A) and 86 mutations (K86A) individually and together(K83A, K86A)**.**

All the proteins were found to have same fold like wild type HU-B. The shape of the spectra is the same for all protein spectra. An error in protein concentration estimation could lead to changes in MRE values. HU contains no tryptophan residues. So, we use variants containing W instead of F in one of the three locations of F residues, and estimate concentrations based on comparison of the raw ellipticity CD signals of the F-containing WT HU, and its W-containing variants. This method can lead to small mis-estimation of concentrations of HU, which would be reflected in the MRE for different samples. Their gel filtration profile also showed main population as dimer. The LPS binding ability of these mutantswas checked by incubating them with LPS in the same concentration used in the previous tryptophan quenching experiments.

Tryptophan quenching experiments showed even single lysine-to-alanine mutants were able to bind to LPS as there was fluorescence quenching. These are shown in Figure 35.

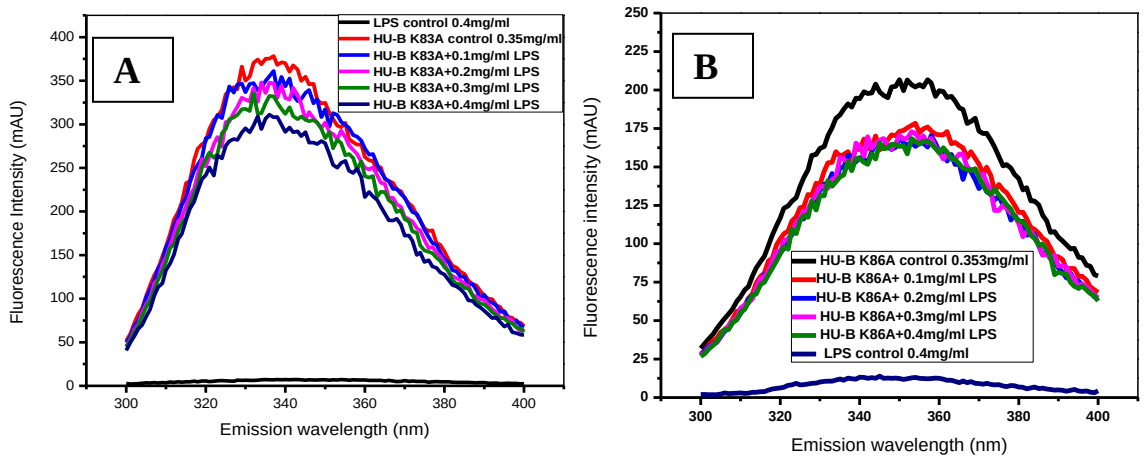


Figure 35: Showing the graph representing tryptophan fluorescence quenching by two single lysine mutants K83A and K86A. The mutations were done on HU-B gene containing tryptophan in place of phenylalanine at 47th amino acid position. All the reaction was performed in PBS (pH.7.4). The LPS solution was sonicated before use to break down the big micelles. The protein concentration was 0.35 mg/ml for all mutants. The concentration of the LPS used for titration reactions are written inside each panels. ***Panel A:*** Showing the LPS mediated tryptophan quenching of K83A**.** ***Panel B***: Quenching experiment for K86A mutant.

Quenching was observed in all tryptophan mutants, suggesting that the regions involved in these lysine mutants was involved in LPS binding. From these quenching experiments, we can assume that HU can bind with LPS through canonical DNA binding sites facilitated by the involvement of loops, and to non-canonical site through lysine residues. To prove the involvement of the loops and lysines together, lysine mutations of loop-replaced HU need to be done. The involvement of non-canonical site lysine residues can be further probed by making lysine double

mutant K83, 86A in HU-B F47W mutant, and probing its binding with LPS by quenching experiments.

3. 1. 8. Conclusion regarding HU's role in biofilm formation :

HU is exuded from bacterial cells either by lysis or through dedicated secretion systems along with DNA (e.g., through the T4SS system which gets activated when bacteriareach stationary phase in culture). The exuded HU gets immobilized on the bacterial cell surface through electrostatic interactionwith negatively charged LPS. The HU protein, due to its mutimeric nature, tends to crosslink bacterial cellsand the capacity for such cross-linking is higher in HU-B as compared to HU-A owning to its forming higher order multimers than HU-A. The binding sites for LPS on HU were studied thoroughly by making loop-replacement and lysine-replacement mutants. As the possibility of LPS binding to HU through the canonical DNA-binding site is low (given that DNA is likely to occupy such sites for all DNA-bound HU), the DNA-binding sites other than the canonical DNA binding site were also explored, to examine whether these are also capable of binding to LPS. The involvementof loops in LPS binding was explored by making loop-replacement mutants, which showed binding with LPS but very hardly any binding with DNA. The presence of other non-specificLPS binding site on HU was probed by looking at the conserved charge pockets which can interact with opposite charges irrespective of conformation. The presence of conserved lysines 83, and 86(forming a non-canonical DNA binding site, along with Phe47) caused us to think about the involvementof these residues in LPS binding. The clue of the involvement of K83, and K86, in LPS binding came from the tryptophan fluorescence quenching studies on F47K and F79W, which are two previously-created mutants of HU-B in which Phe residues are replaced with Trp residues. The position of tryptophan in F47W is at the non canonical DNA binding site, exposed to the solvent, and the position of the tryptophan in F79W is buried somewhere under the canonical DNA binding site (covered by it). Tryptophan in F47W showed remarkable tryptophan fluorescence quenching upon titration with LPS, presumably owing to binding facilitated by K83 and K86. After looking for the presence of charged amino acids along the non-canonical DNA binding site,a conserved sequence of

KALKD amino acids in both the HUs was found near the C-terminal helix. Of these conserved amino acids, K83 and K86 projected their side chains over tryptophan in F47W mutant hence making them an ideal candidate for mutational analysis. The binding of K83 and K86 lysine mutants with LPS was checked by fluorescence quenching studies done with HU-B having F47W mutation. The quenching studies on mutantsK83A and K86A showed no reduction in quenching making the involvement of only the non-canonical DNA binding site in LPS also questionable. The involvement of the both the lysines in LPS binding can be further probed by making the double mutant of K83,86A on HU-B F47W background.LPS can bind with HU through non-canonical DNA binding sites with the additional involvement of loops. To prove the differential involvement of the canonical loop and the non-canonical lysine cluster in terms of LPS binding, lysine mutations of loop replaced HU need to be done to examine whether binding to LPS is completely abolished in such mutants. In summary, HU could interact with LPS through the non-canonical site, consisting of lysines (K83, K86) or through a pair of lysines present on the canonical DNA binding site, consisting of the partially-unstructured loops containing beta-hairpins. The binding of HU with LPS through the involvement of the full canonical DNA binding site (involving both loops) is unlikely, as there is no DNA-like helical structure present in LPS or, indeed, the presence of any structure capable of adopting a DNA-like structure or volume. Thus, we think that HU interacts with LPS only through appropriate electrostatic interactions between the sugar-phosphate moieties on LPS, and a pair of positively charged amino acids on HU. Such pairs are present on both the canonical and non-canonical DNA binding sites. For a schematic explaining this, please see the figure 36, in which Panel A shows how the lipid A headgroup of LPS can interact with one loop from amongst the two loops constituting the canonical DNA-binding site, using one charged phosphate group at each end of the lipid A headgroup of LPS. Panel B shows how each end of the lipid A headgroup might be bound to a pair of lysines at the non-canonical DNA-binding site, spanning two HU dimers in an HU multimer.In our opinion, any multimeric protein displaying a large number of positively charged amino acids on its surface will display a certain amount of binding to LPS and to cells. We think that with HU there is the additional possibility of the operation of some molecular recognition (in terms of the placement of lysines at a distance that is appropriate for binding to canonical and non-canonical DNA binding sites; as shown in the figure). However, at present we have no evidence for this, or for our view that HU's binding to LPS is made somewhat specific

by the positioning of its lysine residues. However, we agree that other proteins can play this role. Additional important points are (1) that HU is one of the most abundant proteins in E.coli, (2) that there is evidence that HU is present in biofilms, and outside the cells, presumably in a state bound to extracellular DNA, and (3) that there is evidence that HU is secreted by bacteria to the cell surface.

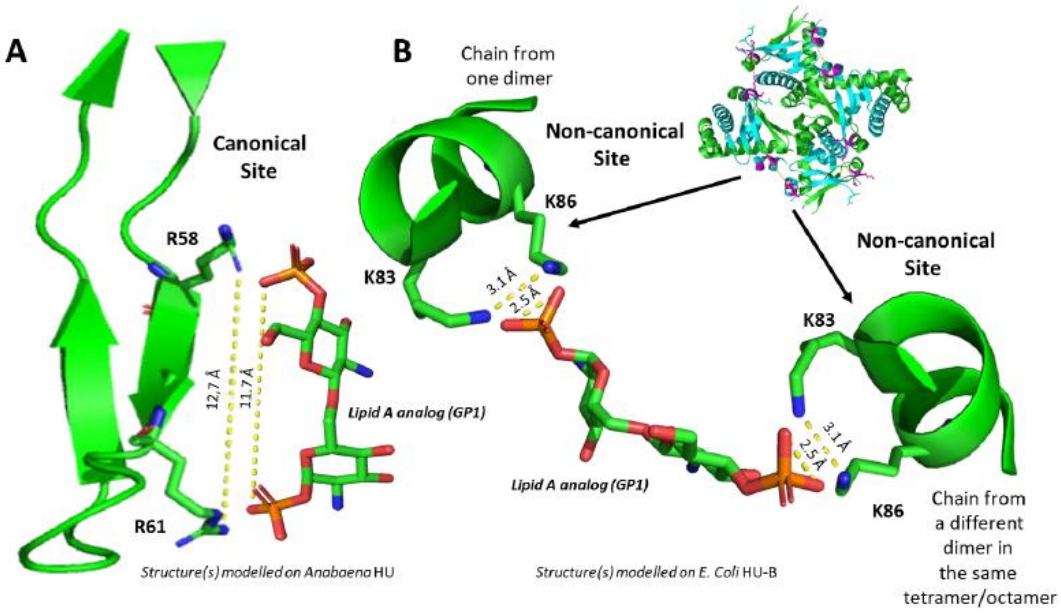


Figure 36: Showing the possible modes of HU's interaction with LPS through the involvement of arginines (R58 and R61) on canonical ***Panel A*** and lysines (K83 and K86) present on non-canonical DNA binding site ***Panel B.***

3.2. ***SECTION 2 : Studies of the homodimeric interface of HU-A and HU-B***

3.2. 1. The HU dimeric-interface survives even at 4 molar urea concentrations (Cross-linking experiments)

The occurrence of subunit exchanges in HU was probed by investigating the strength of the dimeric-interface strength in both the HU homologs, HU-A and HU-B, to assess the feasibility of populations of HU2A or HU2B turning into HUAB. To assess the strength of the interface, the proteinswere incubated overnight (at concentrations of ~ 0.2mg/ml) with different concentrations of urea, and this was followed by attempted cross-linking with 0.1% glutaraldehyde, to fix and

examine the existence of dimers at different concentrations. It may be noted that the samples were run on SDS-PAGE which is expected to separate all monomers that are not covalently cross-linked. Therefore, dimers that do not become subject to glutaraldehyde-mediated crosslinks are not detected during gel electrophoresis, and this is not a technique for estimating the relative abundances of dimers (note : the structure of HU is such that it is difficult to conceive of the existence of structured monomers, and all HU polypeptide chains are expected to be dimeric). On the other hand, if crosslinked chains are still observed in the presence of very high concentrations of urea, after overnight incubation, it must be concluded that dimers exist even at such high urea concentrations, at least to the extent observed on the gel (as crosslinked dimers) since there is no possibility of any monomers becoming crosslinked through chance collisions at concentrations as low as ~ 0.2 mg/ml. This having been clarified, let us turn to the data. The samples were - run on SDS-PAGE, and the results are shown in Figure 37.

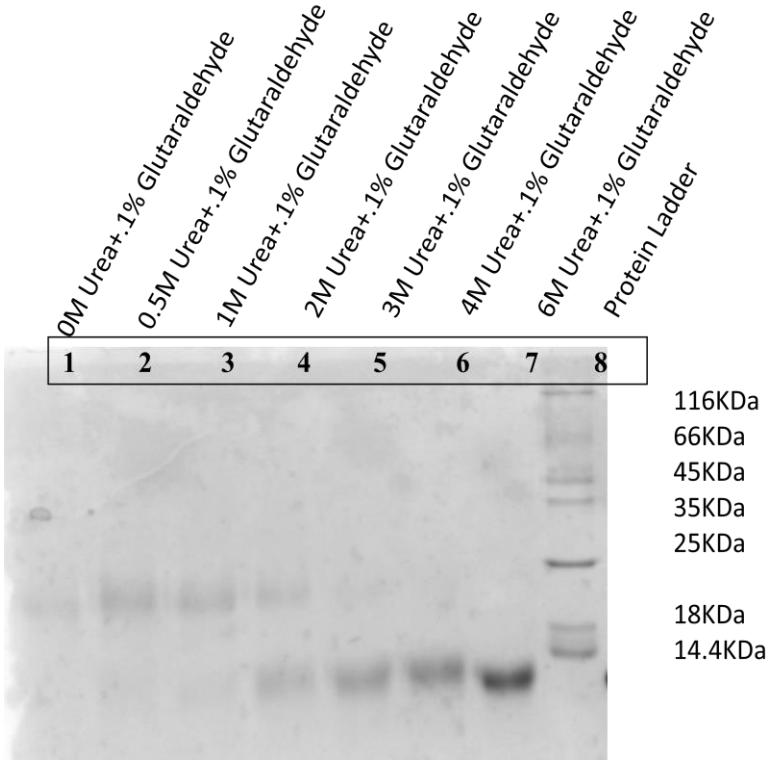


Figure 37: Glutaraldehyde cross-linking of HU-A incubated at different molar concentrations of urea.

The figureshows clear evidence of the entire HU-A population being dimeric in the absence of urea, with almost no monomer seen. It also shows clear evidence of the survival of HU-A dimers-even at 2M urea concentration,whereas at higher urea concentrations increasing fractions of the population of dimers dissociate into monomersthat are no longer subject to glutaraldehyde-mediated crosslinking, as seen in the lanescorresponding to the use of 3 M urea concentration, and with higher urea concentrations.

Similar experiments were conducted with glutaraldehyde crosslinking attempted after incubationof HU-B with various concentrations of urea. as shown in Figure 38. Here too, it is evident that the entire HU-B population exists as dimers, tetramers and higher species when there is no urea present, or when the urea is present at a concentration of 1 M. Thus, HU-B existspredominantly in dimeric, tetrameric, and higher oligomeric states, as seen- in the first and second lanes of the gel shown in Figure 38.

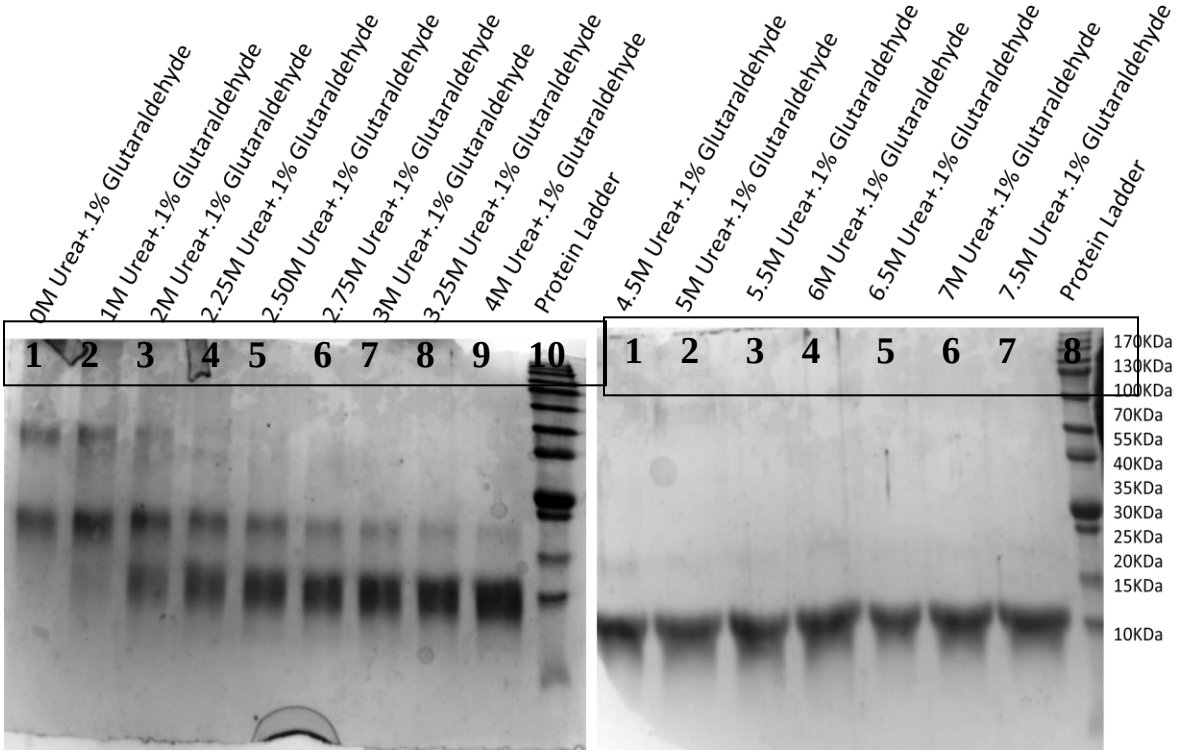


Figure 38: Glutaraldehyde cross-linking of HU-B at different molar concentrations of urea.

With increasing urea concentration, there is a disappearance of the tetrameric HU-B population, as evidenced by the gradual reduction in intensity of the band corresponding to the tetrameric population with increasing urea concentrations as seen -in the lanes of the SDS-PAGE gel corresponding to urea concentrations of 2 M. Above this concentration, all the way up to 4 M urea, the dimeric population survives, but dimishes in amount, with compensation seen in the form of increasing amounts of monomeric HU-B. Above 4M concentration of urea,–it appears that there is complete dissociation of the HU-B dimers into monomers, as glutaraldehyde treatment is unable to capture any dimeric species. As already mentioned, such data indicates that there is persistence of the dimeric interface even at concentrations as high as 2 M urea (for HU-A) and 4 M urea (for HU-B) although the dimeric or tetrameric populations captured with glutaraldehyde are likely to underestimate the actual persistence of the dimeric interface, since – probabilistically speaking – every dimer, or higher oligomeric species, is unlikely to be captured with gluraraldehyde cross-linking. The persisting dimers that do not undergo cross-linking would register as monomers in this method of assessment. That being the case, it may still be argued that the persistence of the HU-B dimeric interface even at 4M urea concentration suggests that the interface is very strong, and it makes the hypothesis of subunit exchange among HU homologs quite questionable and also raises serious questions about the earlier claims of HU-B being unstable.

Further, the oligomeric status of the HU-B in PBS buffer without urea, and in7.5 M urea concentrations was checked by dynamic light scattering (on a Malvern Zetasizer instrument). HU-B was found to form dimers(~3 nm diameter) and linearly-associated tetramers or octamers (~11 nm diameter) along with completely-unfolded extended chains(with an average ~100 nm length, or diameter), as evident from the presence of distinct peaks corresponding to each of these sizes.

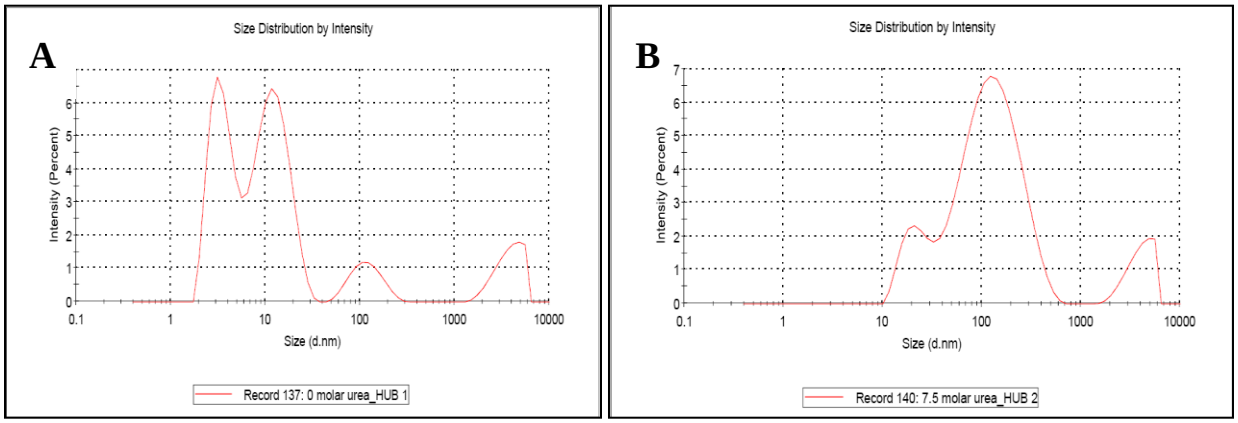


Figure 39: Panel A: Size distribution graph of HU-B under non-denaturating condition in PBS buffer. ***Panel B:*** The size distribution of same protein in 7.5 M urea concentration.

The sizes are shownplotted in term of intensity fraction, rather than weight fraction, where a very small quantity of large ~100 nm aggregates wouldappear to dominate the size distribution spectrum, because scattering increases non-linearly with increasing particle size. However, no aggregates of intermediate size are observed. This set of spectra informs us about the HU-B dimer being the most abundant population followed by the tetramer, or octamer. Upon incubation with 7.5 M urea, the entire population undergoes denaturation, with some octamer surviving and with all chains fully-extended. It is unusual to find that denatured HU has a size of ~100 nm. The observation can only be explained if denatured HU consists of chains in multimers (e.g., octamers) that remain associated despite substantial denaturation. While we have found that the two monomers of HU do remain associated even in 4 M urea in which the chains are substantially unfolded, we have no explanation for how this can happen even in 7.5 M urea. Perhaps, as the referee correctly suggests, the weigh-fraction data would shed light on this.

3.2.2. Partial unfolding precedes subunit dissociation

After doing the glutaraldehyde cross-linking, which suggests that dimers and higher forms survive even in the presence of substantial concentrations of urea, we decided to find-out whether the dissociation of chains in HU is preceded by partial unfolding of subunits, or whether the subunits dissociate directly into folded monomers without undergoing any unfolding prior to dissociation.In gel filtration experiments, partially unfolded dimers (i.e., dimers with an unfolded

α-helical core), and fully-folded monomers would be expected to elute at different elution volumes. To examine this, protein samples were incubated overnight at different molar concentration of urea and chomatographed on a Superdex-75 column (GE Healthcare). In the absence of any denaturants, HU elutes at an elution volume of 12 ml on superdex-75 column as seen in Figure 39 panel A. The partial unfolding of dimmers could be expected to increase the overall hydrodynamic volume of the protein(s), leading to a shift of the elution peak toward lower elution volume(s).

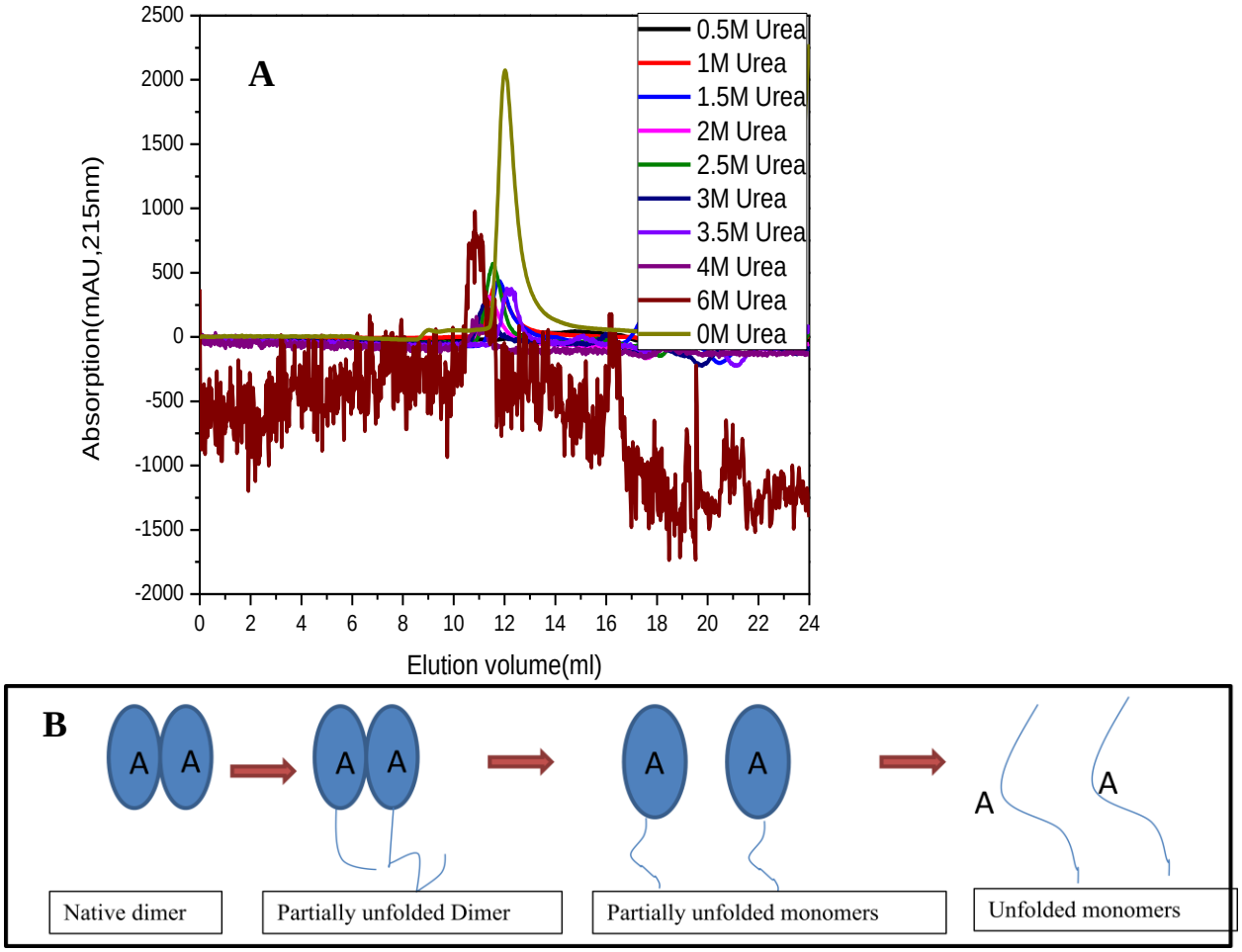


Figure 40: PanelA: Gel filtration chromatograms of HU-A incubated overnight at different molar concentrations of urea obtained after running on Superdex-75 10/300 GL column (GE healthcare). ***Panel B*:** Schematic showing the behaviour of HU dimer dissociation after analysing the chromatographs of HU-A elution profiles in panel A in the same figure.

There was gradual decrease in the elution volume till 2.5M urea concentration. On further increase in urea concentration, there was a gradual increase in the elution volume suggesting dissociation of partially unfolded dimers into partially unfolded monomers. These partially unfolded monomers finally unfolded into the completely unfolded monomers. The schematic of full unfolding and dissociation is shown in Figure 40 panel B. Higher (multi-molar) concentrations of urea used for protein unfolding are associated with higher absorption of ultraviolet light of 215 nm wavelength and, therefore, to noise in chromatograms when the detector's saturation limit is approached. The data is noisy for the highest concentration of urea used (6 M) in most of the baseline region, but the elution is clearly seen without noise. The chromatograms for all the other urea concentrations are not noisy. The evidence of glutaraldehyde cross-linking after incubating HU in different molar concentrations of HU suggests the presence of a strongdimeric interface in both homomeric HU-A and HU-B dimeric species, and this leads us to question whether subunit exchange to form heterodimers can occur. Notably, subunit exchange is hinted-at by other groups that describe HU-AB heterodimers. To prove whether dimer dissociation occurs with/without unfolding and dissociation of subunits involving the HU dimer interface, it was necessary to perform the size exclusion chromatography experiment with urea incubated HU-A and HU-B homodimers.

3.2.3. On-column unfolding to check dimer dissociation

The dissociation of HU-A was checked by immobilizing HU-A through N-terminal 6X-Histidine tag on Ni-NTA agarose beads and applying a gradient of Urea on Akta Purifier 10 instrument (GE). The assumption before performing this experiment was that some of the HU dimers will bind with Ni-ions through single histidine tag (as both the monomers have6xHis histidine tags)which after dissociation will be detected through UV-absorption. The HU used in this experiment was an HU-A double mutant, containing tryptophan at position 47 and 79 (F47W, F79W) in place of phenylalanine (courtesy Dr.Kanika Arora), hence making it detectable at280nm. The protein after immobilization on Ni-NTA beads was dissociated by applying a gradient of urea (100% is equivalent to 8M urea) as shown in Figure 40.

HU-A F47W, F79W mutant (Mut 1,3) unfolding

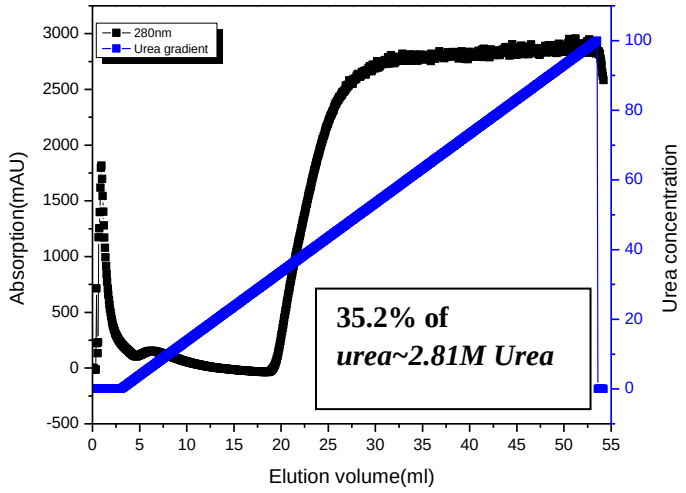


***Figure 41*:** On-column dissociation of HU-A in presence of a urea gradient. The blue line represents the gradient of Urea which start at 4 ml . The conc. of Urea used for this experiment was 8 M which is achieved at 100% gradient after reaching 55 ml elution volume. The figure is also showing the concentration of Urea required for HU dimer dissociation , which is ~2,81M.

After immobilization, the column was run on 0% B (only on PBS buffer) to remove unbound protein from the column which is evident by the presence of UV-280 nm absorption peak before 5 ml. The absorption was zero till 20ml which corresponds to 35.2% of urea concentration marking this point as the minimum concentration of urea required for HU dissociation. The slope kept on increasing till the end of the flow due to dissociation of more and more monomers from the dimers at higher urea concentration. The concentration of urea obtained from this experiment was found in conjunction with HU-A dissociation observed with glutaraldehyde cross-linking. The assumption for these experiments was that some of the HU dimers would bind with Ni-NTA through single Histidine tag and come into the elution after dissociating form Ni-NTA bound HU subunit.

3.2.4. HU-B urea denaturation

The trend of unfolding and dissociation of HU-B was similar to HU-A except the dissociation of HU-B dimers occurred at higher molar concentration of urea (~3.5 M) as compared to ~2M in case of HU-A. The results were found in accordance with glutaraldehyde cross-linking gel of HU-B shown in the Figure 42, where the dimeric population ceased to appear after ~4M concentration of urea.

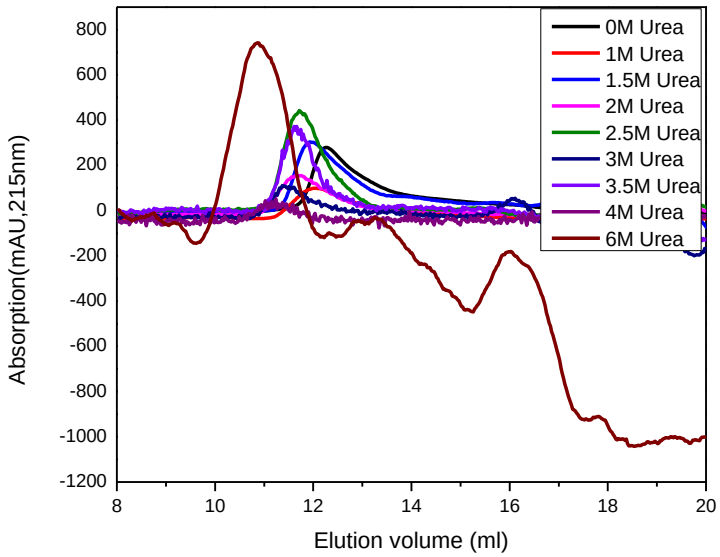


Figure 42: Size exclusion chromatogram of HU-B in the presence of different molar concentrations of urea.

3.3.5. Heterodimer formation: 6x-Histidine tagged HU-B is affinity purified with Tag-RFP-HU-A lacking 6xHistag.

After failed attempts toreplicateresults already reported in the literature (of heterodimer formation by subunit exchange *in-vitro*), and due to our findings regarding the stability of the HU-HU dimeric interface (especially in regard of HU-B), we decided to examine whether heterodimer formation takes place *in-vivo.* We did this by producing a fluorescent protein-tagged HU (Tag-RFP-HUA) in the same bacterial cells in which an affinity-tagged HU-B (6xHis-HU-B) protein was being produced through co-transformation of two plasmids encoding the two subunits. Specifically, the co-transformation into *E. coli* BL21(DE3) pLyS* was made by transforming the *pET23-a* vector containing Tag-RFP-HU-A, and pET-28a vector containing Venus-HU-B and then selecting for transformantsthroughtriple antibiotic selection (for the two vectors, and for the strain). The co-transformant was made by transforming one plasmid at a time and then making the cells competent for acceptance of a second plasmid. The tag-RFP-HU-A transformant lacked any affinity tag (i.e., it was untagged) and the probability of our obtaining any fluorescence from Tag-RFP-HU-A in the elution after affinity chromatography involving 6xHis-HU-B would necessarily involve co-purification of Tag-RFP-HU-A, possible only through heterodimer formation between HU-A and HU-B. The schematic for the construction of co-transformant is shown in Figure42.Protein expression was induced by IPTG induction of positive clones for both protein constructs at an O.D. at 600 nm of 0.6. The supernatant after sonication and centrifugation was loaded onto a Ni-NTA column. Protein was purifiedusing the same method that was used for wild type HU. The eluted HU was expected to have either homodimers of HU-B (HU-B2) and/orheterodimers(HU-AB), since homodimers of HU-A would lack the 6xHis affinity tag and would be unlikely to be co-purified, unless they were able to interact with HU-B dimers. In other words, the only possibilities of HU-A appearing in the elution fraction would be either (a) because of the HU-A polypeptide's association with the His-tagged HU-B polypeptide, as a partner in dimerization, or (b) because of the HU-A dimer's association with the HU-B dimer, as a partner in higher-order multimerization. The purified protein was checked-for by monitoring the presence of an emission from the RFP fluorophore attached to HU-A, in populations of HU-B purified through Ni-NTA's affinity for its 6xHis tag.

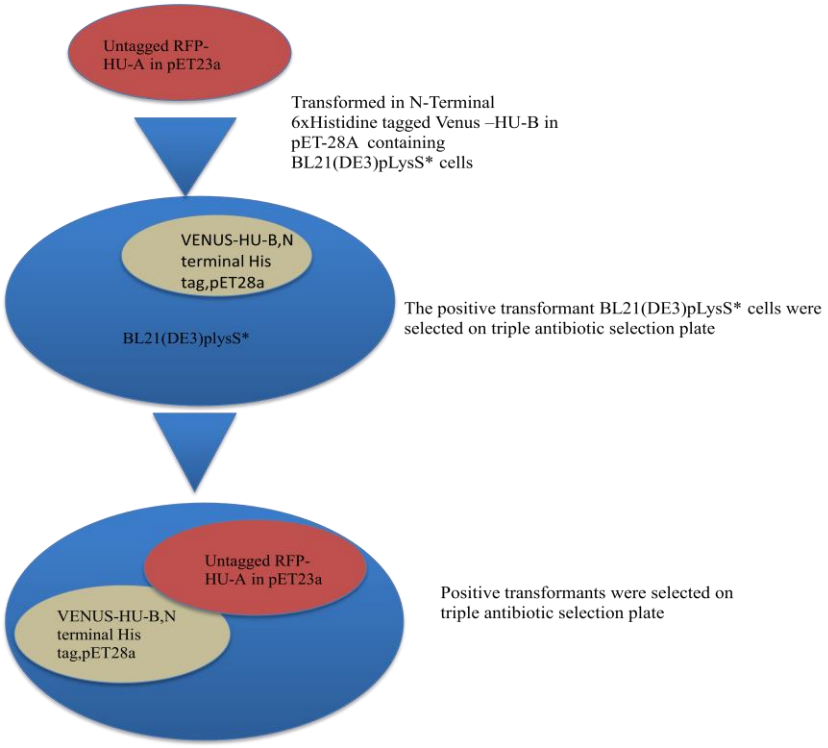


Figure 43: Showing the schematic for the construction of co-transformant *BL21(DE3)pLyS** cells. The red oval represents the plasmid pET-23a containing untagged RFP-HU-A to be transformed in to *BL21(DE3)pLyS** cells containing Venus HU-B in pET-28a (N–terminal 6xHistidine tagged) . The clones were selected on triple antibiotic selection plate followed by purification of Untagged RFP-HU-A along with N-terminal 6xHistidine tagged Venus-HU-B via Ni-NTA affinity purification.

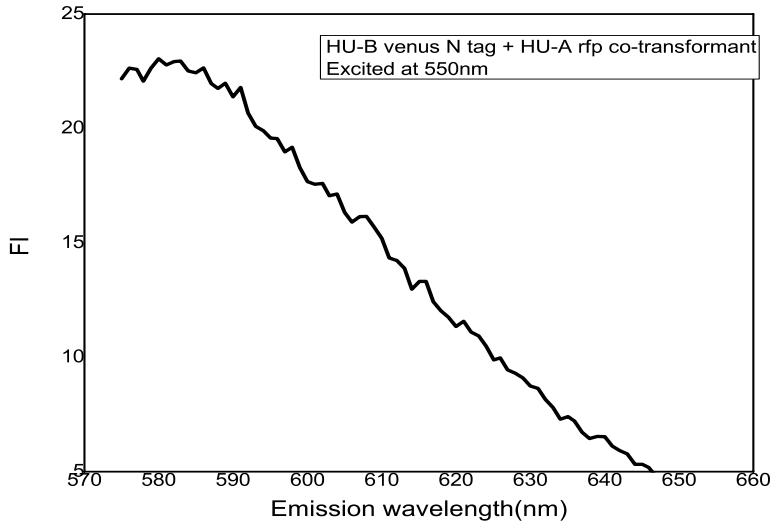


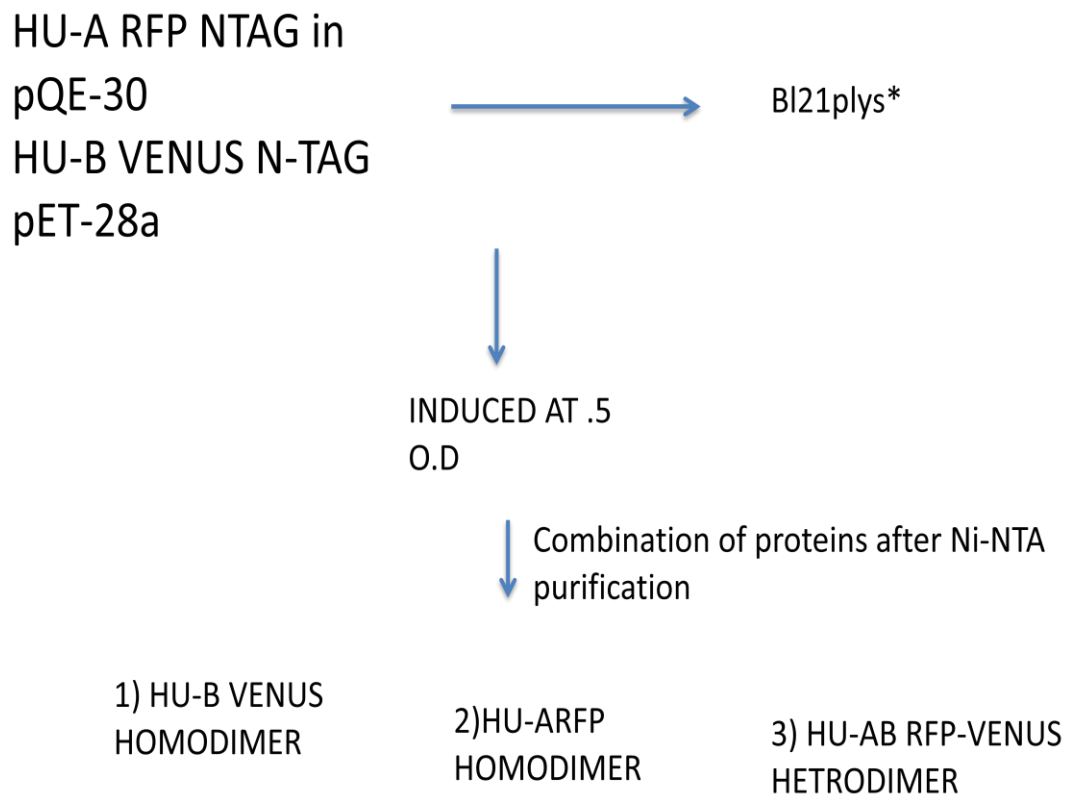
Figure 44: Showing the fluorescence emission of tag-RFP-HU-A in elution fraction after passing the double transformant lysate through Ni-NTA affinity purification.

The emission was checked by exciting Tag-RFP-HU-A at 550nm which is the excitation maximum of RFP (with excitation and emission slitsset at 5nm). The presence of fluorescence indicated the presence of HU-A in association with HU-B,as shown in the Figure 44.However, the experiment did not allow us to distinguish whether heterodimers of HU-A and HU-B had formed along with homodimeric populations of HU-A and HU-B, or whether such homodimers of HU-A and HU-B had undergone association. A caveat that bears mentioning is the issue of compatibility of plasmids. the rationale for the use of the two plasmids is that we wished to create mRNA for both HU-A and HU-B in the same cell, to facilitate formation of both homodimers and heterodimers (given HU's need to form dimers; because of the large hydrophobic surface present on each natively structured monomer that would need to be buried through dimerization). Formation of heterodimers would only be possible if mRNA coding for HU-A and HU-B were to be available in comparable numbers, produced by different plasmids, utilizing different antibiotic resistance markers, to allow for double-selection to maintain both plasmids. We admit, however, that the use of similar replication origins could affect copy numbers of pET-23a (Amp) and pET-28a (Kan). We submit that the experiment was confirmatory, and not quantitatively regulated, and that we used available plasmids presenting

different antibiotic resistance marker options (which could both be maintained inside bacterial cells under suitable antibiotic selection pressure).

3.2.6. 6xHis-tagged Venus-HU-B engages in FRETwith co-purified RFP-HUA

The heterodimer formation was checked by co-tranforming *BL21(DE3)pLyS** cells with pET-23a, 28a plasmids containing 6xHis-tagged Tag-RFP-HU-A and 6xHis-tagged Venus-HU-B. The protein was purified by Ni-NTA affinity chromatography with the same protocol and buffers used for wildtype HU purification. In this case the elute would be the mixture of all three types of dimeric population, i.e., homodimers of HU-A, and HU-B, and the heterodimer HU-AB. The steps followed for the construction of double transformant are shown in Figure 45. The steps were similar to the previous one.



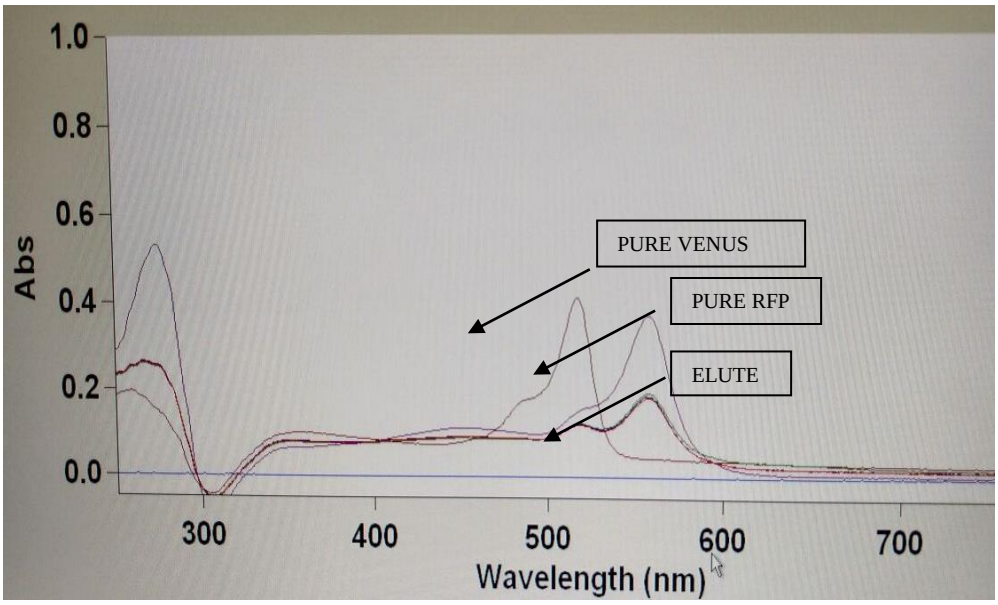


Figure 45: Showing the absorption spectra used for estimation of the concentration of pure tag-RF, pure-venus and the elute containing mixture of all HU dimers, including the RFP and Venus tagged species.

The concentration of pure Venus-HU- and Tag-RFP-HU-A is more than their concentration in eluted fraction after Ni-NTA purification, which is evident by the intensity of absorption peak for each fluorophore as shown in the Figure 45.

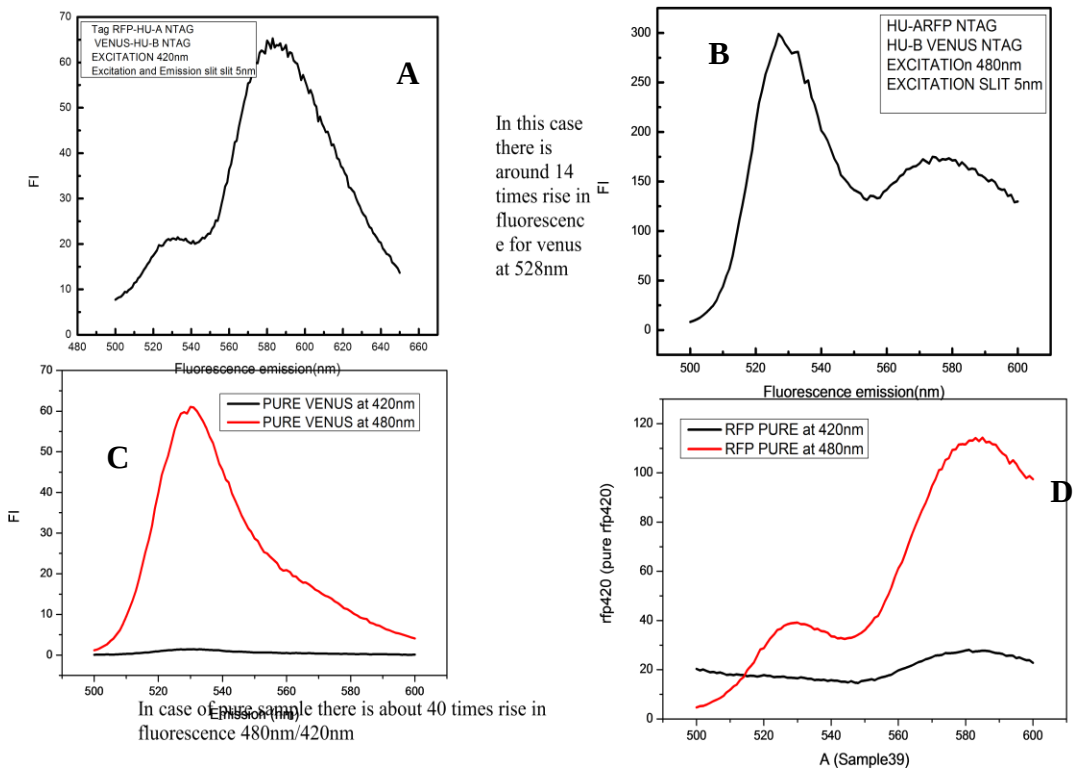


Figure 46: Showing the emission profile of eluted protein fraction at 420 nm (***Panel A***) and 480nm (***Panel B***). ***Panels C and D*** represents emission spectra for pure Venus and tag-RFP proteins.

The FRET between RFP and Venus was employed to study hetero-HU formation (whether hetero assembly of monomers, or hetero assembly of homodimers).The emission maximumfor Venus is 535 nm which overlaps with the excitation spectra of Tag-RFP as shown in Panel C of Figure 46. Along with its 535 nm excitation maximum, Venus also shows some excitation in the 480nm region which is absent for RFP. Exciting pure Venus and RFP pure proteins at 420 nm showed very little fluorescence emission as compared to excitation at 480nm. The comparison of RFP emission for pure and eluted protein showed an enhanced fluorescence emission for eluted protein despite having lowermolar concentration of the RFP fluorophore, as compared to the pure RFP preparation. The increase is explained if there is energy transfer from Venus fluorescence to RFP through FRET. The presence of Venus emission peak explains the presence of Venus-HU-B homodimers in the solution. These results suggest that HU-A and HU-B interact, either at the level of monomers or dimers.

3.2.7. Discussion: HU exists as two isoforms i.e. HU-A and HU-B in *E.coli* cells. Geometrically the genes encoding for HU-A and B are positioned distantly on bacterial chromosomes and are expressed during different time periods, HU-A homodimers appears in bacterial cytoplasm during early log phase followed by the expression of HU-B in early and mid-exponential phases. During the late exponential phase HU-A and B has been shown to form heterodimers which becomes predominant population of HU thereafter.

The formation of the heterodimer has been shown to alter the HU mediated compaction by changing the oligomerization status of both the HU homodimers. The oligomeric status of HU hetrodimer has been reported to be somewhat between HU-A (predominantly dimer) and HU-B (exists as dimer, tetramers, hexamers and octamers) and it seems like the formation of hetrodimers neutralize the dimeric HU-A for the higher compaction of bacterial nucleoid during stationary phase.

The well documented mechanism for heterodimer formation reported in literature is by subunit exchange between individual HU homodimers. The formation of HU heterodimers by subunit exchange can only happen if the dimeric interfaces HU homodimers are not tightly packed against each other. Structurally, HU dimer can be divided into two structural components where the alpha helices forms the body of the dimer and the pair of antiparallel beta strands contributed by both the monomers along with beta hair pin loops form the DNA binding region. The thermal denaturation of the HU has been shown to follow the three state unfolding pathways with the unfolding of alpha helical core preceded by the unfolding of DNA binding region. The glutaraldehyde cross-linking of HU-A and B proved the presence of HU as dimers (dimer interface) upto ~3M urea concentration, making the formation of heterodimer formation by subunit exchange impossible. As, the alpha helical core of HU is less stable compared to the region containing a network of antiparallel beta strands the subunit dissociation without partial unfolding of HU dimers are impossible. The speculation was proved by the gel filtration data for both the HU homodimers under differ molar concentration of urea, where it was observed that the homodimers dissociates via formation of partially unfolded (cross-linkable) dimeric HU intermediate states. Apart from in-vitro studies, the studies done in-vivo by co-transforming *E.coli* cells by plasmids contacting HU-A and HU-B suggest that the hetrodimers are not the only population present in the bacterial cells during stationary phase. The formation of

heterodimers seems by the chance meeting of HU monomers into the bacterial cytoplasm. In the FRET experiment between RFP-HU-A and Venus-HU-B (although they don't not form FRET pairs) the RFP emission ratio was more in the elution containing populations of both HUs suggesting the presence of HU heterodimers along with homodimers.

3.3 *SECTION 3: Creation of a new HU-inspired DNA-binding protein domain*

3.3.1. Creating a new monomeric DNA binding protein by fusing HU derived beta finger loops in tandem from *Escherichiacoli* and *Thermus thermophilus*

The 3D structure of HU from all bacterial species, in spite of sequence dissimilarity, shares common folds, with canonical DNA binding sites made up of beta hair pin loops and anti parallel beta strands. The alpha helical core of HU seems independent of DNA binding canonical sites and shows dual temperature unfolding maxima with core unfolding preceded by the unfolding of canonical DNA binding region.

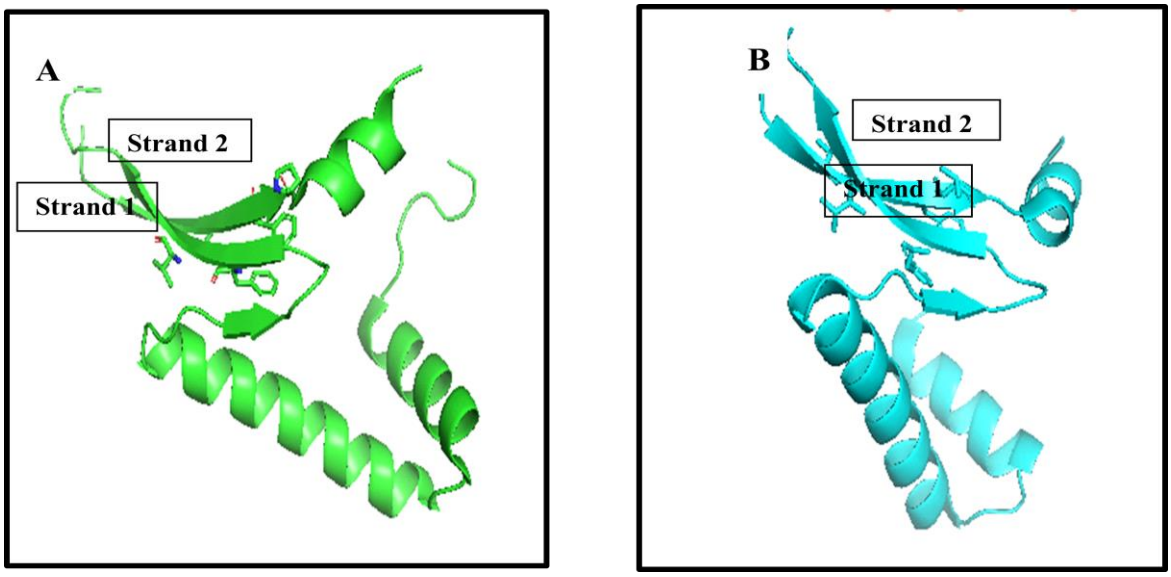


Figure 47: Panel A: Showing the structure of *Thermus thermophilus* HU (PDB ID: 5EKA) having structural homology with *E.coli* HU-A (PDB ID: 1MUL) in ***Panel B***; additionally, an N-terminal extension present exclusively in *T. thermophilus* HU.

The beta hairpin loops are unstructured in the absence of DNA and mobility makes these loops invisible in the crystal structures of HU solved without DNA present. The loops, in absence of DNA,can be called as intrinsically disordered regions (IDR) of HU. The beta hairpin loops are rich in positively charged lysine and arginine residues which, due to electrostatic repulsions, makeinter-loop interactions impossible. The canonical DNA binding region of HU consists of the beta hair pin loops themselves and four anti parallel beta strandsthat bind specifically with

helical DNA (and/or secondary structure containing RNA) by wrapping the beta hair pin loops around the minor grove and positioning its anti-parallel beta strands in major groove as shown in Figure 48.

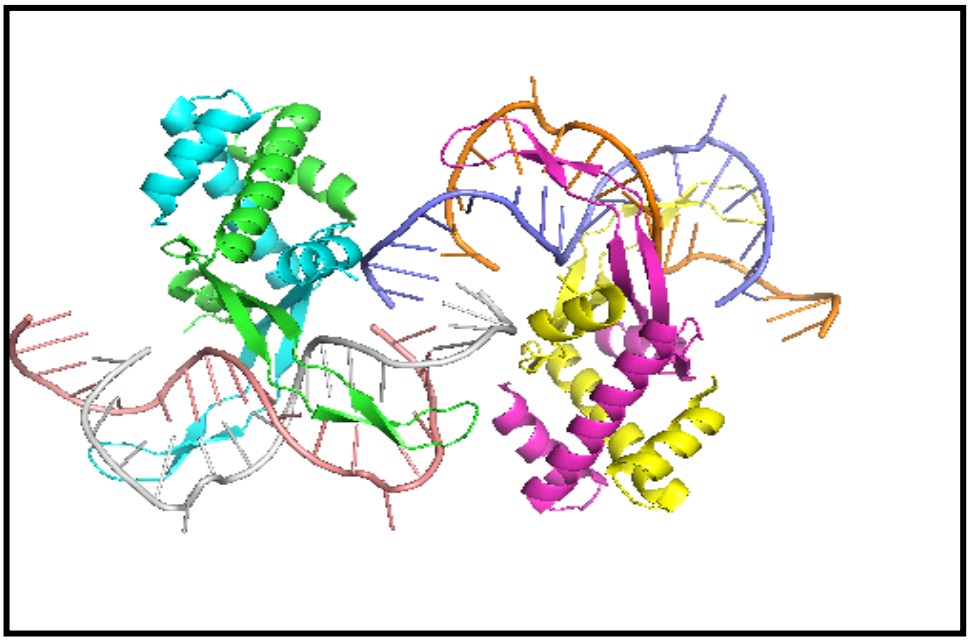


Figure 48: Shows the binding of *Anabaena* HU (PDB ID: **1P51**) with DNA, in which the beta-hairpin loops wrap around the minor goove with the antiparallel beta strands positioned at the major groove.

The core helical region of HU shows minimal contacts with anti parallel beta strands and the hydrophobicity of the amino acids in the anti-parallel beta strands is satisfied by the hydrophobic residues from the antiparallel beta strands from the other subunit.In the literature, it has been reported that the hydrophobic contacts in the helical region are very few as compared to those in the canonical DNA bindingregion. Therefore, we thought of making a fusion protein of the *E.coli* canonical DNA binding region lacking a helical core. To satisfy the hydrophobicity associated with the beta strands of each monomer, we decided to fuse the canonical DNA binding sites of two monomers in tandem.In this case, the hydrophobicity of the non polar residues along the anti-parallel beta strands would be satisfied by the residues from the anti-parallel beta strands of the same polypeptide chain, making the molecule monomeric in nature, although mimicking the DNA binding regions of HU, for the most part. The construction of this

genetic constuct initially failed because of limilations with PCR amplification and fusion of DNA sequences from the same gene (presumably owing to homologous recombination). Then we proceeded with the fusion of DNA binding canonical sites from *E.coli*K12 and *T.thermophilus* HU because these bacterial strains were available in our lab. The beta hairpin loops of *T.thermophilus* HU is longer and contains more positively charged amino acids than the corresponding regions of *E.coli* HU, with largely the same types of amino acids present in anti-parallel beta strands. The sequence and the region from where the two proteins were fused are shown in Figure 48. Before proceeding with chimera construction, the crystal structures of both the proteins were superimposed to calculate RMSD between them. Both the proteins were nicely superimposed with the RMSD value of 1.07 Å.

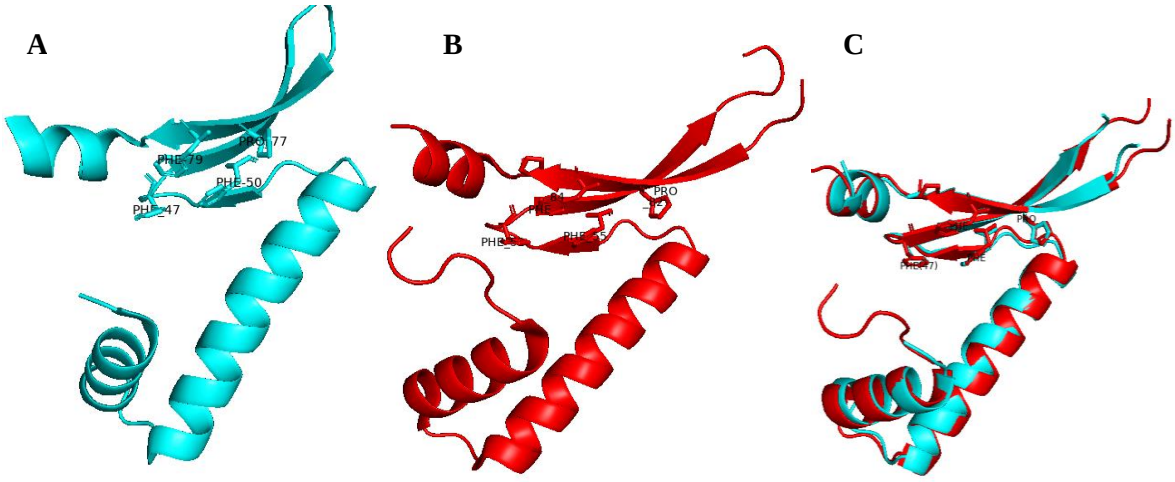


Figure 49: Panel A: Ribbon structure for HU-A (PDB ID: 1mul) of *E.coli*. ***Panel B:*** Thermus Thermophilus (PDB ID: 5eka) HU, figure 2A,B. Superimposed HUs shown in the ***Panel C***. The magnified superimposed canonical site showing the perfect alignment of conserved hydrophobic amino acid residues figure 2C.

3.3.2. Structure prediction of the (final tandem-HU-DßF) fusion DNA binding protein

Robetta software which predicts structure on energy minimization of each amino acid was used to predict the structure of the fusion DNA binding protein. The modelled protein structure was docked with 20 mer DNA using NP docking software (nucleotide protein docking). The docking image shows a DNA protein assembly similar to wild type HU, where beta hairpin loops are wrapped around the DNA backbone from two sides and stabilizing anti-parallel beta strands with

the major grove of DNA. The binding arrangement where both the arms of protein are stabilized by DNA backbone supports its preferences for helical DNA structure.

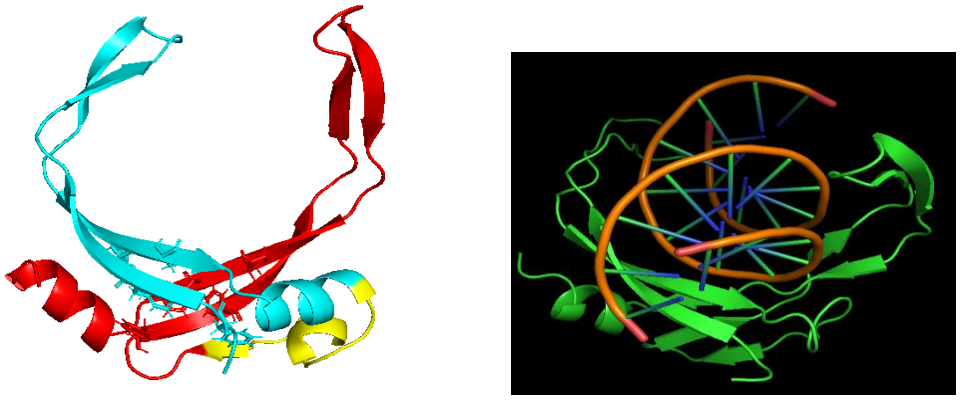


Figure 50: Panel A: Showing the modelled structure of tandem-HU-DßF(DNA binding fingers) using ab-initio protein structure prediction software Robetta.. ***Panel B:*** Protein docked with 20 mer DNA.

3.3.3. Purification of the fusion DNA binding protein (tandem-HU-DßF)

For the cloning of DNA binding region of two HU the strategies of splicing by overlap extension, followed by restriction digestion and ligation were used. The region beyond the red mark of both the HU was used for making fused canonical DNA binding site. The fused gene using SOE PCR was cloned in pET-23a between *Nde-1/Xho-1* restriction sites containing C-terminal Histidine tag.The plasmid was transformed in Rosetta(DE3) cells whose specifications are provided in the general material and methods chapter. The protein was expressed after induction with 1mM immidazole and was purified by Ni-NTA affinity purification. The DNA binding region of *E.coli* HU was kept at the N-terminal followed by the DNA binding region of *T. thermophilus* HU at C-terminal.

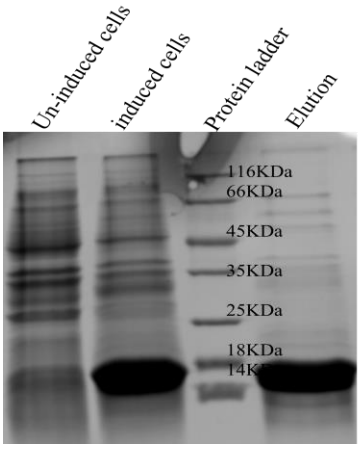
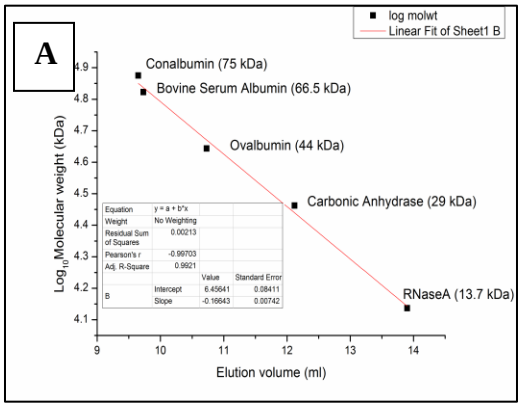


Figure 51: Showing the SDS-PAGE image of the Ni-NTA affinity purified protein with most of the protein coming in the elution fraction.

There was significantly highprotein expression of the HU-derived double beta hairpin loop construct after IPTG induction as shown in Figure 51. The cells were induced at $O.D._{600}$ of 0.6. The protein was found showing anomalous mobility with the purified protein band diplaying poorer mobility (i.e., higher apparent molecular weight) than the calculated size of ~12kDa.

3.3.4. Size exclusion chromatography studies

The oligomeric nature of the protein was checked by size exclusion chromatography by loading protein in Superdex-75 10/300 GL (GE Healthcare) column on Akta purifier instrument. The absorption was measured at 215nm as the fusion protein lacks any tryptophan.



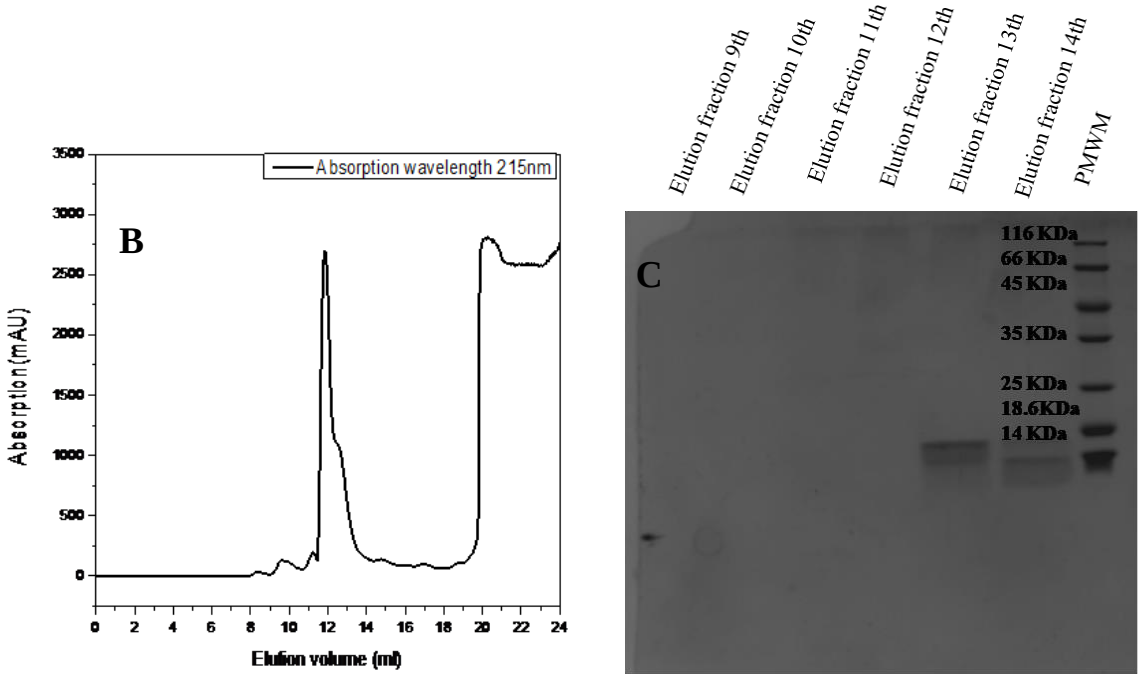


Figure 52: Panel A: Calibration curve of the Superdex-75 10/300GL column***. Panel B:*** Showing the size exclusion chromatogram of protein monitored at 215 nm absorbance. ***Panel C:*** The SDS-PAGE image of the eluted fractions with the presence of protein in 13th and 14th fraction.

From the gel filtration chromatogram, it is clear that most of the protein exists as a homogeneous population of molecules eluting at 12ml elution volume, corresponding to a size of ~13 kDa. A small fraction of molecules exists in other conformation whose presence is indicated by an existence of a hump on the major elution peak. As the fusion protein is made up of two long unstructured beta hair pin loops which can attain any shape in the solution, it is difficult to calculate molecular mass of the protein by size exclusion chromatography. It is clear from the gel filtration chromatogram that majority of the molecules of the protein attains conformation without any evident attendant aggregation. with the hydrodynamic analysis suggestive of monomericity if one considers all the data together, i.e., gel filtration, dynamic light scattering, and gluraraldehyde crosslinking.

3.3.5. Dynamic light scattering studies

The size of the protein molecules and population size distribution was estimated by analysis on the Wyatt QELS DLS instrument. The monoexponential decay of the correlation function is indicative of the homogeneity of molecular conformations in protein solution as shown in the Panel A of the Figure 53. The correlation function represents distribution of scattering intensity with time interval and is used to calculate diffusion constant which finally gives rise to hydrodynamic radius of the molecule.

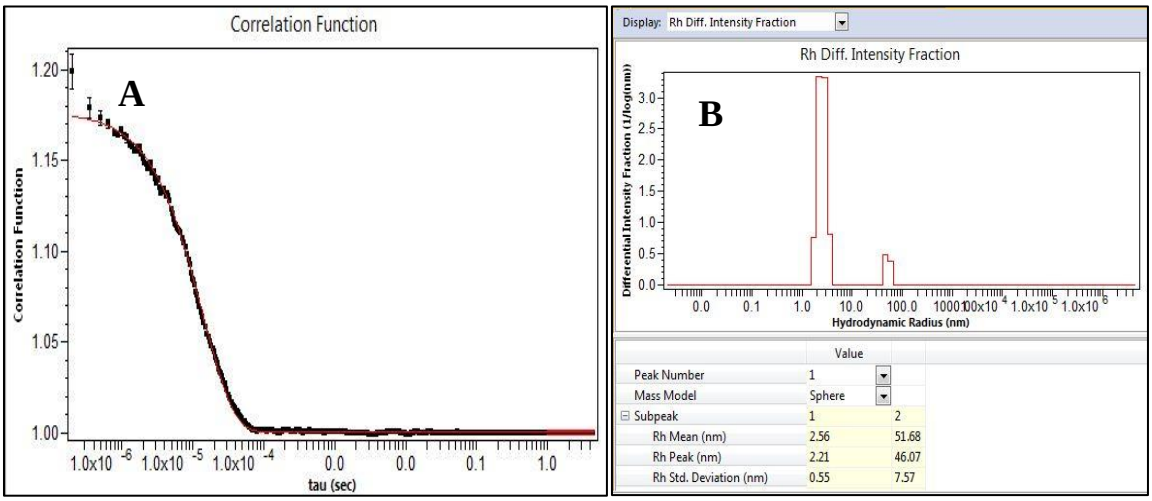


Figure 53: Showing the correlation function distribution of fusion DNA binding protein along with the size distribution of the molecules in term of weight fraction

The molecule size is plotted in term of weight fraction with an average size of 2.56nm radius for majority of the population as shown in the panel B of the Figure 53.Very minorfractions of the molecules exist as aggregates or higher-sized oligomers.

3.3.6. Glutaraldehyde cross-linking of the fusion

The protein fractions after size exclusion chromatography were cross-linked with different concentration of glutaraldehyde to check their oligomeric status. The protein in 14th elution fraction was even incubated with LPS (Lipopolysaccharide) and crosslinked with glutaraldehyde.

The appearance of big aggregates on the interface of stacking and separation gel is indicative of protein LPS crosslinked aggregates which are unable to enter inside gel suggesting association between these molecules.

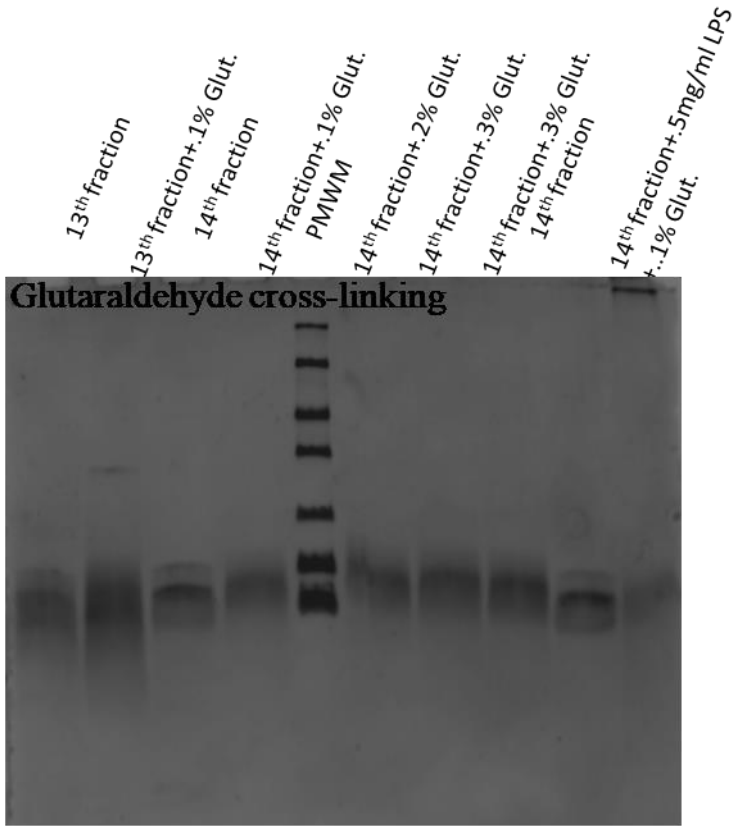


Figure 54: Showing the glutaraldehyde cross-linked SDS-PAGE gel of elution fraction along with cross-linking with LPS shown in the last lane of SDS-PAGE. As LPS crosslinked proteins are forming big aggregates which are unable to enter into the gel, suggesting their binding with LPS like wild type and Loop replaced HUs.

The protein exists as single population even at higher concentration of glutaraldehyde. The two diffused bands in each lane can arise due to proteolysis of the beta hair pin loops on protein molecules which are associated through hydrophobic interaction among antiparallel beta strands. One possibility is that dimerization occurs upon DNA. Cross-linking was not tried on DNA. The two monomers that dimerize to create functional HU have many additional regions that allow the HU dimer to form tetramers and octamers. However, in the new molecule, these regions have been removed. Therefore, what is left is just the regions that can wrap around DNA and these (the two loops) are genetically linked into the same polypeptide chain. We cannot think of any

means by which this construct would form higher-order associations (except by aggregation, which was not observed).

3.3.7. Circular dichroism spectroscopy and protein thermodynamic stability

The secondary structure of the protein was calculated by monitoringdifferential absorption of right and left circularly polarized light on a Chirascan (Applied Photophysics) CD instrument. The CD spectrum shows maximum differential absorption for random coil followed by alpha helix and beta sheets. The structure of the protein as predicted by Robetta is dominated by the random coil which results in maximum CD signal in the random coil region. The absorption due to a unit peptide bond in a beta strand is approximately 4-8 times lower than that of a peptide bond in an alpha helix or a randomly coiled region.

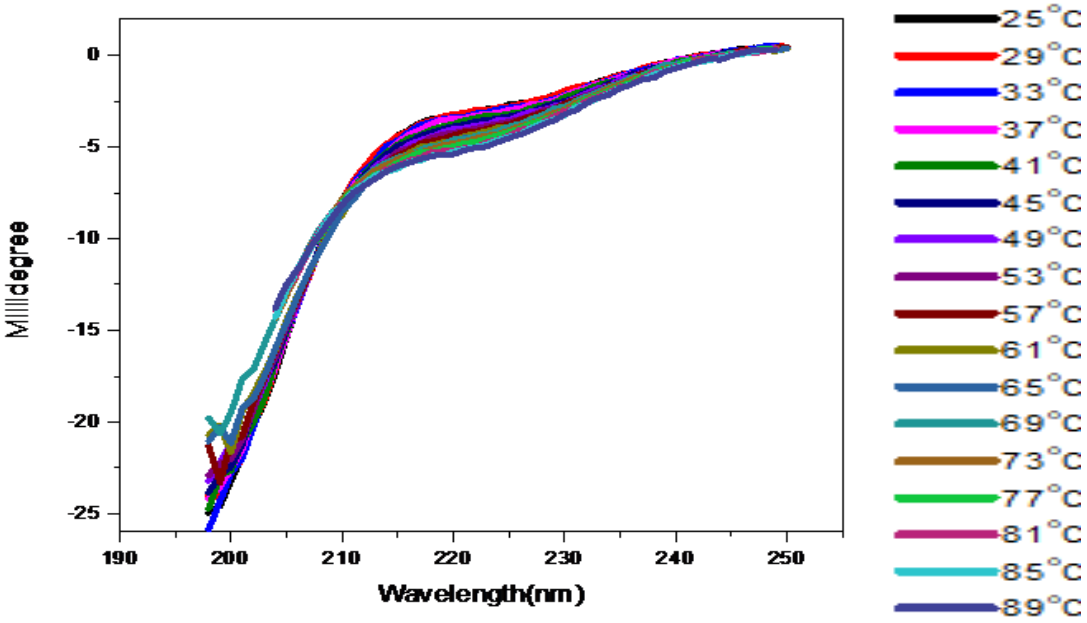


Figure 55: CD spectra showing the variation in secondary structure of fusion protein with over a wide temperature range. The data was plotted in terms of millidegre (raw ellipticity) as the concentration of protein was unknown.

The spectra show a dip for random coil at 195nm along with some absorption in 222 nm and 216nm regions which is indicative of alpha helix and beta strands. The spectra were collected over a temperature range of 25ºC to 90ºC with a temperature interval of 3ºC. The spectral shape remains uniform even at high temperature, establishing high thermodynamic stability of the protein as shown in the Figure 55. In summary, the circular dichroism (CD)spectrum indicates alpha helical, beta strand and random coil content in the fusion construct. The predicted structure for the model obtained using the Robetta software contains ~50 % random coil content, ~35 % beta sheet content and 15 % alpha helical content for the construct. The CD spectrum has been collected between 250 and 198 nm and suggests maximum contribution fromrandom coils to the predicted structure with other secondary structures also present (the highest negative mean residue ellipticity (MRE) is at 198 nm, and there is a negative band feature in the216-218 nm range, indicative of beta sheet content, as well as a negative band feature at 222 nm ,indicative of some alpha helical content.

3.3.8. Kinetic stability of the fusion protein

The kinetic stability of the fusion double beta finger protein was checked at 90ºC by keeping protein at same temperature for 8 minutes. The spectra were collected after 2 minutes time interval as shown in the Figure56.

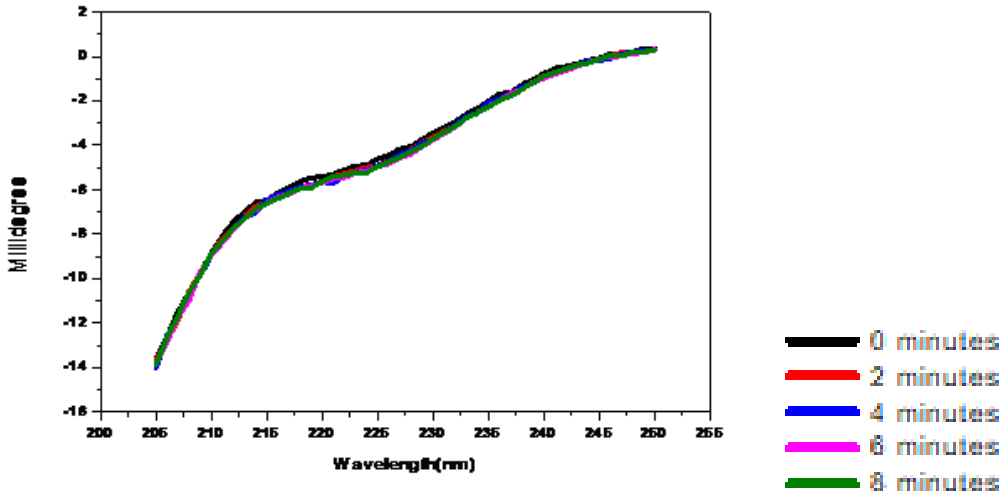


Figure 56: Showing the kinetic stability of the protein by monitoring protein heated at 90ºC for 8 minutes.

The protein possesses high kinetic thermal stability which is visible from completely unaffected spectra in spite of heating.

3.3.9. Electrophoretic mobility shift assay (EMSA)

The substrate specificity of the fusion protein towards double and single stranded DNA was checked by EMSA on 2% agarose gel. In EMSA, the binding of the protein with DNA will affects its mobility through agarose gel by combinatorial affect of DNA backbone neutralization and formation of cross-linked DNA molecules of bigger sizes by protein molecules. The smearing of DNA bands is indicative of association between these molecules as shown in the first four lanes of the agarose gel in Figure57. It showed no binding of the protein with single stranded 20 mer DNA in last four lanes of agarose gel (present after DNA ladder). There was no change in the control 20 mer double and well as single stranded DNA with bands appearing sharply at their size positions. The concentration of double stranded and single stranded DNA was 10 and 20 µM respectively.

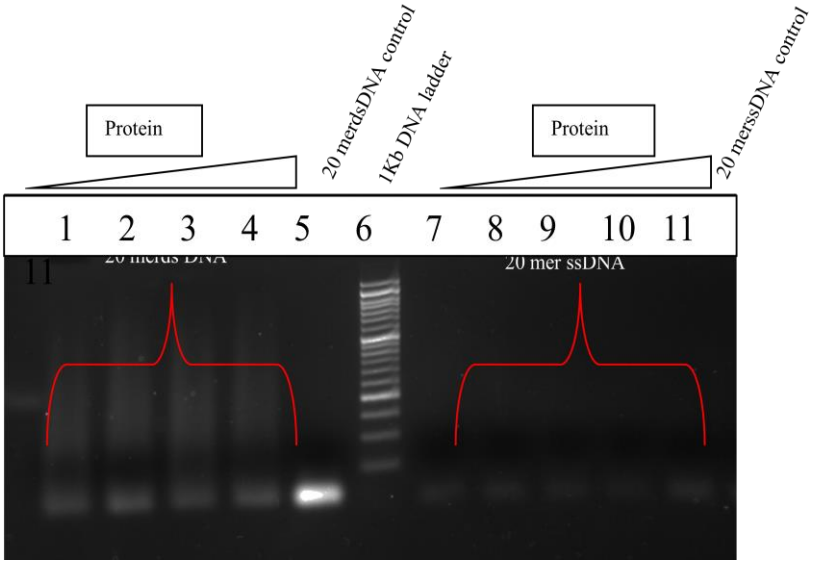


Figure 57: Showing the EMSA agarose gel of the fusion DNA binding protein incubated with double and single stranded DNA. The presence of smearing in the first 5 lanes are an indicative of association between protein and dsDNA .

The formation of the canonical DNA binding site in fusion DNA binding proteins, like in wild type HU protein, has made it specific for double stranded DNA. The binding of fusion DNA

binding protein with double stranded DNA proves it's folding like wild type HU protein. HU's binding to DNA seems to occur in multiple modes with different levels of loading due to binding of HU to HU, and also binding of HU to DNA, and also bending of DNA by HU. Therefore, smearing can arise from HU bound to DNA and HU bound to HU.

3.3.10. Bio layer interferometry studies for protein DNA interaction

The binding of fusion DNA binding protein with 20 mer double stranded and single stranded DNA was checked by biolayer interferometry studies. In this method, the protein was immobilized on Ni-NTA biosensor using 6X-Histidine tag which was further used for interaction studies with DNA. The formation of DNA layer over immobilized protein will change the interference pattern of incident light which will be reflected as interaction peak in sensogram. The interaction of molecules is reflected as sensogram peak.

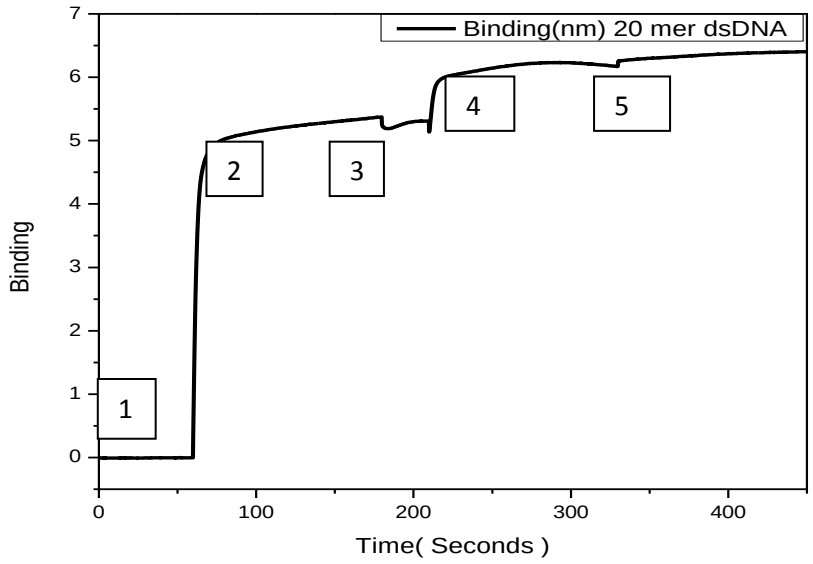


***Figure 58**:* Showing the sensogram of fusion DNA binding protein with 20 mer double stranded DNA. The steps followed during data collection are labelled as 1 to 5.

Figure 58 shows a sensogram of protein binding with 20 mer double stranded DNA. The step 1 in the sensogram represents the initial baseline taken by dipping Ni-NTA probe in reaction buffer. The second step represents the binding of the protein to the Ni-NTA biosensor tip. This

step is known as association. Third step represents the baseline of the protein bound Ni-NTA biosensor tip by dipping in the reaction buffer. Fourth step resents binding of DNA with protein immobilized on NI-NTA biosensor and the step is known as association. The fifth step is dissociation where the biosensor tip is dipped in reaction buffer to estimate the binding affinity of protein towards DNA.

In Figure 58, it is evident from the fourth step that fusion DNA binding protein interacts with double stranded 20 mer DNA with no dissociation even after dipping probe in buffer in the fifth step of sensogram suggesting strong association between protein and DNA. This interaction is stronger than the interaction between HU-LPS as there was dissociation of attached LPS from HU in buffer.

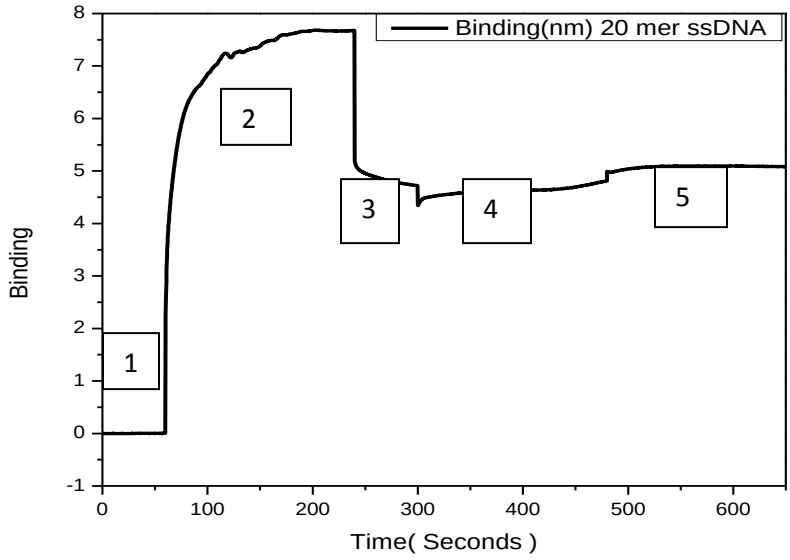


Figure 59: Showing the sensogram of fusion DNA binding protein with 20 mer single-stranded DNA. The steps followed during data collection are labelled as 1 to 5, similarly to Figure 58.

Figure 59 shows no association between fusion DNA binding protein and 20 mer single-stranded DNA. The signal intensity in the second Ni-NTA protein association step went beyond 5 response units with the same concentration of the protein used for binding studies with 20 mer double stranded DNA. The response units came back to 5 in third step during baseline collection.

The flat line in 4 step indicates no association between protein and single stranded DNA. This suggests that the protein is a potential candidate for making a fusion with DNA polymerase. The DNA binding should be specific toward double stranded DNA to prevent dilution of polymerase by binding with single and double stranded DNA.

3.3.11. Discussion: The three state or two transitions in the unfolding of HU homodimers (the homodimers remains associated with each other even after unfolding of the alpha helical core) paved way for the creation of a novel DNA binding protein. After satisfying hydrophobicity among antiparallel beta-strands (coming from the same polypeptide chain in fusion) and mimicking natural hydrophobic interactions in wild type HU (where hydrophobic interactions are satisfied through interaction among hydrophobic residues in antiparallel beta strands coming from other subunit), the molecule formed was monomeric in nature. The resultant protein was largely intrinsically disordered and used unstructured beta hairpin loops for interactions with DNA. The molecule showed specific binding towards double stranded DNA like wild type HU where the binding through beta hairpins loops are specific towards dsDNA and RNA forming secondary structures.

References

1. Dengler, V., et al., *An Electrostatic Net Model for the Role of Extracellular DNA in Biofilm Formation by Staphylococcus aureus.* J Bacteriol, 2015. **197**(24): p. 3779-87.
2. Ali Azam, T., et al., *Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid.* J Bacteriol, 1999. **181**(20): p. 6361-70.
3. Balandina, A., D. Kamashev, and J. Rouviere-Yaniv, *The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids.* J Biol Chem, 2002. **277**(31): p. 27622-8.
4. Hammel, M., et al., *HU multimerization shift controls nucleoid compaction.* Sci Adv, 2016. **2**(7): p. e1600650.
5. Das, D., S. Sidiq, and S.K. Pal, *A simple quantitative method to study protein-lipopolysaccharide interactions by using liquid crystals.* Chemphyschem, 2015. **16**(4): p. 753-60.