FOLDING, STABILITY AND DNA BINDING OF *E. COLI* HU AND ITS ENGINEERED VARIANTS

KANIKA ARORA

A Thesis submitted for the partial fulfillment

of the degree of

Doctor of Philosophy

Indian Institute of Science Education and Research Mohali

Department of Biological Sciences

Mohali-140306

November, 2015

CERTIFICATE

The work presented in this thesis has been carried out by me under the supervision of Professor Purnananda Guptasarma at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or 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.

Date: November 6, 2015

Place: Mohali **(Kanika Arora)**

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

(Prof. Purnananda Guptasarma)

Supervisor

Dedicated To My Family

ACKNOWLEDGEMENTS

"Knowledge is in the end based on acknowledgement."- Ludwig Wittgenstein

This thesis arose out of the five years of my Ph.D research conducted in Prof. Purnananda Guptasarma's lab. Many people have contributed to the completion of this thesis, directly or indirectly. It is a pleasure to thank all the people who have helped me in this journey.

I owe my heartfelt thanks to my mentor, Professor Purnananda Guptasarma for providing me his consistent guidance, supervision and motivation throughout the duration of this work . His scientific approach, out of the box novel ideas, logical reasoning and articulation have been an inspiration. He has been a great teacher, not only in field of science, but life also. Every discussion with him whether scientific or otherwise was a pleasurable learning experience and has helped me grow as a researcher and as a person. He has been a great support during the tough times and taught me how to trouble-shoot with both professional as well as personal problems. Sir, during this period, I have learnt a lot from you that I will remember throughout my life. I owe you much more than you will ever know.

I am thankful to the Director IISER Mohali, Professor N. Satyamurthy for providing us the excellent research facilities to work in the institute.

I wish to warmly thank my doctoral committee members Dr. Rachna Chabba and Dr. Kausik Chattopadyay for their valuable comments and suggestions.

I extend my heartfelt thanks to all my seniors Dr. Neeraj and Dr. Uzma who provided me the initial training and taught me how to function in the lab. I have learnt a lot from both of them. I would also like to thank all the other present and past lab members Dr. Satyaprakash, Dr. Arpana, Dr. Javed, Dr Prerna, , Sukhdeep Sir, Nitin, Pallavi, Prince, Bhisham for their kind support and always keeping conducive lab environment. One learns more when one teaches it to others. It has been a great experience working with and training the fifth year MS students of the lab Kiran and Abhishek.

I am thankful to my friends Dr. Anil, Dr. Shailesh, Aman, Zulkifli, Avinash, Shruti, Shambhu, Bhupinder and Mayank for being wonderful company and extending their help whenever required.

There are many people outside IISER, who have been instrumental in making me what I am and it is a pleasure to thank them for all the 'big' and 'small' things, they have done for me. Most importantly, I remain ever grateful to my family; my parents, my sister Heena, my brother Puru and my husband Dominic who have

always been a source of love, support and strength in my life. I would also like to thank my friends Avantika, Harsha and Shweta for their friendship and support over the years. I thank all the teachers of my school days, graduation and post-graduation for inspiring and encouraging me.

My sincere appreciation goes to all IISER faculty members, library and instrumentation staff, whose help and support was always there when needed.

I sincerely thank all of them whom I have unintentionally left out; who at some point of time, in some way or the other, have contributed in this journey.

The financial support by UGCand CPSDE is duly acknowledged

Kanika Arora

List of Publications

1. Dhaunta N**, Arora K**, Chandrayan SK, Guptasarma P. Introduction of a thermophile-sourced ion pair network in the fourth beta/alpha unit of a psychophile-derived triosephosphate isomerase from Methanococcoides burtonii significantly increases its kinetic thermal stability. **Biochim Biophys Acta**. 2013 Jun;1834(6):1023-33.

2. Kumar S, Sharma P, **Arora K**, Raje M, Guptasarma P. Calcium binding to Beta-2-microglobulin at physiological pH drives the occurrence of conformational changes which cause the protein to precipitate into amorphous forms that subsequently transform into amyloid aggregates. **PLoS One**. 2014 Apr 22;9(4):e95725.

3. **Arora K,** Mangale S, Guptasarma P. Single cell-level detection and quantitation of leaky protein expression from any strongly-regulated bacterial system. **Anal Biochem**. 2015 Jun: 484: 180-182.

4. **Arora K,** Guptasarma P. High affinity DNA-protein interactions and the difficulty of removing DNA from *Escherichia coli* HU during its purification: Implications and consequences of survival of HU-DNA complexes after bacterial cell lysis (Manuscript submitted).

5. **Arora K,** Thakur B, Kumari A. Guptasarma P. Structure and stability of subunit and subunit interface of *E. coli* HU-A and HU-B homodimers: Implications in subunit swapping (Manuscript under preparation).

6. **Arora K,** Thakur B, Guptasarma P. Studies on folding, association and DNA binding of natural and engineered (simulant) HU-AB heterodimer (Manuscript under preparation).

7. **Arora K,** Guptasarma P. DNA binding of *E. coli* HU: Introduction of an intrinsic fluorescence probe to parameterize binding (Manuscript under preparation).

Presentation and Abstracts

1. Presented a poster entitled "New Insights into DNA binding and dimerization dynamics of *E. coli* histone like protein HU" at ASBMB annual meeting held in conjugation with Experimental Biology-2015, Boston, MA (March 28-April1, 2015)

2. Advanced Microscopy and Imaging technique, by DSS Imagetech and Photometrics in IISER Mohali (May 28-30, 2013).

3. Mini-Symposium on Ubiqutin systems and cellular processes, Department of Biological Sciences, IISER Mohali (Oct17-18, 2013).

4. 82nd Annual Meeting of Society of Biological Chemists, University of Hyderabad, Hyderabad (Dec 2-5, 2013).

Synopsis

DNA compaction influences many critical cellular processes. In eukaryotes, histone protein molecules that wrap DNA around themselves are largely responsible for DNA compaction. In prokaryotes, there are no histones. Instead a number of other proteins exist to perform similar functions of compaction and these are known as "histone-like proteins" or "nucleoid associated proteins (NAPs)". Many different classes of NAPs are expressed during different phases of prokaryotic cell growth and in response to different environmental conditions. Different bacteria contain different set of NAPs, with only HU being ubiquitously present in all bacteria. Some bacteria such as *Escherichia coli* contain two different isoforms of HU, named HU-A and HU-B. These isoforms are differentially expressed during growth of the bacterial culture and they form both homodimers (HU-A2, HU-B2) and the hetero-dimer, HU-AB, with relative concentrations of different dimeric species at different times being somehow responsible for different grades of DNA compaction. Although, HU-A and HU-B share high sequence and structural similarity, they have different binding affinities for different DNA substrates and therefore, different effects on DNA compaction. The role of the hetero-dimer and its mechanism of formation are not well understood. We have tried to understand HU homodimerization, heterodimerization and DNA binding through studies of *E. coli* HU and its engineered variants.

Overall, the focus of this thesis is to (i) understand DNA binding by HU under different conditions, (ii) determine the differential overall stabilities of HU-A & HU-B homodimers and the HU-AB hetero-dimer, and compare these to the stability of dimer interface, (iii) perform studies on various HU mutants and use fluorescently tagged HU to physically highlight DNA in microscopic studies and quantitate leaky expression in otherwise tightly regulated expression systems. The thesis is divided into 7 chapters.

The first chapter consists of an introduction of HU and the questions addressed in the thesis.

The second chapter describes the details of the materials and methods used.

The third chapter entitled *"**Examination of over-expressed HU and attempts to purify DNA-free untagged and affinity tagged HU**"* describes a new protocol for purification of HU under native conditions, without any attendant DNA contamination. Given HU's high DNA-

binding affinity, we discovered that upon being purified under native conditions, HU pulls down DNA as well as other DNA-binding proteins along with itself. We also found that some of these are known DNA-binding proteins while numerous others (including low abundance proteins) are not currently known to be either themselves DNA-binding in nature, or known to bind to other DNA-binding proteins, suggesting that this could be a way to identify new proteins associated either directly, or indirectly, with bacterial DNA. DNA binding by HU was then examined under various different conditions to identify those under which HU releases all bound DNA and any associated proteins, to allow purification. We found that DNA is released from HU either when the structure of the protein is disturbed using chaotropic agents, or when ionic-interactions between DNA-protein complexes are destabilized by the use of very high salt concentrations, with the latter being preferable to the former, since no unfolding-refolding is involved (MS under preparation).

The fourth chapter of the thesis, entitled ***"Characterization of structure and stability of affinity tagged E. coli HU-A and HU-B"*** focuses on the study of the differential stabilities (inter and intra-subunit) of the two HU homodimers. The significant findings of the chapter are: ***a)*** HU (both HU-A and HU-B) is capable of refolding from a completely unfolded state after thermal denaturation, with the refolded protein retaining DNA-binding characteristics; ***b)*** HU-B is more thermally and thermodynamically stable than HU-A, displaying higher Tm and Cm values as well as higher ΔG and ΔH values associated with thermal unfolding; ***c)*** The dimer interface is very stable in both HU-A and HU-B, with the molecule remaining dimeric even in a substantially (but incompletely) unfolded state, suggesting that dissociation of HU monomer subunits follows (rather than precedes) the unfolding of the bulk of each HU subunit in the dimer (MS under preparation).

In the fifth chapter entitled ***"Studies on folding, association and DNA binding of HU-BA fusion"***, we describe the construction of a fusion protein in which the two HU subunits, HU-B and HU-A were genetically fused to each other with HU-B preceding HU-A, with an 11 amino-acids long 'serine-glycine' linker separating the two fusion partners. The fusion construct was made in the hope that it would allow us to physically 'simulate' the HU-AB heterodimer in the form of an "HU-B-HU-A fusion" forcing only the hetero-meric interface to form under conditions disallowing inter-chain interactions that would otherwise allow homo-meric interactions. The construct was made because it is difficult to isolate pure HU-AB heterodimers.

The construct was tested for structural content and DNA-binding ability, and found to be qualitatively like the HU-A and HU-B homodimers. The stability of the hetero-dimer was determined and compared to that of the homo-dimers, HU-A and HU-B. Thermodynamic and chemical stability parameters of the construct suggest that heterodimers form (and that hetero-dimer association is thermodynamically driven) because of the greater stability of HU-B-HU-A, in comparison with HU-A or HU-B (MS under preparation).

In the sixth chapter, entitled ***"DNA binding: Introduction of an intrinsic fluorescence probe into HU",*** we introduce aromatic residues into HU-A and HU-B which naturally lack both tyrosine and tryptophan residues. Intriguingly, the absence of tryptophan is evolutionary conserved across all HU sequences in different bacteria. We created multiple tryptophan-insertion mutants (replacing phenylalanine residues) of both HU-A and HU-B to generate the novel scope for florescence spectroscopic analysis of DNA-binding. Mutants incorporating Trp were characterized and compared to the wild type protein in respect of their structural properties and stability. DNA binding studies were carried out with selected mutants to calculate dissociation constant values for HU-A and HU-B binding to a synthetic 4-way junction (cruciform) DNA. Using these different mutants, residue specific structural changes on DNA binding were also examined by monitoring differential levels of solvation of different tryptophan residues in the presence and absence of DNA. We also tentatively explored the effect of tryptophan insertion upon DNA integrity, in respect of the likelihood of photosensitized oxidative modifications occurring when Trp residues are present in close proximity to DNA over long durations (MS under preparation).

In the seventh chapter, entitled ***"Studies on fluorescently tagged HU",*** through the use of fluorescently tagged HU molecules we designed a system to highlight DNA in microscopic images of bacteria, and to quantitate leaky expression in tightly regulated bacterial expression systems by fluorescence imaging. The fusion proteins (RFP-HU-A or Venus-HU-B) were found to bind to the *E. coli* nucleoid, localizing the fluorescence signal and making it possible to detect molecules expressed in the absence of the inducer, IPTG, thus allowing quantitation of leaky expression both statistically and at the level of single cells (Arora K, Mangale S, Guptasarma P.. Anal Biochem 2015).

Abbreviations

%	Percent
mg	Milligram
M	Molar
ml	Milliliter
min	Minute
nM	Nanomolar
µM	Micomolar
mM	Millimolar
kDa	Kilo dalton
CD	Circular dichorism
DSC	Differential scanning calorimetry
PCR	Polymerase chain reaction
EMSA	Electrophoretic mobility shift assay
4WJ	Four way-junction

Contents

CHAPTER 1

General Introduction of HV And Review of Literature

Contents

Genomic DNA is much longer in size (centimeters to meters), in comparison with the cell in which it is contained (a few micrometers) [1,2]. This raises the need for compaction and condensation of otherwise highly negatively charged, stiff and elongated DNA into compact form. Compaction is achieved by the coordinated action of several factors which induce and regulate the super-coiling of DNA, and by the "architectural" DNA-binding proteins [3,4]. Unlike, the regulatory DNA binding-proteins which bind to DNA only transiently, the architectural DNA-binding proteins remain bound to DNA and neutralize its negative charge. These proteins also control DNA topology by altering its shape through bending, folding, wrapping and bridging, leading to compaction of DNA by formation of loops and coils [1,5]. The folded form of the genome associated with various architectural DNA-binding proteins is referred to as "chromatin". In eukaryotic cells, histones are the main architectural DNA-binding proteins which self-assemble to form octameric structures that act as ‘protein-beads’ around which the genomic DNA is wrapped and condensed within the nucleus [6,7]. However, in bacteria and archaea, the genome is not constrained in any particular compartment and lies freely floating inside the cytoplamic environments of cells, as a compact object known as the ‘nucleoid’. The prokaryotic genome, because of its functional equivalence to the eukaryotic nucleus, is referred to as the "nucleoid" [8,9]. The proteins which perform the analogous function of DNA compaction in prokaryotes, although strictly speaking they may or may not serve to condense DNA, are the "histone-like proteins" also known as "nucleoid-associated proteins" (NAPs) [10].

The *E. coli* nucleoid is composed of approximately 4.6 million base-pairs (Mb) of circular DNA, compacted (over a 1000 times) and organized into a central core surrounded by more than 10,000 loops of super-coiled DNA [11,12,13]. This nucleoid is associated with several RNA molecules and more than a hundred classes of DNA-binding proteins which are found within, and also coat, the entire nucleoid [14,15]. The structure of the nucleoid is very ordered, and apparently well organized, and yet also very dynamic [3,16,17]. The level of super-coiling (the topology) of DNA affects processes like replication, control of transcriptional noise and suppression of cell to cell viability [18,19,20]. The ordered architecture of the genome plays a vital role in regulation of gene expression, and any perturbation in the overall super-helicity of chromatin changes the expression profile of the cell, which can be highly lethal [21,22]. The

reverse is also true, i.e., changes in the expression profiles of proteins, during certain stress conditions lead to changes in the overall structure of the nucleoid [23]. This suggests that the ordered organization of the nucleoid is critical for the normal growth of bacterial cells. Of all the factors that regulate nucleoid compaction, the NAPs play a very significant role by physically interacting with DNA and controlling its topology and super-helicity.

1.1. Nucleoid-associated proteins (NAPs)

There are several classes of NAPs present in prokaryotes which are highly diverse in their structure as well as function. NAPs show growth phase-dependent variations in their expression (Figure 1) and hence they regulate the dynamics of the nucleoid and its structure [2,24]. There are a few other NAPs which are not expressed in normally-growing cells, and are only induced under stressful conditions [25,26]. The most studied NAPs include: *HU* (histone-like) proteins that bind to DNA non-specifically, *H-NS* (histone-like nucleoid structuring) protein that play roles both in nucleoid compaction and transcription repression by binding specifically, with high affinity, to AT rich double-stranded DNA (dsDNA) [10], *IHF* (integration host factor) which binds to specific sequences on DNA and forms U-turns in DNA around the binding-site [27], *Fis* (factor for inversion stimulation), *CbpA/B* (curved DNA-binding protein), *DnaA* (DNA-binding protein A), *Dps* (DNA-binding protein from starved cells), *Lrp* (leucine-responsive regulatory protein), *Hfq* (host factor for phage Qb), *IciA* (inhibitor of chromosome initiation A), *StpA* (suppressor of td2 phenotype A) and *Dan* (DNA-binding protein under anaerobic conditions) [26,28].

Based on their localization inside the cell, the NAPs have been categorized into two classes; A) *major NAPs* that are uniformly distributed throughout the cell (HU, IHF, H-NS, StpA, Dps) and B) *regulatory NAPs* that are present only at specific loci inside the nucleoid (CbpA/B, SeqA, Fis, IciA) [29]. Based on their specificities of DNA binding, NAPs can be categorized into two classes; 1) *non- specific* DNA binding proteins (CbpA, Dps, Hfq, H-NS, HU, IciA, and StpA) and 2) *specific* DNA binding proteins (IHF, DnaA).

Changes in the concentration of NAPs occur as the cell grows, and these changes act to induce local structural transitions in the DNA. The intracellular concentrations of a few of these NAPs can be as high as 10μM during different phases of growth [24,26,30]. HU and Fis are the

most abundant of NAPs during the log phase of growth and their concentrations decrease during the stationary phase. Conversely, IHF and Dps are induced when cells enter the stationary phase. Dan and Dps on the other hand are expressed under anaerobic stress [31]. Promiscuity in the DNA binding properties of NAPs makes them excellent regulators of gene expression. Besides their role in compaction of DNA, these proteins are also involved in the regulation of several genes [28]. These proteins can show either similar or antagonistic effects on the expression levels of various genes [10,30].

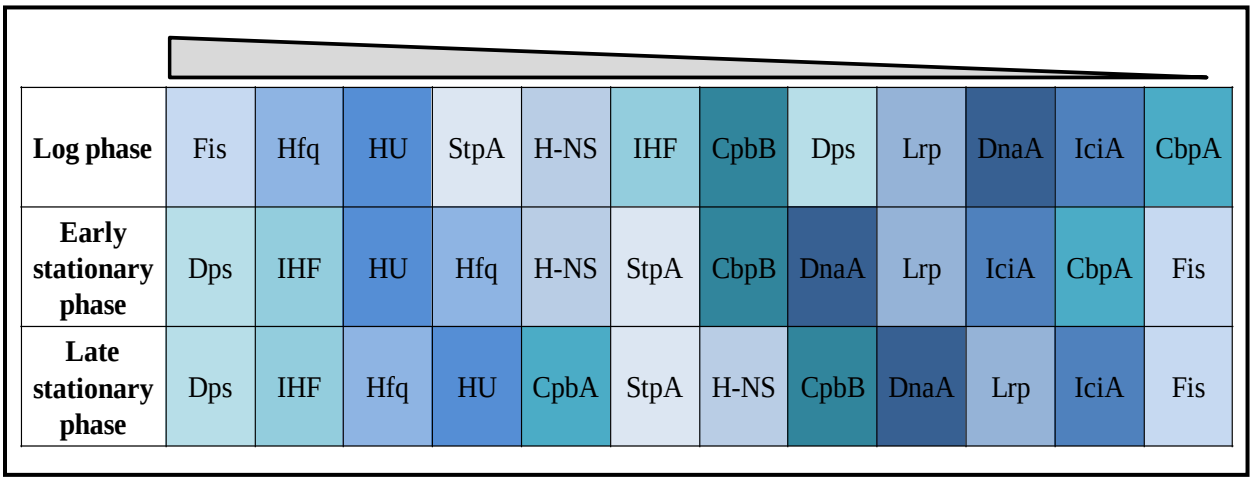


Figure 1: Growth phase dependent variation of concentration of different nucleoid associated proteins [26].

Unique aspects of HU: There are several properties of HU which make it unique, amongst the NAPs. Of all nucleoid-associated proteins, only HU is ubiquitously present in all eubacteria and it is the most abundant NAP which binds to DNA and also manages to bend it [21]. It has been demonstrated that for some bacterial species, upon deletion of one class of NAP, other NAPs are up-regulated to compensate for the deletion. Deletion of HU from *E. coli* is not lethal but the mutants show multiple growth defects [32,33] but in a few species like *Bacilus subtilis* [34] and *Mycobacterium tuberculosis* [35], which contain fewer NAP's (5 in *Mycobacterium tuberculosis* in contrast to 12 in *E. coli*).

1.2. HU: An introduction

HU was identified in *E. coli* U-93 as a factor named 'U' (from U-93) which can stimulate the *E. coli* RNA polymerase-mediated transcription of bacteriophage λ DNA [36,37]. This protein was later found to show DNA binding properties similar to histones, but the protein's sequence did

not show any homology to eukaryotic histones [38]. In any case, the protein was named HU, i.e. histone-like protein derived from U-93 bacteria [37,39]. In a parallel set of experiments, two isoforms of HU were identified as contaminants in preparations of ribosomes, and were named NS1 and NS2 [40]. Yaniv *et al* renamed these isoforms, HU1 and HU2 [41,42]. The names more commonly used now are HU-A or HU-α, for the protein originally-called HU2, and HU-B or HU-β for the protein originally-called HU1 [37]. It is important to note that HU protein shares high sequence similarity with the integration host factor, IHF. Unlike HU though, which binds to all kinds of DNA, IHF recognizes a specific sequence on DNA; however, the two share a very similar mode of binding.

HU is a 90-99 amino acids-long, alkaline, dimeric, histone-like protein which is capable of binding to DNA, bending it, and wrapping DNA around itself while associating into multimeric higher-order forms [36]. It belongs to the DNAB II family of DNA binding proteins. Unlike other proteins of this family, HU is ubiquitous in all eubacteria. It is present in very high intracellular concentrations of around 12,000-50,000 dimers/cell [24,43,44] and is the most abundant of nucleoid-associated proteins [28]. Besides eubacteria, it is also present in a few species of archaebacteria [45], bacteriophages [46,47], animal viruses [48], and dinoflagellates [49].It is also present in mollicutes, which have very small genome sizes, with HU being the only NAP which has been identified in these organisms so far [50]. HU has also been identified in eukaryotes, in organelles such as chloroplasts and mitochondria, which are thought to have a prokaryotic origin [51,52]. Sequence alignment of HU proteins from various sources (Figure 2) shows that HU is a highly conserved protein. HU from some organisms like *Deinococcus radiodurans* contains an additional N-terminal domain which shows sequence similarity with histone H1. These residues are not necessary for DNA binding but they do appear to modulate the DNA binding properties of the protein [53].

1.3. HU isoforms

HU is a highly conserved protein. It is mostly encoded by a single gene. However, in *Enterobaceriaceae* and *Vibrionaceae*, it is encoded by two genes, e.g., as we have already seen, *Escherichia coli* produces two isoforms of HU, HU-A and HU-B[54], and in species such as *Serratia marcescens* [55] *and Salmonella typhimurium* [56] also HU exists as two highly homologous isoforms. In *E.* coli, the two HU isoforms are encoded by closely related genes,

hupA and *hupB*, which encode HU-A and HU-B, respectively. On the *E. coli* genome, *hupA* is located at position 90 minutes [57] and *hupB* is located at position 10 minutes [58] from

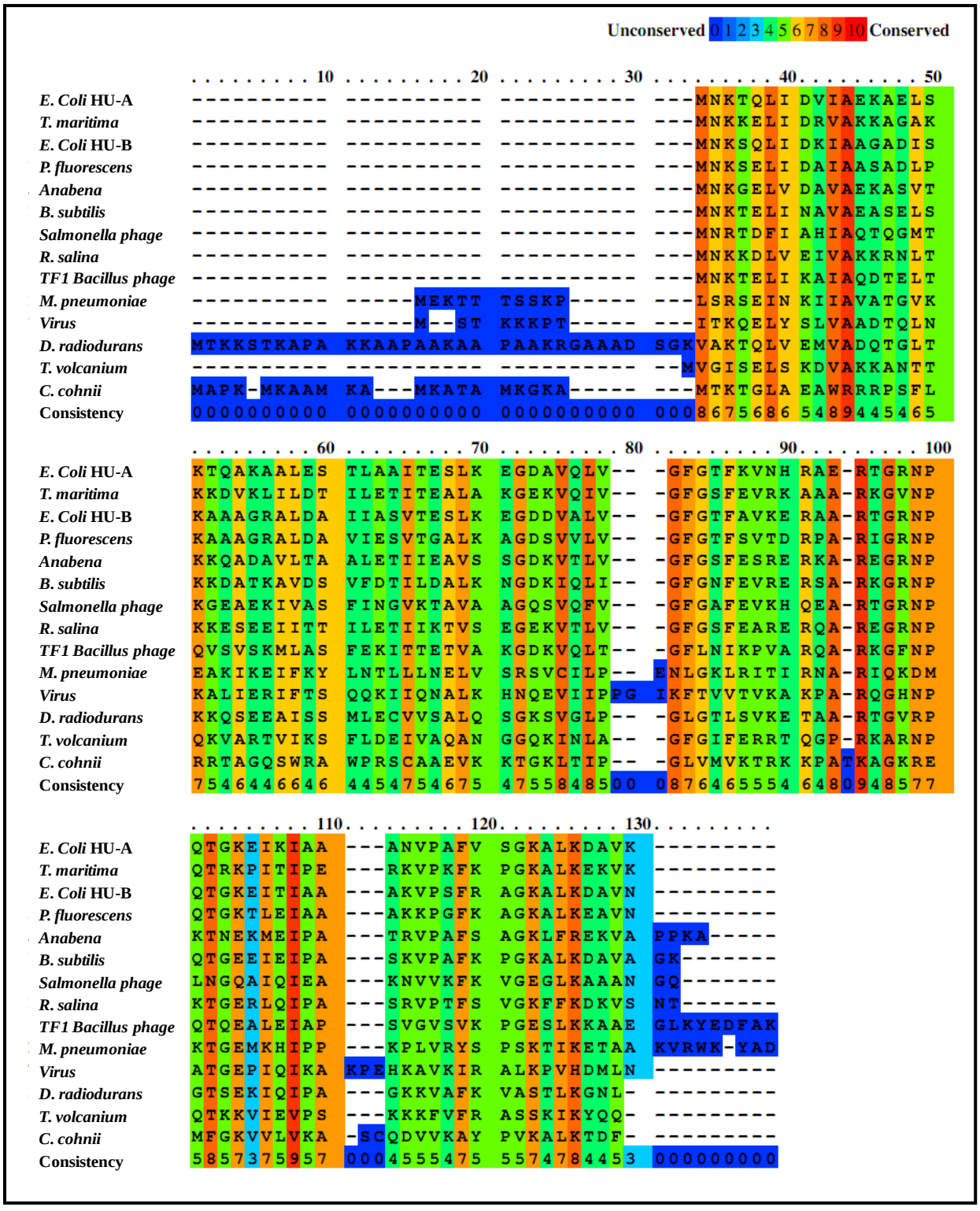


Figure 2: Sequence alignment of HU from various organisms.

the origin of replication. The gene, *hupA*, is expressed from a single promoter, but *hupB* expression can occur from three different promoters (P2, P3 and P4). These promoters actively transcribe in a non-coordinated manner during different phases of growth, to regulate the levels of HU-B expression [59]. It has been reported that both HU-A and HU-B negatively auto-regulate the expression levels of each other and that the *hupB* gene is less sensitive to being repressed by expression of the *hupA* gene[42].

Both HU-A and HU-B are dimeric proteins and can exist either as homo-dimers (HU-A2 and HU-B2) or hetero-dimers (HU-AB). As is seen with the other NAPs, the concentration of HU is also growth phase-dependent [59]. During the cell cycle, along with the changes in the overall total concentrations of HU, the relative concentrations of HU-A and HU-B, and hence, of the different dimeric forms, also change significantly. The HU-A2 homo-dimer is most abundant during the early log phase, followed by the HU-AB hetero-dimer population predominating during the late exponential and stationary phase [37]. Besides this, the relative concentrations of HU and H-NS are also critical for the normal growth of cells. The ration of HU: H-NS is 2.5 during the exponential phase of growth, but decreases to 1.0 during stationary phase, because of a decrease in the amount of HU. It is known that H-NS is a better compactor of DNA that HU; hence, the decrease in the HU: H-NS ratio probably explains the greater compaction of the nucleoid during the stationary phase [60].

1.4. HU: structural features

Molecular structures have been solved for *E. coli* HU (2O97, 1MUL), *B. stearothermophilus* (1HUU), *T. maritima* (1B8Z) and *Anabaena* (1P51) [61,62,63,64,65]. HU from all of these organisms shows very high sequence and structural similarity (Figure 2). As Figure 3A shows, the amino-terminal section of the protein consists of two α-helices connected by a turn (shown in blue) and the carboxy-terminal section has one α-helix (shown in aqua color). The remainder of the protein forms a three-stranded β-sheet structure (shown in green) which includes a β-ribbon extension in the middle (shown in red). This β-ribbon extension is reported to be disordered in the absence of bound DNA [64] and only becomes structured in the presence of DNA. It binds to the minor groove of DNA. In an HU dimer, the two subunits come together to form a compact α-helical body that is capped by β-sheets that extend into two β-ribbon arms (Figure 3B).

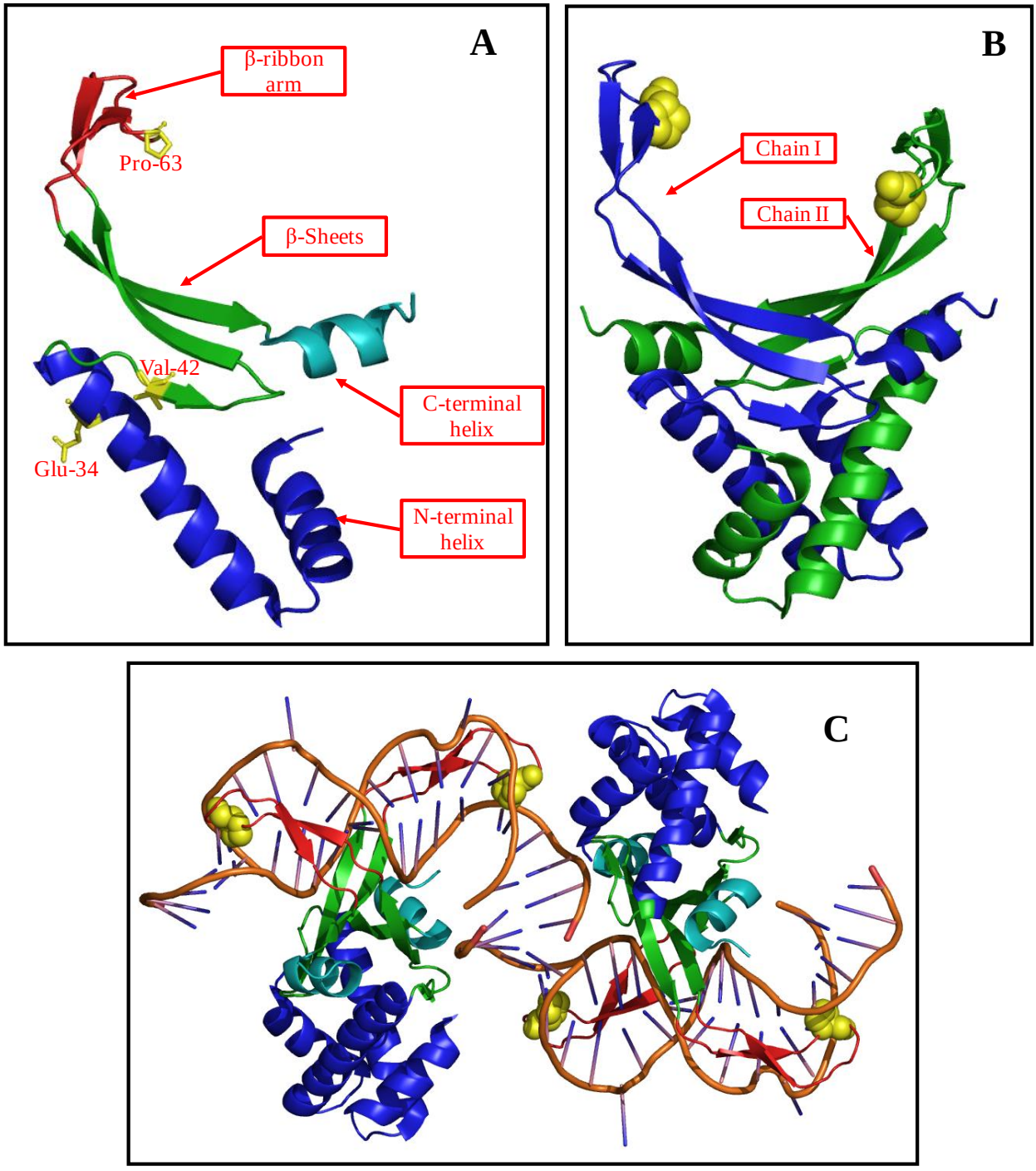


Figure 3: ***Panel A***- Structure of *Anabaena* HU monomer (1P51), showing various structural elements. ***Panel B***- Structure of *Anabaena* HU dimer. ***Panel C***- DNA bend introduced by binding of *Anabaena* HU dimers to DNA.

There is a highly conserved residue, Pro63, at the tip of two β-strands which plays a critical role in the interaction of HU with the minor groove of DNA. Based on analyses of sequence conservation, residues Gly-15, Glu-34 and Val-42 have been pointed out to be important for the thermo-stability of HU from thermophile organisms [66]. These residues are present in the loop connecting the two N-terminal helices, and hence they probably stabilize the dimer. In *E. coli* HU, of these three residues, Glu-34 and Val-42 are present but Gly-15 is missing.

1.5. DNA binding

Upon binding to DNA, both Pro63 residues in the β-ribbon extensions of the two subunits introduce kinks into DNA (Figure 3C) which are stabilized by the formation of salt bridges between DNA and proximal positively charged residues, and this leads to the bending of DNA [30,47,67,68]. The distance between two kinks introduced by the two Pro63 residues and the surface salt bridges determine the size of the binding site. DNA bending in general is an energetically un-favorable process and hence, the preferred substrates for HU are pre-bent DNA i.e., DNA molecules with imperfections such as base-pair mismatches or nicks, and cruciform DNA [30,69]. It is probably thus that the affinity of HU for pre-bent DNA is much higher (low nM) [70] than for linear DNA (in μM) [71,72,73]. May be for this reason, for the co-crystallization of DNA and HU from *Anabaena*, mis-matched DNA was used as substrate [65]. Due to structural distortions induced in DNA upon HU binding, the process of DNA binding by HU is enthalpically unfavorable but is entropically favorable.

Despite high structural conservation, HU proteins from various organisms show large variations in (i) size of the binding site, (ii) affinity for different DNA substrates and (iii) the bend angle introduced upon binding. The binding site for HU can either be as low as 9 bp, or as high as 34 bp, for HU protein from different organisms and their different DNA binding modes. For *Anabaena* HU, the binding site is 17-19 bp in length, and for *T. maritima* HU the binding site is greater than 35 bp in length. The bend angles introduced by binding of HU protein from different organisms have been reported to vary from 109-135°. Similarly, binding affinities of HU proteins can vary from 200-2500nM for pre-bent DNA [16,26,39,68,74,75,76,77]. The differences in the size of the binding site and in binding affinity, owe to differences in the number of basic residues on the surfaces of different HU proteins which stabilize DNA-protein interactions at the interface through the formation of salt bridges. As previously mentioned, HU

protein binds to linear DNA with very low affinity but it can bind to the following substrates with high binding affinities:

- Kinked DNA with distortions like single stranded breaks and mismatch [39,78]
- Nicked DNA [79]
- Negatively supercoiled DNA [80]
- Cruciform DNA [78,79,81]
- RNA [82,83]

1.6. Different DNA binding modes

The role of HU in DNA condensation is widely accepted and explored. Upon deletion of an HU encoding gene, nucleoid unfolding is observed. But, over-expression of HU does not show any effect on compaction of nucleoid [84]. On the other hand, for a similar NAP, H-NS which is also known to cause nucleoid compaction, over-expression leads to a significant over-compaction of the nucleoid [85]. Surprisingly, the DNA compaction properties of HU are antagonistic to those of H-NS i.e., HU can release the DNA compaction induced by binding of H-NS to DNA [86]. In the light of all this, the DNA compaction property of HU becomes questionable. However, it has now been demonstrated that HU has two different DNA binding modes and can have different effects on DNA condensation, depending on the relative protein concentrations [87] and also on different salt concentrations [88].

a) At low protein concentrations (<100 nM), HU dimer binds to DNA and bends it, which is reflected as a decrease in the persistence length of DNA.

b) At higher protein concentrations (>100 nM), HU molecules arrange helically around the DNA and form rigid filaments which does not reflect in any change in the persistence length of DNA. Under these conditions of binding, negligible DNA condensation is observed [87]. Only after prolonged incubation DNA-protein rearrangements happen to lead to a decrease in DNA persistence length [89].

This suggests that HU can lead to DNA compaction and condensation even at very low protein concentrations (less than 100 nM). So the question that arises is why the intracellular concentrations of HU are so high (10 mM)! It is postulated that by changing HU concentrations during the cell cycle (and hence by exploiting the two different binding modes of HU binding to

DNA), dynamics between the rigid and compact DNA is maintained inside the cell [90].It is commonly thought that variations in HU's binding modes and affinities facilitate HU's role in maintaining the structure and dynamics of prokaryotic chromatin during different phases of bacterial growth, and in modulating diverse processes such as replication, transcription, and recombination.

1.7. Pleotropic effects of HU: nucleoid compaction and more

In *hupAB* deletion mutants, the nucleoid structure is unfolded and supercoiling is also reduced [32,41]. HU protein, upon binding to DNA *in vitro* leads to compaction of DNA into nucleosome like structures[91]. But, HU does much more than just dynamic modulation of nucleoid structure [92]. HU-DNA binding acts as a global transcription regulator and has been shown to regulate 353 genes (corresponding to 229 operons) either by co-operating with the known transcription regulators of the genes or by direct physical interactions with the DNA [93].

- ***Replication***

Low concentrations of HU have been shown to induce *in vitro* replication[94]. The first and very critical step in DNA replication is formation of a pre-replication complex that opens up DNA at OriC. HU is found associated with the pre-replication complex, along with other components, and it stimulates replication by bending DNA [95]. HU recruitment to the complex is facilitated by interaction of HU with DnaA [96]. HU upon binding to the complex, assists in binding of DnaB to the pre-priming complex [97]. But if the concentration of HU exceeds that required to restrain super coiling, or when HU is absent, the initiation of replication can proceed through the formation of only an R-loop [98]. Besides this, bacteriophages like CTXφ also requires HU for replication initiation [99].

- ***Transcription:***

It is well known and established that in eukaryotes, organization of DNA over histones is associated with transcriptional regulation. Evidences suggest that there exists a correlation between nucleoid compaction patterns and transcription regulation in prokaryotes also. It has been shown that a mutant of HU (HU-A E38K, V42L) with different binding characteristics than the wild type HU besides causing changes in the extent of nucleoid condensation, also causes significant changes in the transcription patterns of cell [44,100]. HU can regulate transcription,

either by directly stimulating the activity of T7 polymerase and by stabilizing mRNA *in vivo*[101] or by bending DNA to facilitate repression or de-repression of various operons. Binding of HU to the *gal* operon, induces bending of DNA and helps in recruitment of GalR to bipartite *gal* operators [102,103,104]. GalR specifically interacts with HU, recruiting it to a specific site, *hbs,* between two *gal* operators [105,106].The role of HU in the *proU* operon has also been studied [107].

- ***DNA repair and recombination***

HU proteins can bind to cruciform DNA such as repair and recombination intermediates like DNA invasions, DNA over-hangs and DNA forks with high affinities. These intermediates are stabilized upon binding of HU and are protected from cellular exonucleases [71]. In addition to this, HU stimulates DNA repair by facilitating the RecA-dependent repair pathway[108]. This makes HU critical for repair and recombination, and makes *hup* deletion strains highly sensitive to γ and UV-irradiations [109,110]. Besides this, HU is considered to be one of the crucial factors responsible for radio tolerance in radiation resistant organisms (like *Deinococcus radiodurans* and *Kineococcus radiotolerans*) [21,30].

- ***Transposition***

HU plays a significant role in site specific recognition of DNA by other proteins to form higher order nucleoprotein complexes [37]. Studies on HU deletion mutants show that HU is important for the formation of protein-DNA complexes required for transposition of bacteriophage *Mu* and transposon *TnJO*[32].

- ***Regulation of DNA super-coiling***

HU stabilizes and constrains supercoils by binding to DNA loops. Normal intracellular concentrations of HU are sufficient to modulate topoisomerase-I activity. In the absence of HU (in HU deletion strains), the relaxation efficiency of topoisomerase-I increases, suggesting that HU could have roles in maintaining super-helicity. The deletion of *hup* is also associated with the accumulation of suppressor mutations [32] inside cells, specifically in *gyrB* [111]. Decrease in super-coiling occurring as a result of HU deletion is balanced by suppressor mutation in gyrase, pointing towards the mutually antagonistic actions of gyrase and HU. Recently it has been shown that HU from *Mycobacterium tuberculosis* physically interacts with the

topoisomerase and helps in maintaining the super helical density by having stimulatory effects at lower concentrations where as by interfering at higher concentrations [112].

- ***Other functions***
  - *hin*-mediated gene inversion [113]
  - Incorporation of stationary phase adaptive mutations (by hetro-dimer) [114].
  - Extracellular functions: HU homologs have also been reported to be involved in few extracellular functions like cell adhesion [115] and in eliciting immune responses [116]. and in maintenance of structural integrity of the bacterial biofilms [117].

1.8. Regulation of HU expression

As mentioned earlier, both the absolute and relative concentrations of the isoforms of HU, as well as of its various homo- and hetero-dimeric forms, tend to vary with the growth phase of cells in culture. As different dimeric forms of HU have different binding affinities for various DNA substrates, this could be a mechanism of fine tuning nucleoid compaction. HU-A and HU-B are expressed from different promoters and their mechanisms of regulation are also different (although they do show some auto regulatory effects). Deletion of either *hupA* or *hupB* alone does not reflect in any significant effects on the growth of cultures[33], as the deletion of one of these genes upregulates the expression of the other [42,118]. On the other hand, upon overexpression of one of the two isoforms, the mRNA levels of the other are observed to be down-regulated [33] but such overexpression is well tolerated and compatible with transcription and cell growth [84]. Total absence of HU in cells, through creation of a double deletion mutant of both *hupA* and *hupB* leads to several abnormalities in cells. On deletion of one of the HU isoforms, the other can perform the similar functions, but deletion of both isoforms has many effects on the physiology of *E. coli* cells. A few of these are mentioned below:

- Non-uniform partitioning of chromosomes, leading to the formation of a few anucleate cells [33] due to accumulation of background secondary mutations, also in *min* operon [119].
- Cells with abnormal nucleoids [120]
- Formation of mini cells [119]
- Cells showing increased sensitivity to cold shock and becoming highly fragile [32]
- Immotile cells, because of loss of flagella [121]

- Increased sensitivity to UV and ionizing irradiations [108,109,110]
- Altered outer membrane protein composition [122]

1.9. The scope of present study

Most bacteria (unlike *E. coli*) have only one isoform of HU that shares sequence similarity with HU-B of *E. coli* and, therefore, only a single kind of dimer is formed. This makes it interesting to study the significance of the existence and regulation of three structurally similar but functionally distinct dimeric forms of HU in *E. coli*. We have explored the structural and functional properties of HU to (i) understand DNA binding by HU under different conditions, (ii) determine the differential overall stabilities of HU-A & HU-B homodimers and the HU-AB hetero-dimer, and compare these to the stability of dimer interface, (iii) perform studies on various HU mutants and (iv) use fluorescently tagged HU to highlight DNA in microscopic studies and quantitate leaky expression in otherwise tightly regulated expression systems. For each of these studies, we designed and created different native, mutant and variant clones of HU.

1.10. Rationale for construction and expression of native, mutant and variant clones of HU

Native forms of HU: A total of 6 native forms of *E. coli* HU-A and HU-B were cloned in various vectors, and expressed from the respective hosts. C-terminally 6xHis-tagged, N-terminally 6xHis-tagged and untagged forms of both HU-A and HU-B were expressed, purified and characterized. The following questions were addressed using these forms of HU.

1. *To examine whether HU can be expressed with affinity tag and which tags are viable*:
hup genes were cloned to express HU proteins containing either a C-terminal 6xHis-tag or an N-terminal 6xHis-tag in fusion. Untagged HU isoforms were also cloned as controls to examine whether the fusion of 6xHis-tag at either terminus interferes with the structural or functional properties of HU. The pET23a vector contains a sequence encoding a 6xHis-tag at the 3'-end of the multiple cloning site (MCS); hence upon expression in BL21Star(DE3)[pLysS] cells, C-terminally 6xHis-tagged HU isoforms were over-expressed by inducing cells. By using a primer encoding stop codon between the gene sequence and the restriction site, untagged forms of HU

were expressed from pET23a vector transformed into BL21Star(DE3)[pLysS] cells. For expressing HU encoding N-terminal6xHis-tag, *hup* genes were cloned into pQE30 vectors, which encode a 6xHis-Tag at the 5'-end of the MCS, leading to expression of N-terminally 6xHis-tagged isoforms. pQE30 vectors containing N-terminally 6xHis-tagged HU isoforms were expressed either from XL1Blue cells in a constitutive manner, or from M15 cells in a regulated manner, i.e., with expression occurring only after induction.

2. *To explore the role of HU as a DNA carrier:*

HU has been demonstrated to be a key component of DNA-containing biofilms [117] and as an efficient and effective carrier (into eukaryotic cells, through a process akin to transformation) of synthetic DNA drugs. This suggests that HU has high DNA binding affinities and can carry DNA outside and between bacterial cells. Given HU's high DNA-binding affinity, we wanted to see if HU purified under native conditions copurifies with DNA. Using N-terminally 6xHis tagged HU and affinity purification, the propensity of HU to carry DNA with it in bound form under various conditions (that are known to break DNA-protein interactions) was explored in detail.

3. *To check if the co-eluting DNA pulls out other DNA-binding proteins also during affinity purification:*

After having established that HU can carry DNA when purified under native conditions, we next wanted to see if HU purification is associated with contamination of other DNA-binding proteins. For that purpose, affinity purification of N-terminally 6xHis-tagged HU isoforms was performed under native conditions. Mass spectrometric analysis of the co-eluting proteins was done to identify if they are known DNA binding proteins.

4. *To design a single step purification method to purify DNA-free but DNA-binding competent HU:*

N-terminally 6xHis tagged HU isoforms were purified under many different conditions to find conditions under which HU can be purified free of DNA and other contaminating proteins but retains its DNA binding competence and structural integrity.

5. *To examine how easily HU can be unfolded and refolded, and whether the refolded form is DNA binding competent:*

It is reported that HU is a heat-stable protein. Affinity tagged forms of HU were used to examine how HU can be unfolded by heat and also whether it unfolds and can be refolded. The DNA-binding competence of the refolded protein was also examined. Refolding of HU both after thermal and chemical denaturation was also studied. Both dimerization propensity and DNA binding competence were examined for the refolded proteins.

6. *To compare stabilities of the dimer interface and the overall chain stability:*

As discussed earlier, HU-A and HU-B besides forming homo-dimers also form HU-AB hetero-dimers. The mechanism of chain switching to form hetero-dimers from homo-dimers is not well understood. To understand the mechanism of chain switching, we looked at the differential stabilities (inter and intra- subunit) of the homo-dimeric (N-terminally tagged) forms of HU.

7. *To examine issues relating to subunit exchange in homo and hetero-dimers:*

To validate whether hetero-dimerization of HU-A and HU-B can occur *in vitro*, in absence of DNA, a genetic fusion of HU-A and HU-B was constructed. In the fusion protein HU-B and HU-A were placed next to each other with HU-B preceding HU-A, and with an 11 amino-acids long 'serine-glycine' linker separating the two fusion partners. The fusion construct was made in the hope that it would allow us to physically 'simulate' the HU-AB heterodimer in the form of an "HU-B-HU-A fusion" forcing only the hetero-meric interface to form under conditions disallowing inter-chain interactions that would otherwise allow homo-meric interactions. The construct was made because it is difficult to isolate pure HU-AB heterodimers (given the extreme similarities of the properties of HU-A and HU-B). The construct was tested for structural content and DNA-binding ability, and found to be qualitatively just like the HU-A and HU-B homodimers. The stability of the hetero-dimer was determined and compared to that of the homo-dimers, HU-A and HU-B.

Mutant forms of HU: A total of five single substitution mutants of HU-A and HU-B and one double substitution mutant of HU-A were cloned. The mutants made were HU-A F47W, HU-A F51W, HU-A F79W, HU-B F47W, HU-B F79W and HU-A F47W F79W. The following points describe the rationale(s) for construction of different HU mutants.

8. *Introduction of Tryptophans in place of Phenylalanine to create spectroscopic (fluorescent) handles, or probes, for monitoring DNA binding*:

HU-A and HU-B naturally lack both tyrosine and tryptophan residues. Intriguingly, the absence of tryptophan is evolutionary conserved across all HU sequences in different bacteria. To generate the scope for fluorescence spectroscopic analysis of DNA-binding, we created multiple tryptophan-insertion mutants (replacing phenylalanine residues) of both HU-A and HU-B. Mutants incorporating Trp were characterized and compared to the wild type protein in respect of their structural properties and stability. DNA binding studies were carried out with selected mutants to calculate dissociation constant values for HU-A and HU-B binding to a synthetic 4-way junction (cruciform) DNA using fluorescence spectroscopy.

9. *To examine effect of introduction of Trp residue in mutants:*

We also tentatively explored the effect of tryptophan insertion upon DNA integrity, in respect of the likelihood of photosensitized oxidative modifications occurring when Trp residues are present in close proximity to DNA over long durations

Visibly fluorescent variants of HU: Two different visibly fluorescently tagged variants of HU, HU-A-RFP and HU-B-Venus were cloned and expressed, with the following objectives in mind:

10. To microscopically examine whether HU is present in extracellular biofilms:

Fluorescently tagged HU-B protein (HU-B-Venus) was cloned in the pQE30 vector which contains T5 promoter and can be identified and expressed by *E. coli*T5 polymerase. HU-B-Venus cloned in pQE30 was transformed into the biofilm-forming MG1655 strain, and localization of the fluorescent protein and effects of over-expression of HU-B-Venus on cell morphology were studied using fluorescence microscopy coupled with DIC imaging.

11. To examine whether HU can carry a fluorescent protein (FP) as payload while retaining dimeric nature and DNA binding competence:

HU tagged to fluorescent protein can be used to image bacterial nucleoid. We wished to explore the effects of tagging a small protein like HU (10 kDa) to an approximately 25 kDa big fluorescent protein on the dimerization propensity and DNA binding property of the smaller fusion partner.

12. *To quantitate levels of leaky expression from differentially regulated vector-host systems:*
Fluorescently tagged HU molecules were used to design a system to highlight DNA in microscopic images of bacteria, and to quantitate leaky expression in tightly regulated bacterial expression systems by fluorescence imaging. The fusion proteins (RFP-HU-A or Venus-HU-B) contain a DNA binding protein and a fluorescent protein such that the fusion protein is capable of binding to the *E. coli* nucleoid and hence, localizing the fluorescence signal. These constructs were used to detect molecules expressed in the absence of the inducer, IPTG, followed by quantitation of leaky expression both statistically and at the level of single cells.

1.11. References

[1] J.A. Alberts B, Lewis J, et al. , Molecular Biology of the Cell., 4th ed., Garland Science, New York, 2002.

[2] R.L. Ohniwa, K. Morikawa, J. Kim, T. Ohta, A. Ishihama, C. Wada, K. Takeyasu, Dynamic state of DNA topology is essential for genome condensation in bacteria, The EMBO journal 25 (2006) 5591-5602.

[3] R.T. Dame, The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin, Mol Microbiol 56 (2005) 858-870.

[4] R. de Vries, DNA condensation in bacteria: Interplay between macromolecular crowding and nucleoid proteins, Biochimie 92 (2010) 1715-1721.

[5] A. Anuchin, A. Goncharenko, O. Demidenok, A. Kaprelyants, Histone-like proteins of bacteria (review), Applied biochemistry and microbiology 47 (2011) 580-585.

[6] R.D. Kornberg, Chromatin structure: a repeating unit of histones and DNA, Science 184 (1974) 868-871.

[7] F. Thoma, T. Koller, A. Klug, Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin, The Journal of cell biology 83 (1979) 403-427.

[8] C. Robinow, E. Kellenberger, The bacterial nucleoid revisited, Microbiological reviews 58 (1994) 211.

[9] A. Popławski, R. Bernander, Nucleoid structure and distribution in thermophilic Archaea, Journal of bacteriology 179 (1997) 7625-7630.

[10] C.J. Dorman, K.A. Kane, DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes, FEMS Microbiol Rev 33 (2009) 587-592.

[11] H. Delius, A. Worcel, Electron microscopic visualization of the folded chromosome of Escherichia coli, Journal of molecular biology 82 (1974) 107-109.

[12] The PyMOL Molecular Graphics System, Schrödinger, LLC.

[13] R. Kavenoff, O.A. Ryder, Electron microscopy of membrane-associated folded chromosomes of Escherichia coli, Chromosoma 55 (1976) 13-25.

[14] R.L. Ohniwa, Y. Ushijima, S. Saito, K. Morikawa, Proteomic analyses of nucleoid-associated proteins in Escherichia coli, Pseudomonas aeruginosa, Bacillus subtilis, and Staphylococcus aureus, PLoS One 6 (2011) e19172.

[15] S.F. Lee, M.A. Thompson, M.A. Schwartz, L. Shapiro, W.E. Moerner, Super-resolution imaging of the nucleoid-associated protein HU in Caulobacter crescentus, Biophys J 100 L31-33.

[16] S.S. Broyles, D.E. Pettijohn, Interaction of the Escherichia coli HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch, J Mol Biol 187 (1986) 47-60.

[17] E. Toro, L. Shapiro, Bacterial chromosome organization and segregation, Cold Spring Harb Perspect Biol 2 (2010) a000349.

[18] C.J. Dorman, Genome architecture and global gene regulation in bacteria: making progress towards a unified model?, Nat Rev Microbiol 11 (2013) 349-355.

[19] A. Travers, G. Muskhelishvili, DNA supercoiling - a global transcriptional regulator for enterobacterial growth?, Nat Rev Microbiol 3 (2005) 157-169.

[20] J.M. Vilar, L. Saiz, DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise, Curr Opin Genet Dev 15 (2005) 136-144.

[21] M. Toueille, B. Mirabella, P. Guerin, C. Bouthier de la Tour, S. Boisnard, H.H. Nguyen, L. Blanchard, P. Servant, A. de Groot, S. Sommer, J. Armengaud, A comparative proteomic approach to better define Deinococcus nucleoid specificities, J Proteomics 75 (2012) 2588-2600.

[22] M. Gellert, K. Mizuuchi, M.H. O'Dea, T. Itoh, J.I. Tomizawa, Nalidixic acid resistance: a second genetic character involved in DNA gyrase activity, Proc Natl Acad Sci U S A 74 (1977) 4772-4776.

[23] S. Rimsky, A. Travers, Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins, Curr Opin Microbiol 14 (2011) 136-141.

[24] T. Ali Azam, A. Iwata, A. Nishimura, S. Ueda, A. Ishihama, Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid, J Bacteriol 181 (1999) 6361-6370.

[25] C.J. Lim, S.Y. Lee, J. Teramoto, A. Ishihama, J. Yan, The nucleoid-associated protein Dan organizes chromosomal DNA through rigid nucleoprotein filament formation in E. coli during anoxia, Nucleic acids research 41 (2013) 746-753.

[26] T.A. Azam, A. Ishihama, Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity, J Biol Chem 274 (1999) 33105-33113.

[27] K.K. Swinger, P.A. Rice, IHF and HU: flexible architects of bent DNA, Curr Opin Struct Biol 14 (2004) 28-35.

[28] D.F. Browning, D.C. Grainger, S.J. Busby, Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression, Curr Opin Microbiol 13 (2010) 773-780.

[29] T.A. Azam, S. Hiraga, A. Ishihama, Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid, Genes Cells 5 (2000) 613-626.

[30] A. Grove, Functional evolution of bacterial histone-like HU proteins, Curr Issues Mol Biol 13 (2011) 1-12.

[31] J. Teramoto, S.H. Yoshimura, K. Takeyasu, A. Ishihama, A novel nucleoid protein of Escherichia coli induced under anaerobiotic growth conditions, Nucleic Acids Res 38 (2010) 3605-3618.

[32] O. Huisman, M. Faelen, D. Girard, A. Jaffe, A. Toussaint, J. Rouviere-Yaniv, Multiple defects in Escherichia coli mutants lacking HU protein, J Bacteriol 171 (1989) 3704-3712.

[33] M. Wada, Y. Kano, T. Ogawa, T. Okazaki, F. Imamoto, Construction and characterization of the deletion mutant of hupA and hupB genes in Escherichia coli, J Mol Biol 204 (1988) 581-591.

[34] S. Fernández, F. Rojo, J.C. Alonso, The Bacillus subtilis chromatin-associated protein Hbsu is involved in DNA repair and recombination, Molecular microbiology 23 (1997) 1169-1179.

[35] T. Bhowmick, S. Ghosh, K. Dixit, V. Ganesan, U.A. Ramagopal, D. Dey, S.P. Sarma, S. Ramakumar, V. Nagaraja, Targeting Mycobacterium tuberculosis nucleoid-associated protein HU with structure-based inhibitors, Nat Commun 5 4124.

[36] J. Rouviere-Yaniv, F. Gros, Characterization of a novel, low-molecular-weight DNA-binding protein from Escherichia coli, Proc Natl Acad Sci U S A 72 (1975) 3428-3432.

[37] K. Drlica, J. Rouviere-Yaniv, Histonelike proteins of bacteria, Microbiol Rev 51 (1987) 301-319.

[38] B. Laine, D. Kmiecik, P. Sautiere, G. Biserte, M. Cohen-Solal, Complete amino-acid sequences of DNA-binding proteins HU-1 and HU-2 from Escherichia coli, Eur J Biochem 103 (1980) 447-461.

[39] B. Castaing, C. Zelwer, J. Laval, S. Boiteux, HU protein of Escherichia coli binds specifically to DNA that contains single-strand breaks or gaps, J Biol Chem 270 (1995) 10291-10296.

[40] T. Suryanarayana, A.R. Subramanian, Specific association of two homologous DNA-binding proteins to the native 30-S ribosomal subunits of Escherichia coli, Biochim Biophys Acta 520 (1978) 342-357.

[41] J. Rouviere-Yaniv, N.O. Kjeldgaard, Native Escherichia coli HU protein is a heterotypic dimer, FEBS Lett 106 (1979) 297-300.

[42] K. Kohno, M. Wada, Y. Kano, F. Imamoto, Promoters and autogenous control of the Escherichia coli hupA and hupB genes, J Mol Biol 213 (1990) 27-36.

[43] M.D. Ditto, D. Roberts, R.A. Weisberg, Growth phase variation of integration host factor level in Escherichia coli, J Bacteriol 176 (1994) 3738-3748.

[44] S. Kar, E.J. Choi, F. Guo, E.K. Dimitriadis, S.L. Kotova, S. Adhya, Right-handed DNA supercoiling by an octameric form of histone-like protein HU: modulation of cellular transcription, J Biol Chem 281 (2006) 40144-40153.

[45] F. Orfaniotou, P. Tzamalis, A. Thanassoulas, E. Stefanidi, A. Zees, E. Boutou, M. Vlassi, G. Nounesis, C.E. Vorgias, The stability of the archaeal HU histone-like DNA-binding protein from Thermoplasma volcanium, Extremophiles 13 (2009) 1-10.

[46] E.P. Geiduschek, G.J. Schneider, M.H. Sayre, TF1, a bacteriophage-specific DNA-binding and DNA-bending protein, J Struct Biol 104 (1990) 84-90.

[47] A. Grove, Surface salt bridges modulate DNA wrapping by the type II DNA-binding protein TF1, Biochemistry 42 (2003) 8739-8747.

[48] J.G. Neilan, Z. Lu, G.F. Kutish, M.D. Sussman, P.C. Roberts, T. Yozawa, D.L. Rock, An African swine fever virus gene with similarity to bacterial DNA binding proteins, bacterial integration host factors, and the Bacillus phage SPO1 transcription factor, TF1, Nucleic Acids Res 21 (1993) 1496.

[49] J.T. Wong, D.C. New, J.C. Wong, V.K. Hung, Histone-like proteins of the dinoflagellate Crypthecodinium cohnii have homologies to bacterial DNA-binding proteins, Eukaryot Cell 2 (2003) 646-650.

[50] S. Levitskiy, A. Sycheva, D. Kharlampieva, J. Oberto, D. Kamashev, M. Serebryakova, S. Moshkovskii, V. Lazarev, V. Govorun, Purification and functional analysis of recombinant Acholeplasma laidlawii histone-like HU protein, Biochimie 93 (2011) 1102-1109.

[51] J.-F. Briat, S. Letoffe, R. Mache, J. Rouviere-Yaniv, Similarity between the bacterial histone-like protein HU and a protein from spinach chloroplasts, FEBS letters 172 (1984) 75-79.

[52] T. Kobayashi, M. Takahara, S.-y. Miyagishima, H. Kuroiwa, N. Sasaki, N. Ohta, M. Matsuzaki, T. Kuroiwa, Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids, The Plant Cell 14 (2002) 1579-1589.

[53] S. Ghosh, A. Grove, The Deinococcus radiodurans-encoded HU protein has two DNA-binding domains, Biochemistry 45 (2006) 1723-1733.

[54] J. Oberto, K. Drlica, J. Rouviere-Yaniv, Histones, HMG, HU, IHF: Meme combat, Biochimie 76 (1994) 901-908.

[55] J. Oberto, J. Rouviere-Yaniv, Serratia marcescens contains a heterodimeric HU protein like Escherichia coli and Salmonella typhimurium, J Bacteriol 178 (1996) 293-297.

[56] N.P. Higgins, D. Hillyard, Primary structure and mapping of the hupA gene of Salmonella typhimurium, J Bacteriol 170 (1988) 5751-5758.

[57] Y. Kano, K. Osato, M. Wada, F. Imamoto, Cloning and sequencing of the HU-2 gene of Escherichia coli, Mol Gen Genet 209 (1987) 408-410.

[58] Y. Kano, S. Yoshino, M. Wada, K. Yokoyama, M. Nobuhara, F. Imamoto, Molecular cloning and nucleotide sequence of the HU-1 gene of Escherichia coli, Mol Gen Genet 201 (1985) 360-362.

[59] L. Claret, J. Rouviere-Yaniv, Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival, J Mol Biol 273 (1997) 93-104.

[60] T.A. Azam, A. Iwata, A. Nishimura, S. Ueda, A. Ishihama, Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid, Journal of bacteriology 181 (1999) 6361-6370.

[61] I. Tanaka, K. Appelt, J. Dijk, S.W. White, K.S. Wilson, 3-A resolution structure of a protein with histone-like properties in prokaryotes, Nature 310 (1984) 376-381.

[62] E. Borgmann, F. Kreienbring, M. Meinl, J. Wunsche, [Protein digestibility and the absorption of amino acids in various segments of the digestive tract of pigs. 3. Results of the fractionation of ileum chyme after feeding various rations], Arch Tierernahr 29 (1979) 419-435.

[63] S.W. White, K.S. Wilson, K. Appelt, I. Tanaka, The high-resolution structure of DNA-binding protein HU from Bacillus stearothermophilus, Acta Crystallogr D Biol Crystallogr 55 (1999) 801-809.

[64] R. Boelens, H. Vis, C.E. Vorgias, K.S. Wilson, R. Kaptein, Structure and dynamics of the DNA binding protein HU from Bacillus stearothermophilus by NMR spectroscopy, Biopolymers 40 (1996) 553-559.

[65] K.K. Swinger, K.M. Lemberg, Y. Zhang, P.A. Rice, Flexible DNA bending in HU-DNA cocrystal structures, EMBO J 22 (2003) 3749-3760.

[66] E. Christodoulou, C.E. Vorgias, The thermostability of DNA-binding protein HU from mesophilic, thermophilic, and extreme thermophilic bacteria, Extremophiles 6 (2002) 21-31.

[67] A. Grove, T.C. Saavedra, The role of surface-exposed lysines in wrapping DNA about the bacterial histone-like protein HU, Biochemistry 41 (2002) 7597-7603.

[68] E. Kamau, N.D. Tsihlis, L.A. Simmons, A. Grove, Surface salt bridges modulate the DNA site size of bacterial histone-like HU proteins, Biochem J 390 (2005) 49-55.

[69] L. Jen-Jacobson, L.E. Engler, L.A. Jacobson, Structural and thermodynamic strategies for site-specific DNA binding proteins, Structure 8 (2000) 1015-1023.

[70] C.I. Vitoc, I. Mukerji, HU binding to a DNA four-way junction probed by Forster resonance energy transfer, Biochemistry 50 (2011) 1432-1441.

[71] D. Kamashev, J. Rouviere-Yaniv, The histone-like protein HU binds specifically to DNA recombination and repair intermediates, EMBO J 19 (2000) 6527-6535.

[72] V. Pinson, M. Takahashi, J. Rouviere-Yaniv, Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA, J Mol Biol 287 (1999) 485-497.

[73] K.K. Swinger, P.A. Rice, Structure-based analysis of HU-DNA binding, J Mol Biol 365 (2007) 1005-1016.

[74] E. Bonnefoy, J. Rouviere-Yaniv, HU and IHF, two homologous histone-like proteins of Escherichia coli, form different protein-DNA complexes with short DNA fragments, EMBO J 10 (1991) 687-696.

[75] C. Chen, S. Ghosh, A. Grove, Substrate specificity of Helicobacter pylori histone-like HU protein is determined by insufficient stabilization of DNA flexure points, Biochem J 383 (2004) 343-351.

[76] D. Esser, R. Rudolph, R. Jaenicke, G. Bohm, The HU protein from Thermotoga maritima: recombinant expression, purification and physicochemical characterization of an extremely hyperthermophilic DNA-binding protein, J Mol Biol 291 (1999) 1135-1146.

[77] S. Ghosh, A. Grove, Histone-like protein HU from Deinococcus radiodurans binds preferentially to four-way DNA junctions, J Mol Biol 337 (2004) 561-571.

[78] A. Pontiggia, A. Negri, M. Beltrame, M.E. Bianchi, Protein HU binds specifically to kinked DNA, Molecular microbiology 7 (1993) 343-350.

[79] D. Kamashev, A. Balandina, J. Rouviere-Yaniv, The binding motif recognized by HU on both nicked and cruciform DNA, EMBO J 18 (1999) 5434-5444.

[80] H. Shindo, A. Furubayashi, M. Shimizu, M. Miyake, F. Imamoto, Preferential binding of E. coli histone-like protein HUα to negatively supercoiled DNA, Nucleic acids research 20 (1992) 1553-1558.

[81] E. Bonnefoy, M. Takahashi, J.R. Yaniv, DNA-binding parameters of the HU protein of Escherichia coli to cruciform DNA, Journal of molecular biology 242 (1994) 116-129.

[82] M. Macvanin, R. Edgar, F. Cui, A. Trostel, V. Zhurkin, S. Adhya, Noncoding RNAs binding to the nucleoid protein HU in Escherichia coli, J Bacteriol 194 (2012) 6046-6055.

[83] A. Balandina, D. Kamashev, 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 277 (2002) 27622-27628.

[84] V. McGovern, N.P. Higgins, R.S. Chiz, A. Jaworski, H-NS over-expression induces an artificial stationary phase by silencing global transcription, Biochimie 76 (1994) 1019-1029.

[85] R. Spurio, M. Dürrenberger, M. Falconi, A. La Teana, C.L. Pon, C.O. Gualerzi, Lethal overproduction of the Escherichia coli nucleoid protein H-NS: ultramicroscopic and molecular autopsy, Molecular and General Genetics MGG 231 (1992) 201-211.

[86] R.T. Dame, N. Goosen, HU: promoting or counteracting DNA compaction?, FEBS letters 529 (2002) 151-156.

[87] J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, R.T. Dame, Dual architectural roles of HU: formation of flexible hinges and rigid filaments, Proc Natl Acad Sci U S A 101 (2004) 6969-6974.

[88] B. Xiao, R.C. Johnson, J.F. Marko, Modulation of HU-DNA interactions by salt concentration and applied force, Nucleic Acids Res 38 (2010) 6176-6185.

[89] B. Kundukad, P. Cong, J.R. van der Maarel, P.S. Doyle, Time-dependent bending rigidity and helical twist of DNA by rearrangement of bound HU protein, Nucleic Acids Res 41 (2013) 8280-8288.

[90] G. Nir, M. Lindner, H.R. Dietrich, O. Girshevitz, C.E. Vorgias, Y. Garini, HU protein induces incoherent DNA persistence length, Biophys J 100 (2011) 784-790.

[91] J. Rouvière-Yaniv, M. Yaniv, J.-E. Germond, E. coli DNA binding protein HU forms nucleosome-like structure with circular double-stranded DNA, Cell 17 (1979) 265-274.

[92] D. Sagi, N. Friedman, C. Vorgias, A.B. Oppenheim, J. Stavans, Modulation of DNA conformations through the formation of alternative high-order HU–DNA complexes, Journal of molecular biology 341 (2004) 419-428.

[93] J. Oberto, S. Nabti, V. Jooste, H. Mignot, J. Rouviere-Yaniv, The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction, PLoS One 4 (2009) e4367.

[94] N.E. Dixon, A. Kornberg, Protein HU in the enzymatic replication of the chromosomal origin of Escherichia coli, Proceedings of the National Academy of Sciences 81 (1984) 424-428.

[95] R.S. Fuller, B.E. Funnell, A. Kornberg, The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites, Cell 38 (1984) 889-900.

[96] S. Chodavarapu, M.M. Felczak, J.R. Yaniv, J.M. Kaguni, Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin, Molecular microbiology 67 (2008) 781-792.

[97] T.A. Baker, B. Funnell, A. Kornberg, Helicase action of dnaB protein during replication from the Escherichia coli chromosomal origin in vitro, Journal of Biological Chemistry 262 (1987) 6877-6885.

[98] K. Skarstad, T.A. Baker, A. Kornberg, Strand separation required for initiation of replication at the chromosomal origin of E.coli is facilitated by a distant RNA--DNA hybrid, EMBO J 9 (1990) 2341-2348.

[99] E. Martinez, E. Paly, F.X. Barre, CTXphi Replication Depends on the Histone-Like HU Protein and the UvrD Helicase, PLoS Genet 11 (2015) e1005256.

[100] S. Kar, R. Edgar, S. Adhya, Nucleoid remodeling by an altered HU protein: reorganization of the transcription program, Proc Natl Acad Sci U S A 102 (2005) 16397-16402.

[101] P. Morales, J. Rouviere-Yaniv, M. Dreyfus, The histone-like protein HU does not obstruct movement of T7 RNA polymerase in Escherichia coli cells but stimulates its activity, J Bacteriol 184 (2002) 1565-1570.

[102] T. Aki, H.E. Choy, S. Adhya, Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor, Genes Cells 1 (1996) 179-188.

[103] T. Aki, S. Adhya, Repressor induced site-specific binding of HU for transcriptional regulation, EMBO J 16 (1997) 3666-3674.

[104] Y.L. Lyubchenko, L.S. Shlyakhtenko, T. Aki, S. Adhya, Atomic force microscopic demonstration of DNA looping by GalR and HU, Nucleic Acids Res 25 (1997) 873-876.

[105] S. Kar, S. Adhya, Recruitment of HU by piggyback: a special role of GalR in repressosome assembly, Genes Dev 15 (2001) 2273-2281.

[106] D.E. Lewis, S. Adhya, In vitro repression of the gal promoters by GalR and HU depends on the proper helical phasing of the two operators, J Biol Chem 277 (2002) 2498-2504.

[107] D. Manna, J. Gowrishankar, Evidence for involvement of proteins HU and RpoS in transcription of the osmoresponsive proU operon in Escherichia coli, J Bacteriol 176 (1994) 5378-5384.

[108] Q.-M.Z. Miyabe, Y. Kano, S. Yonei, I, Histone-like protein HU is required for recA gene-dependent DNA repair and SOS induction pathways in UV-irradiated Escherichia coli, International journal of radiation biology 76 (2000) 43-49.

[109] F. Boubrik, J. Rouviere-Yaniv, Increased sensitivity to gamma irradiation in bacteria lacking protein HU, Proc Natl Acad Sci U S A 92 (1995) 3958-3962.

[110] S. Li, R. Waters, Escherichia coli strains lacking protein HU are UV sensitive due to a role for HU in homologous recombination, Journal of bacteriology 180 (1998) 3750-3756.

[111] M. Malik, A. Bensaid, J. Rouviere-Yaniv, K. Drlica, Histone-like protein HU and bacterial DNA topology: suppression of an HU deficiency by gyrase mutations, J Mol Biol 256 (1996) 66-76.

[112] S. Ghosh, B. Mallick, V. Nagaraja, Direct regulation of topoisomerase activity by a nucleoid-associated protein, Nucleic acids research 42 (2014) 11156-11165.

[113] R.C. Johnson, Mechanism of site-specific DNA inversion in bacteria. , Curr. Opin. Genet. Devel 1 (1991) 412-416.

[114] A.B. Williams, P.L. Foster, The Escherichia coli histone-like protein HU has a role in stationary phase adaptive mutation, Genetics 177 (2007) 723-735.

[115] M.W. Stinson, R. McLaughlin, S.H. Choi, Z.E. Juarez, J. Barnard, Streptococcal histone-like protein: primary structure of hlpA and protein binding to lipoteichoic acid and epithelial cells, Infect Immun 66 (1998) 259-265.

[116] K. Pethe, V. Puech, M. Daffe, C. Josenhans, H. Drobecq, C. Locht, F.D. Menozzi, Mycobacterium smegmatis laminin-binding glycoprotein shares epitopes with Mycobacterium tuberculosis heparin-binding haemagglutinin, Mol Microbiol 39 (2001) 89-99.

[117] A. Devaraj, S.S. Justice, L.O. Bakaletz, S.D. Goodman, DNABII proteins play a central role in UPEC biofilm structure, Mol Microbiol (2015).

[118] K. Kohno, K. Yasuzawa, M. Hirose, Y. Kano, N. Goshima, H. Tanaka, F. Imamoto, Autoregulation of transcription of the hupA gene in Escherichia coli: evidence for steric hindrance of the functional promoter domains induced by HU, J Biochem 115 (1994) 1113-1118.

[119] A. Jaffe, D. Vinella, R. D'Ari, The Escherichia coli histone-like protein HU affects DNA initiation, chromosome partitioning via MukB, and cell division via MinCDE, J Bacteriol 179 (1997) 3494-3499.

[120] A.M. Dri, J. Rouviere-Yaniv, P.L. Moreau, Inhibition of cell division in hupA hupB mutant bacteria lacking HU protein, J Bacteriol 173 (1991) 2852-2863.

[121] S. Nishida, T. Mizushima, T. Miki, K. Sekimizu, Immotile phenotype of an Escherichia coli mutant lacking the histone-like protein HU, FEMS Microbiol Lett 150 (1997) 297-301.

[122] E. Painbeni, M. Caroff, J. Rouviere-Yaniv, Alterations of the outer membrane composition in Escherichia coli lacking the histone-like protein HU, Proc Natl Acad Sci U S A 94 (1997) 6712-6717.

CHAPTER 2

General Materials and Methods

Contents

2.1. Materials

2.1.1. *Strains and vectors used*

Strain	Genotype	Used for
1. *E. coli* XL1 Blue	*recA1 endA1 gyrA96 (nal^R) thi-1 hsdR17($r_k^- m_k^+$) supE44 relA1 lac [F'proABlacIqΔ(lacZ)M15 Tn10 (Tet^r)]* • Derived from *E. coli* K12 strain. • Nalidixic acid resistance • Tetracycline resistance (from F plasmid)	• Used as cloning host for all the constructs. • As expression of some constructs cloned in pQE30 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 λ prophage DE3.
	• IPTG inducible lac UV5 promoter.
	• pLysS plasmid encodes
	  • Chloramphenicol resistance gene.
	  • T7 phage lysozyme (inhibitor for T7 polymerase) which reduces expression from transformed T7 promoter containing plasmids when not induced
	• Mutation in RNaseE gene, (involved in mRNA degradation) offering enhanced mRNA stability for protein expression.
		Used as expression host for all genes cloned in pET vectors.

3. *E.* coli M15	*F- strthi$^-$ lac$^-$ ara$^+$ gal$^+$ mtl$^-$ recAuvr$^+$ lon$^+$Φ80 ΔlacM15 KmR*
	1. *E. Coli* K12 derived strain.
	2. Contains pREP4 which encodes
	  • lac repressor-presence of trans-lac repressor ensures tight regulation.
	  • Kanamycin resistance gene
		• Used as expression host for genes cloned in pQE30 vector.

2.1.2 Chemicals and Kits

Reagents used in this study were of analytical grade, and obtained from commercial sources. Restriction/modification enzymes and molecular biological reagents were obtained from New England Biolabs (NEB), USA or Fermentas. Protein molecular weight markers were purchased from Fermentas. Plasmid mini-prep kits, gel extraction kits, PCR purification kits, Plasmid midi-prep kits, Ni-NTA agarose spin column and Ni-NTA Agarose/Super flow used in this study were obtained from Qiagen, USA. All other fine chemicals were obtained from Sigma Chemicals, USA.

2.1.3 Media

Luria Broth (LB)

Component	Amount for LB
Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Agar (For LB agar plates)	2%
pH	7.4
Total volume	1 L

The media was sterilized by autoclaving (15 psi for 15 minutes at 121 °C).

2.1.4 Antibiotics

Ampicilin, Kanamycin, Chloramphenicol and Tetrcycline used in this study were procured from Sigma Chemicals, USA. For all antibiotics,1000X stocks were prepared as follows:

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

Stock solutions of antibiotics were sterilized through 0.22μM Millipore filters. Stocks were stored as aliquots at -20°C.

2.1.5. Buffers used for molecular biology work

2.1.5.1 Buffer for preparation of chemical competent cells

Calcium chloride	60 mM
Glycerol	15 % v/v
PIPES	10 mM
pH	7

The solution was sterilized by passing through 0.22 µm filter followed by autoclaving and stored at 4 °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

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

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

Ethidium bromide	0.1 g
Deionized water	10 ml

The stock solution was stored in amber color vial/bottle at 4 °C till further use.

2.1.5.5 TE buffer (In deionized water)

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

2.1.6. Buffers and solutions SDS-PAGE

2.1.6.1 Acrylamide

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

2.1.6.2 Lower Tris (4X), pH 8.8

Tris	18.17 g
10 % SDS	4 ml
pH	6.8
Total volume	100ml

2.1.6.3 Upper Tris (4X), pH 6.8

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

2.1.6.4 5X Sample loading buffer

Tris.Cl (pH 6.8)	0.15 M
SDS	5 %
Glycerol	25 %
β-mercaptoethanol	12.5 %
Bromophenol blue	0.06 %
Total Volume	10 l

2.1.6.5 Laemmli buffer **[1]**

Tris buffer	3.00 g
Glycine	14.4 g
SDS	1 g
Total volume	100ml

2.1.6.6 Gel staining solution

Methanol	40 %
Glacial acetic acid	10 %
Coomassie Brilliant Blue R-250	0.1 %
Deionized water	50

2.1.6.7 Gel destaining solution

Methanol	40 %
Glacial acetic acid	10 %
Deionized water	50 %

The composition of the stacking and resolving gels was the same as specified [2,3].

2.2 Methods

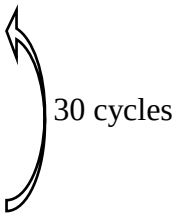
2.2.1 Polymerase chain reaction (PCR)

Standard PCR reactions with following composition were set:

	Component	**Supplier**	**Stock concentration**	**Final concentration**
1.	Template		Variable	1-2 ng of plasmid
2.	Forward Primer	IDT	10 mM	0.5 mM
3.	Reverse Primer	IDT	10 mM	0.5 mM
4.	dNTPs	NEB	250 mM	250 µM
5.	Polymerase(Taq, Vent, Deep Vent, Phusion)	NEB	2 units/µl	0.02 units/µl
6.	Buffer	NEB	10X (for Taq and vent) 5X (for Phusion)	1X
7.	$MgSo_4$ (optional)	NEB	100 mM	2-14 mM

All PCR reactions were carried out in Eppendorf PCR machine. The following PCR program was used:

	Steps	Temperature	Time
1.	Initial denaturation	95 °C	• 3 min (Vent/ Deep Vent) • 30 sec (Phusion)
2.	Denaturation	95 °C	• 30 sec (Vent/ Deep Vent) • 10 sec (Phusion)
3.	Annealing	40-65 °C	• 30 sec
4.	Extension	72 °C	• 1 min/kb (Vent/ Deep Vent) • 30 sec/kb (Phusion)
5.	Final extension	72 °C	• 10min



The PCR reactions were electrophoresed on agarose gel for size analysis.

2.2.2 Agarose gel electrophoresis

The separation or fractionation of DNA fragments was done by running them on 0.8% (for plasmids) and 1.0% (for PCR products) agarose. Gels were prepared by dissolving the required amount of agarose in 1X TAE. Ethidium bromide (0.5 μg/ml) was supplemented for visualizing DNA on an UV transilluminator. 6X gel loading buffer was added to DNA samples at a final concentration of 1X prior to loading onto the gel. Electrophoresis was carried out in 1X TAE buffer at ~ 8V/cm or 90 volts. 100 bp and 1 kb ladders were used as markers for calculating the size of DNA fragments from their relative mobility.

2.2.3 Purification of DNA bands from agarose gel

After electrophoresis, the gels were visualized on trans-illuminator and the band of interest was excised out of the gel and chopped into small pieces. Qiagen gel extraction kit was used to elute DNA out of the gel. The protocol used was as follows:

1. **Solubilization:** Gel pieces were weighed and dissolved in solubilization and binding buffer - Buffer QG (100 μl/ 100mg of gel pieces). Incubation was done at 50 °C till complete dissolution of the agarose gel was achieved.
2. **Binding:** The dissolved agarose solution containing the DNA was then poured onto a QIAquick spin column (provided by the manufacturer) to allow the adsorption of DNA onto the silica gel matrix.

3. **Washing:** This was followed by washing with wash buffer PE (contains ethanol) to remove impurities.
4. **Elution:** DNA was finally eluted in either elution buffer (10 mM Tris.Cl, pH 8.5) or in autoclaved distilled water.

2.2.4 Direct PCR purification

For PCR reactions giving neat single band of correct size (checked by agarose gel), for better yield, the PCR reactions were directly cleaned using Sigma PCR Clean up kit. The procedure included following steps:

1. **Binding:** 3 volumes of column wash buffer were added to the PCR reaction and was loaded onto the column supplied with the kit to allow adsorption of DNA onto the silica gel matrix.
2. **Washing:** This was followed by washing with wash buffer (containing ethanol) to remove impurities.
3. **Elution:** DNA was eluted in either elution buffer (10 mM Tris.Cl, pH 8.5) or in autoclaved distilled water.

2.2.5 Quantification of DNA

The purified DNA fragments (either after gel purification of after PCR purification) were quantified either by running on the agarose gel and visually comparing them with the various bands of the ladder depicting a definite amount of DNA in a definite amount of ladder being loaded, or by measuring 260 nm absorbance using a Nano drop spectrophotometer. This was done by putting 1 µl of DNA on the analysis probe of Nano drop spectrophotometer after setting up the baseline with water or TE depending upon which was being used for storing DNA. Absorbance of DNA was measured at 260 nm. An absorbance of 1 at 260 nm was considered equivalent to a concentration of 50 ng/ml double stranded DNA [2]. The purity of DNA was confirmed by measuring the ratio of OD_{260}/OD_{280} . Purified DNA should have a OD_{260}/OD_{280} ratio of around 1.8-2.0.

2.2.6 Restriction digestion

After quantification, the PCR and the plasmid DNA were digested using specific restriction endonucleases procured from Fermentas. The following reaction (general description) was set up:

1.	Template	~200 ng of PCR ~1 µl of Plasmid
2.	Fast digest buffer (10X)	1X
3.	Restriction enzyme	1 µl
4.	Water	For volume make-up

The reaction was incubated at 37˚C for not more than 30 mins. The reaction was run on agarose gel, gel purified and quantified.

2.2.7 Ligation

Digested and purified PCR and plasmid DNA were then used to set up a ligation reaction using T4 DNA Ligase or Quick Ligase (from NEB). An insert:vector ratio of 3:1 was used. The following reaction was set up:

1.	Vector	50 ng
2.	Insert	x ng
3.	Buffer (10X in case of T4 DNA ligase, 2X in case of Quick ligase)	1X
4.	Ligase	1 µl
5.	Water	For volume make-up

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

The reaction was incubated at 24°C for two hours or 16°C for 20 h in case of T4 DNA Ligase and at 25 °C for 10-15 minutes in case of Quick Ligase.

2.2.8 Preparation of E. coli competent cells

1 A single colony of *E. coli* was inoculated into 5 ml LB media and grown overnight

2 The culture was re-inoculated into 200ml of fresh LB medium (dilution 1:100) and grown to early log phase (A_{600} of 0.3-0.4).

3 The cells were chilled on ice for 15 min, centrifuged at 1600x g for 7 min at 4^0Cinprechilled centrifuge tubes. Cells were kept on ice at all subsequent steps during competent cell preparation.

4 The supernatant was discarded and cells were resuspended in 20 ml of ice cold60 mM calcium chloride solution, after which they were again centrifuged at 1100xg for 5 min at 4^0C.

5 The supernatant was again discarded and step 4 was repeated.

6 The cell pellet obtained in step 5 was resuspended in 20 ml of ice-cold calcium chloride solution and kept on ice for 30 min.

7 The cells were again centrifuged at 1100x g for 5 min at 4^0C and supernatant was carefully discarded after which the cells were resuspended in 4 ml of ice-cold calcium chloride solution.

8 Finally aliquots of 100 µl were made from the suspension obtained above and were used immediately or stored at -80^0C for later use.

2.2.9 Transformation

1. The competent cells were thawed on ice for 15 minutes.
2. Ligation mixture was added to the competent cells and mixed gently on ice. Mixture was incubated on ice for 15-20 minutes.
3. To give heat shock, cells were incubated at 42 °C for 90 seconds in water bath(Amersham Pharmacia Multitemp III).
4. After giving heat shock, immediately 1 ml of sterile LB media was added to the cells. Cells were allowed to grow at 37 °C for 1 hour in an incubator shaker.
5. Cells were centrifuged at 5000 rpm for 2 minutes and the supernatant was discarded.
6. The pelleted cells were resuspended in 100 µl of fresh media then plated on the LB agar plates supplemented with appropriate antibiotics.
7. The plates were incubated overnight at 37 °C and were observed for the colonies or the transformants the next day.

2.2.10 Plasmid DNA purification

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

1. **Cell growth:** Few of the transformants from the plates were picked and inoculated into 5ml of LB media supplied with the appropriate antibiotics. The cells were grown overnight so as to achieve saturation.
2. **Peletting and resuspension:** Cells were peletted at 13,000 rpm for 1 minute. The supernatant was discarded and 250 µl of P1 buffer (resuspension buffer) was added to the tube to resuspend the cells.
3. **Lysis:** Cells were lysed by adding 250 µl of P2 buffer (lysis buffer) and gently mixing it by inverting up and down.
4. **Neutralization**: 350 µl of N3 buffer (neutralization buffer) was added and mixed gently. The precipitated solution was centrifuged at 13,000 rpm for 10-15 minutes.
5. **Loading:** Supernatant was loaded onto the columns (supplied in the kit) and was spun at 13,000 rpm for 1 minute. Flow through was discarded.
6. **Washing:** Column was washed with 750 µl of PE buffer (wash buffer) and spun at 13,000 rpm for 1 minute. An empty spin of 1 minute at max speed was given to remove residual alcohol.
7. **Elution:** Finally DNA was eluted in either elution buffer (10 mM Tris.Cl, pH 8.5) or in autoclaved distilled water.

2.2.11 Screening of the transformants

Two step screening for positive clones was done prior to sequencing.

1. **Colony PCR:** For colony PCR, colonies were picked and suspended into PCR reaction mixture aliquoted into PCR tubes. For colony PCR, Taq polymerase (from NEB) was used. Vector specific primers (T5promoter Forward and T5 Reverse for pQE vectors and T7promoter Forward and T7 Reverse for pET vectors) were used for amplification. A normal PCR reaction as mentioned above was run and the reaction was run on agarose gel and size of the amplified band was observed. The clones with right size amplification were inoculated in LB media for plasmid purification.

2. **Restriction Digestion:** Following the isolation of plasmid DNA from the cultures of selected clones, the plasmids were checked for the integration of insert into the vector. For this, the plasmid DNA was double digested with the respective restriction enzymes by the method as described above. The digested plasmid was run on agarose gel, to check for fall out of insert of the correct size. If the insert of the expected size was observed, the clone was sent for sequencing.

2.2.12 Expression of recombinant proteins in E. coli

For expression of the proteins from the genes cloned in pET series vectors, the sequenced plasmids were transformed in BL21star(DE3)plysS cells and genes cloned in pQE30 the sequenced plasmids were transformed into the *E. coli* M-15 cells. Genes cloned in pQE vector can also be expressed in *E. coli* XL1 Blue cells, but in XL1 Blue cells, there is a constitutive expression of protein, so *E. coli* M-15 cells were used for regulated overexpression. Expression checking included the following steps:

1. **Primary culture:** cells were inoculated into 5 ml LB media supplemented with antibiotics and were incubated at 37ºC with shaking at 220 rpm.
2. **Secondary culture:** 1% overnight grown cultures were inoculated into fresh LB media. Cells were allowed to grow till mid log phase ($O.D._{600}$ 0.6)
3. **Induction:** Cells were induced using 1 mM IPTG and were further grown for 4-5 hrs. Approximately 1 ml of both uninduced (prior to induction) and induced cultures (after 2 and 4 hrs) were each harvested.
4. **SDS- PAGE:** Harvested cells were boiled for 5 mins in 50 µl of SDS-PAGE sample loading buffer. These samples were then analyzed on an SDS-PAGE to check for the expression of recombinant proteins.

To check the presence of protein in the soluble fraction, cells were lysed under non-denaturing condition. 1 ml of induced cultures were harvested and lysed by sonication in presence of non-denaturing buffer (lacking 8 M urea or 6 M Gdm.HCl). The lysed cells were centrifuged at high speed for separation of the cell debris from the soluble fraction. The cell debris and soluble fraction separately boiled with 10 µl of SDS PAGE sample buffer and analyzed on an SDS-PAGE to check for the presence of protein either in cell debris or in the soluble fraction.

2.2.13 Protein mini preps to check binding from small expression cultures

Binding of 6Xhis tagged proteins was checked using Ni-NTA spin columns using the supplier's (Qiagen) protocol as described below:

1. **Lysis:** 5-10 ml of induced cells were centrifuged and resuspended in 400 µl of Buffer B. The cells were lysed by sonication and the lysate was spun at 15000g for 20-30 minutes.
2. **Loading:** The supernatant was loaded onto a pre-equilibrated (with 600 µl of Buffer B) Ni-NTA spin column, and spun at 700 g for 2 min.
3. **Washing:** The spin column was then washed twice with 600 µl of wash buffer (Buffer C) at 700g for 2 min.
4. **Elution:** The bound protein was eluted with 2 X 200 µl of elution buffer (Buffer E).
5. **SDS-PAGE:** All fractions were collected, boiled with sample buffer and analyzed on SDS-PAGE.

2.2.14 Freeze stock/ glycerol stock (15 %):

For glycerol stocks, 1500 µl of an overnight grown culture (37 °C, 220 rpm for ~16 hours) obtained by inoculating a single colony into sterile LB media supplemented with appropriate antibiotics was mixed with 500 µl of 60 % glycerol (autoclaved), so as to make a final solution of 15 %. The stock was stored at -80 °C.

2.2.15 Electrophoretic mobility shift assay (EMSA)

EMSA was carried out using 0.5% Agarose gels prepared in 1X TAE solution. EtBr was supplemented to the gel for visualization of DNA. Protein samples were mixed with the required amount of DNA and incubated at room temperature for 5 minutes. 1X DNA loading dye was added to the sample and was loaded onto the gel. Gel was run for 30-45 minutes at 60V and imaged using a Bio-rad gel documentation system.

2.2.16 Trypsin digestion

Reagents used:

1. Acetonitrile solution (ACN)
2. 100mM Ammonium bicarbonate (ABC) stock solution: Diluted to 25mM working solution.

3. 30% ACN in 25mM ABC
4. 200mMTributylphosphine (TBP) stock in 1-methyl-2-pyrrolidone: Diluted to 20mM using 25mM ABC
5. 0.5M Iodoacetamide (IAA) stock solution in water: Diluted to 40mM with 25mM ABC
6. 50% ACN
7. Trypsin reaction buffer: 9% ACN in 40mM ABC
8. Trypsin solubilization buffer: 1mMHCl
9. 10µg/µl Trypsin in Trypsin solubilization buffer

In Gel Digestion Protocol

1. Band of interest was cut out from the gel very carefully and chopped into small pieces.
2. Gel pieces were dehydrated with 40-100µl of 1M Acetonitrile (ACN). This step was repeated 3 times.
3. Gel pieces were destained using 40-100µl 30% ACN solution containing 25mM Ammonium bicarbonate (ABC) by incubating the samples on shaking at 30^0C for 30minutes. This step was also repeated 3 times.
4. Destained gel pieces were further dehydrated by incubating with 40µlACN solution for few minutes. After which ACN was removed and samples were allowed to dry.
5. 100µl of 20mM TBP was added to the gel pieces and incubated in thermo mixer at 37^0C for 15mins.
6. Excess of TBP was removed and 100µl of 40mM IAA was added and incubated at 37^0C for 30mins.
7. Supernatant was removed and gel pieces were washed with 200µl of 25mM ABC for 15minutes at 37^0C. This step was repeated twice.
8. Third and final wash was given with 200µl 50% ACN solution in 25mM ABC for 15minutes at 37^0C.
9. Supernatant was removed and gel pieces were allowed to dry.
10. To the dried gel pieces, 50µl of Trypsin reaction buffer were added and 1µl of 10µg/µl of Trypsin prepared in Trypsin solubilization buffer. Samples were incubated at 37^0C in thermo-mixer.

11. After 1hour of incubation, 1µl of Trypsin was added again and incubated at 37^0C overnight.
12. Supernatant was collected after centrifuging samples at 9500rpm for 20sec and used for mass specrometric analysis.

2.3. References

[1] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680-685.

[2] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning : a laboratory manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbour, New York, USA (1989).

[3] J. Sambrook, D. Russel, Molecular cloning : a laboratory manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA (2001).

CHAPTER 3

Examination of over-expressed HV and attempts to purify DNA-free untagged and affinity tagged HV

Contents

3.1. Introduction

The bacterial nucleoid-associated protein, HU, binds to DNA with binding constants varying over a wide range (200 nM to 2500 nM) [1]. Differences in DNA-binding have been observed for: *(a)* HU from different bacterial sources, in respect of binding to the same forms of DNA [2];*(b)* differently associated forms of HU from the same bacterial source (e.g., homo- or hetero-dimers of HU-A and HU-B from *Escherichia coli*), in respect of binding to the same forms of DNA [3]; *(c)* different types of DNA substrates, varying in length, topology or conformation (e.g., relaxed or supercoiled plasmids, linear DNA of different lengths, synthetic four-way junctions/ cruciforms etc.), in respect of binding to the same form of HU from the same source [4,5,6]; as well as *(d)* different relative concentrations of HU and DNA, since HU displays multiple, and cooperative, modes of DNA-binding that result in greater binding as the HU concentrations are increased over DNA concentrations [7].

HU has very recently been demonstrated to be a key component of DNA-containing biofilms formed by uropathogenic *Escherichia coli* [8]. HU from a thermophile, *Thermotoga maritima* and recombinant modified HU from *Bifidobacterium longum* have also used as potential DNA drug-delivery vehicles [9] whereas the exact mechanism of its crossing the eukaryotic cell membrane with bound DNA has not yet been established, it does appear that HU-DNA complexes can enter cells, and also that HU transiently acts to protect DNA from degradation[10].HU has also been associated with lateral gene transfer (as a protein playing a role opposite to that played by DNA-bridging protein H-NS)[11]. Such reports suggest that HU can act as a carrier of DNA in extracellular spaces or, at the very least, that HU can bind to DNA present in extracellular spaces (eDNA).

In this chapter, we have explored the role of HU not as a genome organizer, but as a "carrier" of DNA between the cells and as a "protector" against the action of nucleases which has recently gained implications in human health and health-related applications. The questions that we have focused on are: Can HU act as a "carrier" and "protector" of DNA, playing roles in lateral gene transfer, and biofilm formation? Can molecules of HU exiting dead bacterial cells, carry along with them fragments of bound genomic, or plasmid, DNA (as well as other DNA-binding proteins which are attached to such DNA) into the microbial extracellular environment? Can the DNA bound to HU (and the associated DNA-binding proteins) contaminate the HU

purifications if HU is purified under non-denaturing conditions? What would be the best method to remove the contaminating DNA to yield native, dimeric, DNA-free and DNA-binding competent HU protein for biochemical and biophysical studies? For a better understanding of the above, we have explored many different methods for the purification of HU, to identify and standardize the best and worst method(s) for purification of DNA-free and DNA-binding-competent HU under native conditions. Naturally, such studies would also shed light on the best and worst conditions for the survival of HU-DNA interactions, with implications for lateral DNA transfer.

We started by examining if HU carries along some *E. coli* DNA when purified through non-denaturing affinity purification involving a 6xHis affinity tag from *E. coli* cells over-expressing the protein, in absence of salts, chaotropic agents or detergents specifically present to facilitate any dissociation of DNA bound to HU protein. Following this we explored ten different conditions to purify wild-type, untagged and tagged variants of both HU-A and HU-B and we have characterized the proteins purified in terms of their DNA binding and structural characteristics.

Earlier, HU has been purified by other researchers through the following protocols: (i) IMAC (i.e., immobilized metal affinity chromatography of HU protein fused genetically with a 6xHis tag), followed by ion-exchange chromatography using the strong cation exchanger, SP sepharose[12], (ii) ammonium sulfate fractionation and precipitation of HU, followed by ion-exchange chromatography using the weak cation exchanger, CM sepharose, followed by one additional step of ammonium sulfate precipitation, and finally a hydrophobic interaction chromatography step involving phenyl sepharose[6], (iii) ion-exchange chromatography using the strong cation exchanger SP sepharose to purify HU, followed by a second round of cation exchange chromatography using Mono S resin, finally followed by gel filtration chromatography[13] or (iv) ammonium sulfate fractionation and precipitation, followed by strong cation exchange chromatography using SP sepharose, followed by a further step of heparin sepharose(affinity) chromatography [14], (v) affinity purification including multiple washing steps and using low pH, mild detergent and high salt in buffers [15]. Most of the above methods use multiple steps of purification, suggesting that a single purification step is not sufficient to obtain reasonably pure HU protein. We have used the following conditions to

remove DNA bound to HU protein: *(a)* use of very high temperatures (e.g., 90 °C) prior to purification, to effect precipitation of all other proteins in the bacterial cell lysate (which are prone to thermal aggregation), to enrich the population of HU which is likely to undergo facile thermal unfolding and refolding without any precipitation, since HU is largely helical in structure (and most such proteins generally refold quite well); *(b)* use of non-specific endonucleases to degrade the contaminating DNA co-purifying with HU in bound form; *(c)* use of mild or high concentrations of denaturing/chaotropic agents to unfold the partially or completely to remove the bound DNA, followed by refolding of the protein;*(d)* use of mild concentrations of non-ionic detergents, to break DNA-protein interactions; *(e)* use of high concentrations of salts, so as to increase the ionic strength of the solution to a point at which DNA may be expected to dissociate naturally from HU without, however, necessarily causing any salting out of the HU protein itself. We have purified both HU-A and HU-B using the above mentioned approaches.

3.2. Materials and methods specific to this chapter

3.2.1. Cloning and expression of E. coli HU-A and HU-B

To obtain high yields of purified HU protein, 6xHistidine affinity tagged (both N-terminally tagged and C-terminally tagged) as well as untagged variants of both HU-A and HU-B were cloned and produced. The genes encoding *E. coli* HU-A (*hupA*) and HU-B (*hupB*) were amplified from *E. coli* genomic DNA by *in situ* colony PCR using primers mentioned in Table 1. To introduce a 6xHistag at the C-terminal, *hupA* and *hupB* genes were cloned between NdeI and XhoI restriction sites of the pET23a vector which encodes a 6xHistag at the C-terminus of the multiple cloning site (MCS). PCR cloning was done using primers 1 and 2 for HU-A and primers 3 and 4 for HU-B. To clone HU without the 6xHistag, *hupA* and *hupB* genes were cloned between the NdeI and XhoI restriction sites of pET23a but a stop codon was introduced at the end of the gene using the reverse primer 5 for HU-A and primer 6 for HU-B. The sequence verified pET23a plasmids containing the inserts were transformed into BL21Star(DE3)pLysS cells (Novagen) for protein expression. To introduce a N-terminal 6xHis tag, *hup* genes were cloned into the pQE30 vector (Qiagen) that contains a 6xHistag at the N-terminus of the MCS. The *hupA* gene was cloned between the BamHI and SmaI restriction sites, using primers 7 and 8 and the *hupB* gene was cloned between the BamHI and HindIII restriction sites using primers

9and 10. Sequence verified plasmids containing the inserts were transformed into M15[pREP4] cells (Qiagen) for protein expression.

	Primer	5'→3'
1.	HU-A NdeI Forward	AGCTACTCATATGAACAAGACTCAACTGATTGATG
2.	HU-A XhoI Reverse	GAATACTCTCGAGCTTAACTGCGTCTTTCAGTGC
3.	HU-B NdeI Forward	AGCTACTCATATGAATAAATCTCAATTGATCGACAAG
4.	HU-B XhoI Reverse	GAATACTCTCGAGGTTTACCGCGTCTTTCAGTG
5.	HU-A XhoI UT Reverse	GAATACTCTCGAGTTACTTAACTGCGTCTTTCAGTGC
6.	HU-B XhoI UT Reverse	GAATACTCTCGAGTTAGTTTACCGCGTCTTTCAGTG
7.	HU-A Bam HI Forward	AGCTACTGGATCCATGAACAAGACTCAACTGATTG
8.	HU-A SmaI Reverse	GAATACTCCCGGGTTACTTAACTGCGTCTTTCAATG
9.	HU-B BamHI Forward	AGCTACTGGATCCATGAATAAATCTCAATTGATCG
10.	HU-B HindIII Reverse	GAATACTAAGCTTTTAGTTTACCGCGTCTTTCAGT

Table 1: List of primers used for cloning of affinity-tagged and untagged variants of HU-A and HU-B.

3.2.2. Purification of 'untagged' naturally-expressed HU by heat: Wild-type *E. coli* cells were grown up to $O.D._{600}$ of 1; cells were harvested and re-suspended in 20mM phosphate buffer and sonication was done to lyse the cells. Supernatant after sonication was heated at 90^0C for 15, 30, 45 and 60 minutes, respectively.

3.2.3. Purification of 'untagged' naturally-expressed HU by ion exchange chromatography: As HU is a highly basic protein, cation-exchange chromatography was used for its purification. The supernatant after sonication of *E. coli* cells in 20 mM phosphate buffer was loaded on a cation exchange chromatography column (GE Healthcare Mono S; bed volume 1ml). HU protein was eluted by increasing the concentration of NaCl gradient from 0 mM to 500 mM over 40 column volumes (CV) using GE-Akta purifier system. Protein elution was monitored by recording 215 nm (because HU lacks Trp residues and displays poor absorbance at 280 nm) and also at 280 nm, to identify fractions with poor 280 nm absorbance and high 215 nm absorbance. Fractions were collected and electrophoretically analyzed on a 15% acrylamide SDS-PAGE to test purity.

3.2.4. Purification of 'untagged' over-expressed HU-A and HU-B by ion-exchange chromatography: Un-tagged HU-A and HU-B were over expressed by IPTG induction form BL21Star(DE3)pLysS cells. Induced cells were harvested and lysed by sonication in phosphate buffer saline (PBS). Soluble HU was recovered in the lysate after centrifugation. The lysate was loaded onto a manually-packed column containing SP sephadex (strong cation exchanger) resin (Bio-rad). HU protein was eluted by applying NaCl gradient from 150 mM (in PBS) to 1 M using GE-AKTA purifier system. As described previously, protein elution was monitored by recording 215 nm and 280 nm absorbance and the fractions were collected and analyzed by 15% acrylamide SDS-PAGE.

3.2.5. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC: C-terminal and N-terminal His-tagged HU-A and HU-B were expressed from BL21star(DE3)pLyS and M15[pREP4] cells, respectively, after induction with IPTG. Induced cells were harvested and purification was performed using Ni-NTA affinity IMAC chromatography. Cells (5ml/g of harvested cells) were resuspended in lysis buffer containing 10 mM imidazole and 1 mg/ml lysozyme in PBS and incubated on ice for 30 minutes. Cells were then lysed by sonication and the soluble protein was recovered in the supernatant (lysate) after centrifugation and loaded on the IMAC column pre-equilibrated with PBS containing 10 mM imidazole, followed by washing with PBS containing 20mM imidazole. HU protein was eluted using PBS containing 250 mM imidazole. To remove imidazole, over-night dialysis was performed against PBS buffer. For further purification, gel filtration chromatography was performed, using a 24 ml Superdex-75 (GE) column.

3.2.6. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC in presence of nucleases: To enzymatically degrade DNA bound to HU, a non-specific nuclease, Benzonase (Merck, Cat. no.1.01653.0001) was used. Approximately, 1 unit/ ml of Benzonase was added in the lysis buffer (PBS + 10 mM imidazole) and incubated at 37^0C for 1 hour to allow Benzonase to act on and cleave DNA into smaller fragments following which affinity purification was done.

3.2.7. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC under conditions of partial unfolding (0.5M urea): Low concentrations of urea were used to partially unfold the protein, to release it from DNA. Cells were lysed by sonication in PBS containing 0.5M urea (pH 8), followed by incubation at 37^0C for 5-6 hours to allow unfolding to happen, following which

the lysate was loaded onto a pre-equilibrated Ni-NTA IMAC column and washed extensively to remove non-specifically bound proteins by decreasing the pH to 6.3. This was followed by elution of bound HU at pH 4.5. Purified protein was extensively dialyzed against PBS to allow refolding of protein subjected to partial unfolding to release DNA.

3.2.8. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC under conditions of complete unfolding (6 M urea): Cells were lysed by over-night incubation in 6 M urea (pH 8) to completely denature the protein to facilitate dissociation of HU from DNA. The lysate after centrifugation was loaded onto a pre-equilibrated Ni-NTA IMAC column. The column was washed by lowering the pH to 6.3 and HU protein was eluted by further lowering the pH to 4.5. Protein was refolded by extensively dialysis against PBS.

3.2.9. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC in the presence of detergent (1.0 % Triton X-100): The mild neutral detergent Triton X-100 was added to remove DNA from HU during cell lysis. Approximately 1 % Triton X-100 was added to PBS containing 10 mM imidazole during the lysis step, followed by washing of Ni-NTA bound protein with 20 mM imidazole and elution was done with 250 mM imidazole.

*3.2.10. Purification of '6xHis-tagged' over-expressed HU-A and HU-B by IMAC in the presence of salt (1 M NaCl):*High concentration (1 M) of the salt, NaCl, was used to disrupt DNA-HU interactions which are primarily ionic, through the action of the counter-ions produced through NaCl dissociation. NaCl was added to the lysis buffer to achieve a concentration of 1 M. Following loading on protein on the Ni-NTA IMAC column, washing was done with 1 M NaCl and 20 mM imidazole. The HU protein was then eluted using buffer containing 250 mM imidazole and 1 M NaCl.

3.3.11. SDS-PAGE, agarose DNA electrophoresis and Acid Urea Triton (AUT) PAGE: Proteins were separated on 15% SDS-PAGE using standard methods. The presence of DNA in purified protein samples was visualized using 1% agarose DNA gel electrophoresis, using gels and conditions which are also now standard methods, requiring no introduction or reference. Proteins co-eluting with HU were analyzed for the presence of basic proteins by a somewhat uncommon technique called Acid-Urea-Triton polyacrylamide gel electrophoresis (AUT-PAGE) on a 7.5% polyacrylamide gel containing 25% acetic acid , 2.5 M urea and 1% Triton X-100 [16]. Protein samples were mixed with 2X-loading dye containing acetic acid, urea, triton X-100 and

phenolphthalein. The gel was run in reverse polarity mode, from positive to negative electrode for 5-6 hours.

3.2.12. Mass spectrometric analysis: Mass spectrometric analysis of the co-eluting non-specific bands was performed to identity proteins contaminating the HU purification. Proteins were separated on SDS-PAGE, and gel pieces corresponding to individual bands, or groups of non-resolved bands, were excised and subjected to in-gel tryptic digestion using mass spectrometry-grade trypsin. Digested peptides obtained were pooled and analyzed using an LC-coupled Waters Synapt G2-S HDMS ESI-Q-TOF mass spectrometer. Peptides were resolved using a C-18 column (Waters) prior to mass spectrometry.

3.3. Results

3.3.1. Purification of naturally occurring HU

In *E. coli* cells, HU (i.e., HU-A plus HU-B)is expressed at concentrations as high as 10 mM[2]; however, the actual concentrations of both HU-A and HU-B vary during different phases of bacterial growth. The HU-A homodimer dominates during the log phase, followed by the accumulation of HU-AB heterodimers during the late log phase and finally, HU-B homodimer during the stationary phase. We first attempted to purify native HU from wild-type *E. coli* cells without over-expression of HU. The aim was to first purify both HU-A and HU-B together in all of the above-described dimeric forms. We used following two methods:

3.3.1.1. Enrichment by heat induced coagulation of other proteins: Most *E. coli* proteins precipitate when the cell lysate is heated at high temperatures for several tens of minutes [17,18]. HU is a small helical protein and it has been reported that HU can bind to DNA even after it has been exposed to 100 ^{0}C [19,20,21]. This suggests that the protein should be able to undergo refolding after thermal denaturation (without aggregating) during heating, and cooling. Thus, we heated the *E. coli* lysate at 90^{0}C for 15, 30, 45 and 60 minutes, respectively to precipitate the bulk of the cytoplasmic populations of all other *E. coli* proteins. After heating and protein precipitation, the pellet obtained and the supernatant were loaded on SDS-PAGE as shown in Figure 1. From the figure, it is evident that the supernatant was enriched with protein of approximately 10 kDa size (likely to be HU) but the yields were very low.

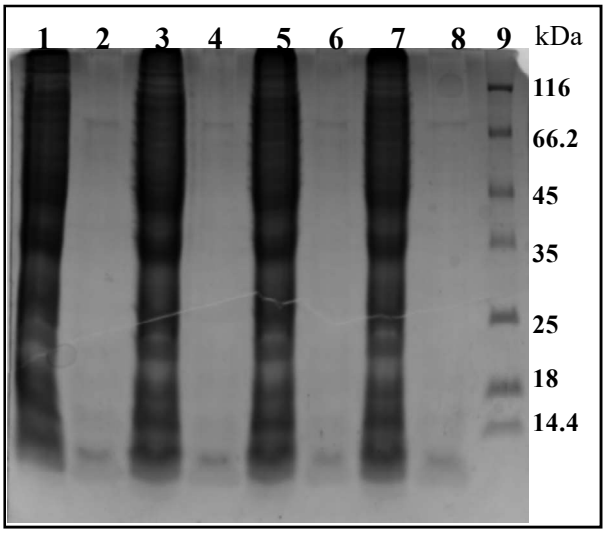


Figure 1: Pellet and supernatant of *E. coli* cell lysate heated at 90 ^{0}C for different time intervals. Lane 1- pellet after 15 minutes heating, Lane 2- supernatant after 15 minutes heating, Lane 3- pellet after 30 minutes heating, Lane 4- supernatant after 30 minutes heating, Lane 5- pellet after 45 minutes heating, Lane 6- supernatant after 45 minutes heating, Lane 7- pellet after 60 minutes heating and Lane 8- supernatant after 60 minutes heating, Lane 9- protein molecular weight marker.

3.3.1.2. Ion exchange chromatography of whole cell lysate: Most *E. coli* proteins have isoelectric points (pIs) that lie in the near-neutral or acidic range, with pI values less than 7 [22].In contrast, as expected (because it is a DNA-binding protein), HU is basic in nature and has high pI values of 9.57 for HU-A, and 9.69 for HU-B, respectively. The positive charges on HU-A and HU-B at pH 7 are 3.86 and 3.77, respectively, as predicted by Vector NTI software. Cation exchange chromatography is widely used for purification of basic DNA binding proteins, and also for HU purification as already mentioned in the introduction. Thus, a strong cation exchanger, Mono S (GE), was next used to purify the whole HU population from cell lysates as shown in Figure 2A. Both HU-A and HU-B have no tryptophan or tyrosine residues; hence, they absorb no significant light of 280 nm wavelength, but they both show absorbance at 215 nm owing to the absorbance of peptide bonds. As Figure 2A shows, HU elution fractions displaying absorbance at 215 nm, but not at 280 nm, were selected, concentrated using 3.0 kDa cut-off Microcon filters (Millipore) and run on SDS-PAGE, as shown in Figure 2B. HU could thus be purified, as seen in the lane 5 corresponding to fraction 10, but the level of purity obtained through single step of ion-exchange purification was not sufficient for further experiments, given the high level of contaminant proteins seen, and the fact that the yields were much lower than could be expected from HU over-expressing cultures. Thus we decided to work next with HU over-expressing cultures.

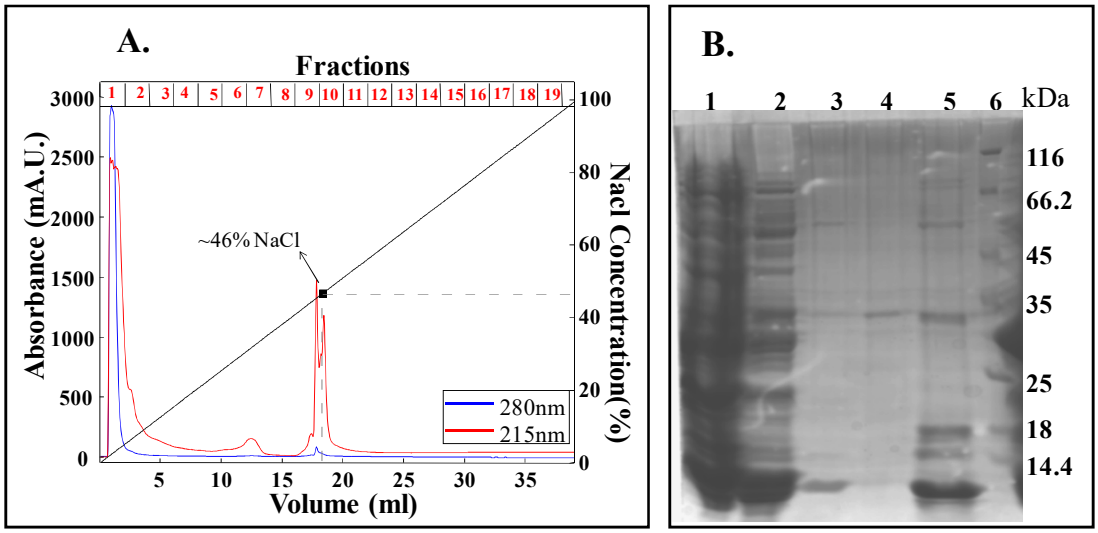


Figure 2: Panel A: Chromatogram for ion exchange chromatography from *E. coli* cell lysate showing elution profile of cell lysate monitored at 280 nm (blue) and 215 nm (red), NaCl gradient is represented as black line corresponding to left Y-axis and fractions are labeled on top. ***Panel B:*** Fractions from cation exchange chromatography on SDS-PAGE, fractions 1,2,7,9 and 10 in lanes 1-5 and protein molecular weight marker in lane 6, respectively.

3.3.2. Purification of over-expressed HU-A and HU-B lacking any affinity tag

Large quantities of protein are required for extensive characterization, and this requirement is usually achieved through recombinant over-expression of any protein. Therefore, to purify HU-A and HU-B separately and with better yields, cloning was performed to over-express the proteins, initially without any affinity tags, to try and improve the purification and yield through ion-exchange chromatography.

3.3.2.1. Purification of untagged HU-A and HU-B by ion-exchange chromatography: BL21Star(DE3)pLysS cells engineered to over-express untagged HU-A and HU-B were grown, harvested, lysed through sonication and proteins were recovered in the soluble cytoplasmic fraction. The proteins were purified using SP Sephadex cation exchange resin, using a salt gradient, and found to elute at ~400 mM NaCl. The proteins thus obtained were checked for purity on SDS-PAGE. Lane 2 in Figure 3A shows the purity obtained for HU-A, and lane 4 in Figure 3B shows the purity obtained for HU-B. A visual comparison of these ion-exchange purifications with those of non-overexpressed HU (in Figure 2) reveals a distinct improvement in the apparent purity, which can probably be ascribed to the fact that there is more HU protein available in the cytoplasmic fraction after over-expression, to undergo binding to the ion

exchange resin. We conclude that the consequent competition reduces the binding of other proteins which are not over-expressed, and this improves the overall purity obtained.

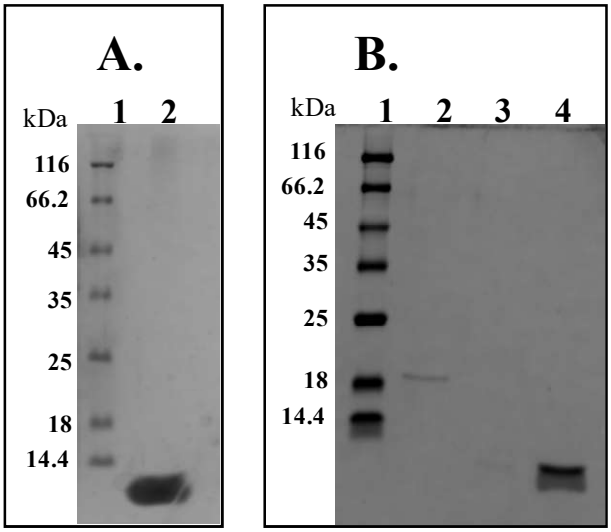


Figure 3: SDS-PAGE showing cation exchange purification profiles of untagged HU proteins. ***Panel A***: purified HU-A protein in lane 2 ***Panel B:*** purified HU-B protein in lane 4.

3.3.3. Purification of over-expressed HU-A and HU-B bearing 6xHis affinity tags

3.3.3.1. Purification of C-terminally His-tagged HU-A and HU-B by affinity chromatography: HU-A and HU-B incorporating 6xHis affinity tags at their C-termini were produced and recovered in soluble form after lysis of cells engineered to over-express these proteins from the pET23a vector, after induction with IPTG. IMAC purification was performed under non-denaturing conditions according to standard protocols using Ni-NTA agarose resin. Both HU forms bound to Ni-NTA agarose beads in the presence of 10 mM imidazole, and were eluted by increasing the imidazole concentration to 250mM. As can be seen in lane 5 of Figure 4A (for HU-A) and lane 5 of Figure 4B (for HU-B), there are several non-specific proteins that are co-purified with both HU-A and HU-B proteins. However, significantly, the protein underwent precipitation after incubation, or storage for some hours. SDS-PAGE studies indicated that the precipitated protein fraction contained the HU-A and HU-B proteins (data not shown), indicating that the presence of the 6xHis affinity tag at the C-terminal probably affected the structural stabilities of these proteins, and caused them to aggregate with time.

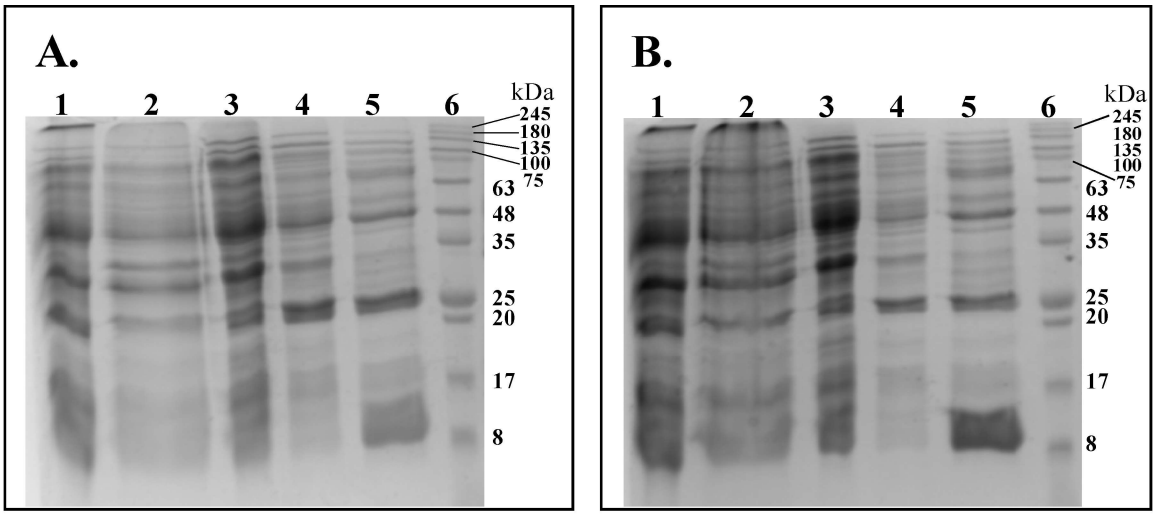


Figure 4: Ni-NTA purification profile of C-Tagged HU-A ***Panel A*** and HU-B ***Panel B.*** In both the panels, *Lane 1* corresponds to pellet after sonication, *Lane 2* corresponds to supernatant after sonication i.e. lysate, *Lane 3* corresponds to flow through from Ni-NTA column, *Lane 4* corresponds to wash, *Lane 5* corresponds to the elution fraction and *Lane 6* corresponds to the protein molecular weight marker.

3.3.3.2. Purification of N-terminally His-tagged HU-A and HU-B by affinity chromatography: N-terminally 6xHis-tagged HU-A and HU-B cloned in the pQE30 vector were over-expressed in M15 cells. Affinity-tagged HU forms (hereafter referred to as N-tagged HU) were recovered in the cytoplasmic fraction after cell lysis, and purified using procedures similar to those used for purifying the C-terminally tagged HU, but these purified fractions too contained many non-specific proteins as observed with the C-terminally 6xHis tagged HU protein.

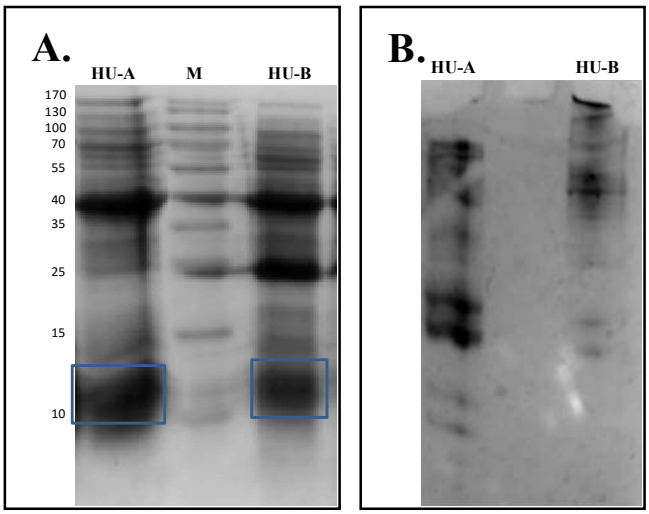


Figure 5: Panel A: SDS-PAGE profiles of HU-A and HU-B nominally-purified by IMAC chromatography, shown along with molecular weight markers. ***Panel B:*** AUT-PAGE profiles of HU-A and HU-B. All lanes are marked directly on the figure.

3.3.3.3. Comparison of N-terminally and C-terminally 6xhis tagged HU-A and HU-B: To summarize, (1) both the N- and C-tagged forms of HU happened to be purified along with many co-purifying contaminant proteins, rather than as pure protein bands; although the literature does not explicitly mention this, this tendency for contamination probably explains why other researchers have combined affinity purification with other methods of purification [12], instead of using simple single-step affinity purification. (2) the purified C-tagged protein underwent aggregation and precipitation, but the N-tagged protein was highly soluble and showed no signs of precipitation. It is difficult to understand the differences in stability between the N-tagged and C-tagged forms from a structural viewpoint. In the crystal structure of the HU-AB hetero-dimer, the C-termini of both HU-A and HU-B appears to be exposed to the solvent and distal to any inter-subunit interface [23].

3.3.4. Presence of protein and DNA contaminants in affinity-purified HU preparations and identification of the contaminants

The purity of purified HU was determined by subjecting it to SDS-PAGE analysis. From affinity-purified samples of HU-A, and HU-B proteins (Figure 4A, 4B and 5A) it is evident that the HU-A and HU-B preparations are highly contaminated by numerous other proteins.

3.3.4.1. Confirmation that many of the contaminating proteins are basic in nature: The affinity purified HU-A and HU-B preparations were electrophoresed on an Acid-Urea-Triton poly acrylamide gel (AUT-PAGE). AUT PAGE is used to resolve basic proteins in native form, based on the presence of net positive charge, and is often used to resolve DNA-binding proteins [24]. As can be seen in Figure 5B, many proteins that contaminate the HU-A and HU-B preparations also migrate into the AUT gel. This indicates that these could be other DNA binding proteins as only proteins carrying net positive charge should enter the AUT-PAGE and get resolved. Unless one postulates either (1) that these other proteins are bound to DNA which is bound to HU, or (2) that these other proteins are not directly bound to DNA, but are instead bound to other proteins which are bound to the DNA that is bound to HU, it is not possible to explain why affinity-purified HU preparations should be contaminated by such other (especially basic) proteins.

3.3.4.2. Confirmation of the presence of DNA in purified HU: The affinity-purified, but impure, HU-A and HU-B samples were further subjected to gel filtration for confirmation of their

dimeric status, since HU binds to DNA as a dimer. Following this further step of gel filtration, performed in an analytical mode to examine quaternary structural status, rather than to attempt any further purification, the HU protein samples were subjected to DNA agarose gel electrophoresis on a 1 % agarose gel containing ethidium bromide, to investigate the presence of DNA in the HU preparation. In Figure 6, DNA can be clearly seen to be present in the purified HU-A protein in lane 1 of the gel, in the proximity of the well in which the HU protein was loaded. Further, in lane 1, in addition to the large DNA stuck in the well, one can also see small fragmented pieces of DNA (of sizes < 100bp) towards the bottom of the lane. This incontrovertibly establishes that HU-A which purified through affinity chromatography from HU-A over-expressing lysed *E. coli* cells does indeed contain bound DNA. The same result holds true for HU-B (data not shown).

3.3.4.3. Proteomic analysis and identification of protein contaminants in HU preparations: Mass spectrometric analysis was performed to identify the contaminating protein bands, after excising all contaminating bands and performing in-gel tryptic digestion. Bands 1-4 for HU-A and 5-10 for HU-B as marked in Figure 7 were excised, digested and were pooled and used for further analysis. Following MS-MS data collection on a WATERS ESI-Q-TOF G2-S HDMS mass spectrometer, proteomic analysis was carried out using the protein linked global server (PLGS) software to analyze data with 10 ppm tolerance. The results are presented in Table 2.

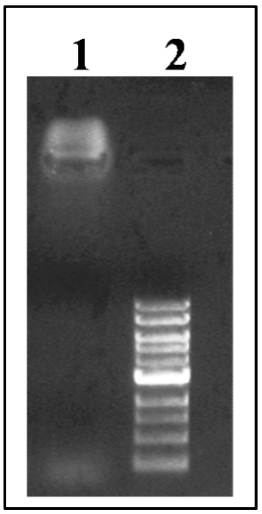


***Figure 6*:** DNA agarose gel profile of purified HU protein containing contaminating DNA (lane 1), along with a 100- bp DNA molecular weight marker ladder, with bands every 100 bp (lane 2).

S. no.	Accession no.	Description	mW (Da)	pI	PLGS Score	Abundance rank [7]	Ref.
1	P02342	DNA-binding protein HU-alpha	9529	10.1235	13184.56	36	[8]
2	P02341	DNA-binding protein HU-beta	9219	10.2393	3659.856	26	[8]
3	O24728	Host factor-I protein	6898	10.1045	1597.869		[8]
4	P06987	Histidine biosynthesis bifunctional protein	40252	5.7188	3047.892		
5	P06975	Ferric uptake regulation protein	16784	5.6528	1763.754	546	[9]
6	P77581	Succinylornithine transaminase	43638	5.8931	2966.061		
7	P00574	DNA-directed RNA polymerase alpha chain	36489	4.7813	3199.135	115	[10]
8	P37192	Tagatose biophospahte aldolase	30792	5.9	793.8		
9	P03023	Lactose operon repressor	38540	6.4	642.7		[11]
10	P30856	FKBP-type peptidyl-prolyl cis-trans isomerase	20839	4.6934	694.0863	126	
11	P45473	Hypothetical acetyltransferase	18522	4.3726	405.2241		
12	P02934	Outer membrane protein A precursor	37177	5.9678	325.0825	61	
13	P04790	Triosephosphate isomerase	26954	5.5679	255.8274	62	
14	P32665	Glycerol dehydrogenase	38687	4.604	1091.83		
16	P06996	Outer membrane protein C	4.343	4.4	1652	159	
17	P17169	Glucosamine--fructose-6-phosphate aminotransferase	66721	5.4785	6385.536	286	
18	P06139	60 kDa chaperonin	57161	4.6494	2968.975	40	
20	P00575	DNA-directed RNA polymerase beta chain	150537	4.9629	1300.699	132	[12]
21	P02996	Elongation factor G	77401	5.0713	1105.982	44	
22	P04475	Chaperone protein dnaK	68941	4.6406	979.9931	57	[13]
24	P21599	Pyruvate kinase II	51193	6.2256	623.1518	483	
25	P00391	Dihydrolipoamide dehydrogenase	50525	5.7334	618.5297	67	
26	P09373	Formate acetyltransferase	85171	5.6016	423.8096	90	
27	P08324	Enolase	45495	5.1636	392.119	43	
28	P00483	Chloramphenicol acetyltransferase	25646	5.9063	14351.06		
29	P03020	Catabolite gene activator	23625	8.436	3704.857	238	[14]
30	P08330	Ribose-phosphate pyrophosphokinase	34065	5.0625	328.7793	180	
31	P30177	Hypothetical protein ybiB.	35026	6.3999	63.9118	699	
32	P09453	Ribosomal protein alanine acetyl tranferase	16599	4.4	1217.4		
33	P17963	ADP L gycero D manno heptose 6 epimerase	34871	4.6	4873.1	213	
34	P76143	Putative aldolase yneB	31872	6.0879	335.0959		

Table 2: List of proteins (including HU-A and HU-B) which appear in samples of affinity purified HU-A or HU-B. For each protein mentioned, its accession number, descriptive name, molecular weight, isoelectric point, PLGS score (indicating the confidence with which the proteomic identification was made), and abundance rank in *E. coli* (indicating the protein's position amongst a list of proteins of determined abundance [7]) are mentioned; wherever available, a reference to the protein's identification as a DNA-associated protein is also included.

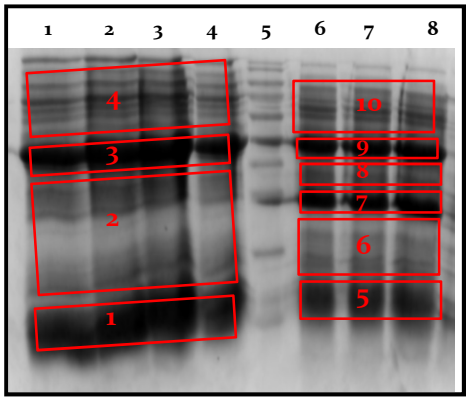


Figure 7: SDS-PAGE showing the bands used for proteomic analysis. Lane 1-4 contain HU-A and lane 6-8 contain HU-B elution fractions.

From Table 2, it is evident that quite a few of the identified proteins are known DNA binding proteins (highlighted in grey). In addition, there are numerous other proteins seen which have not been previously reported to have any DNA-binding potential or characteristics. Some of these proteins are basic in nature, while others are acidic. Many of these additional proteins are present in low abundance in *E. coli*, and the abundance ranks and DNA-binding characteristics, wherever applicable [7-14] are mentioned in Table 2.

3.3.5. Further purification of N-tagged HU by removing DNA and bound non-specific proteins

Since the purified HU protein appears to co-purify with many contaminating proteins, we examined the identities of these proteins and found that quite a few of them are DNA-binding proteins; we also confirmed that the purified HU co-purifies with bound DNA which can be visualized on agarose gels during gel electrophoresis. Thus, we explored various approaches to dissociate HU from DNA while retaining DNA-binding capability in the purified HU. All experiments described below were performed with cells over-expressing the N-tagged forms of HU-A and HU-B.

3.3.5.1. Removing DNA through degradation (by the nuclease, Benzonase): Non-specific nucleases are sometimes used during protein purification, to avoid DNA contamination. These nucleases, e.g., Benzonase act on the available DNA and fragment it into small pieces. In the present instance, assuming that the contaminating non-specifically co-purified proteins are bound to the DNA which is bound to HU, it may be assumed that efficient Benzonase-based degradation of DNA would simultaneously eliminate co-purification with HU of both DNA and

other proteins bound to such DNA. To examine this possibility, cell lysates were treated with Benzonase, prior to standard IMAC purification of HU-A and HU-B. Figure 8 shows that Benzonase treatment did not eliminate the co-purified non-specific contaminant bands. The probable reason is that that Benzonase could not access DNA wrapped up with HU protein.

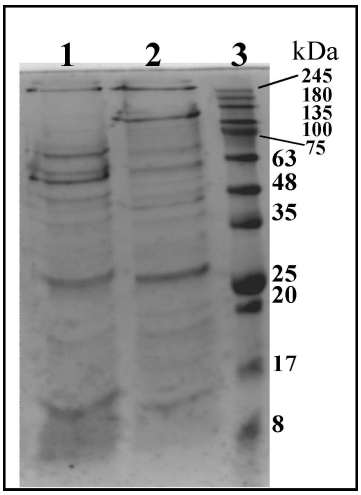


Figure 8: SDS-PAGE showing protein elutions for HU-A (lane 1) and HU-B (lane 2) after purification using Benzonase, protein marker is loaded in lane 3.

3.3.5.2. Removing DNA through 'partial unfolding' of HU: Based on their structural stability, different regions of a protein can behave differentially to the presence of chaotropic agents. Very low concentrations of a denaturing agent such as urea could potentially cause dissociation of DNA from HU without fully unfolding HU. We lysed cells in low concentrations of urea (0.5 M, 1.0 M, 1.5 M and 2.0 M) with the hope that this would remove bound DNA from HU-A an HU-B, based on the assumption that low concentrations of urea would be insufficient to unfold HU protein but sufficient to disrupt DNA-protein interactions, especially in the DNA-associating interface region of HU which is known to be unstructured in the absence of DNA. The protein eluted after Ni-NTA chromatography was then allowed to be rid of the urea through dialysis and subjected to SDS-PAGE analysis.

We found that urea concentration as low as 0.5 M was sufficient to remove all other co-eluting proteins, allowing us to recover pure HU-A and HU-B protein bands on SDS-PAGE (Figure 9A). The HU-A and HU-B proteins were then loaded on an agarose gel containing the dye ethidium bromide (EtBr) which stains DNA and is detectable through fluorescence. The DNA gel shown in Figure 8B further demonstrates that both HU-A (lane 2) and HU-B (lane 4)

proteins are free from any DNA contamination, unlike protein purified in absence of urea as seen in Figure 6.

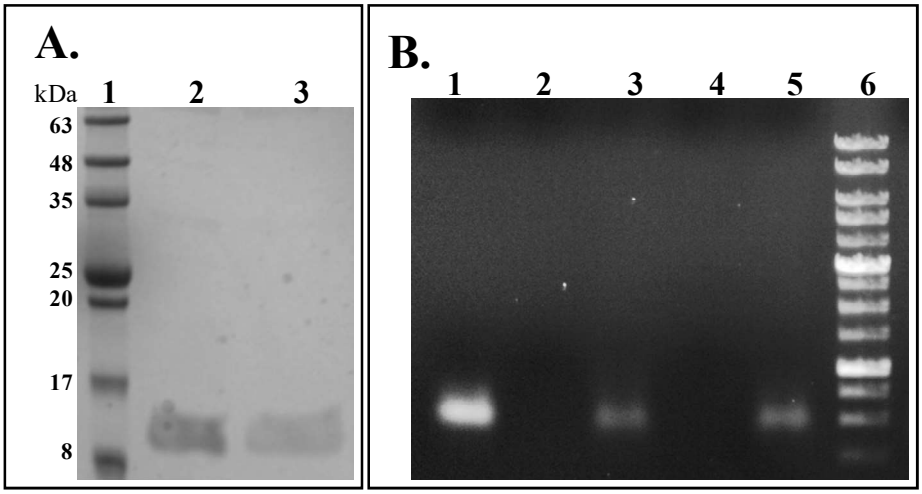


Figure 9: Panel A:SDS-PAGE showing pure protein in elution fractions of HU-A and HU-B in lane 1 and lane 2, respectively, after Ni-NTA chromatography in mild denaturant. ***Panel B*** shows that refolded HU-A protein loaded in lane 2 and HU-B protein in lane 4 are free of DNA contamination and cannot bind to DNA and are not able to show shift in DNA band: HU-A plus DNA in lane 3 and HU-B plus DNA in lane 5 and control containing only DNA in lane 1, 100bp marker in lane 6.

Disappointingly, the dialysed HU proteins were incapable of binding to DNA as seen in lanes 3 and 5 of Figure 9B, as no electrophoretic mobility shift is seen compared to the control synthetic 4-way junction DNA band. To summarize, use of even 0.5 M urea was sufficient to remove DNA (and bound proteins) from HU and to facilitate purification of contaminant-free protein. However, upon refolding, the protein could not refold and was not DNA-binding competent.

3.3.5.3. Removing DNA through 'complete unfolding' of HU: HU-A and HU-B were subjected to complete unfolding (denaturation) in 6M Urea, by lysing bacterial cells in 6 M urea. The idea was to attempt refolding of HU protein to DNA-binding-competent structure from a completely unfolded state, since it was evident that the protein could not return to a DNA-binding-competent state from a partially-structurally-perturbed state created through exposure to 0.5 M urea. Therefore, following IMAC purification, HU-A and HU-B were refolded by dialyzing out the 6 M urea, and the results were examined by SDS-PAGE and agarose gel electrophoresis.

Upon complete unfolding, as expected, all the co-purifying contaminant proteins were eliminated, as seen in Figure 10A, for HU-A (lane 2) and HU-B (lane 3). Likewise, as seen in Figure 10B, the refolded protein also does not contain any contaminating DNA (lane 2 for HU-

A, and lane 4 for HU-B), establishing that exposure to 6 M urea completely removes DNA and protein contaminants from HU. Upon addition of the refolded protein to synthetic 4-way junction DNA, a significant mobility shift was seen in the form of a smear in lane 3 (HU-A) and lane 5 (HU-B), respectively, corresponding to HU-DNA complexes of different sizes. The refolded, DNA-binding-competent HU proteins were then loaded onto a 24 ml Superdex 75 column (GE healthcare) gel filtration column to examine whether they were dimeric, after refolding. Disappointingly, both HU-A and HU-B proteins eluted at the void volume (i.e., 8 ml) of the gel filtration column, as seen in Figure 10C, indicating that the protein refolded from 6 M urea had associated into higher oligomers of HU which are DNA-binding-competent.

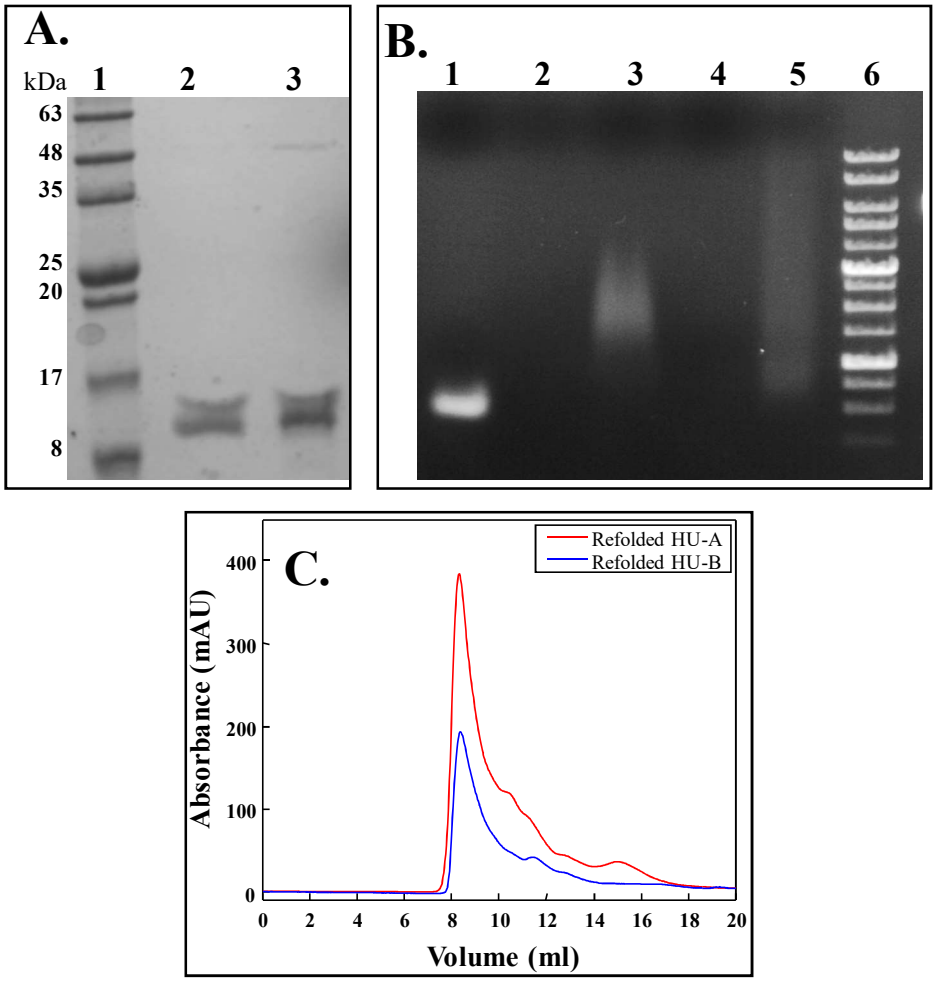


Figure 10: Panel A: SDS-PAGE of urea purified HU-A and HU-B in lane 1 and lane 2, respectively. ***Panel B:*** EMSA gel showing that refolded HU-A and HU-B protein are free of DNA contamination (lane 2 and lane 4, respectively). On addition of protein to DNA, a clear shift in mobility of DNA is observed in lane 3 for HU-A and lane 5 for HU-B. Lane 6 contains 100 bp ladder. ***Panel C:*** Gel filtration profiles of HU-A (red) and HU-B (blue) on 24 ml Superdex 75 column.

3.3.5.4. Removing DNA through the use of a detergent (Triton X-100): Triton X-100 is a nonionic polyoxyethylene detergent. 1% Triton X-100 was used during cell lysis to disrupt protein-DNA interactions, aiding the purification of HU protein as described earlier for the purification of HU protein from sources other than *E. coli*, e.g., *Thermotoga maritima* [25] and *Thermoplasma volcanium* [26].

Use of 1% triton X-100 removed all non-specific protein bands and contaminating DNA from HU, as can be seen in Figure 11A, and Figure11B, respectively. Furthermore, HU-A and HU-B proteins also retained their property of being able to bind to DNA, as can be seen in Figure 11B, in that they were able to induce gel band shifts for the synthetic 4-way junction DNA. Encouragingly, upon loading on a 24 ml Superdex 75 column, HU-A and HU-B proteins eluted at ~12ml, which corresponds to the elution expected for HU dimer.

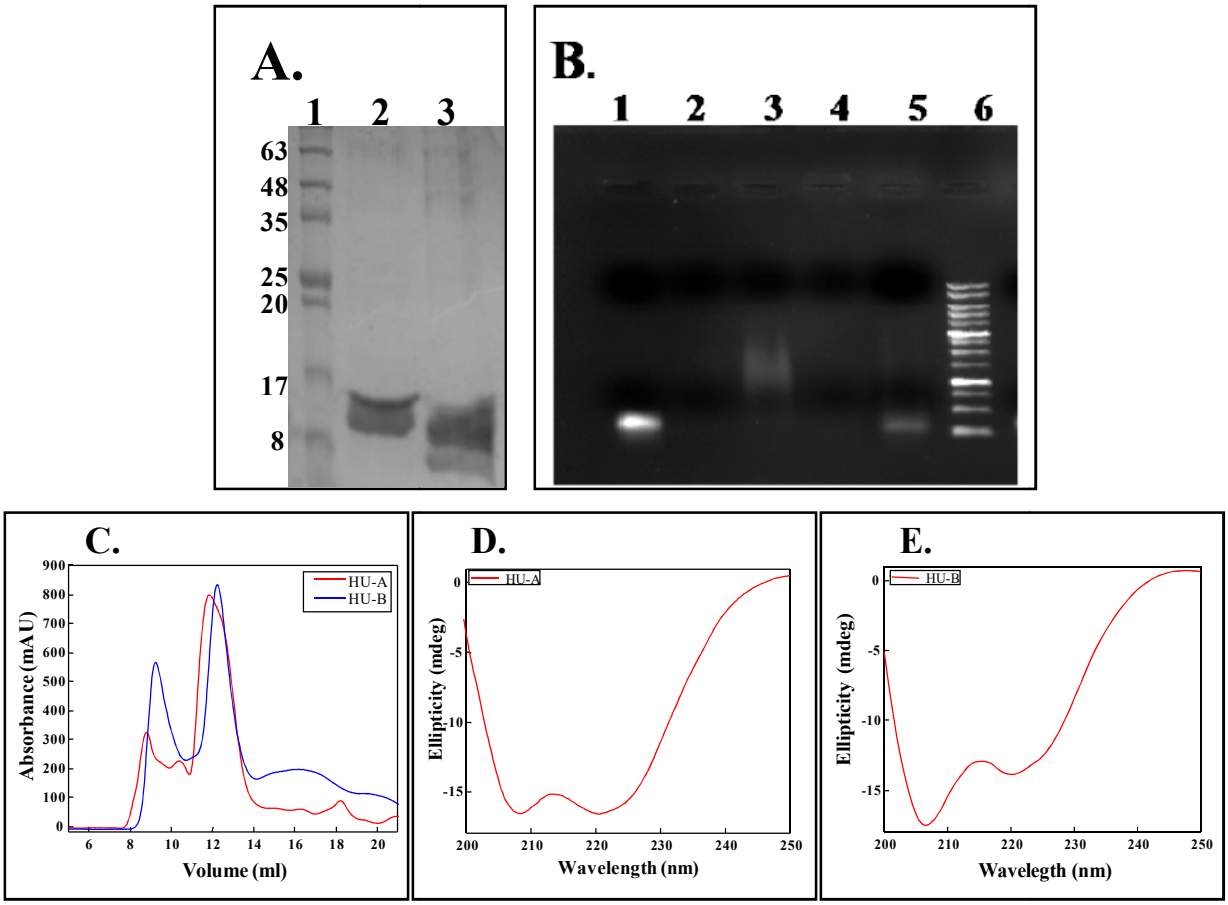


Figure 11: Panel A: SDS-PAGE of HU-A and HU-B purified in presence of Triton-X 100 in lane 1 and lane 2, respectively. ***Panel B:*** DNA free purified protein HU-A (lane 2) and HU-B (lane 4), lane 3 and lane 5 show DNA binding of HU-A and HU-B. ***Panel C:*** Gel filtration chromatograms of HU-A (red) and HU-B (blue) purified using Triton X-100 ***Panel D*** and ***Panel E*** show far UV CD spectra for HU-A and HU-B, respectively.

But disappointingly, as seen in Figure 11C, in addition to the dimeric population, there was also a significant population of higher molecular weight species (soluble aggregates) in both HU-A and HU-B preparations (chromatograms in black and red, respectively), although the major population eluted at an elution volume corresponding to dimeric HU. To estimate whether this protein had folded to helical structure, we also collected the far-UV CD spectra of HU-A (Figure 11D) and HU-B (Figure 11E). Satisfyingly, both proteins displayed the characteristic features of helical proteins, in the form of the prominent negative (ellipticity) bands at 208 and 222 nm. To summarize this part, the detergent (1 % Triton X-100) helped to remove contaminating DNA and proteins, and also to obtain DNA-binding-competent HU as well as largely dimeric HU, but with a significant fraction of soluble aggregates of HU in the purified protein population.

3.3.5.5. Removing DNA through use of salt (NaCl): The interactions stabilizing DNA-protein complexes are mainly ionic interactions, involving positively charged side chains on DNA-binding proteins and negative charges associated with the phosphate groups on DNA. Addition of salt offers 'counter-ions' to satisfy the phosphate groups on DNA by proving Na^+ counter-ions and the positively-charged protein with Cl^- counter-ions, and tease the protein away from the DNA. In a sense, this like dissociating bound protein from an ion-exchange column by using a gradient of salt concentrations, with the dissociated salt ions serving to separate the protein from the resin.

Whereas most purification methods for HU thus far have used ion-exchange chromatographic methods, apparently the real reason for performing ion-exchange was to differentially elute basic proteins from an ion-exchange resin, while probably also (unwittingly) managing to dissociate some DNA away from HU protein through the action of counter-ions (although this was neither the stated intent, nor reported to occur).

To explore these possibilities, 1M sodium chloride was used in the lysis buffer to try and dissociate DNA from bound HU (and other) proteins without effecting any unfolding. This was followed by washing with 1 M (or 2 M) NaCl for several bed volumes after the binding of the HU protein to the Ni-NTA column. Encouragingly, as can be seen in Figure 12A and Figure 12B,the use of 1 M NaCl during bacterial cell lysis and IMAC purification resulted in the complete removal of all contaminating proteins as well as all bound DNA. Also, the protein purified by this method was DNA-binding competent, as assayed through the electrophoretic

mobility shift assay involving change in the mobility of the band corresponding to the synthetic 4-way junction DNA (Figure 12B). Perhaps most encouragingly of all, HU-A could be seen to have been purified entirely as a dimer (Figure 12C). HU-B showed some higher order oligomers but these were smaller than the soluble aggregates observed until now with other methods. This is to be expected, since it has been reported that HU-B does form tetramers and octamers in solution [23]. Both proteins are well structured, displaying CD spectra characteristic of proteins of a largely helical nature (Figure 12D and Figure 12E).

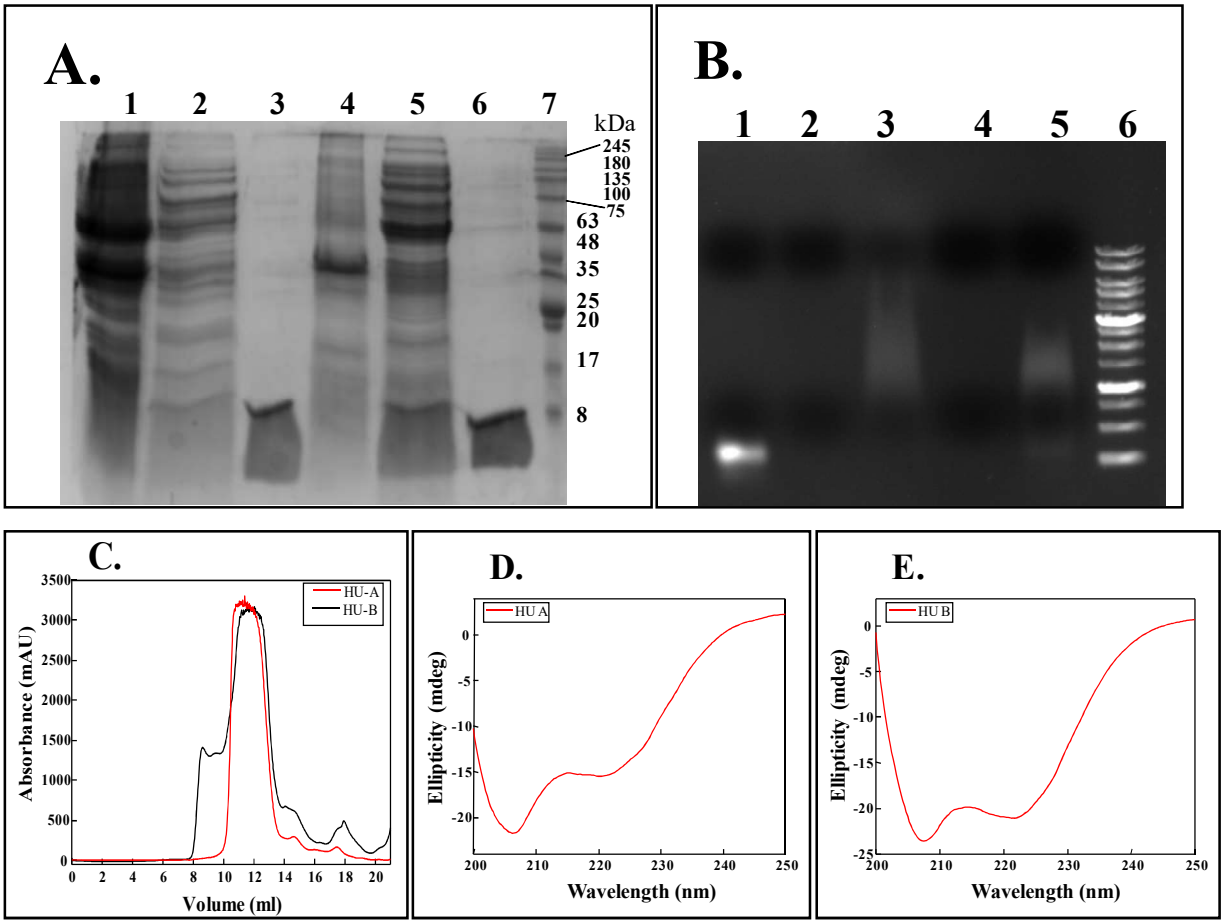


Figure 12: Panel A: SDS PAGE for purification of HU-A and HU-B in presence of 1 M salt. Lane 1- HU-A Pellet, Lane 2 HU-A FT, Lane 3 HU-A Elution, Lane 4 HU-B Pellet, Lane 5 HU-B FT, Lane 6 HU-B Elution and Lane M Marker. ***Panel B:*** DNA free purified protein HU-A (lane 2) and HU-B (lane 4), lane 3 and lane 5 show DNA binding of HU-A and HU-B. ***Panel C:*** Gel filtration chromatograms of HU-A (red) and HU-B (black) purified using NaCl ***Panel D*** and ***Panel E*** show far UV CD spectra for HU-A and HU-B, respectively.

3.4. Discussion and conclusions

It is well-known that biofilms contain large amounts of extracellular DNA (eDNA) originating from lysed and dead bacteria [8]; thus, it would not be surprising to find DNA-binding proteins such as HU (which is the most abundant and non-specifically binding NAP in bacteria) associated with the eDNA. Our results demonstrate that HU carries DNA as cargo out of the cell. Although it is equally conceivable that free HU binds to DNA already present in biofilms after the DNA had already formed the biofilm, upon leaving the cell without any pre-bound DNA. However, given our results, this latter possibility seems unlikely, and it is more likely that HU enters biofilms together with the DNA exiting cells. We have demonstrated this incontrovertibly by showing that the affinity-tagged HU protein does indeed carry fragments of DNA when it is purified from lysed *E. coli* cells under non-denaturing conditions. We have also demonstrate that this 'cargo' DNA carries with it numerous additional proteins that contaminate the HU purification. Mass spectrometric investigation of these 'contaminant proteins' establishes that some of them are known DNA-binding proteins, and others have no previously known DNA-binding characteristics. Interestingly, certain proteins of the latter variety (i.e., unknown DNA-binding potential)are known to be otherwise present in low abundance in *E. coli*, suggesting that sheer numbers cannot explain such contamination. Therefore, it appears either that such proteins tend to be co-purified because they possess DNA-binding potential, or because they happen to be bound to other known DNA-binding proteins. This suggests that this could be used as a method to identify new DNA-binding proteins, and/or new information regarding proteins natively bound to other DNA-binding proteins.

One of our methods for the removal of DNA from N-tagged HU, involving nuclease treatment, established that DNA bound to HU is resistant to degradation by nucleases. This holds important portents for the likelihood of survival of DNA in the extracellular milieu, owing to steric protection by bound HU. Most of the other methods employed – involving the use of denaturants - served to establish incontrovertibly that any attempt at removing DNA through the effecting of conformational alterations in HU do, of course, result in the effective removal of DNA, but create forms of HU which are either not DNA-binding-competent, or able to bind to DNA but only by adopting non-native states (such as soluble aggregates). Use of non-ionic detergents also resulted in populations of soluble aggregates of HU.

Finally, we established that of all strategies attempted to purify DNA (and associated proteins) free HU, the one that produce the highest yields of soluble, DNA-free, and DNA-binding-competent dimeric HU is the use of salt (1M NaCl) in the lysis buffer. Conventionally, HU is purified through a combination of gel filtration and ion-exchange chromatography following multi-stage ammonium sulfate fractionation and precipitation (4-6). Under these conditions, owing to the high ionic strengths used while using salt both during ammonium sulfate fractionation and ion exchange chromatography, it is to be expected that DNA would be dissociated from HU, and that neither DNA nor any DNA-associated proteins would be detected after purification. While affinity (IMAC-based) purification of HU has also been previously attempted (3),such purification too was followed by ion-exchange chromatography. Presumably, this is why contamination of DNA and DNA-associated proteins has never been previously reported for preparations of HU or any other NAP.

3.5. References

[1] E. Bonnefoy, J. Rouviere-Yaniv, HU and IHF, two homologous histone-like proteins of Escherichia coli, form different protein-DNA complexes with short DNA fragments, EMBO J 10 (1991) 687-696.

[2] A. Grove, Functional evolution of bacterial histone-like HU proteins, Curr Issues Mol Biol 13 (2011) 1-12.

[3] V. Pinson, M. Takahashi, J. Rouviere-Yaniv, Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA, J Mol Biol 287 (1999) 485-497.

[4] T.A. Azam, A. Ishihama, Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity, J Biol Chem 274 (1999) 33105-33113.

[5] C. Chen, S. Ghosh, A. Grove, Substrate specificity of Helicobacter pylori histone-like HU protein is determined by insufficient stabilization of DNA flexure points, Biochem J 383 (2004) 343-351.

[6] S. Ghosh, A. Grove, Histone-like protein HU from Deinococcus radiodurans binds preferentially to four-way DNA junctions, J Mol Biol 337 (2004) 561-571.

[7] J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, R.T. Dame, Dual architectural roles of HU: formation of flexible hinges and rigid filaments, Proc Natl Acad Sci U S A 101 (2004) 6969-6974.

[8] A. Devaraj, S.S. Justice, L.O. Bakaletz, S.D. Goodman, DNABII proteins play a central role in UPEC biofilm structure, Mol Microbiol (2015).

[9] D. Esser, H. Amanuma, A. Yoshiki, M. Kusakabe, R. Rudolph, G. Bohm, A hyperthermostable bacterial histone-like protein as an efficient mediator for transfection of eukaryotic cells, Nat Biotechnol 18 (2000) 1211-1213.

[10] A. Mukherjee, A.O. Sokunbi, A. Grove, DNA protection by histone-like protein HU from the hyperthermophilic eubacterium Thermotoga maritima, Nucleic Acids Res 36 (2008) 3956-3968.

[11] C.J. Dorman, K.A. Kane, DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes, FEMS Microbiol Rev 33 (2009) 587-592.

[12] E.V. Ram, R. Naik, M. Ganguli, S. Habib, DNA organization by the apicoplast-targeted bacterial histone-like protein of Plasmodium falciparum, Nucleic Acids Res 36 (2008) 5061-5073.

[13] B. Xiao, R.C. Johnson, J.F. Marko, Modulation of HU-DNA interactions by salt concentration and applied force, Nucleic Acids Res 38 (2010) 6176-6185.

[14] E. Christodoulou, C.E. Vorgias, Cloning, overproduction, purification and crystallization of the DNA binding protein HU from the hyperthermophilic eubacterium Thermotoga maritima, Acta Crystallogr D Biol Crystallogr 54 (1998) 1043-1045.

[15] O. Pellegrini, J. Oberto, V. Pinson, M. Wery, J. Rouviere-Yaniv, Overproduction and improved strategies to purify the threenative forms of nuclease-free HU protein, Biochimie 82 (2000) 693-704.

[16] J. Rouviere-Yaniv, N.O. Kjeldgaard, Native Escherichia coli HU protein is a heterotypic dimer, FEBS Lett 106 (1979) 297-300.

[17] G. Lauer, E.A. Rudd, D.L. McKay, A. Ally, D. Ally, K.C. Backman, Cloning, nucleotide sequence, and engineered expression of Thermus thermophilus DNA ligase, a homolog of Escherichia coli DNA ligase, J Bacteriol 173 (1991) 5047-5053.

[18] D. Esser, R. Rudolph, R. Jaenicke, G. Bohm, The HU protein from Thermotoga maritima: recombinant expression, purification and physicochemical characterization of an extremely hyperthermophilic DNA-binding protein, J Mol Biol 291 (1999) 1135-1146.

[19] A. Bahloul, F. Boubrik, J. Rouviere-Yaniv, Roles of Escherichia coli histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts, Biochimie 83 (2001) 219-229.

[20] K. Drlica, J. Rouviere-Yaniv, Histonelike proteins of bacteria, Microbiol Rev 51 (1987) 301-319.

[21] M. Lammi, M. Paci, C.L. Pon, M.A. Losso, A. Miano, R.T. Pawlik, G.L. Gianfranceschi, C.O. Gualerzi, Proteins from the prokaryotic nucleoid: biochemical and 1H NMR studies on three bacterial histone-like proteins, Adv Exp Med Biol 179 (1984) 467-477.

[22] S. Adhikari, P.V. Manthena, K. Sajwan, K.K. Kota, R. Roy, A unified method for purification of basic proteins, Anal Biochem 400 (2010) 203-206.

[23] F. Guo, S. Adhya, Spiral structure of Escherichia coli HUalphabeta provides foundation for DNA supercoiling, Proc Natl Acad Sci U S A 104 (2007) 4309-4314.

[24] B. Laine, P. Sautiere, G. Biserte, M. Cohen-Solal, F. Gros, J. Rouviere-Yaniv, The amino- and carboxy-terminal amino acid sequences of protein HU from Escherichia coli, FEBS Lett 89 (1978) 116-120.

[25] E. Christodoulou, C.E. Vorgias, The thermostability of DNA-binding protein HU from mesophilic, thermophilic, and extreme thermophilic bacteria, Extremophiles 6 (2002) 21-31.

[26] F. Orfaniotou, P. Tzamalis, A. Thanassoulas, E. Stefanidi, A. Zees, E. Boutou, M. Vlassi, G. Nounesis, C.E. Vorgias, The stability of the archaeal HU histone-like DNA-binding protein from Thermoplasma volcanium, Extremophiles 13 (2009) 1-10.

CHAPTER 4

Characterization of structure and stability of affinity tagged E. coli HV-A and HV-B

Contents

4.1 Introduction

In *E. coli*, HU is encoded by two closely related genes *hupA* and *hupB* which encode for the two HU isoforms HU-A and HU-B, respectively[1].HU-A is expressed from a single promoter but HU-B expression occurs from three different promoters (P2, P3 and P4)[2]. Both HU-A and HU-B exist in dimeric forms and form either the homo-dimers (HU-A2 and HU-B2) or the hetero-dimer (HU-AB).Concentrations of the three dimeric forms varies with growth of bacteria as the promoters for the two genes actively transcribe at different times during the growth of bacteria. Transcription from the *hupA* promoter starts during the early exponential phase of culture growth but there is no significant transcription from *hupB* promoters; hence, the predominant dimeric form during early exponential phase is the HU-A2 homo-dimer. As cells progress into the late exponential phase of growth, the transcription from *hupA* promoter stops and transcription from *hupB* P2 promoter starts, and *hupB* P3 promoter is activated during the stationary phase [3].During the late exponential phase, when both HU-A and HU-B proteins are present inside the cell, the HU-AB hetero-dimer is the most abundant dimeric population [4]. The existence of HU-B homo-dimer during the stationary phase has not been validated. It is postulated that the main functional dimeric form of HU is the HU-AB hetero-dimer and the hetero-dimerization overwhelms homo-dimerization [5]. However, the mechanism of association of homo-dimers to form hetero-dimers is not well understood.

HU-A and HU-B protein share very high sequence identity (70 %) [6] and sequence similarity (80 %) as is evident in Figure 1a which shows a sequence alignment of HU-A and HU-B proteins, created using the Planine multiple sequence alignment tool [7]. Comparison of the structures of the two proteins (Figure 1b) shows that structurally too, these two proteins show very high similarity and there are very few differences in the C-terminal α-helix and in the β-sheets region [8]. Interestingly, despite the very high levels of sequence and structural similarity, the different dimeric forms of HU show very different binding properties which appear to be used by *E. coli* cells to modulate and regulate nucleoid dynamics during different phases of growth. Pathways of regulation of expression of the two isoforms are also very different [9,10].

HU-A2 homo-dimer and HU-AB hetero-dimer have very similar DNA-binding properties which are very different from the HU-B2 homo-dimer.HU-A2 homo-dimer and HU-AB hetero-dimer bind to cruciform, gapped or nicked DNA with high binding affinities and bind to linear

DNA only in the presence of very low salt concentrations. On the other hand, the HU-B2 homo-dimer has very poor binding affinity for nicked or gapped DNA but it can bind to cruciform DNA with binding affinities comparable to that of HU-A2 homo and HU-AB hetero-dimers[6]. Besides these differences in their binding priorities and affinities, HU dimers also show differences in their quaternary structure, HU-A2 is known to exist only as a dimer whereas HU-B2 has been reported to exist mainly as dimer with tendencies to associate to form higher oligomeric species like tetramers and octamers[8].

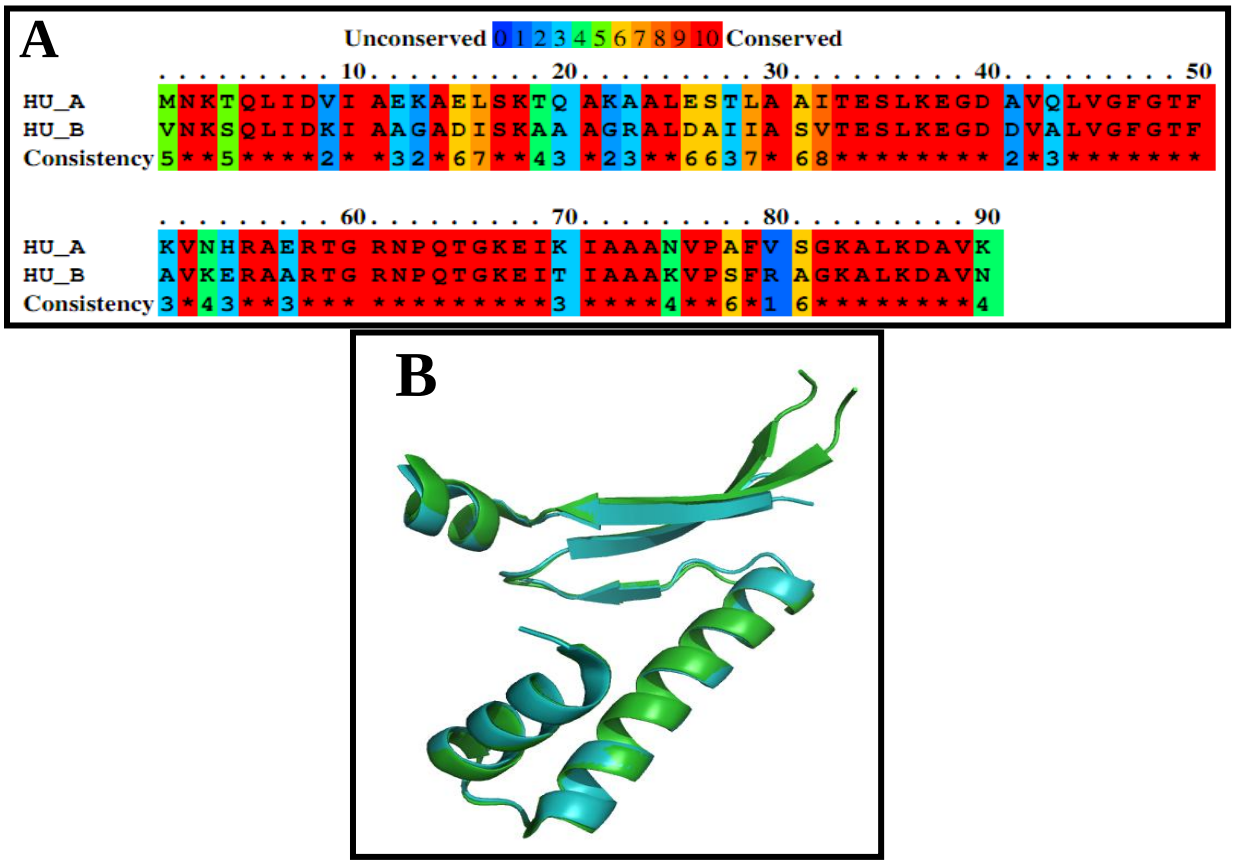


Figure 1: Panel A: Sequence alignment of HU-A and HU-B [11][11][11]done using Praline multiple sequence alignment. ***Panel B:*** Overlay of HU-A (cyan) and HU-B (green) protein structures done using PyMol [11].

Despite the differences in the DNA binding priorities, properties and affinities of the three dimeric forms, deletion of either *hupA* or *hupB* does not significantly affect cell growth or physiology, but deletion of both the genes reflects as significant growth defects in the double deletion mutants[12,13]. This suggests that one of the three dimeric forms (homo-dimer) can compensate for the absence of the other two dimeric forms (the other homo-dimer and the hetero-dimer) but if none of them is present, cell growth, division and metabolism are

significantly affected. It has been shown that deletion of one of these proteins is compensated by up-regulation of the other protein, although the mechanism and regulation of this process are not well understood. This brings attention to the following question: if one of the dimeric forms is both sufficient and necessary for bacterial growth, what is the physiological relevance, or need, of having three different dimeric forms present within *E. coli*? Why are these three differently functioning (DNA-binding) and associating dimeric populations required, if one of them is sufficient? A very similar situation of having two differentially expressed isoforms (forming both homo and hetero-dimers) has also been observed for *a)*H-NS, a histone-like nucleoid-structuring protein which also exists as both homo and hetero-dimers and shows changes in populations of dimeric forms with changes in gene expression patterns [14] and; *b)*an archaeal histone HMf from *Methanothermus fervidus* which also exists in two isoforms HMfA and HMfB, with different biological properties and being expressed during different phases of growth [15].

To explore the role of the different dimeric forms during different phases of *E. coli* growth, and to understand the mechanism favoring formation of the HU-AB hetero-dimer, it is important to study the protein's overall stability and the stability of the dimer interfaces of the individual homo-dimers. In this chapter, we have compared the thermal and chemical stabilities (against chemical denaturants, Gdm.HCl and urea) for HU-A2 and HU-B2 homo-dimers. We have also monitored unfolding kinetics as a function of temperature for these dimeric forms (with and without DNA).In addition, we have investigated the conditions required for dissociation of the dimer and changes in secondary structure occurring conditions promoting subunit dissociation.

4.2 Materials and methods specific to this chapter

4.2.1 Protein purification

Both C-terminally and N-terminally 6xHis tagged forms of HU-A and HU-B were expressed and purified for the studies performed in this chapter. Molar concentrations of salt were used to purify the proteins in pure DNA-binding competent form, as described in the previous chapter. Both HU-A and HU-B lack Tyr and Trp residues making it difficult to quantitate the protein concentrations accurately. To measure accurate protein concentrations for all the experiments, single Trp mutants (structurally and functionally equivalent to the native protein) of HU-A and

HU-B were used (discussed in chapter 6) to make a calibration plot of protein concentration vs. CD signal (for both HU-A and HU-B). Raw ellipticity for a particular concentration was used to calculate the exact concentrations form the calibration curves.

4.2.2 Circular Dichroism spectroscopy:

CD spectra were collected on a Chirascan Spectrophotometer (Applied Photophysics, UK) using a 1 mm path length synthetic quartz cuvette at protein concentration of 0.2 mg/ml (for both HU-A and HU-B). Both the sample and buffer spectra were collected over the range of 200-250 nm. The spectra obtained were corrected for buffer background signal. Thermal denaturation was done by heating the protein samples from 20 ^{0}C to 90 ^{0}C using the spectropolarimeter's peltier block arrangement and data was collected at 2 ^{0}C intervals. Raw ellipticity obtained was converted into mean residue ellipticity [θ] using the following formula:

$$[\boldsymbol{\theta}] = \frac{\boldsymbol{\theta_{obs}}\textbf{(mdeg) X 100 X MRW}}{\textbf{1000 X concentration (mg/ml) X path length (cm)}} \qquad \text{Equation 1}$$

where, MRW is the mean residue weight (molecular weight of protein/ number of amino acids) and θ_{obs} is the raw ellipticiy.

For equilibrium chemical denaturation experiments, the protein samples (0.5 mg/ml) were pre-incubated in various urea and guanidium hydrochloride concentrations over-night and CD spectra were collected on a Biologic MOS-500 instrument using 1mm path length synthetic quartz cuvette.

4.2.3 Differential scanning calorimetry

To examine the melting temperature of HU-A and HU-B, 0.2 mg/ml of the protein samples were heated from 20^{0}C to 100^{0}C at a heating rate of 90 ^{0}C/hour and cooled back at a cooling rate 60 ^{0}C/hour on a VP-DSC instrument from MicroCal. T_m values were calculated after buffer subtraction, normalization and baseline correction using the Origin software supplied with the instrument.

4.2.4 Gel filtration chromatography

HU-A protein was heated at 90 ^{0}C for 5 minutes and allowed to cool back to room temperature. Size-exclusion chromatography was performed for both heat-cooled and unheated protein samples using a 24 ml Superdex 75 ml column (GE healthcare) that had been pre-equilibrated

with PBS. Elution fractions were collected and were analyzed for DNA binding by running them on 0.5 % agarose EMSA gel.

4.2.5 Unfolding kinetic studies and thermodynamic analysis

In order to characterize and compare the thermal and thermodynamic stabilities of HU-A2 and HU-B2 homo-dimers, we measured the temperature dependence of the unfolding rate constants. Time dependent changes in secondary structure at different temperatures were recorded. For all kinetics experiments, a 2 mm synthetic quartz cuvette was used. In advance of time resolved measurements, buffer solution was placed in the spectrometer for 5-10 min in order to reach thermal equilibrium at desired temperatures, following which a small amount of concentrated protein solution was rapidly added to preheated buffer, mixed and the data acquisition was started. The dead time of this manual mixing procedure (~20 s) was taken into account in data analysis. The final protein concentration that was used for both HU-A and HU-B was 25 µM. For HU-DNA binding studies, similar protein concentration of 25 µM was used along with 0.5 µM of DNA (synthetic four-way junction). We performed repetitive measurements every 10 s for a total time period of about 1 hour. Unfolding transitions were monitored by taking the CD signal at 222 nm for samples at different temperatures as a function of time. For all data treatments and data presentations, Origin8.5 software (OriginLab Corp, Northhampton, MA, USA) was used. Fraction folded protein (f_F) was calculated using the following equation:

$$\boldsymbol{f_F = 1 - f_U} \qquad \text{Equation 2}$$

and,

$$\boldsymbol{f_U = (y_F - y)\,/(y_F - y_U)} \qquad \text{Equation 3}$$

Here, f_U – the fractional unfolding, y_F and y_U are the 222 nm CD values of folded and unfolded protein, respectively, whereas y is the observed value of 222 nm CD at any particular point.

According to the transition state theory the apparent free energy change ΔG_{app} between the native state and the transitions state is

$$\boldsymbol{\Delta G_F = \Delta H_F - T\Delta S_F} \qquad \text{Equation 4}$$

which characterizes the unfolding barrier between the native and the unfolded state, and is related to the obtained unfolding rates K_u by the following relation [16].

$$\boldsymbol{k_u = k_o . e^{\Delta G_F/RT}} \qquad \text{Equation 5}$$

In this equation, k_o, the so called pre-exponential factor, represents a largest possible rate constant for a specific chemical reaction in the absence of free energy barriers [17]. For all temperature points, data was fitted with a bi-exponential decay characterized by the unfolding rate constants, k_{u1} and k_{u2}. Further, a simple kinetic model based on the Lumry-Eyring [16] equation was used to compare the unfolding kinetics of different proteins

In order to obtain thermodynamic parameters from the kinetic measurements, unfolding rates were displayed in an Eyring-plot ($\ln(K_u$ /T) vs. 1/T) and the temperature dependence was analyzed using a linear fit (y = -Ax +B). From the slope of the linear fit we obtained ΔH and from the intercept with the y-axis, we were able to determine ΔS [18].

4.2.6 Glutaraldehyde cross-linking

For cross-linking experiments, 0.2 mg/ml of the protein samples that had been pre-incubated in different concentrations of urea were further incubated with 0.1% glutaraldehyde for 5-10 minutes before loading on 15% SDS-PAGE gels.

4.3 Results and discussion

4.3.1 HU-A and HU-B can refold after heat induced unfolding

As mentioned earlier, C-terminal 6xHis-tagged HU-A and HU-B proteins precipitated after some time but N-terminal 6xHis-tagged protein remained soluble. CD studies were done with both C-terminal 6xHis-tagged (before precipitation) and N-terminal 6xHis-tagged HU-A and HU-B.CD spectra for both the N-tagged (Figure 2 Panel A and B) and the C-tagged (Panel C and D) HU-A and HU-B show that both the N-tagged and C-tagged proteins are structured (black trace in Figure 2). Upon heating to 90^0C, the proteins unfolds (red trace) and there is a transformation from a helical structure to a randomly coiled structure (a helix-coil transition). On cooling back to 20^0C, the proteins refolds back into helical structure but a slight loss in raw ellipticity (in mdeg) is observed (blue trace) for HU-A N-tag (Panel A) and HU-B N-tag (Panel B) but a significant loss of structure on refolding is observed in case of HU-A C-tag (Panel C) and HU-B C-tag (Panel D). The difference in extent of structure gained on refolding(after heat induced unfolding) clearly suggests that a 6xHis-tag at C-terminal decreases the refolding efficiency for both HU-A and HU-B proteins.

4.3.2 Heat refolded protein is dimeric and is DNA-binding competent

To check if the structure attained by the refolded protein is native like, N-terminally 6xHis-tagged HU-A and HU-B proteins were heated at 90 ^{0}C for 10 minutes followed by cooling to 20 ^{0}C to allow the refolding to occur. The resultant refolded proteins were analyzed to examine whether the refolded proteins are DNA-binding competent and dimeric like the native protein. As shown in Figure 3A (for HU-A) and Figure 3B (for HU-B), the gel-filtration chromatograms of both the refolded proteins are identical to that of the unheated protein, indicating that the refolded protein is also dimeric and elutes at the elution volume corresponding to the molecular weight of the dimer (~20 kDa), i.e., at approximately 12.5 ml elution volume. The electrophoretic mobility shift assay (EMSA) gel shown in Figure 4 suggests that the refolded proteins also retain the DNA-binding property, in that the synthetic 4-way junction DNA's mobility is retarded by HU binding. The above data suggests that once unfolded by heating, both the HU proteins can refold back to a native/native-like structure competent to perform both DNA binding and dimerization.

4.3.3 HU-B is thermally more stable than HU-A

Thermal melting of the HU-A and HU-B proteins (for both N-terminal 6xHis-tag and C-terminal 6xHis-tag) was done by heating the protein from 20 ^{0}C to 90 ^{0}Cat a controlled rate, and monitoring the change in CD signal (ellipticity at 222 nm). The red traces in Figure 5 represent the 222 nm CD melting curves for all the proteins. Upon heating to 90 ^{0}C, a significant loss in 222 nm ellipticity was observed for all the proteins, i.e., for HU-A N-tag (Panel A), HU-B N-tag (Panel B), HU-A C-tag (Panel C) and HU-B C-tag (Panel D). However, the residual ellipticity suggests that the structure does not unfold completely. As mentioned earlier, on cooling back to 20^{0}C, all HU proteins refold, as reflected by the increase in the ellipticity values (blue trace). Structure of HU-A (both N-tag and C-tag) starts opening up as soon as the temperature is increased to 20^{0}C, but for HU-B (both N-tag and C-tag), there is no significant loss in structure below 30^{0}C. Melting temperature (T_m) values also suggest that HU-B is much more stable than HU-A, with the N-tagged HU proteins being more stable than C-tagged HU proteins. As the C-terminally 6xHis-tagged HU-A and HU-B (a) formed visible precipitates/aggregates after incubation for some time, (b) are unable to completely refold after heat induced unfolding and, (c) show less thermally stability than their respective N-terminally 6xHis-tagged variants, for all

further experiments described we chose to work only with the N-terminally 6xHis-tagged variants of both HU-A and HU-B.

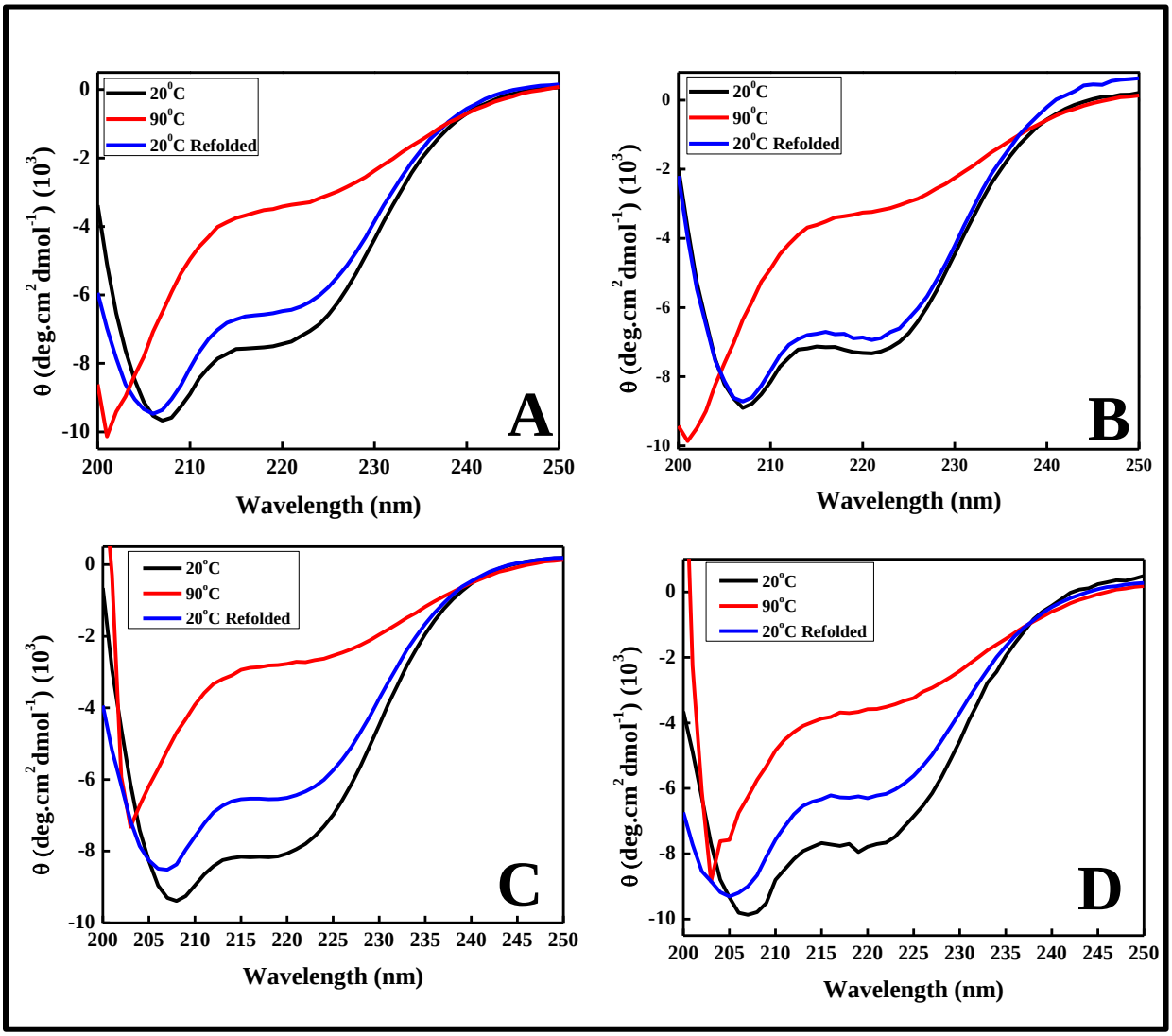


Figure 2: Secondary structures of native protein (black), unfolded protein (red) and refolded protein (blue).***Panel A:*** HU-A N-terminal 6xHis-tagged, ***Panel B:*** HU-B N-terminal 6xHis-tagged, ***Panel C:***HU-A C-terminal His- tagged, ***Panel D:*** HU-B C-terminal 6xHis-tagged**.**

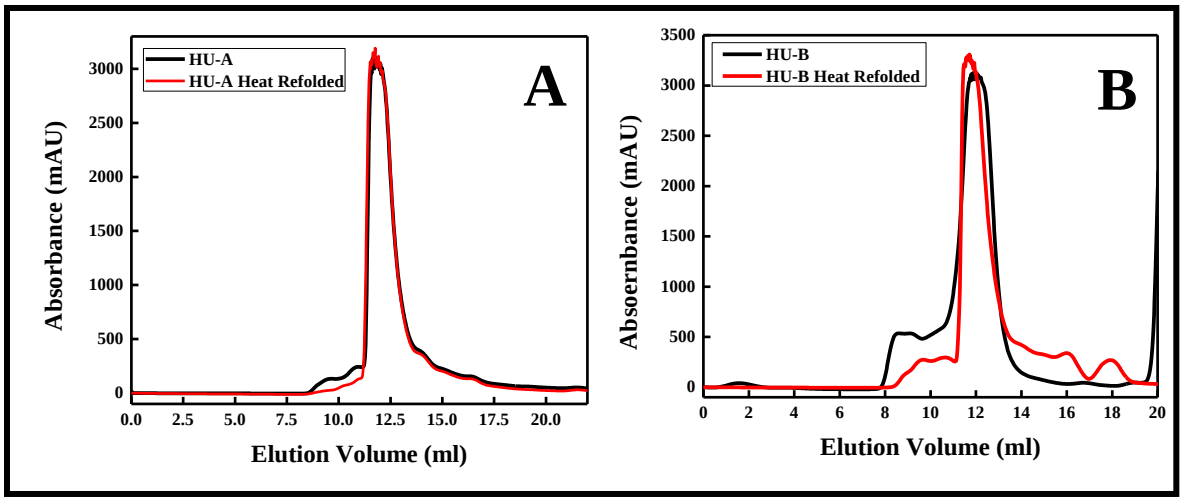


Figure 3: Gel filtration profile of unheated and heated HU-A and HU-B.

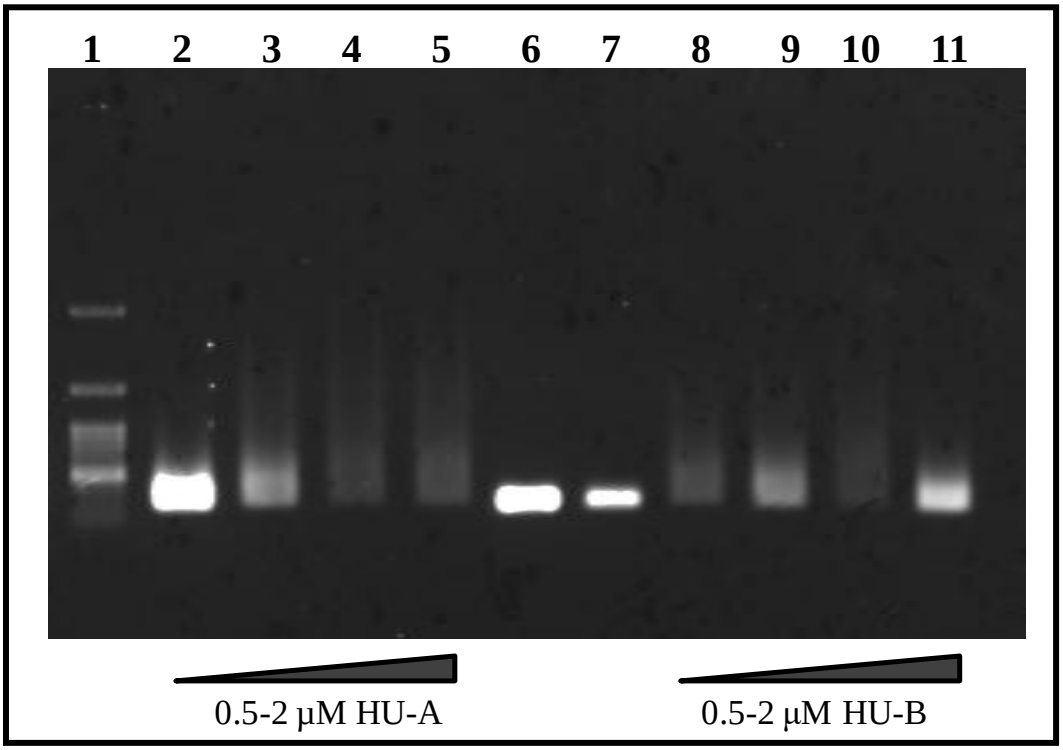


Figure 4: EMSA gel showing binding of refolded HU proteins with DNA. *Lane 1:* 100bp DNA, *Lane 2-5:* HU-A refolded protein plus DNA, *Lane 6:* DNA, *Lane 7:* DNA plus PBS and *Lane 8-11:* HU-B refolded protein plus DNA.

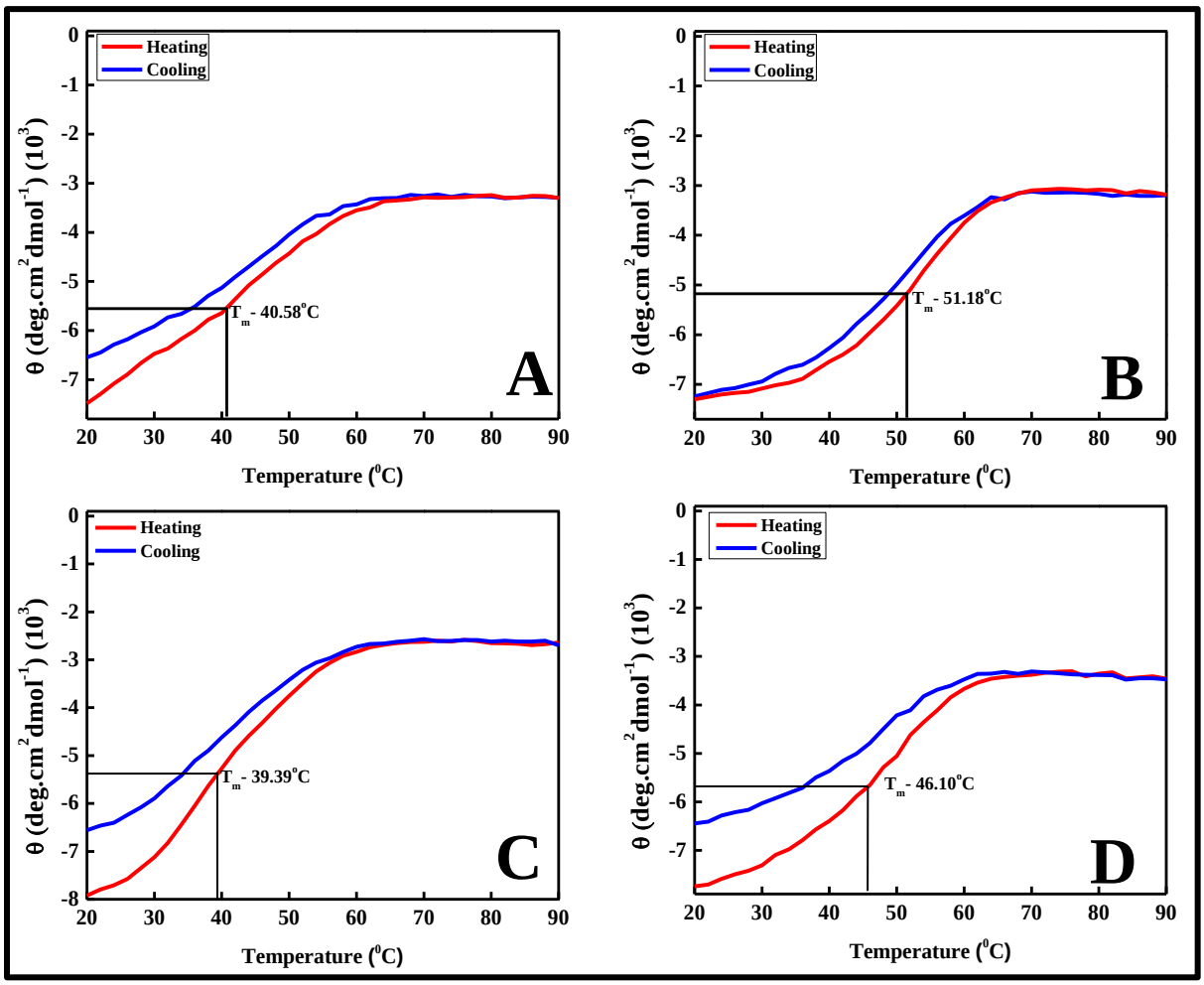


Figure 5: Heat induced changes (unfolding –red and refolding – blue) in secondary structures of HU-A N-terminal 6xHis-tagged ***(Panel A),*** HU-B N-terminal 6xHis-tagged ***(Panel B),*** HU-A C-terminal 6xHis-tagged ***(Panel C)*** and HU-B C-terminal 6xHis-tagged ***(Panel D).***

In conformity with the circular dichroism data, the differential scanning calorimetry (DSC) data (as shown in Figure 6) also points towards a higher stability of HU-B than HU-A. Interestingly, the T_m values obtained by DSC analysis are higher (for both HU-A and HU-B) than those obtained from the sigmoidal fitting of CD (ellipticity at 222 nm) melting curves. Theoretically, in circular dichroism, the changes in secondary structure of the protein determine the observed changes in signal and, therefore, the measured T_m values. T_m determined using CD is the temperature at which half of the overall secondary structure of the protein is lost. DSC, on the other hand, monitors the enthalpy changes associated with heat-linked structural changes that occur in the protein sample (in reference to the buffer) involving both secondary and tertiary structural changes. The obtained higher T_m values in the DSC experiments than in the CD

experiments suggests that there probably are some tertiary associations in the protein molecules that are disrupted only at higher temperature values, with no associated changes in secondary structure, i.e., that under conditions causing the bulk of the secondary structure to appear to have been lost, some tertiary structure could still remain as residue (particularly in the interface region between subunits). This concern is addressed at length in the later part of the chapter.

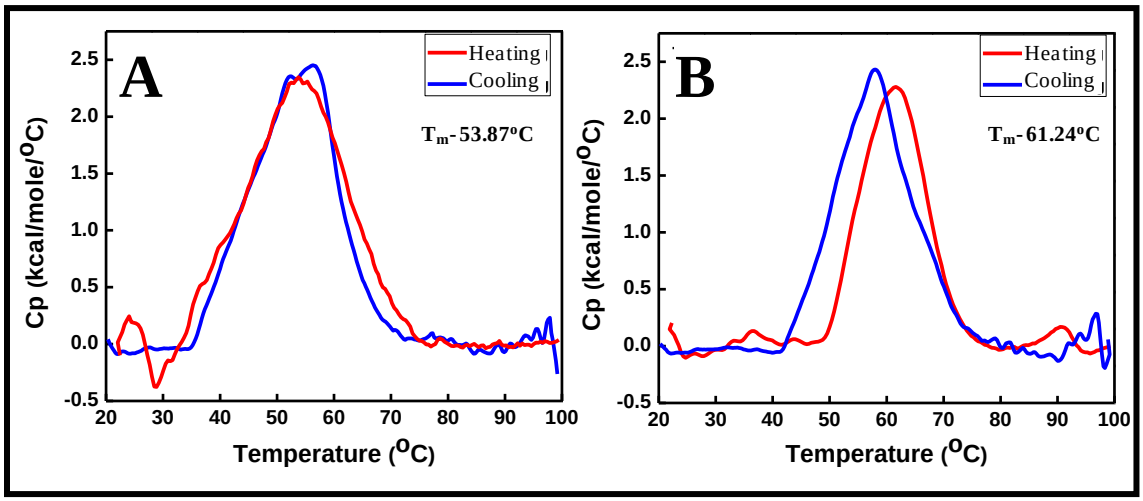


Figure 6**:* Differential scanning calorimetry data for the endothermic structure-melting transitions. ***Panel A*:** HU-A N-terminal 6xHis-tagged and ***Panel B: HU-B N-terminal 6xHis-tagged

4.3.4 HU-B is more stable to chemical denaturation than HU-A

Far UV-CD equilibrium unfolding studies of both HU-A and HU-B were performed in the presence of different concentrations of chaotropic agents; urea as well as guanidium hydrochloride (Gdm.HCl). Protein samples were incubated in 0-8 M urea and 0-6 M Gdm.HCl concentrations at increments of 0.25 M. Far-UV spectra were collected for the afore-mentioned concentrations of urea and Gdm.HCl. Figure 7 shows changes in ellipticity at 222 nm for HU-A with increasing concentration of urea and Gdm.HCl in Panels A and B, respectively. Figure 8 shows far-UV spectra for HU-B in the presence of urea (Panel A) and Gdm.HCl (Panel B). Panel C and D show changes in 222 nm ellipticity for HU-B in presence of urea and Gdm.HCl, respectively. The data was used to calculate the denaturation midpoint (C_m) values by fitting the data into sigmoidal fit. The data suggests that both HU-A and HU-B are very unstable to chemical denaturation and protein structure starts opening up even at very low concentrations of denaturant. Of the two proteins, HU-B is more stable than HU-A as reflected from the C_m values. In both the proteins, structure is almost completely lost in 5 M urea and 2.5 M guanidium hydrochloride.

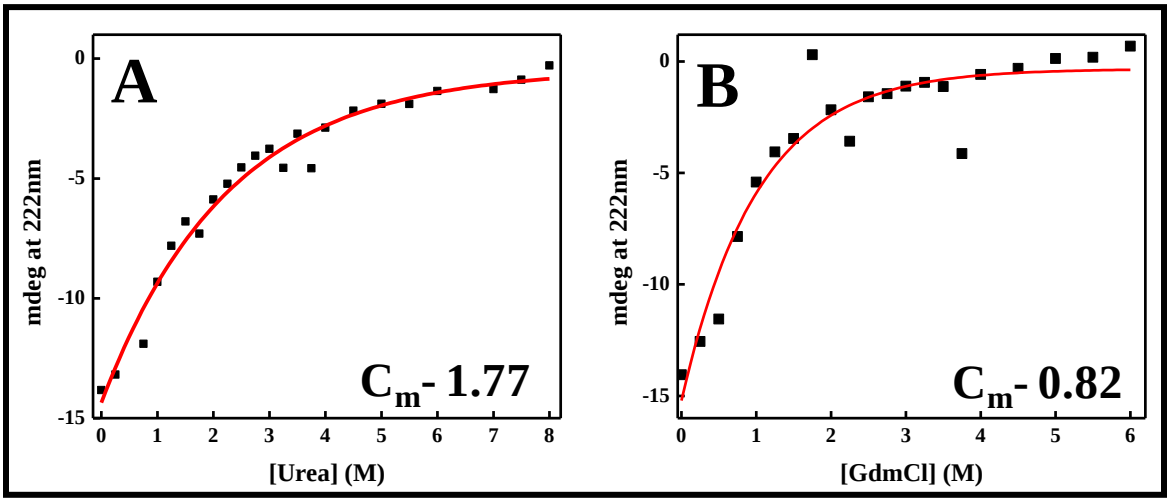


Figure 7*:* mdeg values at different urea and guanidium concentrations for: HU-A N-terminal 6xHis-tagged protein ***Panel A:*** 0-8 M urea ***Panel B*****:** 0-6 M guanidium hydrochloride.

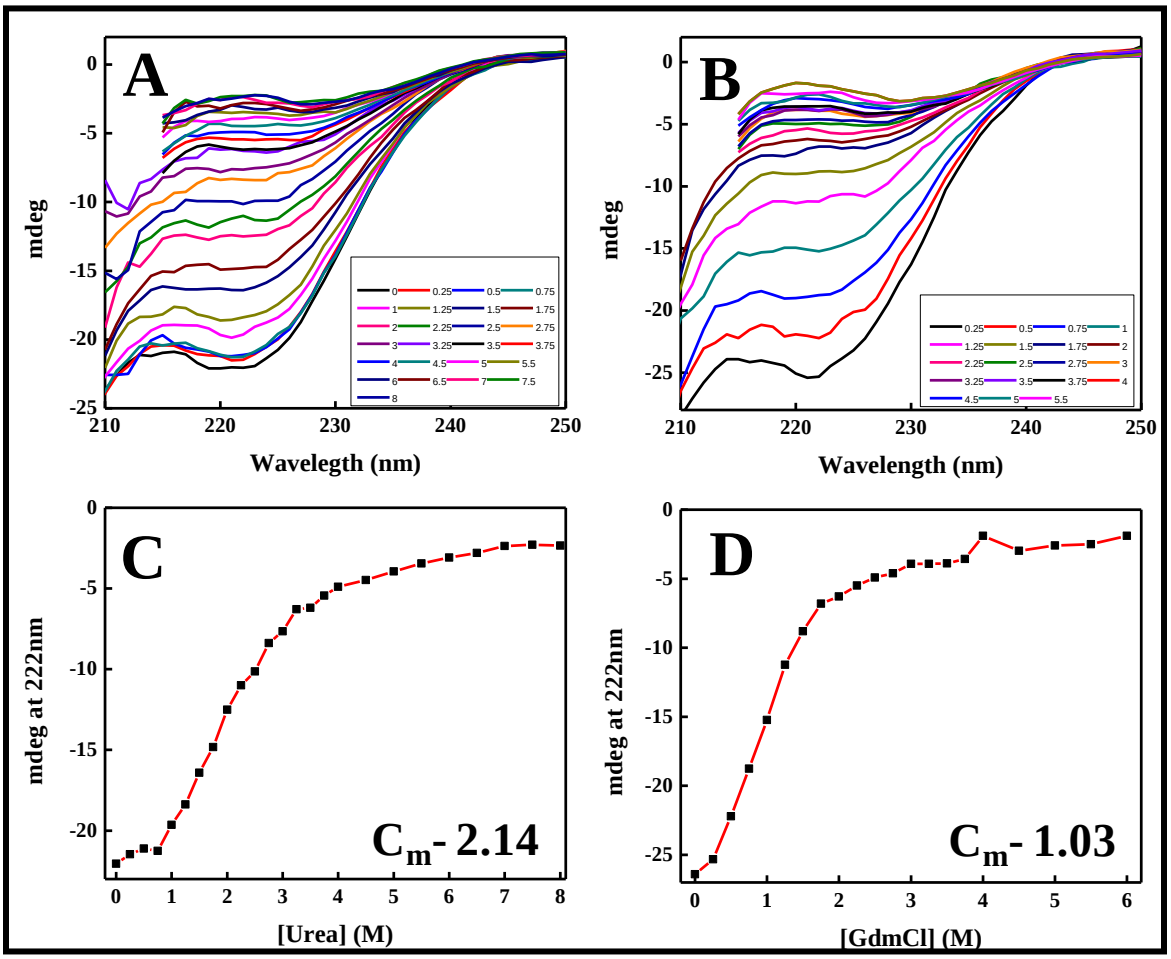
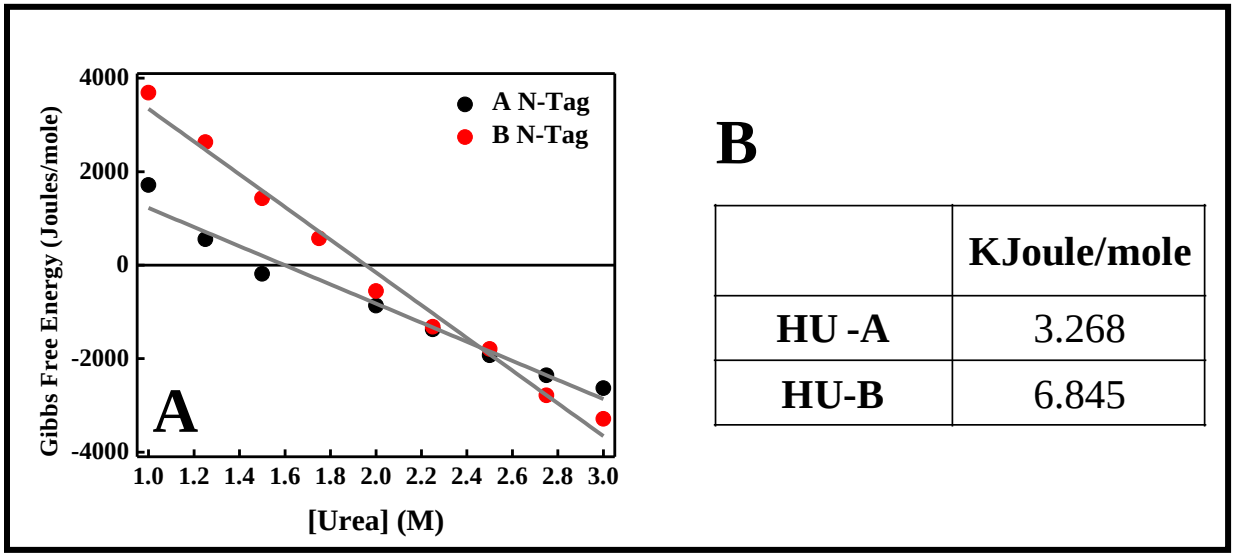


Figure 8: Changes in secondary structure of HU-B N-terminal 6xHis-tagged protein: ***Panel A:*** 0-8 M urea ***Panel B*****:** 0-6 M guanidium hydrochloride. ***Panel C*** and ***Panel D*** represent the changes in mdeg value at 222 nm for the data in panel A and B, respectively.

4.3.5 Calculation of free energy change

From the equilibrium chemical denaturation curves, the fractional unfolding (f_U) and hence the unfolding rate (k_u) and the apparent free energy change (ΔG_{app}) were calculated using equations 2, 3 and 4. Change in Gibb's free energy with respect to the change in urea concentrations was plotted as shown in Figure 9. ΔG_{app} values were calculated by applying linear fit to the data to find the value of Y-intercept corresponding to value 0 for the X-axis. ΔG_{app} values for HU-A and HU-B are 3.268 KJ/mole and 6.845 KJ/mole, respectively. This also points to higher stability of HU-B than HU-A.



B

	KJoule/mole
HU -A	3.268
HU-B	6.845

Figure 9: Panel A: Plot of free energy change vs. urea concentration for HU-A (black circles) and HU-B (red circles). Grey lines represent the fit in both cases. ***Panel B:*** values of free energy change for HU-A and HU-B

4.3.6 Kinetics of thermal unfolding of HU-A and HU-B

Kinetic parameters of thermal unfolding of HU proteins was calculated using the equations 2-5 as mentioned in materials and methods section. Eyring plots were plotted (Figure 10) and the thermodynamic parameters were calculated for same molar concentrations of both HU-A and HU-B proteins. Differences in the stability of HU-A and HU-B are also reflected in differences in activation energies and enthalpy values.

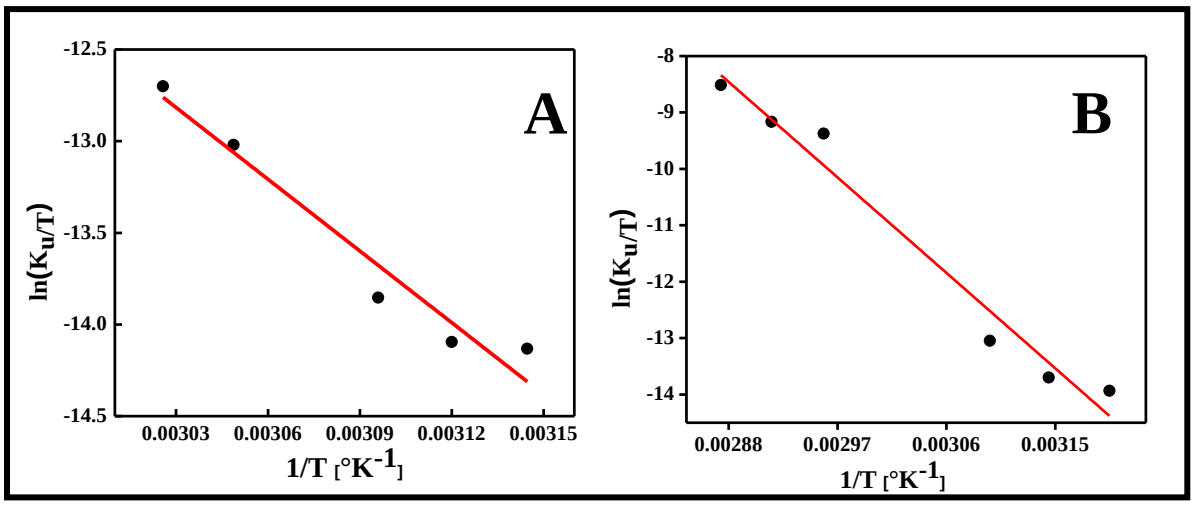


Figure 10: ***Panel A*** and ***Panel B*** show Eyring plots for HU-A and HU-B, respectively.

Protein	E_a (kJ/mol)	ΔH (kJ/mole)	ΔS (J/mole/K)	ΔG_{app} (KJ/mole)	T_m (°C)
HU-A	111	49.93	34.79	10.13	41.59
HU-B	246.46	99.78	83.63	16.15	50.53

Table 1: Thermodynamic parameters obtained from Eyring plots for HU-A and HU-B.

The values of ΔG obtained for both the HU isoforms are extremely low for a dimeric protein. This suggests that the HU dimers are very unstable to thermal denaturation. The same is evident from low T_m and C_m values.

4.3.7 Dissociation vs. Unfolding

Unfolding of a multimeric protein involves two different but not necessarily well-separated processes: (a) dissociation of the multimeric form into the monomers (loss of quaternary structure stabilizing interactions at the subunit interface), and (b) unfolding of the monomers (loss of secondary structure stabilizing interactions, in addition to loss of tertiary structure stabilizing interactions). The simplest unfolding mechanism for a dimeric protein would be a two-state unfolding where the above two processes happen in a concerted manner and the protein dissociates and unfolds in a single step. In more complex mechanisms, involving intermediates, various intermediates like folded monomer or partially unfolded dimer can be present [19]. In

this section, in order to understand the stability of the dimer interface of HU, we have studied the unfolding pathway of HU homo-dimers. Figure 11A shows the structure of *Anabaena* HU dimers bound to DNA showing that the main body of a single DNA-interacting moiety, a dimer is formed by tight associations of the two monomeric halves. Below we have explored the unfolding (chemical) pathway of HU to see if it shows a two-state unfolding or forms one of the intermediates as depicted in Figure 11B.

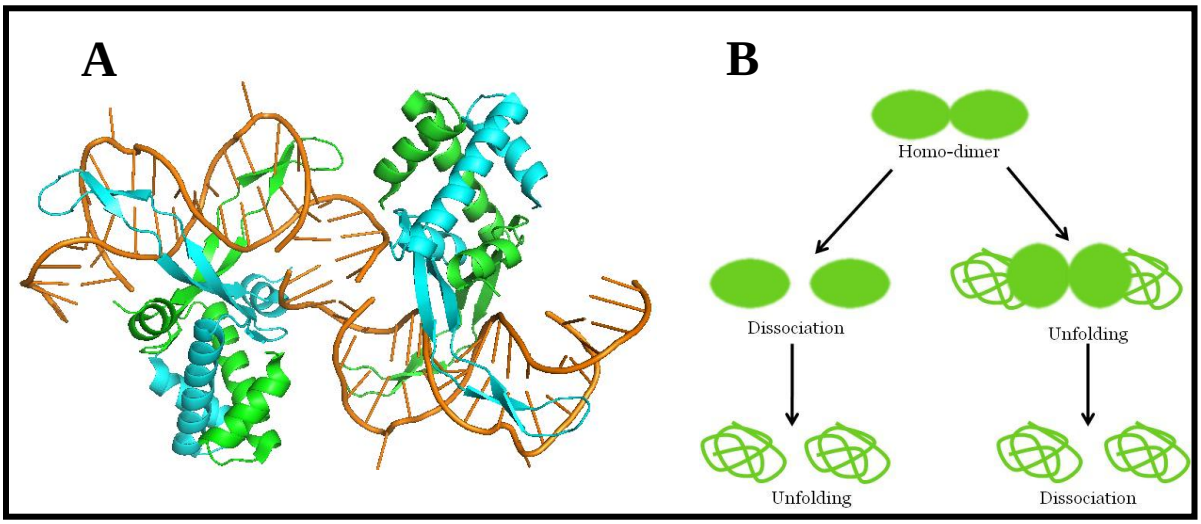


Figure 11: Panel A: Structure of *Anabaena* HU dimers bound to DNA. ***Panel B***: Pathways of unfolding of a dimeric protein.

In order to determine the inter-subunit stability, glutaraldehyde cross-linking was done for both HU-A and HU-B homo-dimers in the presence of different urea concentrations. It can be seen clearly in Figure 12 that even at very high urea concentrations (with respect to the very low C_m values reported in the immediately-preceding section),glutaraldehyde cross-linked dimer can be observed. At these concentrations of urea, 90% of the protein secondary structure is lost (Figure 7A, 8A). This reflects that in case of both HU-A2 and HU-B2, inter sub-unit interactions are retained even when most of the secondary structure of the protein is lost. This suggests that during unfolding pathway of HU protein, unfolding of the protein precedes dissociation of the sub-units and partially unfolded dimeric species are formed prior to complete unfolding.

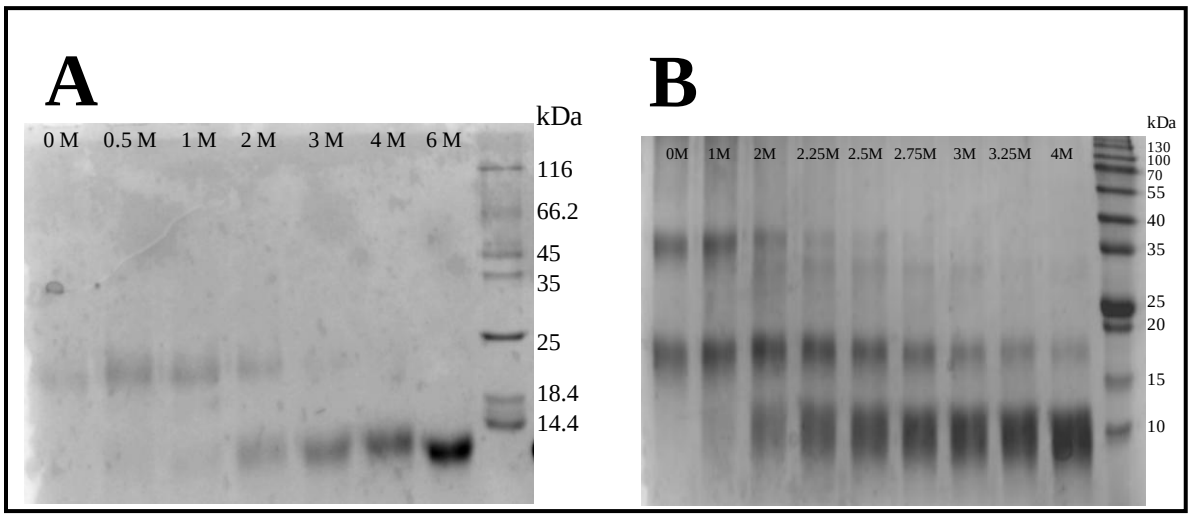


Figure 12: Glutaraldehyde cross-linking in presence of different urea concentrations. HU-A in ***Panel A*** and HU-B in ***Panel B.***

4.4 Discussion

A few misconceptions exist in literature about the HU protein and these continue to be cited. (A) Many papers in the literature mention that the name HU is an abbreviation for Heat-Unstable and is generally referred to as heat-unstable histone-like protein[20,21,22]. (B) Rouviere-Yaniv and co-workers, who were amongst the first few people to purify and characterize HU and also some other groups mention that HU is a heat-stable protein as it has ability to bind DNA even after it has been exposed to heating at 100 ^{0}C [4,23,24]. But, actually, both the above statements are misleading. In HU, the letter 'H' relates to histone-like protein, but the letter 'U' indicates that it was first isolated from U-93 bacteria by Rouviere-Yaniv and coworkers[4]. Our data suggests that although HU does not stand for 'heat-unstable' but HU is indeed a very heat unstable protein and the secondary structure is lost even at very low temperatures (~ 25 ^{0}C for HU-A and ~ 35 ^{0}C and T_m values for loss of secondary structure are also low to be called heat-stable protein (40.58 ^{0}C and 51.18 ^{0}C for HU-A and HU-B, respectively). Further, the ability of HU to bind to DNA even after being exposed to high temperatures comes not from its being very heat stable, but from the fact that HU can refold back to a very native-like structure (Figure 2), which is capable of binding to DNA as well (Figure 4). HU is very unstable to chemical denaturation as well and C_m values in urea are also as low as 1.77 M for HU-A and 2.14 M for HU-B. In terms of stability, of the two dimeric forms, HU-B behaves a little better than HU-A, being more stable.

The best measure of thermodynamic and chemical stability of a protein is the measure of apparent change in free energy (ΔG_{app}) associated with unfolding. ΔG_{app} values calculated from our kinetic experiments also point to the extremely low stability of the HU protein. Thermodynamics of unfolding has been studied in great details for HU from *Thermotoga maritima*. Despite of it being a thermophile derived protein with very high T_m value of 80.5 ^{0}C, the kinetic parameters for unfolding of *T. maritima*-derived HU are also very low [25]. The possible reasons for such low thermodynamic stability of HU proteins could be explained by the presence of a high percentage of unstructured DNA-binding β-arms in DNA-free protein. Residues55-74 in HU-A and residues 56-73 in HU-B are also unstructured in the absence of DNA [8] i.e., approximately 20% of structure in both HU-A and HU-B is unstructured if DNA is not bound to the protein. The fact that this unstructured region does not contribute to heat-changes associated with unfolding of the protein gives a possible explanation for extremely low free energy change values observed. Such unstructured DNA binding domains have now been reported for several other DNA-binding proteins also, including transcription factors [26,27,28]. For such proteins, the unstructured DNA-binding region is postulated to play critical roles in increasing the specificity and stability of protein-DNA interactions and these are hence termed "The Affinity Tuners" [29]. Such unstructured DNA binding interface and the associated low thermodynamic stability of HU proteins explains the ability of HU to bind to various DNA substrates with varying affinities and sheds some light on the role of HU in modulation of nucleoid dynamics inside cells.

Unfolding of a dimeric protein involves breaking of both inter and intra-molecular interactions, which could occur simultaneously in a two step unfolding pathway or in a multi-step unfolding pathway involving formation of intermediates. Formation of an intermediate during unfolding of HU dimer has previously been reported for thermal denaturation monitored using CD and DSC experiments [30]. We do not observe any such intermediate in our CD thermal melt or DSC experiments with either HU-A or HU-B (both N-terminal and C-terminal 6xHis-tagged proteins). We observe a very clear two-state transition for all the variants. This could possibly owe to different purification protocols used or different buffer systems used during the experiments. HU sourced from other organisms under similar experimental conditions also show a typical two-state transition on unfolding [25]. Our thermal melt experiments do not

suggest formation of an intermediate but our glutaraldehyde experiments do suggest formation of a partially unfolded dimer. By NMR studies at higher temperatures Ramstein et al. have also recently shown an intermediate which they also propose to be a partially unstructured dimeric state [31,32]. On the other hand, Vis et al have shown a monomeric intermediate in case of HU derived from *Bacillus sterrothermophillus* [33].

Our data suggests that inter-subunit associations play a significant role in unfolding of HU dimers, since these appear to be retained even in largely unfolded protein. The strong inter sub-unit associations also explain the higher T_m values in DSC experiments than in the CD experiments. So, overall we propose that HU-dimers unfold through formation of a partially unfolded dimeric intermediate, which can refold back after complete unfolding to a DNA binding competent state.

4.5 References

[1] J. Rouviere-Yaniv, N.O. Kjeldgaard, Native Escherichia coli HU protein is a heterotypic dimer, FEBS Lett 106 (1979) 297-300.

[2] K. Kohno, M. Wada, Y. Kano, F. Imamoto, Promoters and autogenous control of the Escherichia coli hupA and hupB genes, J Mol Biol 213 (1990) 27-36.

[3] L. Claret, J. Rouviere-Yaniv, Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival, J Mol Biol 273 (1997) 93-104.

[4] K. Drlica, J. Rouviere-Yaniv, Histonelike proteins of bacteria, Microbiol Rev 51 (1987) 301-319.

[5] R.C. Johnson, M.F. Bruist, M.I. Simon, Host protein requirements for in vitro site-specific DNA inversion, Cell 46 (1986) 531-539.

[6] V. Pinson, M. Takahashi, J. Rouviere-Yaniv, Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA, J Mol Biol 287 (1999) 485-497.

[7] V.A. Simossis, J. Heringa, PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information, Nucleic Acids Res 33 (2005) W289-294.

[8] F. Guo, S. Adhya, Spiral structure of Escherichia coli HUalphabeta provides foundation for DNA supercoiling, Proc Natl Acad Sci U S A 104 (2007) 4309-4314.

[9] L. Claret, J. Rouviere-Yaniv, Regulation of HU alpha and HU beta by CRP and FIS in Escherichia coli, J Mol Biol 263 (1996) 126-139.

[10] E. Bonnefoy, A. Almeida, J. Rouviere-Yaniv, Lon-dependent regulation of the DNA binding protein HU in Escherichia coli, Proc Natl Acad Sci U S A 86 (1989) 7691-7695.

[11] The PyMOL Molecular Graphics System, Schrödinger, LLC.

[12] M. Wada, Y. Kano, T. Ogawa, T. Okazaki, F. Imamoto, Construction and characterization of the deletion mutant of hupA and hupB genes in Escherichia coli, J Mol Biol 204 (1988) 581-591.

[13] O. Huisman, M. Faelen, D. Girard, A. Jaffe, A. Toussaint, J. Rouviere-Yaniv, Multiple defects in Escherichia coli mutants lacking HU protein, J Bacteriol 171 (1989) 3704-3712.

[14] C.M. Muller, U. Dobrindt, G. Nagy, L. Emody, B.E. Uhlin, J. Hacker, Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli, J Bacteriol 188 (2006) 5428-5438.

[15] K. Sandman, R.A. Grayling, B. Dobrinski, R. Lurz, J.N. Reeve, Growth-phase-dependent synthesis of histones in the archaeon Methanothermus fervidus, Proc Natl Acad Sci U S A 91 (1994) 12624-12628.

[16] H. Eyring, The activated complex in chemical reactions, The Journal of Chemical Physics 3 (1935) 107-115.

[17] K.T. Bachmann A. , Kinetic Mechanisms in Protein Folding, in Protein Folding Handbook Wiley-VCH Verlag GmbH, Weinheim, Germany. , 2005.

[18] C. Duy, J. Fitter, Thermostability of irreversible unfolding alpha-amylases analyzed by unfolding kinetics, J Biol Chem 280 (2005) 37360-37365.

[19] J.A. Rumfeldt, C. Galvagnion, K.A. Vassall, E.M. Meiering, Conformational stability and folding mechanisms of dimeric proteins, Prog Biophys Mol Biol 98 (2008) 61-84.

[20] G. Lia, D. Bensimon, V. Croquette, J.F. Allemand, D. Dunlap, D.E. Lewis, S. Adhya, L. Finzi, Supercoiling and denaturation in Gal repressor/heat unstable nucleoid protein (HU)-mediated DNA looping, Proc Natl Acad Sci U S A 100 (2003) 11373-11377.

[21] N.A. Becker, J.D. Kahn, L.J. Maher, 3rd, Eukaryotic HMGB proteins as replacements for HU in E. coli repression loop formation, Nucleic Acids Res 36 (2008) 4009-4021.

[22] L. Czapla, D. Swigon, W.K. Olson, Effects of the nucleoid protein HU on the structure, flexibility, and ring-closure properties of DNA deduced from Monte Carlo simulations, J Mol Biol 382 (2008) 353-370.

[23] A. Bahloul, F. Boubrik, J. Rouviere-Yaniv, Roles of Escherichia coli histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts, Biochimie 83 (2001) 219-229.

[24] M. Lammi, M. Paci, C.L. Pon, M.A. Losso, A. Miano, R.T. Pawlik, G.L. Gianfranceschi, C.O. Gualerzi, Proteins from the prokaryotic nucleoid: biochemical and 1H NMR studies on three bacterial histone-like proteins, Adv Exp Med Biol 179 (1984) 467-477.

[25] J. Ruiz-Sanz, V.V. Filimonov, E. Christodoulou, C.E. Vorgias, P.L. Mateo, Thermodynamic analysis of the unfolding and stability of the dimeric DNA-binding protein HU from the hyperthermophilic eubacterium Thermotoga maritima and its E34D mutant, Eur J Biochem 271 (2004) 1497-1507.

[26] M. Andrabi, K. Mizuguchi, S. Ahmad, Conformational changes in DNA-binding proteins: relationships with precomplex features and contributions to specificity and stability, Proteins 82 (2014) 841-857.

[27] S.H. Khan, R. Kumar, An overview of the importance of conformational flexibility in gene regulation by the transcription factors, J Biophys 2009 (2009) 210485.

[28] T. Sunami, H. Kono, Local conformational changes in the DNA interfaces of proteins, PLoS One 8 (2013) e56080.

[29] D. Vuzman, Y. Levy, Intrinsically disordered regions as affinity tuners in protein–DNA interactions, Molecular BioSystems 8 (2012) 47-57.

[30] J. Ramstein, N. Hervouet, F. Coste, C. Zelwer, J. Oberto, B. Castaing, Evidence of a thermal unfolding dimeric intermediate for the Escherichia coli histone-like HU proteins: thermodynamics and structure, J Mol Biol 331 (2003) 101-121.

[31] N. Garnier, K. Loth, F. Coste, R. Augustyniak, V. Nadan, C. Damblon, B. Castaing, An alternative flexible conformation of the E. coli HUbeta(2) protein: structural, dynamics, and functional aspects, Eur Biophys J 40 (2011) 117-129.

[32] R. Le Meur, K. Loth, F. Culard, B. Castaing, C. Landon, Backbone assignment of the three dimers of HU from Escherichia coli at 293 K: EcHUα2, EcHUβ2 and EcHUαβ, Biomolecular NMR assignments (2015) 1-5.

[33] H. Vis, U. Heinemann, C.M. Dobson, C.V. Robinson, Detection of a monomeric intermediate associated with dimerization of protein Hu by mass spectrometry, Journal of the American Chemical Society 120 (1998) 6427-6428.

CHAPTER 5

Studies on folding, association and DNA binding of HU-BA fusion

Contents

5.1 Introduction

It is now well known that HU plays significant roles in organization and compaction of the nucleoid and hence that it directly (as well as indirectly) regulates the expression of several genes [1,2,3]. It is believed that in *E. coli* where there are two HU homologues and three different kinds of dimeric populations [HU-A2, HU-B2 and HU-AB], the relative abundances of the dimeric forms also play a role in the a regulatory mechanism. By creating deletion strains lacking either HU-A or HU-B or both, it has been shown that HU-A and HU-B have differential effects on *E. coli* growth [3,4,5,6]. A large number of studies have been done to understand the differences in functional properties of HU dimeric forms and significant differences in DNA binding properties of the three dimeric forms have been observed [7,8,9]; however, no attention has been paid to differences in inter-subunit stabilities and overall stabilities of these dimeric forms.

A very interesting and unexplored aspect of HU-focused studies is the mechanism and regulation of dimer subunit switching leading to formation of the hetero-dimer. Theoretically, two mechanisms exist for the same; (1) Co-translational folding and assembly of the two isoforms to form a hetero-dimer, and (2) Sub-unit swapping between two homo-dimeric molecules, one each of HU-A2 and HU-B2, leading to formation of two hetero-dimeric molecules of HU-AB. We have shown in the previous chapter that the inter-dimer interface is very stable in both the homo-dimers (relative to the rest of the molecules) and that the dimer interface interactions are retained even in at urea concentrations as high as 2 M where the protein is ~80-90 % unfolded. This appears to rule out the possibility of the formation of hetero-dimer by simple swapping of folded subunits between the two folded monomers (although it clearly does not rule out the possibility of subunit swapping involving substantial unfolding of subunits during dissociation, followed by reassembly and folding into hetero-dimers). The existing literature suggests that mixing of two homo-dimers *in vitro* spontaneously results in formation of the hetero-dimer, and this appears to rule out the possibility of occurrence of any facile co-translational folding and assembly, resulting in formation of the hetero-dimer [10]. To understand how the two strongly linked subunits of a homo-dimer come together to form the hetero-dimer, we chose to study the thermodynamic properties of the hetero-dimer and compare them to the homo-dimers.

There is a very major problem with HU-dimerization studies, in that that the HU-dimeric forms have identical sizes and very similar pI values making it difficult to discriminate between and purify the three dimeric forms. HU-A and HU-B homo-dimers can be purified to a reasonable degree by over-expressing one of these, since overexpression of one down-regulates the other, resulting in expression of only a single homo-dimeric form, but still a problem exists with purification of pure hetero-dimeric populations. The methods generally used to differentiate between these dimeric forms are:

(1)Acid Urea Triton (AUT) polyacrylamide gel, which is a denaturing gel that is generally used to resolve different forms of histones based on their native positive charge [11,12].

(2) Native denaturing acrylamide gel which is a native gel that allows migration of protein based on its native positive charge [10].

Both the above techniques are qualitative enough to indicate whether the hetero-dimer is present, if one is able to purify the hetero-dimer and fractionate it electrophoretically to reveal both HU-A and HU-B, but these are not very sensitive and quantitative enough to establish that even low amounts of pure homo-dimers of HU-A2 and HU-B2 are not present in a preparation of HU-AB.

To be sure that the population we are studying is pure hetero-dimer, therefore, we created a genetic fusion construct containing both HU-A and HU-B fused to each other with HU-B preceding HU-A (i.e., with HU-B located at the N-terminal end), and with an 11 amino-acids long ‘serine-glycine’ linker separating the two fusion partners. Figure 1B shows the schematic representation of the fusion protein construct used for all the studies described in this chapter. The fusion construct was made in the hope that it would allow us to physically ‘simulate’ the HU-AB heterodimer in the form of an “HU-B-HU-A fusion” forcing only the hetero-meric interface to form under conditions disfavoring inter-chain interactions(owing to higher probability of intramolecular HU-A-HU-B interactions over intermolecular interactions). Also, since the structures of individual subunits are unlikely to remain folded in a native-like structural format upon dissociation (given the nature and structural features, and extent, of the subunit interface), we assumed that there would be a very high probability for intramolecular assembly of HU-A and HU-B chain segments.

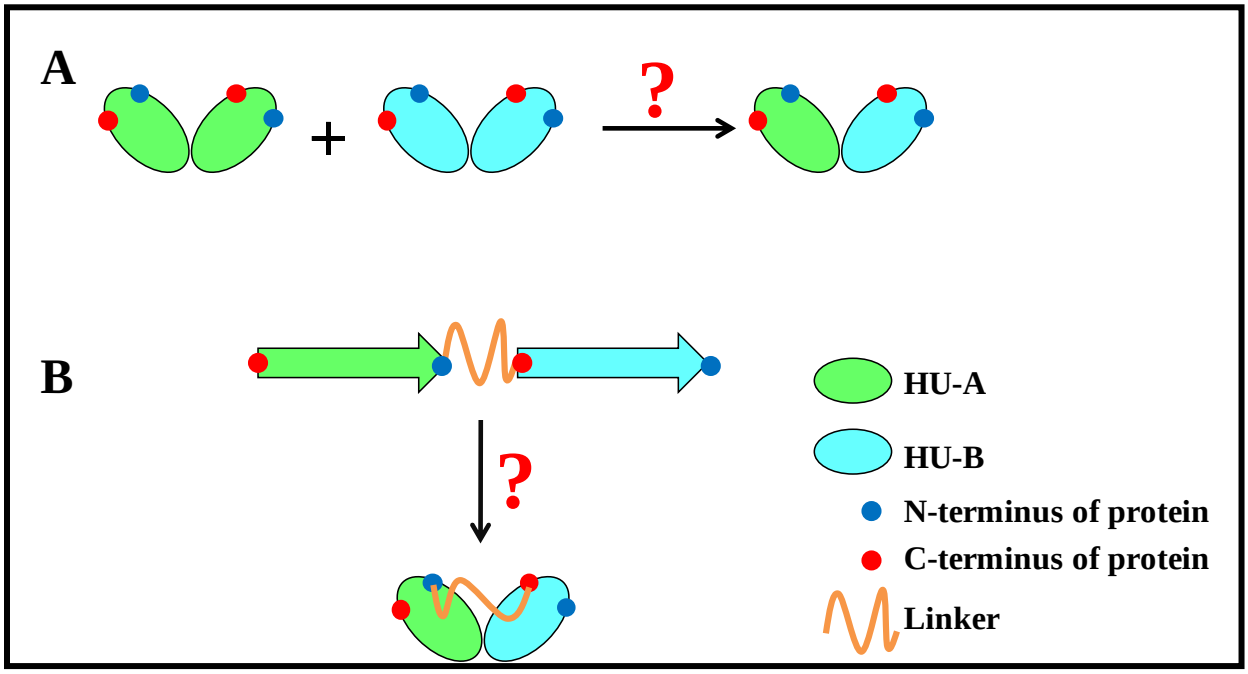


Figure 1: Schematic representation of the rationale of construction of HU-BA fusion. ***Panel A*** shows how HU-A and HU-B dimers align and where the N and C terminal are located in the folded dimeric form. ***Panel B*** shows the components of fusion protein.

5.2 Materials and methods specific to this chapter

5.2.1 Cloning of the fusion construct

In the fusion product, HU-B was fused to N-terminal of HU-A. An 11 residues-long glycine-serine linker (SGGGGSGGGGS) was introduced between the two proteins to allow flexibility of folding at the interface of fusion. The schematic of construction of fusion product is shown in Fig. 2. A few residues of linker were added to C-terminal of HU-B by doing a PCR reaction (PCR1)using the primers HU-B Forward: 5'-AGCTACTGGATCCATGAATAAATCTCAATT GATCG-3' and Linker Reverse: 5'-GCTGCCACCTCCGCCTGAACCTCCTCCACCTGAGTT TACCGCGTCTTTC-3'. By using primers HU-A Reverse: 5'-GAATACTCCCGGGTTACTT AACTGCGTCTTTCAATG-3' and Linker Forward: 5'-AGCGGTGGAGGAGGATCAGGC GGAGGTGGCAGCATGAAC AAGACTCAACTG-3', a few overlapping residues of linker were added to N-terminal of HU-A (by doing PCR-2). The two PCR products were spliced by using HU-B Forward and HU-A Reverse primers in PCR-3. The spliced product was digested and ligated in pQE-30 vector and transformed into XL1-Blue cells, confirmed by sequencing and expressed from the same strain.

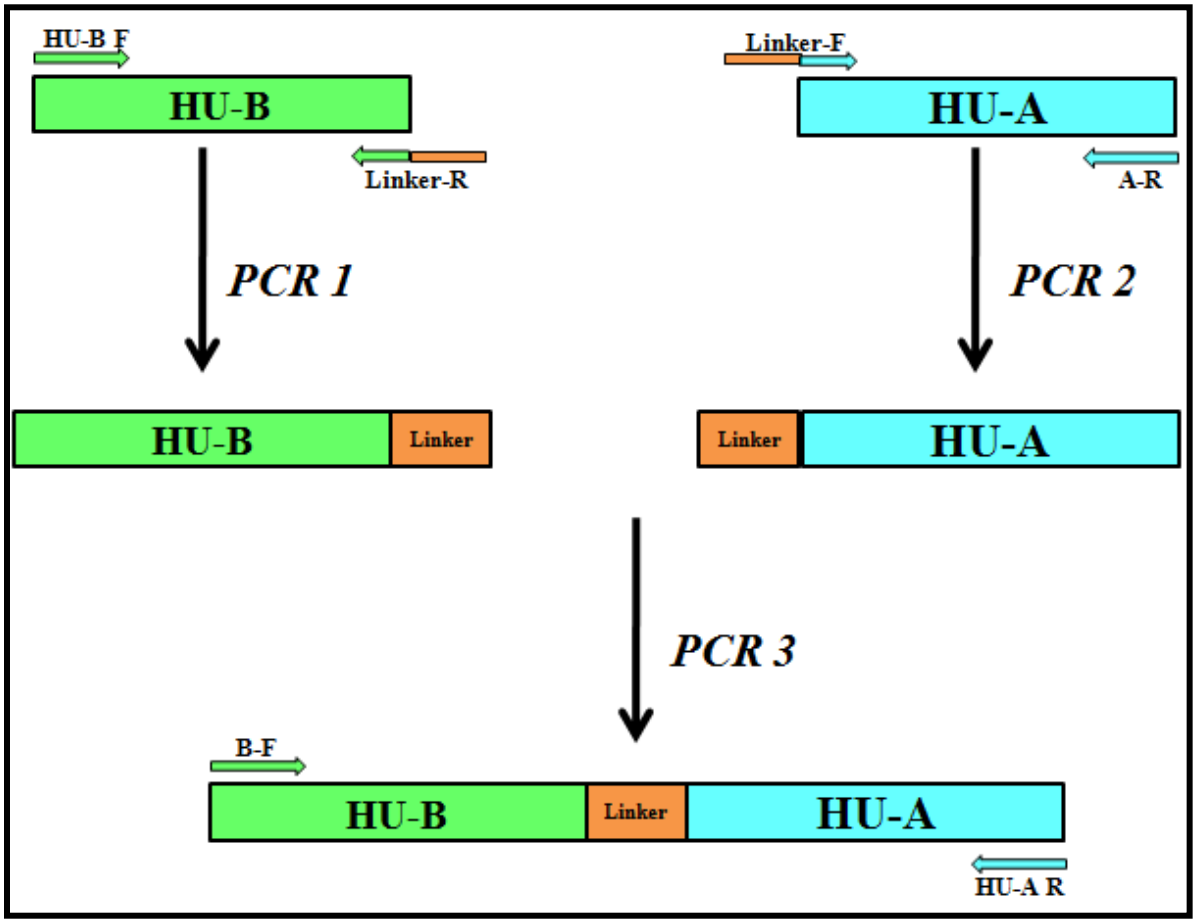


***Figure 2**:* Schematic representation of the cloning methodology used to construct the HU-BA fusion.

5.2.2 Protein purification

The fusion protein was expressed from constitutively expressing XL1-Blue cells and was purified in presence of 1 M NaCl as discussed in chapter 3.

5.2.3 Glutaraldehyde cross-linking

To check the oligomeric status of the fusion protein, 0.5 mg/ml of protein was incubated with 0.01, 0.05 and 1% glutaraldehyde and was incubated for 5 minutes at room temperature before loading on 15 % SDS-PAGE gels.

5.2.4 Circular dichroism studies

CD spectra were collected on Chirascan Spectrophotometer (Applied Photophysics, UK) using a 1 mm path length quartz cuvette and a protein concentration of ~0.5 mg/ml. Both the sample and buffer spectra were collected from 200-250 nm. The spectra obtained were corrected for buffer

background signal. Thermal denaturation was done by heating the protein samples from 20^0C to 90^0C using spectropolarimeter's peltier block arrangement and data was collected at 2^0C intervals. For chemical melting experiments, the protein (0.5 mg/ml) was pre-incubated over-night in respective urea concentrations and CD spectra were collected.

5.2.5 Gibb's free energy calculation

Gibb's free energy calculations were done for the fusion protein by monitoring it's equilibrium chemical denaturation in presence of different urea concentrations. Data was collected and analyzed as described in Chapter 4.

5.3 Results

5.3.1 The fusion product is structured

The first question after making the construct and successfully expressing it and also purifying it was whether the fusion protein of HU-B and HU-A is well-folded? To verify this, the secondary structural content of the protein was examined by far-UV CD spectroscopy. Due to problems with concentration estimation of the fusion product as both the protein lacks both Trp and Tyr residues and single Trp mutants were made either for HU-A or HU-B and not for the fusion protein, the raw CD data was not converted to molar ellipticity values. The far-UV-CD spectra in Figure 3 show that the protein is well-folded and shows a spectra very similar to the HU-A and HU-B dimers suggesting that the fusion product is well-folded.

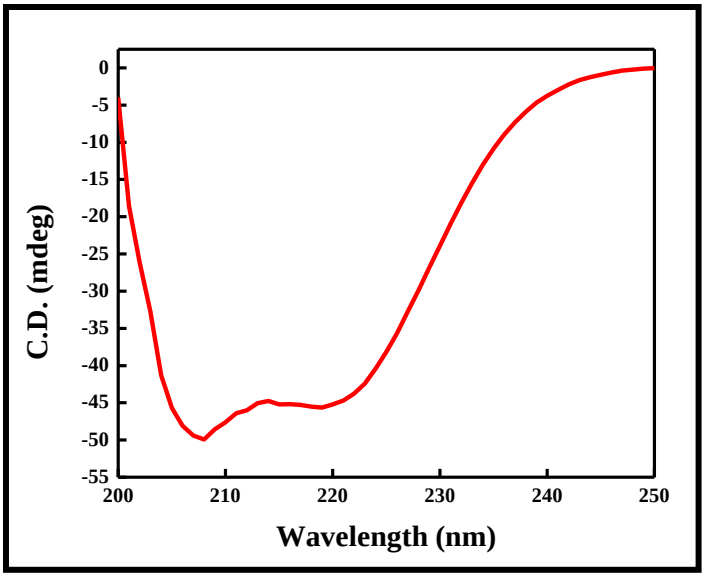


Figure 3: Far-UV C.D. spectra of HU-BA fusion.

5.3.2 The fusion protein is also DNA binding competent

Next, we wanted to examine whether the fusion protein is DNA binding competent. This would tell us whether the protein has folded into native-like conformation or not. As shown in Figure 4, upon binding to DNA (4-WJ), HU-BA fusion leads to a very significant shift in mobility of the DNA which increases with increase in protein concentration, suggesting that the fusion protein is functionally active that is can bind to DNA.

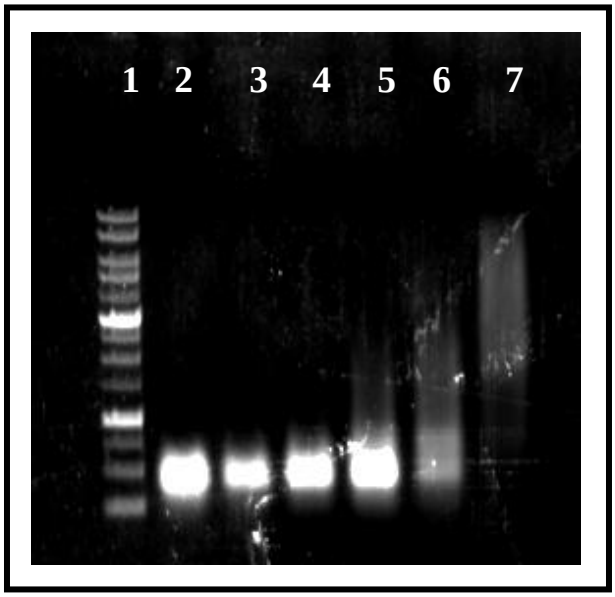


Figure 4: EMSA gel showing binding of HU-BA fusion with 4-WJ. Lane 1 contains the 1 Kb DNA marker, Lane 2 contains the control 4-WJ DNA without protein, Lane 3- Lane 7 contain 4-WJ to which the HU-BA fusion protein had been added in concentrations approximately 5 µM, 10 µM, 15 µM, 20 µM and 25 µM, respectively.

5.3.3 The HU-BA fusion protein is monomeric

We know from our previous experiments that the aggregated form of HU is also DNA binding competent, so we wanted to be sure that the fusion protein is not aggregated. We performed gel filtration chromatography of the fusion protein on Superdex-75, 24 ml column (GE). Figure 5A shows the elution profile of HU-BA fusion from a Superdex-75, 24 ml column and suggests that the fusion protein consisting of HU-B and HU-A chain segments is mainly a monomer and elutes at the same elution volume as dimeric HU, which has been already reported in chapter 3 (elution volume ~12 ml).

Glutaraldehyde cross-linking was also done to know if the HU-BA fusion molecules assemble to form higher-order oligomers. SDS-PAGE for the fusion protein in the absence and presence of 0.1 %, 0.2 % and 0.5% glutaraldehyde is shown in Figure 5B. Lane 1 shows the non

crosslinked HU-BA fusion which corresponds to a ~25 kDa band. Upon cross-linking with glutaraldehyde, the ~25 kDa band shows a higher mobility and shows a band corresponding to ~20 kDa. Increase in mobility could be explained by assumption that extensive intramolecular glutaraldehyde cross-linking is not allowing the molecule to unfold even in the presence of SDS in the sample loading buffer and hence the partially-unfolded molecule migrates faster than the fully-unfolded molecule. Besides the ~20 kDa band, there also is a band at ~40 kDa which should be the dimeric form of the fusion protein.

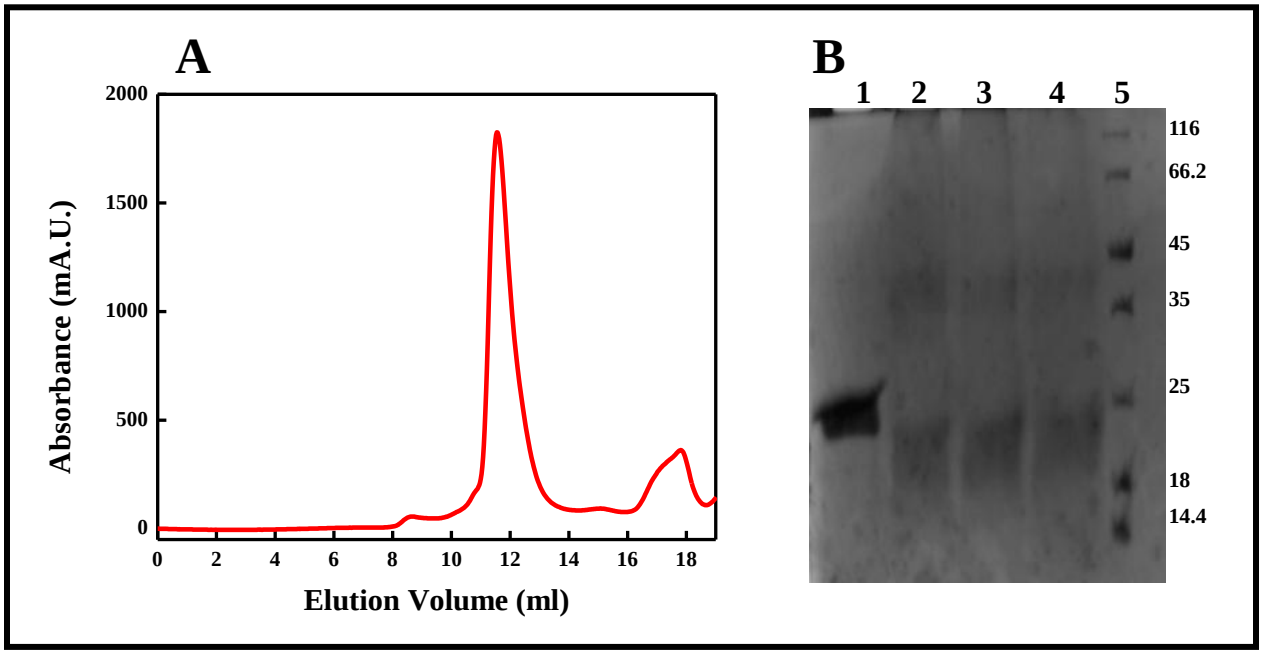


Figure 5: *Panel A:* Gel filtration chromatogram of HU-BA fusion product from Superdex-75, 24 ml column. *Panel B:* SDS-PAGE showing glutaraldehyde cross-linked species of HU-BA fusion. HU-BA in presence of varying amounts of glutaraldehyde Lane 1: 0%, Lane 2:0.1%, Lane 3:0.2 % and Lane 4: 0.5 %.

5.3.4 The fusion protein can be refolded even after heat induced unfolding

We have shown that HU-A and HU-B homodimers can refold back to a very native-like state after heat-induced unfolding. We wanted to see if this can be observed with the fusion protein. For this the protein was unfolded by heating to 90 ^{0}C using the CD instrument's peltier system. A spectrum was collected after 10-15 minutes incubation at this high temperature. Following this, the protein was cooled back to 20 ^{0}C and spectra was collected after 10-15 minutes. Figure 6 shows the comparison of secondary structure content of native (black), unfolded (red) and refolded (blue) HU-BA fusion protein.

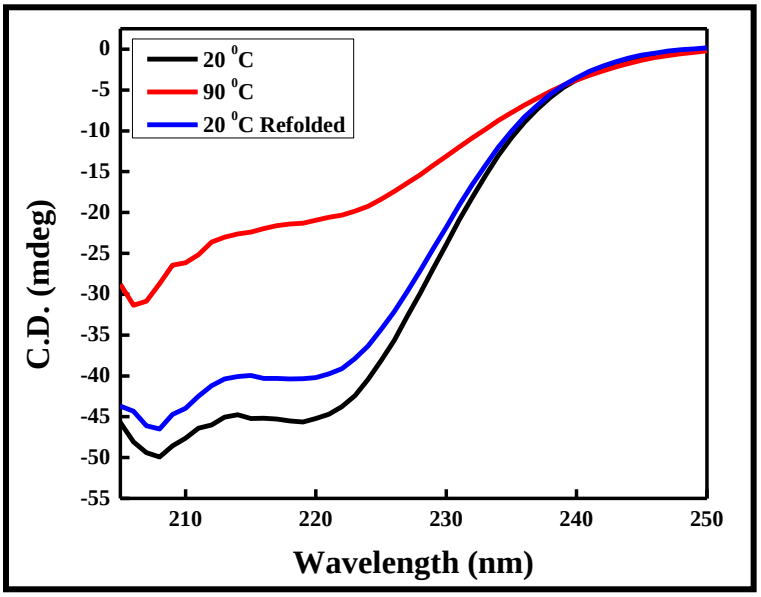


Figure 6:Far UV C.D. spectra of HU-BA fusion protein and 20 ^{0}C (black), unfolding at 90 ^{0}C (red) and refolding when cooled back to 20 ^{0}C (blue)

5.3.5 Two step unfolding of the fusion protein

The fusion protein on melting shows a biphasic transition. Figure 7 shows the changes in Far-UV CD spectra and hence in the secondary structural content of HU-BA fusion protein caused by heating the sample at a constant heating rate of 2 ^{0}C/ minute. Data was fitted to a double sigmoidal curve using the pro-data analysis software supplied with the instrument giving T_m values of 43.69^{0}Cand 71.55^{0}C, respectively. Figure 7 shows the change in secondary structure content (at 222 nm) in black circle and in red is the fitting for double sigmoidal transition. This suggests that there are two independent 'domains' undergoing unfolding, which could be HU-A and HU-B.

5.3.6 Fused (simulated) hetero-dimer is more stable to chemical denaturation

To assess the stability of the fused HU-B-HU-A 'heterodimer-equivalent' construct to chemical denaturation, the protein was incubated overnight in different urea concentrations and loss in structure was monitored by monitoring the changes in CD signal at 222 nm. The ellipticity (mdeg) values for different urea concentrations were plotted and fitted to sigmoidal equation using Origin software to calculate the C_m value. The same data was fitted using equations 3-5 (please see chapter 4) to calculate the ΔG of chemical unfolding, shown in the inset to panel B.

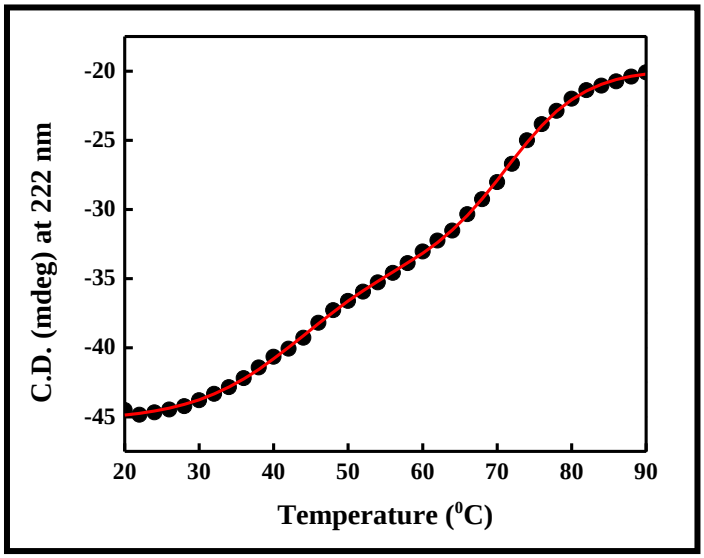


Figure 7:Changes in Far-UV C.D. spectra at 222 nm for the fusion protein, Black circles represent the raw data and red line represents the fit.

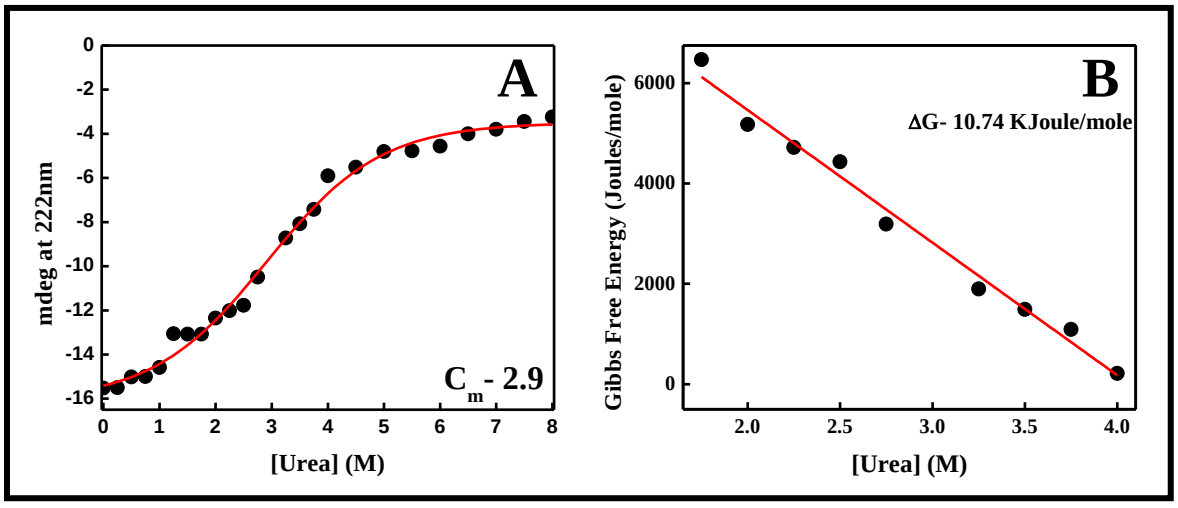


Figure 8: Panel A: Changes in ellipticity at 222 nm as a function of urea concentration ***Panel B:*** straight line fit used to calculate ΔG values. In both curves, the black filled circles the raw data and red line represents the fit.

5.4 Discussion

We created a fusion protein by genetically fusing genes encoding HU-B and HU-A to physically simulate the formation of a hetero-dimer. We next tested the fusion construct for structural content and DNA-binding ability, and compared it to the HU-A and HU-B homo-dimers. The first concern with the fusion construct was the folding of this novel engineered construct. An 11 residues-long glycine-serine linker was introduced between the C-terminal of HU-Band N-terminal of HU-A, so as to allow flexibility at the interface sufficient to favor association of the two chain sections corresponding to HU-B and HU-A. We wanted to see whether the linker allows sufficient conformational flexibility for the two parts of fusion (HU-A and HU-B) to assemble into a functional hetero-dimer. Folding of the fusion construct was checked by looking at the secondary structure of the molecule, which suggested that the newly formed fusion product was indeed folded. Next, we wanted to know if the folded fusion molecule is capable of DNA binding. For this purpose EMSA was performed and we found that the fusion product is DNA binding competent. We further investigated the oligomeric status of the molecule by using gel-filtration chromatography and glutaraldehyde cross-linking. Both gel filtration and cross-linking experiments suggested that the fusion protein is mainly a monomer with a very little dimeric population. This clearly points towards intra-molecular assembly of HU-A and HU-B to form a well-structured and DNA binding competent interface of HU-A and HU-B chain sections.

After having established that the fusion protein is indeed a true folded and DNA binding competent hetero-dimer, we went ahead with characterization of stability of the protein against thermal and chemical unfolding. Table 1 shows the comparison of T_m, C_m and ΔG for HU-A homo-dimer, HU-B homo-dimer and HU-AB hetero-dimer.

	T_m (°C)	C_m (M)	ΔG (KJ/mole)
HU-A	41.59	1.77	3.26
HU-B	50.53	2.14	6.84
HU-AB	T_m1- 43.69 T_m2- 71.55	2.91	10.74

Table 1: Table showing T_m, C_m and ΔG values for the homo and hetero-dimeric HU species.

Unlike, the homo-dimers, the fusion hetero-dimer shows a biphasic unfolding transition upon heating. T_m values calculated for the transitions are 43.69 ^{0}C and 71.55 ^{0}C, respectively, which gives an average T_m of 57.6 ^{0}C. The T_m value of the fused hetero-dimer is much higher than either of the homo-dimers. This points to increased stability of the fusion hetero-dimer in comparison to the homo-dimers. Similarly, C_m and ΔG for the fused HU-AB hetero-dimer are higher than both HU-A homo-dimer and HU-B homo-dimer. The thermodynamics data suggests that the fused hetero-dimer is thermodynamically more stable than either of the homo-dimer. For the actual homo-dimer (without linker), the stability could vary because of the effective concentration of the two constituent subunits is higher in comparison to that for an unlinked homo or heterodimer.

We know from our previous results that the inter dimer interface is very stable in the homo-dimers but still the hetero-dimerization is favored *in vitro* as reported earlier [10]. We propose that, if the unlinked hetero-dimer also displays higher thermodynamic parameters, the increased thermodynamic stability of the hetero-dimer than the homo-dimers could be the favoring factor for hetro-dimerization of HU-A and HU-B.

5.5 References

[1] M. Macvanin, S. Adhya, Architectural organization in E. coli nucleoid, Biochim Biophys Acta 1819 (2012) 830-835.

[2] K. Skarstad, T.A. Baker, A. Kornberg, Strand separation required for initiation of replication at the chromosomal origin of E.coli is facilitated by a distant RNA--DNA hybrid, EMBO J 9 (1990) 2341-2348.

[3] A.I. Prieto, C. Kahramanoglou, R.M. Ali, G.M. Fraser, A.S. Seshasayee, N.M. Luscombe, Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12, Nucleic acids research (2011) gkr1236.

[4] A. Jaffe, D. Vinella, R. D'Ari, The Escherichia coli histone-like protein HU affects DNA initiation, chromosome partitioning via MukB, and cell division via MinCDE, J Bacteriol 179 (1997) 3494-3499.

[5] M. Wada, Y. Kano, T. Ogawa, T. Okazaki, F. Imamoto, Construction and characterization of the deletion mutant of hupA and hupB genes in Escherichia coli, J Mol Biol 204 (1988) 581-591.

[6] N. Goshima, K. Kohno, F. Imamoto, Y. Kano, HU-1 mutants of Escherichia coli deficient in DNA binding, Gene 96 (1990) 141-145.

[7] H. Shindo, A. Furubayashi, M. Shimizu, M. Miyake, F. Imamoto, Preferential binding of E.coli histone-like protein HU alpha to negatively supercoiled DNA, Nucleic Acids Res 20 (1992) 1553-1558.

[8] D. Kamashev, A. Balandina, J. Rouviere-Yaniv, The binding motif recognized by HU on both nicked and cruciform DNA, EMBO J 18 (1999) 5434-5444.

[9] V. Pinson, M. Takahashi, J. Rouviere-Yaniv, Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA, J Mol Biol 287 (1999) 485-497.

[10] J. Ramstein, N. Hervouet, F. Coste, C. Zelwer, J. Oberto, B. Castaing, Evidence of a thermal unfolding dimeric intermediate for the Escherichia coli histone-like HU proteins: thermodynamics and structure, J Mol Biol 331 (2003) 101-121.

[11] J.H. Waterborg, Acid—Urea—Triton Polyacrylamide Gel Electrophoresis of Histones, The Protein Protocols Handbook, Springer, 2009, pp. 251-262.

[12] L. Claret, J. Rouviere-Yaniv, Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival, J Mol Biol 273 (1997) 93-104.

CHAPTER 6

DNA binding: Introduction of an intrinsic fluorescence probe into HV

Contents

6.1 Introduction

The DNA-binding properties of HU have been studied in considerable detail and the subject has also been discussed in considerable detail in the introductory chapter of this thesis. HU shows varying binding affinities for different DNA templates, different DNA/protein ratios and under different experimental conditions[1]. So far, studies of HU binding to DNA rely largely on the electrophoretic mobility shift assay (EMSA) or on single-molecule studies, such as force spectroscopy experiments.

Usefulness of fluorescence spectroscopy: Tryptophan (Trp) fluorescence spectroscopy is widely used for characterization of protein–DNA interactions and serves several advantages over conventional methods used to estimate DNA-protein interactions[2,3,4]. In comparison to 'in-solution' equilibrium fluorescence studies, the conventionally used methods to calculate binding stoichiometry (like EMSA) have a disadvantage, in that methods like EMSA are non-equilibrium methods [5]. Fluorescence spectroscopy gives the advantage of exploring the effect of various environmental conditions on DNA binding. Besides, fluorescence is also a highly sensitive, non-destructive technique and requires very low concentrations of protein [6]. Hence, it serves the advantage of using very low protein concentrations (even in μM) to accurately estimate binding parameters.

Changes in intensity and wavelength: Moreover, Trp fluorescence properties are sensitive to its local environment. Trp residues are excited at 280, or 295, nm and show emission maxima in the range 330-355 nm depending on the extent of solvent exposure and polarity of the solvent surrounding Trp residue, or residues, in the protein[6]. The more exposed the Trp residue, the longer is the wavelength of the emission maximum. As, nucleic acids do not show any fluorescence contribution of their own in the above range, changes in fluorescent properties of DNA-binding proteins are indicative of DNA binding-induced changes in the properties of Trp fluorescence from the DNA-binding protein. Binding of DNA to a protein is monitored by probing the changes in emission intensity, and wavelength of emission maximum, of its Trp residues, for the following reason. The extent of change in fluorescence properties upon binding to DNA depends on the location of Trp residue(s) in the protein with respect to its DNA binding site, and the conformational change introduced in the protein upon DNA binding, or change in the environment of one or more of its Trp residues. DNA binding can lead either to quenching of

Trp fluorescence(decrease in the intensity of fluorescence emission) or in burial of the Trp residue(s) (changes in the protein's emission maximum) or a combination of both.

Changes in quenchability: Fluorescence quenching is directly proportional to the amount of quencher added. Extent of quenching of fluorescence upon formation of a DNA-protein complex is also directly proportional to the amount of DNA bound to the protein. These differences in fluorescence properties on free protein and protein-DNA complex are exploited to calculate the stoichiometry of binding and the equilibrium binding constants. Fluorescence spectroscopy studies can also give an indication of differences in the local environment for different residues in a protein upon DNA binding, especially in proteins containing multiple Trp residues[7,8].

Changes in fluorescence anisotropy: Binding of DNA to a protein results in the formation of a DNA-protein complex with a higher molecular size and hence the rotational diffusion of Trp residues in the protein molecules also undergo changes. Thus, fluorescence anisotropy is also used to study DNA-protein interactions. Fluorescence anisotropy is a measure of extent of rotation of the fluorophore during the excited state lifetime and it depends upon the size of the molecule. Binding of DNA to protein increases its size and hence decreases the extent of rotation of excited fluorophore (Trp residue), leading to an increase in the anisotropy value. Fluorescence anisotropy has been widely used to obtain binding parameters for a large number of DNA-binding proteins and domains [9,10,11,12,13,14].

Lack of Trp and Tyr residues in HU: HU proteins(both HU-A and HU-B) lack both tryptophan as well as tyrosine (Tyr) residues. We created single Trp residue containing mutants, by replacing phenylalanine (Phe) residues at different positions in the natural sequence of the protein. *E. coli* HU (both HU-A and HU-B) contains three conserved Phe residues at positions 47, 51 and 79. To cause minimum structural perturbations due to introduction of a bulky Trp side chain, we chose to replace Phe residues with Trp residues in all positions, and examine the different effects of each introduction to identify mutants suitable for use in DNA-binding studies.

Locations of Phe residues in HU: The structure of the DNA-HU complex is known only for *Anabaena* HU (1P51). Through structure alignment of the known structures of *Anabaena* HU and *E. coli* HU (2O97) it can be seen that the two show very high structural similarity (Figure 2A), such that analyses of the former can be relied up to derive inferences concerning the latter protein. The positions of the three Phe residues is also conserved between the two sequences, and

structures. This suggests that just as in *Anabaena* HU, in *E. coli* HU also the three Phe residues are located near the DNA-binding region of the protein, but pointing away from the DNA-binding region and not engaged in direct van der Waals level contacts with the DNA. Residue F47 is present in the loop connecting β-strand 1 and β-strand 2 and residue F51 and F79 are located on the β-strand 2 and β-strand 3 and point inwards, away from the DNA (Figure 2B). Phe residues of the two dimers come together, pointing towards each other in the core of the molecule, to form an aromatic cluster (Figure 2C).

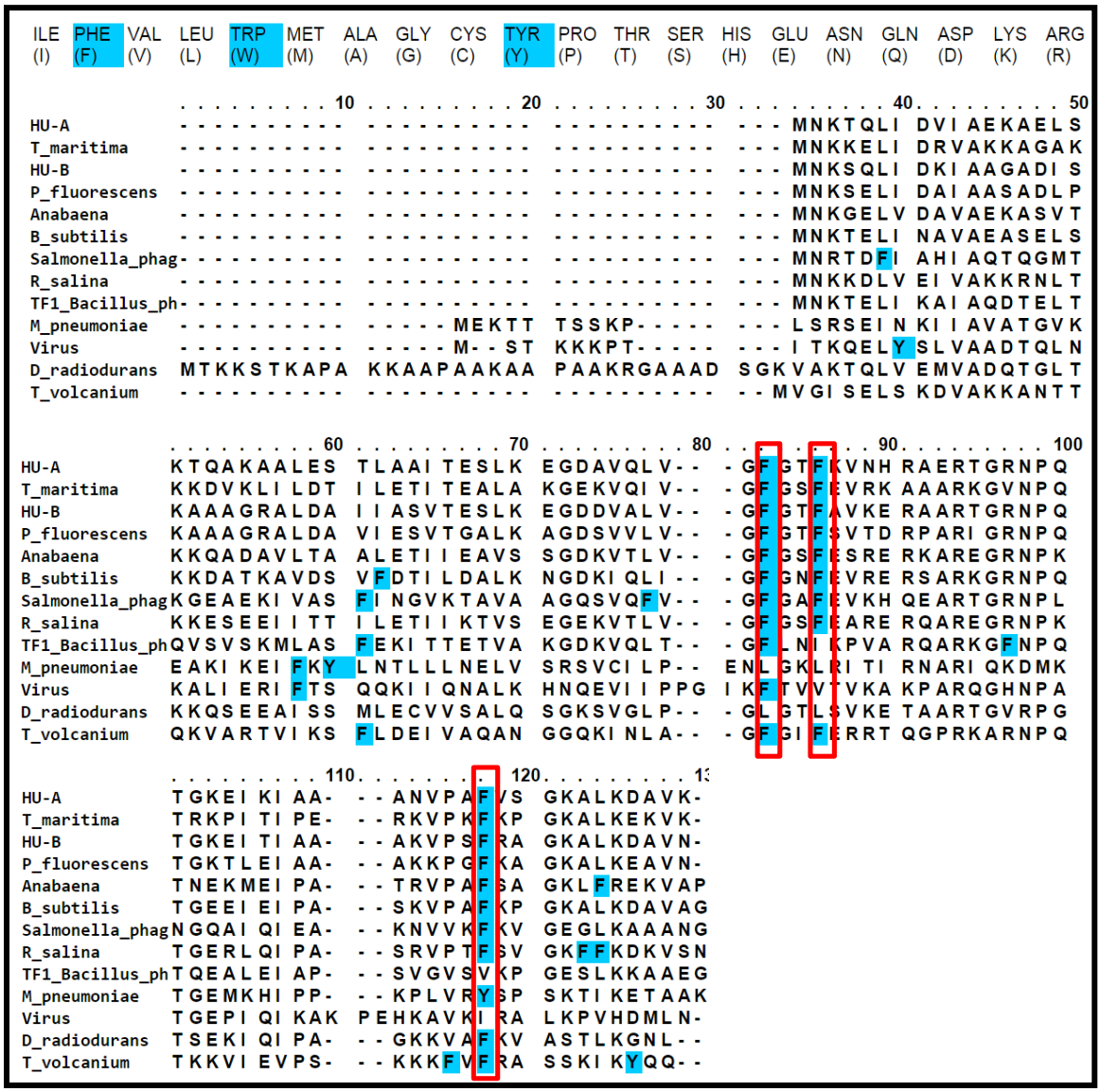


Figure 1: Sequence alignment of HU form various organisms showing that only aromatic amino acid present in most of the HU from various organism is Phenylalanine (F) and Tryptophan (W) and Tyrosine (Y) are missing

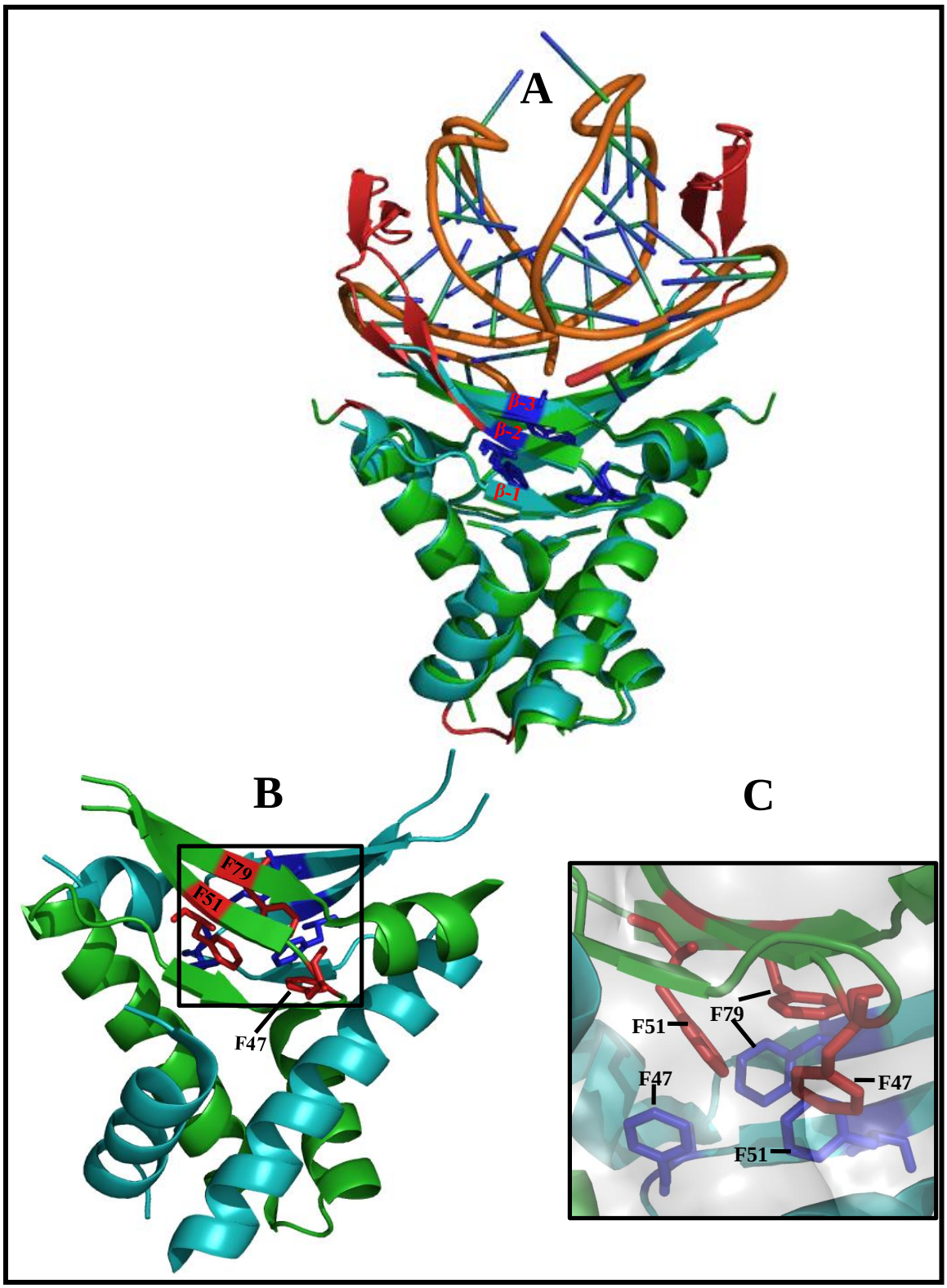


Figure 2: Panel A:Comparison of structures of *Anabaena* HU (1P51) and *E. coli* HU, ***Pane** B:* position of Phe residues in HU chains ***Panel C:*** Arrangement of Phe rings in core of HU.

We chose to replace Phe residues for the following reasons: (a) A Phe-Trp mutation should not cause much structural perturbation as Phe and Trp side chains are both aromatic (one based on a phenyl moiety, and the other on an indole moiety), and the Phe side chain has a methyl group attached to the phenyl ring, whereas the indole of Trp has no extra attached groups. In naturally-occurring proteins, Trp and Phe often substitute each other, and these are considered to be conservative substitution replacement mutations. (b) As, Phe residues are located close to the DNA binding region of HU, introduction of Trp residues in place of Phe would be expected to result not just in Trp fluorescence in HU, but also in differences in the Trp-based fluorescence properties of HU upon DNA binding.

A total of 5 different mutants were created; HU-A F47W, HU-A F51W, HU-A F79W, HU-B F47W and HU-B F79W. All these mutant proteins were purified using the salt-based affinity purification method described in chapter 3. Structural and functional characterization of the mutant proteins was done, and compared to the wild type protein to check if the replacement of Phe residues by Trp residues had caused any structural perturbations in the protein. Only the mutants showing properties (both structural and DNA binding) similar to the wild type protein were used for further studies on DNA binding characterization using Trp fluorescence spectroscopy as a probe.

Further to our intention of creating intrinsically fluorescent HU, there was another reason for our wanting to introduce one or more Trp residues in HU. Evolutionarily speaking, Trp residues are absent in HU proteins from various organisms and also from other DNA-binding proteins that are closely associated with DNA. Eukaryotic histones also lack Trp residues[15,16]. We wanted to understand the reason for the absence of Trp residues in these architectural DNA-binding proteins which are closely associated with DNA. By using the Trp mutants, we explored the effect of tryptophan insertion upon DNA integrity, in respect of the likelihood of photosensitized oxidative modifications occurring when Trp residues are present in close proximity to DNA over long durations. On exposure to ultra-violet light, reactive oxygen species are generated inside the cells, which can read to photo-oxidation of Trp residues. The action of reactive oxygen species (ROS), by attacking the indole ring of the Trp residue, can lead to formation of several well-known tryptophan oxidation products, for example 5-hydroxy-tryptophan (5-HTP), oxindolylalanine (Oia), 3-hydroxyanthranilic acid (3-HAA), N-formyl-kynurenine (NFK),

kynurenine (Kyn), 3-hydroxy-kynurenine (HKyn), and hydroxy-N-formyl-kynurenine. Of these, 3-HAA and HKyn are well known to cause DNA damage and hence act as carcinogens and are associated with many diseases [17,18].

6.2 Materials and methods

6.2.1 Cloning of mutants

For cloning of single Trp mutants of HU-A and HU-B, a combination of splicing by overlap extension PCR (SOE-PCR) and restriction digestion followed by ligation were used. Presence of a KpnI site between residues encoding Phe-47 and Phe-51 was used to omit extra PCR steps. DNA fragments containing mutations F47W and/or F51W, as well as a KpnI site at the 5' or 3' ends, respectively, and other restriction sites (HindIII or BamHI) at the opposite end were digested and cloned into pQE30 vector containing the wild type gene, using KpnI and HindIII sites for F47 W and BamHI and KpnI sites for F51W (Figure 4). To mutate Phe-79, SOE PCR was used as presented in Figure 4. Clones were screened to identify those containing the correctly-sized insert by doing colony PCR using pQE30 promoter (forward) and pQE30 terminator (reverse) primers. Plasmids were purified from clones showing correctly-sized bands in colony PCR and were sequenced to confirm the mutation.

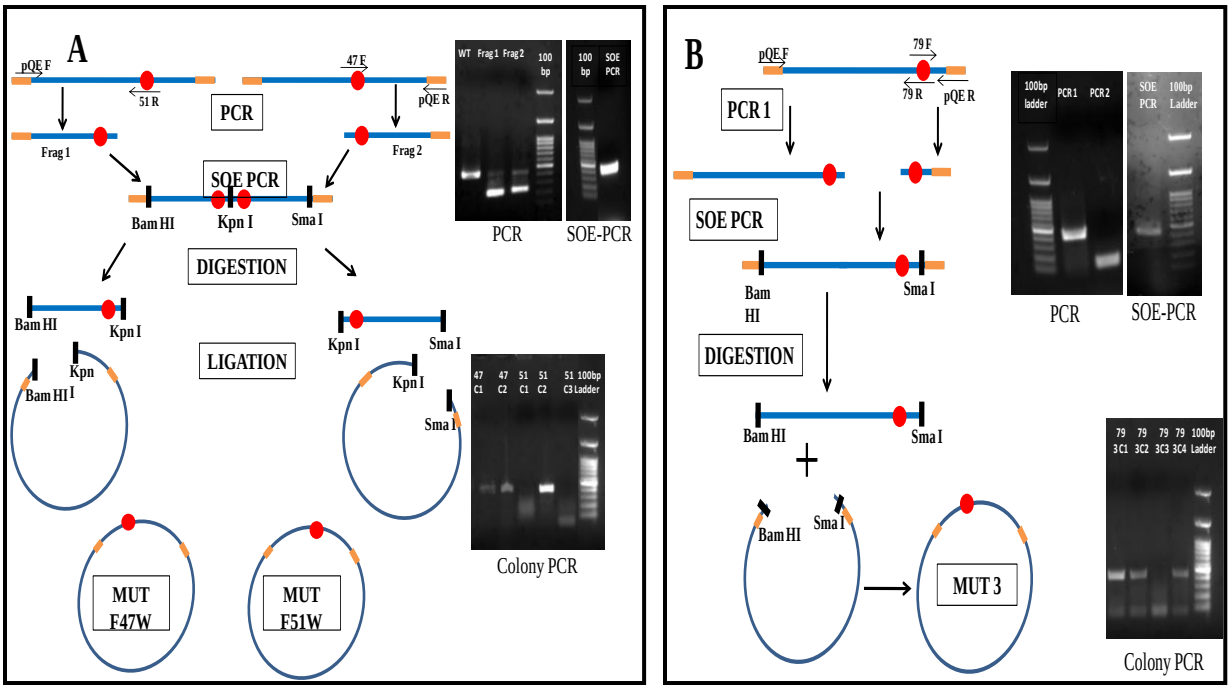


Figure 3: Cloning methodology used for Phe-Trp mutations.

6.2.2 Protein purification and gel-filtration chromatography

Mutant proteins were purified using the salt-based method as described in Chapter 3. Buffer exchange was done to remove imidazole from the Ni-NTA affinity-purified protein by doing gel-filtration chromatography on a Superdex-75 column (24 ml) from GE healthcare. The buffer system used was phosphate buffer saline (PBS) for all the experiments.

6.2.3 Electrophoretic mobility shift assay (EMSA):

To check DNA-protein interactions, the change in the mobility of the DNA-protein complex in reference to DNA alone was analyzed by electrophoresis on ethidium bromide containing 0.5% agarose gels. Gels were imaged using a Bio-rad Eazy Imager gel documentation system. For all EMSA experiments, a synthetic 4-way junction was used as the DNA template. 2 µM 4-way junction was used for all the standard EMSA experiments and the protein concentrations were varied between 0 and 30 µM according to the needs of the experiment. The 4-way junction itself was created by reconstituting the following oligonucleotides (shown below) in the manner described previously [19]:

Strand 1	CCCTATAACCCCTGCATTGAATTCCAGTCTGATAA
Strand 2	GTAGTCGTGATAGGTGCAGGGGTTATAGGG
Strand 3	AACAGTAGCTCTTATTCGAGCTCGCGCCCTATCACGACTA
Strand 4	TTTATCAGACTGGAATTCAAGCGCGAGCTCGAATAAGAGCTACTGT

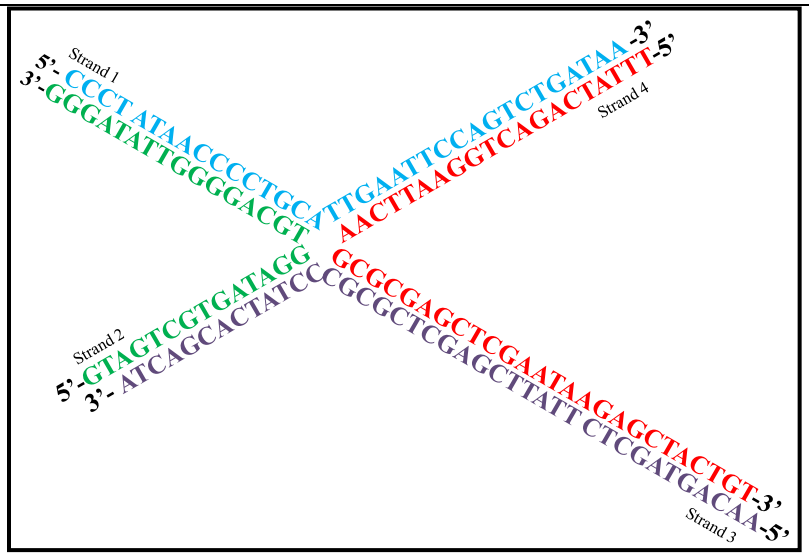


Figure 4:4-way junction resulting from base-pairing of the four oligonucleotides used.

6.3.4 Circular Dichroism studies

CD spectra were collected on a Biologic MOS-500instrument using a 1 mm path length quartz cuvette. For all CD experiments, a protein concentration of 15 µM (for both HU-A and HU-B)and DNA concentration of 1.5 µM was used. Both the sample and buffer spectra were collected in the range of2 00-250 nm and the sample spectra were corrected for buffer background. Thermal denaturation were done on a Chirascan Spectrophotometer (Applied Photophysics, UK)by heating the protein samples from 20^0C to 90^0C using the spectropolarimeter's peltier block arrangement and data was collected at 2^0C intervals. Raw ellipticity obtained was converted into mean residue ellipticity [θ] using the following formula:

$$[\boldsymbol{\theta}] = \frac{\boldsymbol{\theta_{obs}(mdeg)X100XMRW}}{\mathbf{1000Xconcentration(mg/ml)Xpathlength\ (cm)}} \qquad \textbf{Equation 1}$$

Where, MRW is the mean residue weight (molecular weight of protein/ number of amino acids) and θ_{obs} is the raw ellipticiy.

6.3.5 Steady-state fluorescence and anisotropy studies

Fluorescence measurements were performed on a FluoroMax-4 spectrofluorimeter (Horiba Jobin Yvon, NJ). Protein concentration of 15 µM and DNA concentration of 1.5 µMwere used for all the fluorescence experiments. Excitation wavelength was set at 280 nm and excitation and emission slit widths were 0.8 nm and 5 nm, respectively. Trp fluorescence anisotropy measurements were performed by setting λ_{ex} at 280 nm (bandpass 2 nm) and λ_{em} to 350 nm (bandpass8 nm). An integration time of 1 sec was used to obtain a satisfactory signal-to-noise ratio. The steady-state fluorescence anisotropy was estimated using parallel ($I_{||}$) and perpendicular ($I_{\perp}$) intensities and the G- factor as follows [6]:

$$\boldsymbol{r = (I_{||} - I_{\perp}G)/(I_{||} - 2I_{\perp}G)} \qquad \textbf{Equation 2}$$

6.3.6 Fluorescence titration studies

As quenching of protein fluorescence is directly proportional to the concentration of quencher (DNA), quenching of fluorescence can be used to calculate the amount of free and bound protein at a particular DNA concentration. For calculation of binding constant for HU-Trp mutants with 4-way junction DNA, 10 µM of protein sample was titrated with 0.2 µM increments of the 4-way

junction reagent (4-WJ) and the quenching of fluorescence intensity at 340 nm was monitored. To minimize the dilution during titration,100 µM of 4-WJ stock solution was used, leading to addition of only 2 µl of DNA in each titration in a total volume of 1000 µl. Corrections were made to take care of dilution based concentration changes on addition of DNA. Binding constant *k* was calculated from the following equations:

$$Q = \frac{F_0 - F}{F_0} \qquad \textbf{Equation 3}$$

$$\log Q = n \log[DNA] + \log k \qquad \textbf{Equation 4}$$

Here, F_0 is the initial fluorescence intensity at 340 nm and F is the fluorescence intensity after addition of DNA, is the number of binding sites and k is the binding constant[20,21].

6.3.7 Fluorescence quenching experiments

Quenching of Trp fluorescence was monitored by adding increasing amounts of acrylamide (0-250 mM) to 20 µM protein solution or 20 µM protein pre-incubated with 0.2 µM 4-WJ solution. Samples were excited at a wavelength of 280 nm (slit-width 2.5 nm) and emission was monitored at 340 nm (slit-width 5 nm) using the Varian Cary eclipse spectrofluorimeter (Agilent). Data was analyzed using the Stern-Volmer equation [6]:

$$\mathbf{F/F_o = 1 + K_{sv}[Q]} \qquad \textbf{Equation 5}$$

Here, F and F_o are the fluorescence intensities in the presence and absence of quencher, respectively, K_{sv} is the Stern-Volmer quenching constant and [Q] is the molar quencher concentration.

6.3.8 Formaldehyde gel

1 % agarose gels were prepared in MOPS buffer containing 37 % formaldehyde and EtBr solution. Protein samples before and after irradiation were loaded on the gels and electrophoresed. Gels were imaged using Bio-rad gel-doc Eazy imager.

6.3 Results

6.3.1 DNA binding of WT HU

Concentration-dependent binding of WT-HU protein with the 4-way junction was first checked by using EMSA gels. Figures 5A and 5B show binding of different concentrations of HU-A and HU-B proteins, respectively to 2 μM 4-WJ DNA. It is very clear from the EMSA gels that both HU-A and HU-B bind to 4-WJ even at concentrations as low as 0.5 μM and there is a shift in mobility of 4-WJ as compared to the control 4-WJ band. Upon DNA binding, a smear is observed instead of a clean band indicating that DNA molecules are bound by different numbers of protein molecules and this variation in numbers of protein molecules bound to the DNA template is seen as a continuum of bands, i.e. as a smear. The size of the complex formed increases with increase in concentration of protein. The binding saturates at~15 μM for HU-A protein and at 10 μM for HU-B protein. At higher concentrations of HU-B, a lot of DNA was also observed in the wells. This indicates that HU-B unlike HU-A forms some large protein-DNA complexes which get stuck inside the wells and fail to migrate into the 0.5 % agarose gel used for the EMSA experiment.

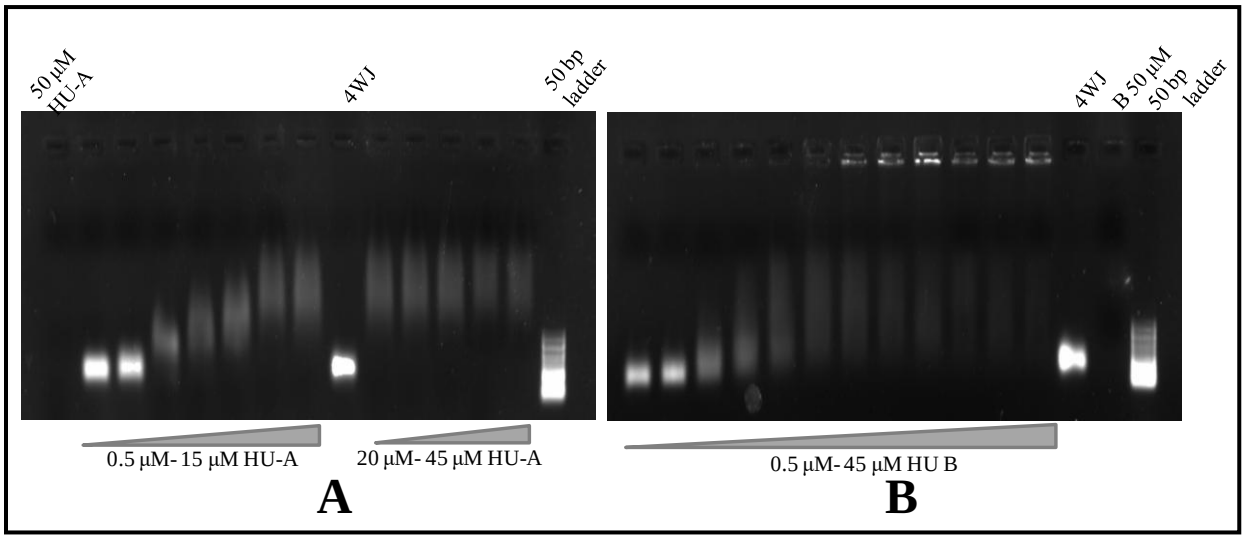


Figure 5: EMSA gel showing binding of HU-A and HU-B with 4-WJ.

6.3.2 The mutant proteins (barring one) are all dimeric

In order to validate that Phe to Trp mutations do not affect the tertiary structure of the mutant proteins, the oligomeric status of the mutant proteins was checked. We have previously shown that HU WT-protein always exists as a dimer; therefore, to check whether the mutants retained the ability to form dimers, gel filtration chromatography was done for the mutant proteins on a GE superdex-75 (24 ml) column. Dimeric WT-protein elutes at approximately 12 ml from a 24 ml bed volume superdex-75 column. Mutant HU-A F47W (red trace in Figure 6A), HU-A F79W (blue trace in Figure 6A), HU-B F47W (red trace in Figure 6B) and HU-BF79W (blue trace in Figure 6B) proteins show peaks near, or at, an elution volume of 12 ml, indicating that these proteins exists mainly as dimers. The fractions corresponding to the dimeric protein were collected and used for all the other experiments. HU-A F51W protein, unlike the WT and the other mutant proteins (magenta trace in Figure 6A), elutes at the void volume indicating that this particular mutation interfered with the structure of the molecule and lead to formation of higher oligomers or soluble aggregates.

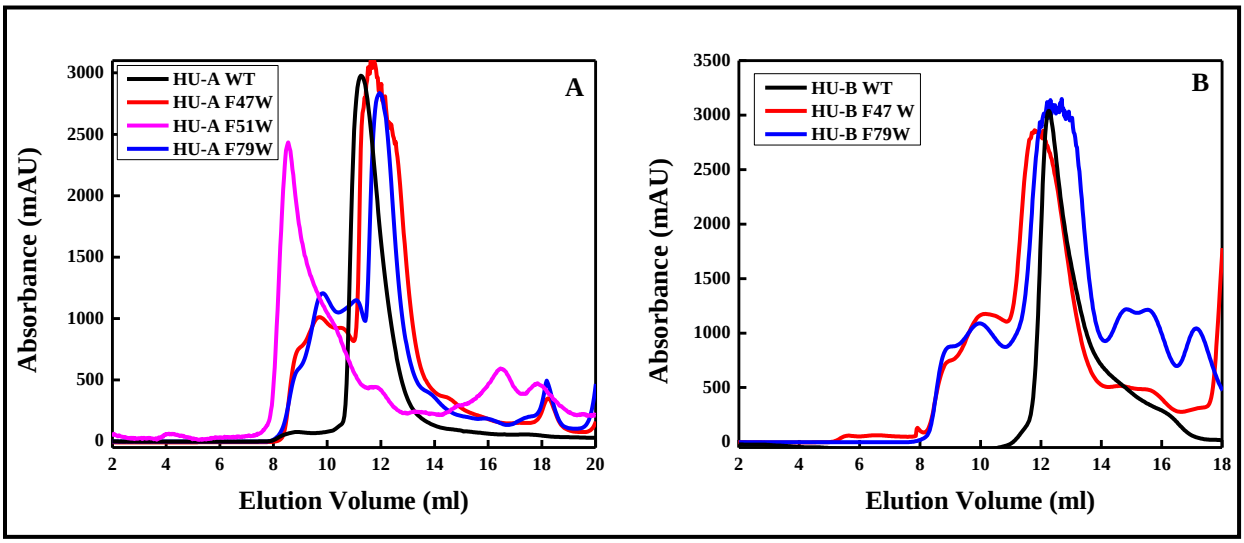


Figure 6: Gel- filtration chromatograms of HU-A (Panel A) and HU-B (Panel-B) on superdex-75 (24 ml) column.

6.3.3 Mutant proteins can refold after thermal denaturation

As described in chapter 4, both HU-A and HU-B proteins can refold after heat induced denaturation. Unfolding and refolding of mutant HU proteins was checked by monitoring the changes in secondary structure of the mutant proteins (222 nm signal). Figure 7 shows that the mutants HU-A F47W (Panel A), HU-A F79W (Panel B), HU-B F47W (Panel C) and HU-B F79W (Panel D) do not unfold completely and the M.R.E. values drop from approximately -7000 to approximately -3000 for all the mutants (red trace in all panels) and the Tm values are approximately, 47 ^{0}C, 46 ^{0}C, 51 ^{0}C and 55 ^{0}C, respectively. Upon refolding (represented by the blue trace), the proteins refold with only a slight effective loss in secondary structure. This suggests that just like the WT-protein, the mutant proteins retain the property of being able to refold. As the HU-A F51W mutant forms soluble aggregates instead of dimers, the refolding studies were not carried out for this particular mutant.

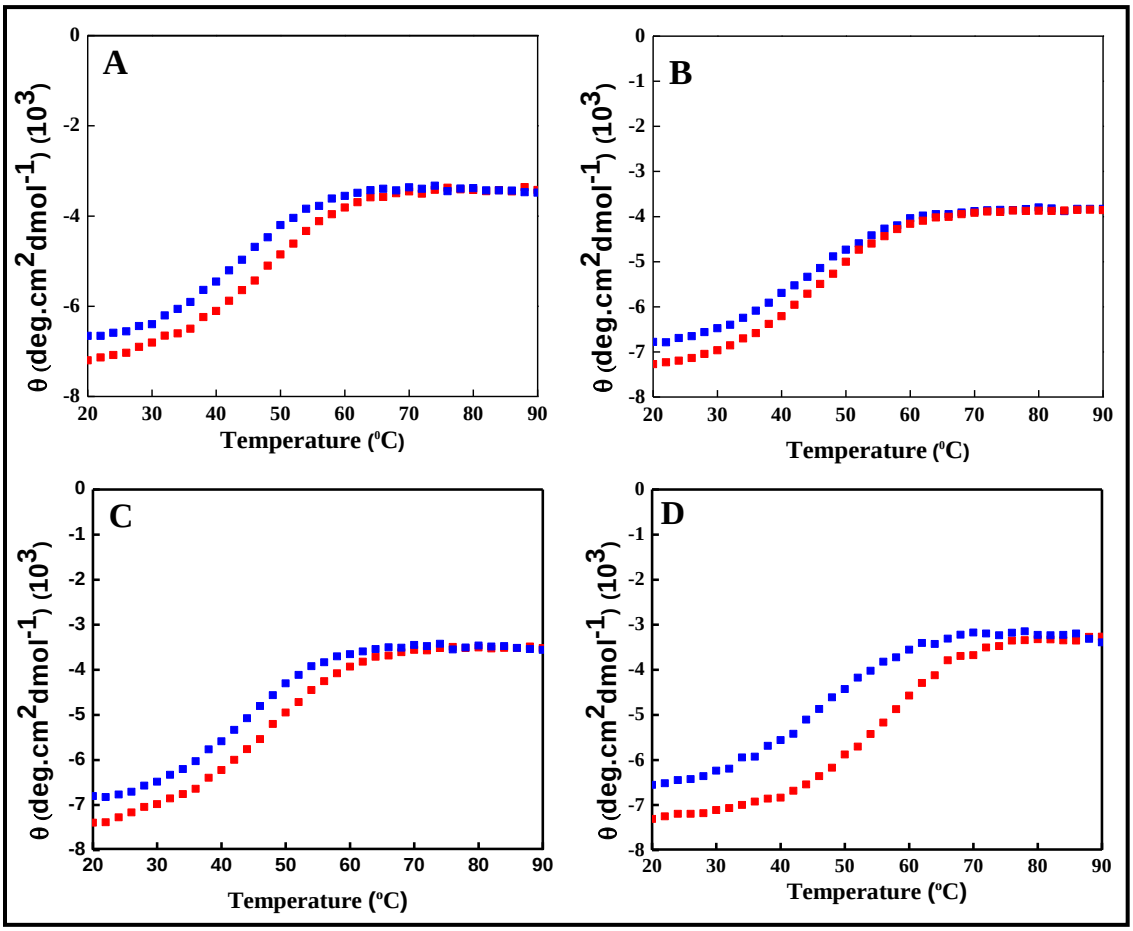


Figure 7: Unfolding (represented in red trace) and refolding (represented in blue trace) of HU-A and HU-B mutants monitored by C.D. ***Panel A*** HU-A F47W, ***Panel B*** HU-A F79W, ***Panel C*** HU-B F47W and ***Panel D*** HU-B F79W.

6.3.4 Mutants are DNA binding competent

DNA-binding of the mutant proteins was checked through EMSA gel experiments. Figure 8 shows the binding of mutant proteins with 4-WJ DNA. The lanes containing only protein were loaded as controls to show that the purified proteins are free of any DNA contamination (with DNA having been removed by the salt-based method developed earlier). Upon addition of these mutant proteins to the 4-WJ, there is a clear shift in the mobility of 4-WJDNA and a smear is observed, indicating that these mutant proteins are indeed DNA binding competent. With increase in concentration of the mutant protein (Figure 9), the size of the DNA protein complex increases just like for the WT-protein. In case of HU-A F47W, upon DNA binding, there is DNA observed in the wells but in the case of HU-B F79W, there are clean bands observed at higher protein concentrations. HU-A F51W protein exists as soluble aggregates (as already mentioned) but the data suggests that the soluble aggregates are also DNA-binding competent, which is quite interesting. A possible explanation which we have not yet followed up is that the mutation transformed the HU-A protein into a form akin to the large multimers normally formed by HU-B at high concentrations (which are also DNA-binding in character, and which help to form tightly compacted nucleoids in the stationary phase).

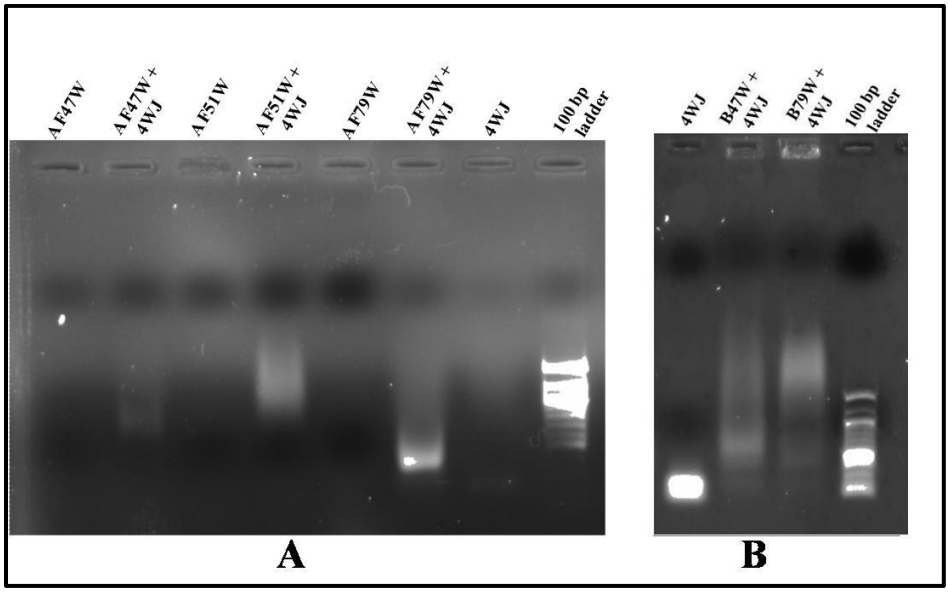


Figure 8: Binding oh HU-A ***(Panel A)*** and HU-B ***(Panel B)*** mutants to 4-WJ

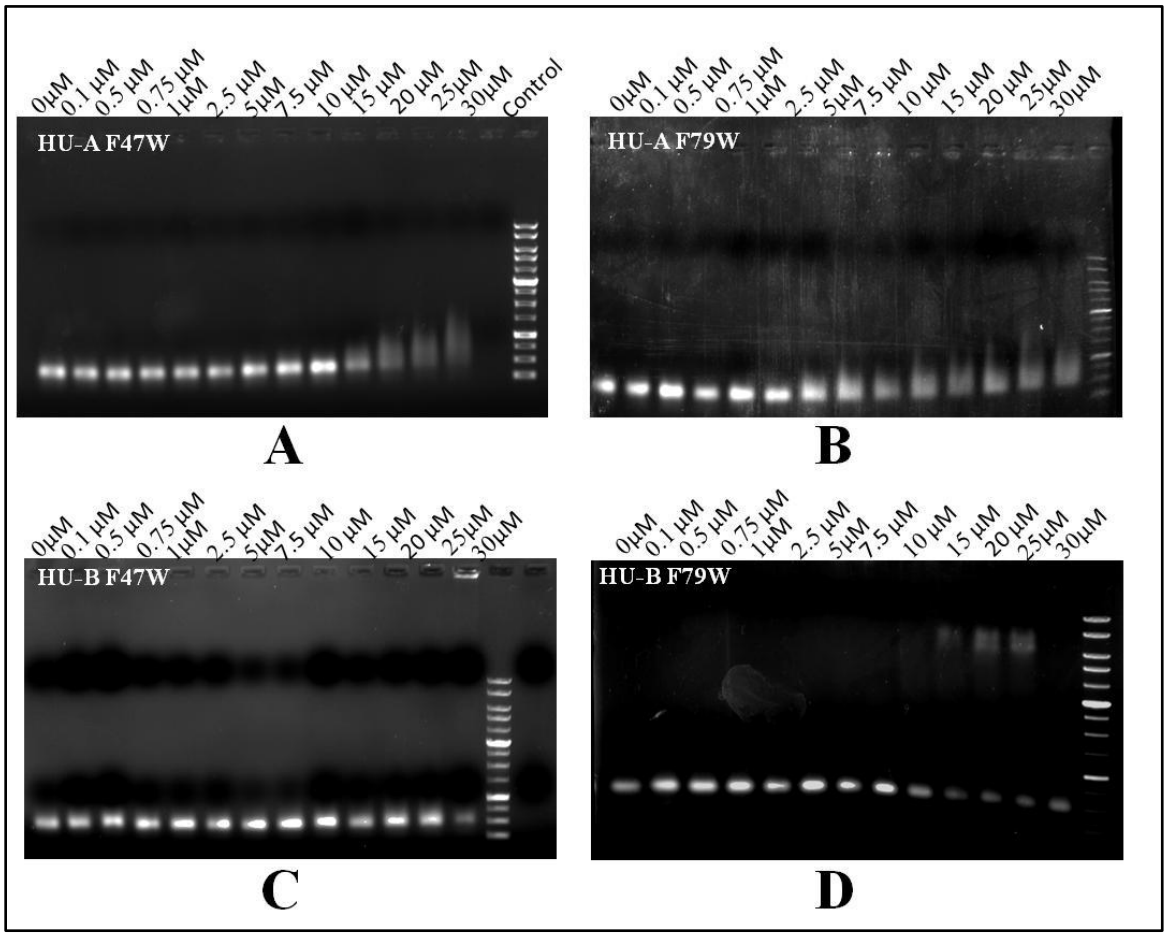


Figure 9**:** Protein concentration dependent binding of HU-A and HU-B mutants to DNA. ***Panel A*** HU-A F47W, ***Panel B*** HU-A F79W, ***Panel C*** HU-B F47W and ***Panel D*** HU-A F79W.

6.3.5 Mutant proteins show changes in CD spectra upon DNA binding

It is known that the β-arm region of HU protein is largely unstructured and becomes structured only in presence of DNA. Structure stabilization, or consolidation/reinforcement, or alteration, occurring within the protein in the presence of DNA was probed by monitoring the changes in secondary structure of protein upon presumptive DNA binding, through addition of DNA (4-WJ) to the mutant HU proteins. CD experiments were done for HU-A F47W (Figure 10A) and HU-A F79W (Figure 10B) mutants. Blue traces show secondary structure of protein in absence of DNA and blue trace shows the changed structure in the presence of DNA. Our data suggests that there is a little stabilization in HU-A F47W protein on addition of DNA, which is not very significant in HU-A F79W.

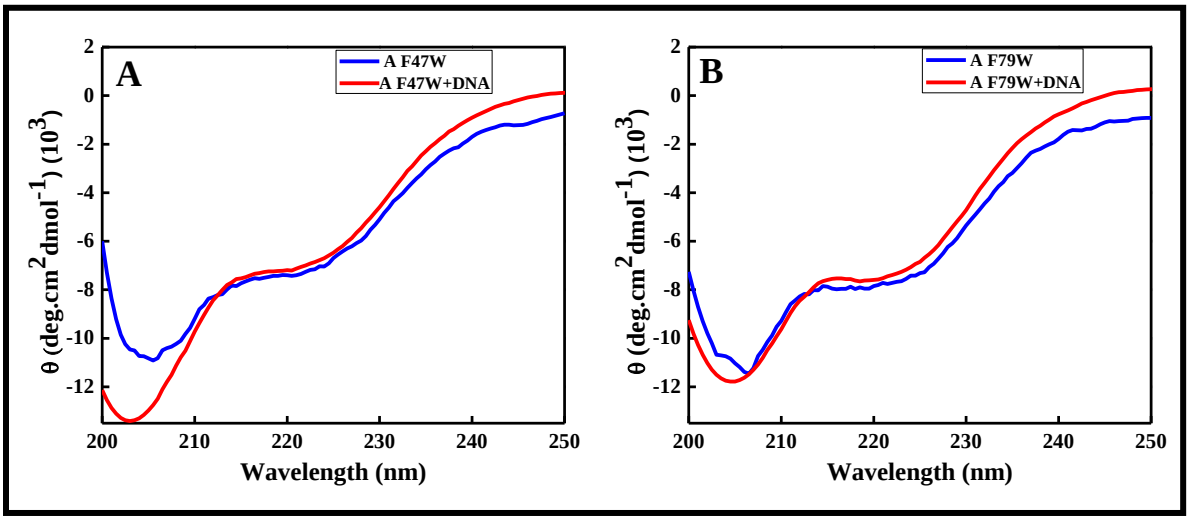


Figure 10: Changes in secondary structure of mutant proteins on DNA binding.

6.3.6 Changes in steady-state fluorescence and anisotropy on DNA binding

Tryptophan fluorescence is sensitive to local environment, within proteins. HU-A and HU-B mutants, F47W (Figure 11A and D), and F79W (Figure 11C and E), show significant quenching of tryptophan fluorescence upon DNA binding. However, there is no change in fluorescence intensity or emission maxima in the case of HU-A F51W, which is satisfying because this result serves as an unintended control experiment. Anisotropy data also validates the observation from fluorescence studies. There is increase in anisotropy for F47W and F79W mutants but the change is not very significant but is reproducible as reflected from the very narrow error range (calculated from 5 different experimental data sets). In brief, both the steady-state fluorescence and the anisotropy data suggest that binding of 4-WJ DNA to HU mutants F47W and F79W leads to changes in fluorescence properties of the respective tryptophan residues. HU-A F51W mutant protein exists in an aggregated form (Figure 6) which is, however, DNA binding competent (Figure 8); intriguingly though, no change in fluorescence properties were observed with this mutant.

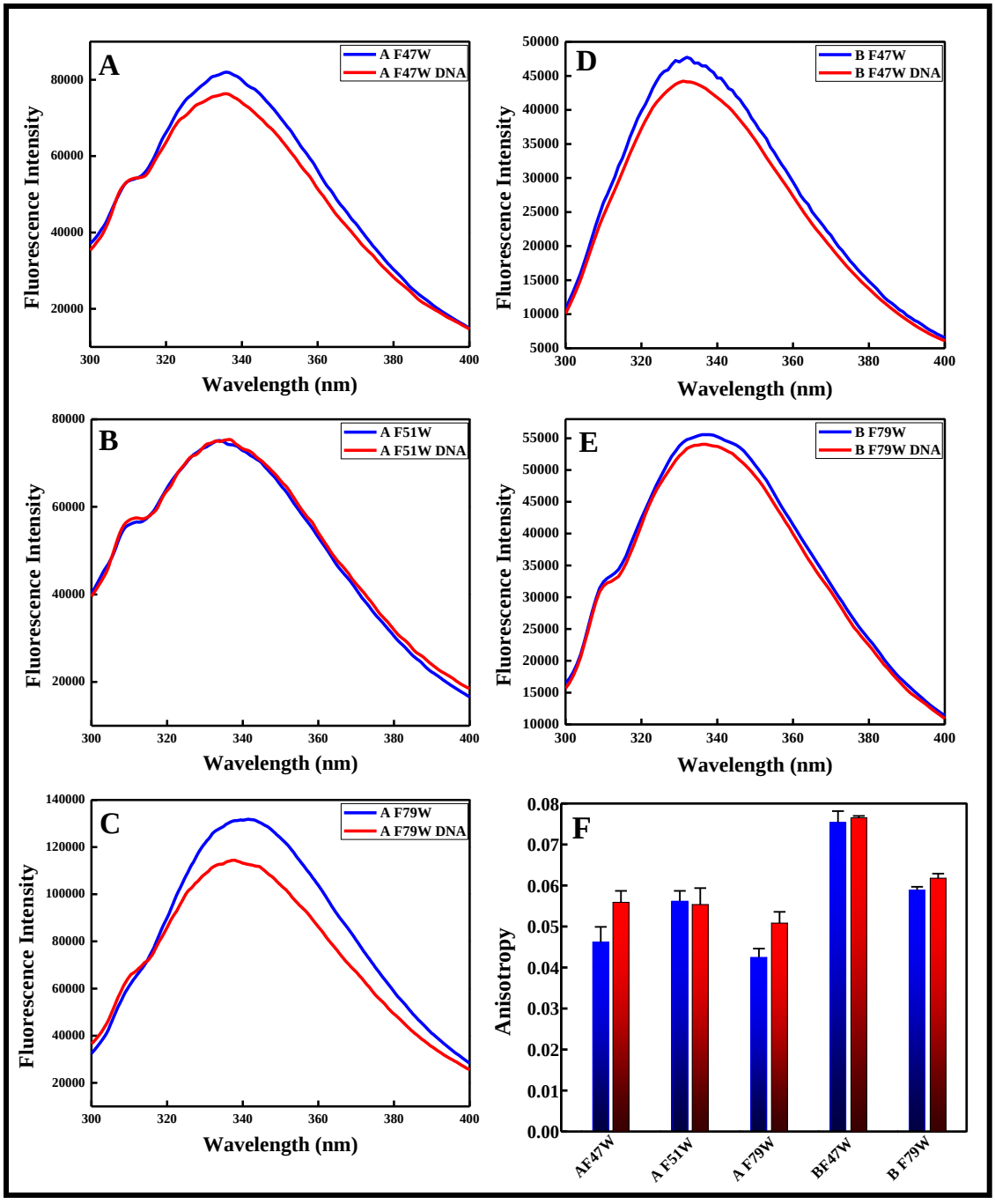


Figure 11: Quenching of Trp fluorescence upon DNA binding for all mutants HU-A F47W ***(Panel A)***, HU-A F51W ***(Panel B)***, HU-A F79W ***(Panel C)***,HU-B F47W ***(Panel D)*** and HU-B F79W ***(Panel E)*** and changes in fluorescence anisotropy upon DNA binding ***(Panel F).***

6.3.7 Binding parameters as calculated from fluorescence titrations

Binding affinities of HU-A and HU-B proteins (and mutants) for 4-WJ DNA were calculated by monitoring Trp fluorescence quenching upon DNA binding for both HU-A F79W and HU-B F79W mutant proteins. A fixed concentration of protein solution was titrated with large incremental additions of highly-concentrated DNA solution, and decrease in the fluorescence intensity was measured for each addition. Notably, changes were observed in the fluorescence intensity values, but not in the emission maxima for both HU-A (Figure 12A) and HU-B mutants (Figure 12C). Fractional decrease in fluorescence was calculated by using equation 3 (as described in Materials and methods section) from which K_d values were calculated by using equation 4, giving values of binding constant 0.015 nM for HU-A and 0.008 nM 7 for HU-B.

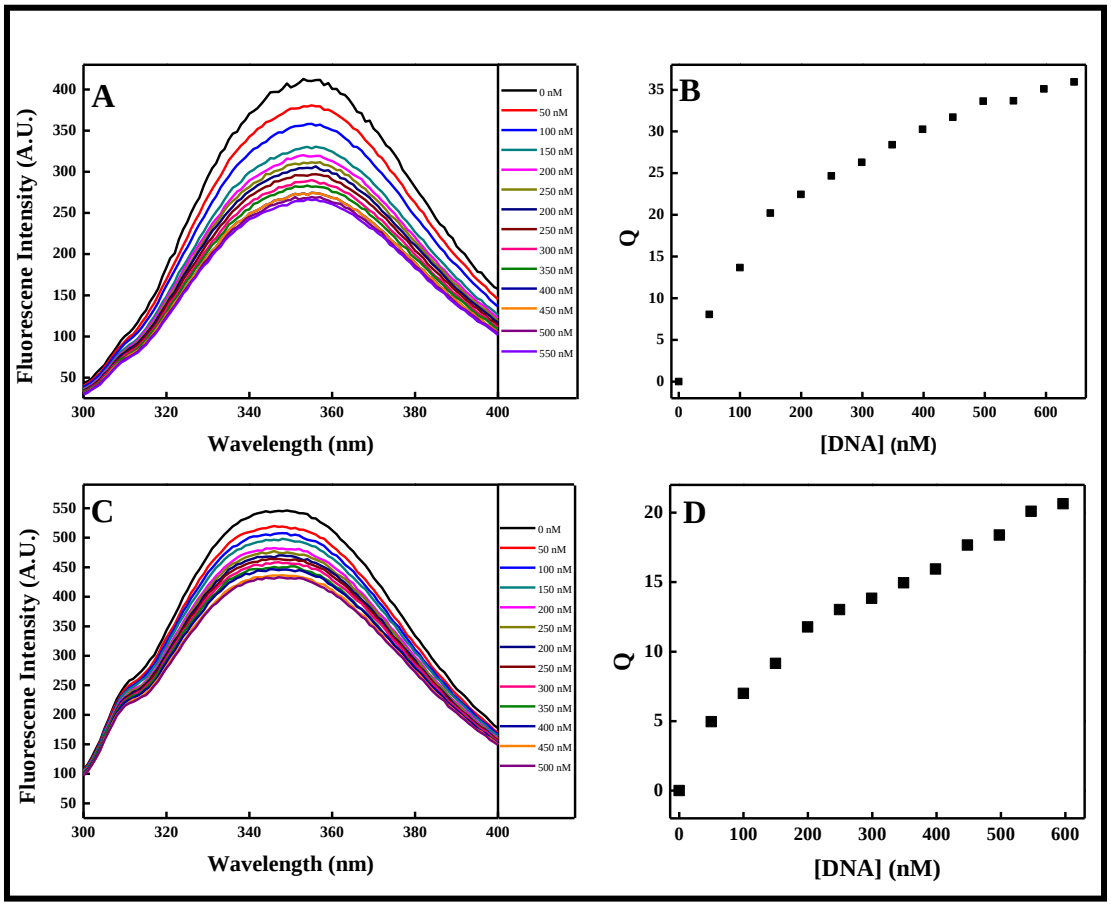


Figure 12: DNA concentration dependent quenching of Trp fluorescence for mutant HU-A F79W***(Panel A, B)*** and HU-B F79W ***(Panel C, D)***

6.3.8 Residue specific changes on DNA binding

Using the different Trp mutants, residue-specific structural changes upon DNA binding were also examined by monitoring changes in the levels of solvation of different tryptophan residues in the presence and absence of DNA. Trp quenching experiments were performed using acrylamide (quencher). The Stern-Volmer plots of HU-A F47W, HU-A F79W, HU-B F47W and HU-B F79W mutants are shown in Figure13. The slope of such a plot provides the Stern-Volmer quenching constant (K_{sv}; Eq. 5; Materials and Methods) which is related to the accessibility of the tryptophan to the quencher. The higher the K_{sv}, the higher is the accessibility of the quencher to the fluorophore, which in this case in a Trp residue in the protein HU. The values of K_{sv} are plotted in Figure 14. Clear evidence of changes in accessibility can be seen, with DNA binding further shielding Trp residues from acrylamide.

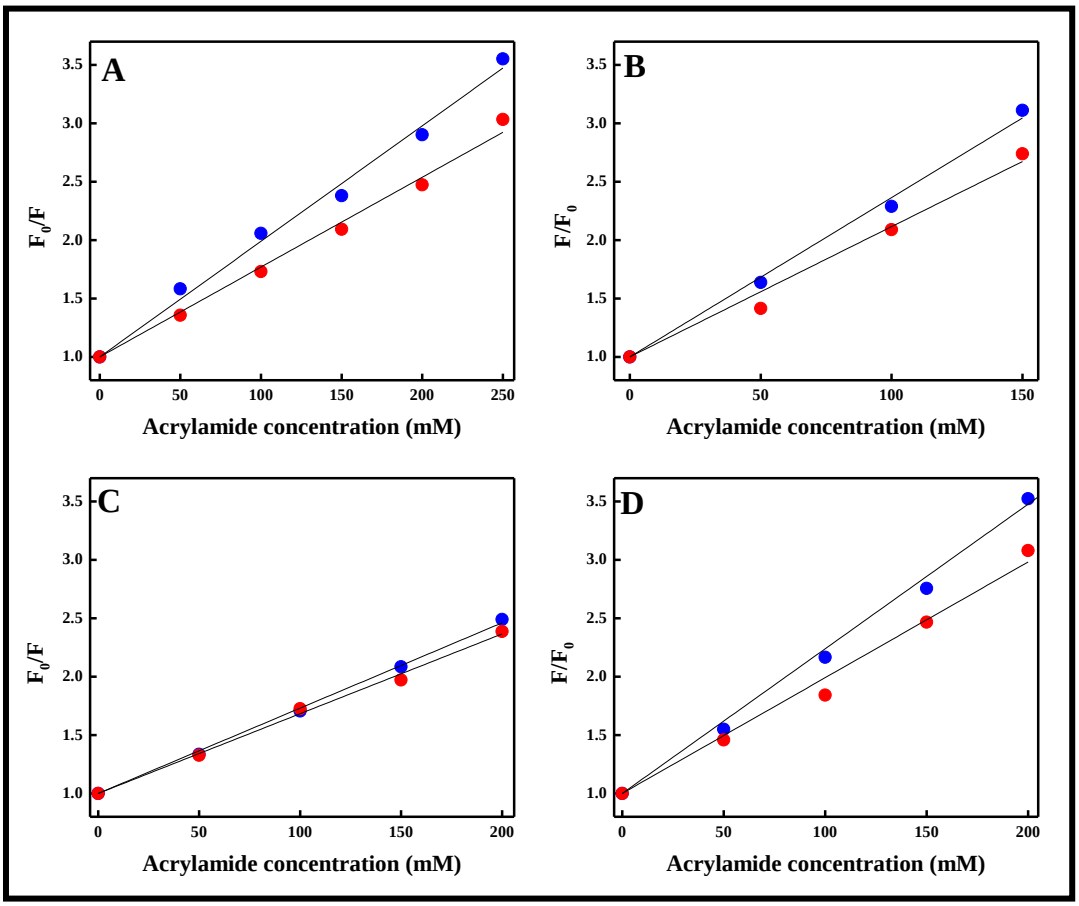
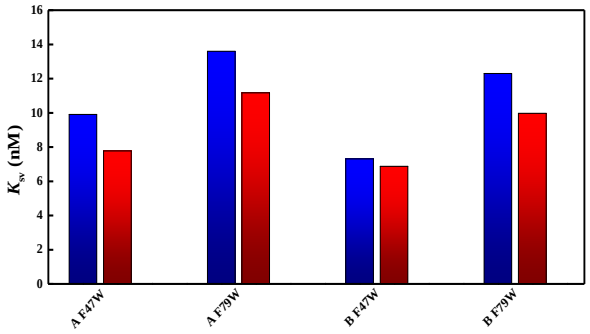


Figure 13: Stern-Volmer plots for DNA-free (Blue trace) and DNA-bound (Red trace) and their corresponding fits. HU mutants. ***Panel A*** HU-A F47W, ***Panel B*** HU-A F79W, ***Panel C*** HU-B F47W and ***Panel D*** HU-B F79W.



Mutant	K_{sv}	Mutant	K_{sv}
HU-A F47W	9.9	HU-B F47W	7.3
HU-A F47W + DNA	7.7	HU-B F47W + DNA	6.8
HU-A F79W	13.6	HU-B F79W	12.3
HU-A F79W + DNA	11.1	HU-B F79W + DNA	9.9

Figure 14:Values of K_{sv} for HU mutants in DNA-free (blue) and DNA bound (red) state.

6.3.9 Examining whether DNA damage is more likely in Trp-containing HU mutants

Circular plasmid DNA was incubated with HU-A F47W mutant protein and was irradiated with ultra violet light for 15 minutes. After the UV exposure, the samples were loaded on a native agarose gel and also on denaturing (formaldehyde) agarose gels. Figure 15A shows that both the control DNA irradiated in the absence of HU mutant protein (lane 1) and the DNA irradiated in presence of HU mutant protein (lane 2) display slightly different mobilities and no DNA damage is observed. But when the same samples were electrophoresed on a formaldehyde gel, corresponding to a single band observed in the control lane (lane 1), two very well separated bands can be seen in the DNA sample irradiated in presence of HU mutant protein (lane 2). We are following up this observation.

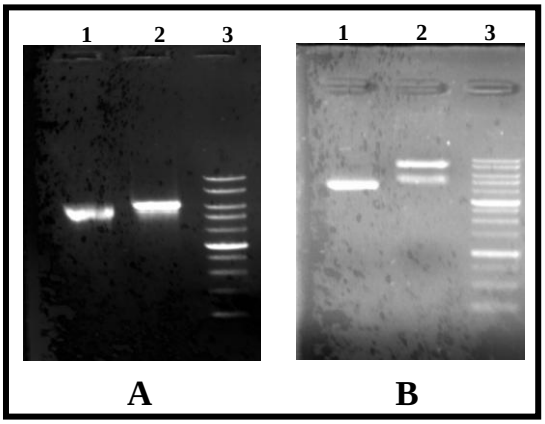


Figure 15: Panel A shows native agarose gel and ***Panel B*** shows formaldehyde agarose gel for irradiated circular plasmid DNA (lane 1), circular plasmid DNA irradiated in presence of HU mutant protein (lane 2) and 1 kb DNA ladder (lane 3).

6.4 Discussion

The main focus of this chapter is to characterize DNA binding of HU by using fluorescence spectroscopy as a tool. For that purpose, we introduced Trp residues into HU-A and HU-B which naturally lack both tyrosine and tryptophan residues. Intriguingly, the absence of tryptophan is evolutionary conserved across all HU sequences in different bacteria. We created multiple tryptophan-insertion mutants (replacing phenylalanine residues) of both HU-A and HU-B to generate the novel scope for florescence spectroscopic analysis of DNA-binding by HU. Mutants incorporating Trp residues were characterized and compared to the wild type protein in respect of their structural properties and stability. DNA binding studies were carried out with selected mutants to calculate dissociation constant values for HU-A and HU-B binding to a synthetic 4-way junction (cruciform) DNA.

In EMSA experiments performed with wild type (WT) HU proteins, a smear of DNA bound to protein is observed upon DNA binding, instead of discrete higher molecular weight band in the gel. Formation of smears in place of clean bands could owe to populations of 4-WJ being bound by different number of HU molecules, as well as to the presence of larger complexes of HU within such DNA-HU complexes (i.e., a situation in which some HU molecules are only bound to other HU molecules and not directly to DNA), especially when the concentration of HU is increased, since HU (especially HU-B) is known to form higher order oligomers (tetramers, octamers and bigger complexes) as protein concentration is increased. In concert with this interpretation, at higher protein concentrations, HU-B but not HU-A protein forms bigger DNA-protein complexes which get stuck in the loading wells of 0.5% agarose gels. This explains that despite having low DNA-binding affinities, HU-B shows bigger DNA-protein complexes.

Characterization of structure and function of *E. coli* HU mutants shows that HU-A and HU-B mutants, F47W, and F79W, are well-structured and are dimeric but that HU-A F51W forms higher oligomers and elutes at the void volume of aSuperdex-75 column. This indicates that mutations of Phe residue to Trp residues at positions 47 and 79 are acceptable, and tolerated, by the protein's structure and the protein structured is not disturbed; however, but mutation of the Phe residue at position 51 to Trp leads to destabilization of structure of the protein leading to formation of higher oligomers which retain some DNA-binding potential (enough to show up in

non-equilibrium EMSA gel experiments, but not in spectroscopic experiments examining, e.g. fluorescence anisotropy). HU mutants F47W and F79W, like the WT protein also displayed the ability to refold after thermal denaturation and retain DNA binding competence after such refolding.

Upon DNA binding, minor changes in the secondary structure of the protein were observed but addition of DNA to the protein lead to significant quenching of the intrinsic Trp fluorescence. Both the steady state fluorescence and anisotropy data suggests that upon addition of DNA to the protein, local environment of the Trp residues undergo alteration. We used these changes in fluorescence properties of the Trp residue(s)to calculate *kd* values for binding of HU-A and HU-B binding to 4-WJ synthetic cruciform DNA. Values suggest that of the two, HU-A has higher affinities to bind to DNA. In solution binding affinities of *E. coli* HU proteins with 4-WJ have not been determined using Trp fluorescence. Our calculated *kd* values for HU-A and HU-B DNA binding lie in the sub nano molar range as that reported for HU-DNA binding [22,23].

Next, we examined the residue-specific structural changes occurring upon DNA binding by monitoring differential levels of solvation of tryptophan residues in different mutants. It is reported that for a totally exposed Trp residue, the *Ksv* value lies near 16 and decreases with the burial of the Trp residue. Our data suggests that on DNA binding, a decrease solvation of the single Trp residue both at position 47 and 79 is observed for both HU-A and HU-B proteins. In absence of DNA, residue 47 in both the proteins is more buried than residue 79 and hence the decrease in *ksv* observed for the residue 47 mutant is less as compared to that for residue 79.

On close comparison of protein sequences of the architectural DNA-binding proteins, we can see that none of these contain Trp residues. We wanted to explore why this absence of Trp is evolutionarily conserved..We explored the effect of tryptophan insertion upon DNA integrity, in Trp residues are present in close proximity to DNA over long durations. Our data suggests that irradiation of a circular DNA bound to HU mutant containing Trp residue, relaxation of DNA is observed by formation of nicks suggesting that these Trp mutants of HU can induce DNA damage when exposed to UV.

6.5 References

[1] D. Kamashev, J. Rouviere-Yaniv, The histone-like protein HU binds specifically to DNA recombination and repair intermediates, EMBO J 19 (2000) 6527-6535.

[2] A.G. Kozlov, R. Galletto, T.M. Lohman, SSB–DNA Binding Monitored by Fluorescence Intensity and Anisotropy, Single-Stranded DNA Binding Proteins, Springer, 2012, pp. 55-83.

[3] T. Heyduk, J.C. Lee, Application of fluorescence energy transfer and polarization to monitor Escherichia coli cAMP receptor protein and lac promoter interaction, Proceedings of the National Academy of Sciences 87 (1990) 1744-1748.

[4] M.P. Brown, C. Royer, Fluorescence spectroscopy as a tool to investigate protein interactions, Current opinion in biotechnology 8 (1997) 45-49.

[5] V.J. LiCata, A.J. Wowor, Applications of fluorescence anisotropy to the study of protein–DNA interactions, Methods in cell biology 84 (2008) 243-262.

[6] J.R. Lakowicz, Principles of fluorescence spectroscopy, Springer Science & Business Media, 2013.

[7] P.L. Privalov, I. Jelesarov, C.M. Read, A.I. Dragan, C. Crane-Robinson, The energetics of HMG box interactions with DNA: thermodynamics of the DNA binding of the HMG box from mouse sox-5, Journal of molecular biology 294 (1999) 997-1013.

[8] R. Favicchio, A.I. Dragan, G.G. Kneale, C.M. Read, Fluorescence spectroscopy and anisotropy in the analysis of DNA-protein interactions, DNA-Protein Interactions, Springer, 2009, pp. 589-611.

[9] F.W. Sevenich, J. Langowski, K. Rippe, V. Weiss, DNA binding and oligomerization of NtrC studied by fluorescence anisotropy and fluorescence correlation spectroscopy, Nucleic acids research 26 (1998) 1373-1381.

[10] S. Titolo, E. Welchner, P. White, J. Archambault, Characterization of the DNA-binding properties of the origin-binding domain of simian virus 40 large T antigen by fluorescence anisotropy, Journal of virology 77 (2003) 5512-5518.

[11] D. Jantz, J.M. Berg, Probing the DNA-binding affinity and specificity of designed zinc finger proteins, Biophysical journal 98 (2010) 852-860.

[12] B. Owen, C. McMurray, Rapid method for measuring DNA binding to protein using fluorescence anisotropy, Protocol Exchange (2009).

[13] T. Heyduk, Y. Ma, H. Tang, R.H. Ebright, Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction, Methods in enzymology 274 (1996) 492.

[14] V. LeTilly, C.A. Royer, Fluorescence anisotropy assays implicate protein-protein interactions in regulating trp repressor DNA binding, Biochemistry 32 (1993) 7753-7758.

[15] A.D. Baxevanis, D. Landsman, Histone Sequence Database: a compilation of highly-conserved nucleoprotein sequences, Nucleic acids research 24 (1996) 245-247.

[16] K.E. van Holde, The proteins of chromatin. I. Histones, Chromatin, Springer, 1989, pp. 69-180.

[17] B.D. Guptasarma P., Hydroxyl Radical Mediated Damage to Proteins, with Special Reference to the Crystallins Biochemistry 31 (1992) 4296-4303.

[18] I.S. Hiraku Y., Oikawa S., Yamamoto K. , Tada S., Nishino K. and Kawanishi S., Metal-mediated oxidative damage to cellular and isolated DNA by certain tryptophan metabolites, Carcinogenesis 16 (1995) 349-356.

[19] S. Ghosh, A. Grove, Histone-like protein HU from Deinococcus radiodurans binds preferentially to four-way DNA junctions, J Mol Biol 337 (2004) 561-571.

[20] M. Shakir, M. Azam, S. Parveen, A.U. Khan, F. Firdaus, Synthesis and spectroscopic studies on complexes of N, N'-bis-(2-pyridinecarboxaldimine)-1, 8-diaminonaphthalene (L); DNA binding studies on Cu (II) complex, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 71 (2009) 1851-1856.

[21] M. Deiana, K. Matczyszyn, J. Massin, J. Olesiak-Banska, C. Andraud, M. Samoc, Interactions of Isophorone Derivatives with DNA: Spectroscopic Studies, PloS one 10 (2015) e0129817.

[22] E. Bonnefoy, M. Takahashi, J.R. Yaniv, DNA-binding parameters of the HU protein of Escherichia coli to cruciform DNA, Journal of molecular biology 242 (1994) 116-129.

[23] C.I. Vitoc, I. Mukerji, HU binding to a DNA four-way junction probed by Forster resonance energy transfer, Biochemistry 50 (2011) 1432-1441.

CHAPTER 7

Studies on fluorescently-tagged HV

Contents

7.1. Introduction

7.1.1 Use of affinity-tagged HU for single cell-based detection and quantitation of leaky expression

The bacterium, *Escherichia coli*, is the most widely used of all microbial cell factories available for the expression and purification of both prokaryotic and eukaryotic proteins [1]. Several factors favor the use of *E. coli* as a cell factory, such as the ease of its genetic manipulation, short generation time, ease of protein recovery, ease of scaling up, and low costs of growth media and downstream processing [1,2]. Recently, it has been pointed out that ~ 68 % of all publications associated with protein expression prior to successful crystallization report the use of *E. coli* as a cell factory [3]. Of course, several issues of concern also remain regarding the use of *E. coli*, e.g., differences in codon usage between genes from different source organisms, problems with oxidative folding in the *E. coli* cytoplasm, absence of post-translational modification machinery in *E. coli*, deposition of proteins as inclusion bodies, and effects of over-expression of heterologous proteins on *E. coli* growth and viability [2]. To address such concerns, promoter-operator-vector-host systems that obviate the most obvious problems continue to be developed, and improved.

This chapter focuses on the promoter-operator-vector-host systems believed to ensure tight regulation of expression of cloned genes. In particular, our concern here is with problems that can potentially arise due to: (i) protein-linked cellular toxicity, and (ii) the burden of protein expression on *E. coli* cell growth and viability. These problems can arise whenever systems presumed to offer tight regulation of expressionactually allow a certain amount of 'leaky expression'. The intensity of the problem varies with the actual amounts of such expression, and the nature of the protein involved. Proteins expressed in *E. coli* can, in principle, be toxic to the organism even at extremely low levels of expression (e.g., 10-50 molecules per cell), affecting growth, cell mass and protein yield, and causing accumulation of background mutations [4]. Such proteins can have deleterious consequences, e.g., if they were to bind with high affinity to *E. coli* proteins present at similarly low levels, and inactivate such proteins [e.g., repressors/activators, cell cycle regulators, cell division components, or components of replisomes]. Tightly-regulated plasmid-based expression systems offer a partial solution to such concerns [2]. In tightly-regulated systems, the promoter of the designated gene is normally not allowed to transcribe, without being specifically induced to do so, through the binding of a repressor molecule to an operator

sequence which keeps the gene repressed. When the gene remains repressed, the host culture is expected to grow to a reasonable optical density without the burden of protein toxicity. Then, when only a few cell divisions remain for the culture to reach stationary phase, an inducer (which enters cells and titrates repressor molecules away from operators) is added to titrate away the repressor and trigger expression of the toxic protein [5].

Several expression systems employ the above approach assuming that the repressor binding to operator sequences completely inhibits expression of the designated gene prior to specific induction of expression. The question we wish to ask here is whether some expression still manages to 'slip through, between the cracks' in such tightly-regulated systems, i.e., whether there is a certain amount of 'leaky expression'. It is necessary to have a system available to examine whether leaky expression does occur, when it is not easy to detect by conventional means. Here, as 'proof of principle', we demonstrate a fluorescence imaging-based method for the detection of leaky expression, applying it to two commonly used promoter-operator-vector-host systems that are believed to tightly regulate expression, both of which use the lac repressor. These systems are: (a) pET-28c-BL21Star(DE3)[pLysS] and (b) pQE-30-M15[pREP4].

The pET(vector)-BL21Star(DE3)[pLysS](host) system is the most widely used promoter-operator-vector-host system in *E. coli* [3]. It uses a highly specific T7 RNA polymerase-T7 promoter and lac repressor-operator system. Cells of BL21Star(DE3)[pLysS] additionally contain the λDE3 lysogen which carries DNA encoding T7 RNA polymerase under a lacUV5 promoter [6]. Lac repressor formed from the native copy of *lac* gene and from plasmid over-expressed *lac* gene binds to the lacUV5 promoter and inhibits formation of T7 polymerase. Upon addition of the inducer, IPTG, which is a allolactose mimic, the lac repressor is titrated away from the operator, leaving the lacUV5 promoter free, and allowing T7 RNA polymerase (and hence the gene of interest transcribed by T7 RNA polymerase) to be expressed. These cells also contain the pLysS vector which contains DNA encoding T7 lysozyme [7]. The pLysS plasmid is already present in the BL21(DE3) cells before the pET vector bearing the gene of interest is transformed into these cells. T7 lysozyme, besides acting on cell wall peptidoglycans, also binds to T7 RNA polymerase and inhibits transcription [8,9,10,11,12], thus ensuring that even if low levels of T7 RNA polymerase are produced, there is no transcription from T7 RNA polymerase-driven promoters when IPTG is absent and T7 RNA polymerase is not over-expressed. In addition, there is a mutation in the gene encoding RNase E (*rne131*) to enhance mRNA stability [13,14]. A further regulation of

expression is achieved by the incorporation of the lac operator sequence downstream of the T7 promoter [15] in pET vectors, which ensures inhibition of expression of the gene of interest even if low amounts of T7 RNA polymerase happen to be expressed and escape degradation by T7 lysozyme.

The pQE (vector)-M15[pREP4] (host) promoter-operator-vector-host system is another very popular system. It uses two modes of operator binding to increase binding affinity, i.e., there are two lac operator sequences located downstream of a strong T5 promoter. Two adjacent lac operator sequences allows lac repressor tetramer binding, leading to higher levels of bound lac repressor in the path of RNA polymerase, and a tighter control on expression prior to induction of gene expression (http://kirschner.med.harvard.edu/files/protocols/QIAGEN_QIAexpressionist_EN.pdf). The pQE30 vector uses the ColE1 origin of replication [16], and it exists in cells in copy numbers estimated to vary between 20-30 [17] and 50-70 copies per cell [18]. To ensure that there are sufficient number of copies of lac repressor molecules available within M15 *E. coli* cells to inhibit transcription from promoter-operator complexes located on all copies of the pQE30 plasmid, an additional plasmid, pREP4, is present in the host strain. From this plasmid, lac repressor is constitutively expressed, to maintain high cellular levels of the repressor. The plasmid, pREP4, is already present in M15 cells before the pQE plasmid bearing the gene of interest is transformed into these cells. Thus, in the absence of the inducer, IPTG, no protein expression is expected to occur within un-induced cells.

To monitor leaky expression using the two tightly-regulated systems described above, we chose to use fluorescence imaging [19], since fluorescence offers high sensitivity. We also used two isoforms of the DNA-binding protein, HU, i.e., HU-A and HU-B, and produced these isoforms in fusion with either a monomeric red fluorescence protein variant (Tag-RFP-T) or a fast-folding variant of yellow fluorescent protein (YFP) known as 'Venus'. As already described in numerous chapters of this thesis, HU is a histone-like protein and is known to bind to DNA and help in the compaction of the *E. coli* genome into the bacterial 'nucleoid'. Fusion genes encoding (a) RFP fused with HU-A, or (b) the yellow fluorescent protein, Venus, fused with HU-B were cloned, respectively, into the pQE-30 and pET-28c vectors described above, for use with M15[pREP4] and BL21Star(DE3)[pLysS] cells.

The version of RFP used to make the fusion protein was Tag-RFP-T. This version of RFP is encoded by a gene codon-optimized for expression in *E. coli.* Tag-RFP-T is bright, pH

stable, highly soluble, monomeric, and known to show no detectable aggregation. The version of YFP fused to HU-B is also encoded from a gene optimized for *E. coli.* Venus is also a soluble, pH stable, weakly dimeric protein with very fast maturation rate [20]. The fusion partners of the fluorescent proteins in the two constructs, i.e., HU-A and HU-B, are also highly soluble proteins which naturally form dimers. We over-expressed and purified the RFP-HU-A and Venus-HU-B fusions through Ni-NTA affinity chromatography, and confirmed through gel filtration chromatography and EMSA that the fusion products are DNA binding-competent, dimeric, and fluorescent, with no detectable aggregation, indicating that the RFP/Venus present at the N-terminus of HU-A/HU-B does not interfere either with either its dimerization or with its DNA binding. Notably, HU over-expression is known to have no global effects on transcription of other genes [21].

Upon expression, HU-A or HU-B proteins, tagged with RFP or Venus, bind to the *E. coli* nucleoid. Binding of fluorescent proteins to DNA within the nucleoid leads to a highly localized fluorescence signal which can be easily detected [22]. It is difficult to detect small amounts of freely diffusing fluorescent molecules against the auto-fluorescence backgrounds of cells [23]; however, binding of HU to DNA, localizes the fluorescence (i.e., concentrates it to one location in three-dimensional space). Hence, even small levels of the fluorescent protein-tagged HU can be detected, when such a fusion protein is expressed under a tightly-regulated promoter which shows leaky expression even when no induction is performed.

7.1.2. Use of affinity tagged HU to study role of HU in biofilms

Infection causing bacteria, generally make biofilms around them which acts to protect the bacteria from extracellular damage. Biofilms are the functional, multicellular bacterial communities held together by polymeric matrix released by death and lysis of the other bacterial cells. The main constituents of biofilm are extracellular polymeric substance (EPS) like polysaccharides, lipids, extracellular proteins and extracellular DNA (eDNA) [24]. EPS causes the bacterial cells to aggregate and provides a protective barrier to make them resistant to the external environmental conditions [24]. Presence of biofilms makes it difficult to clear the bacterial infection either by innate or adaptive immune response [25,26]. It is necessary to disrupt the biofilm matrix to clear the microbial infection making it important to understand the components that hold the biofilms together and the methods of clearance of biofilms [27]. Recently the eDNA present in the biofilms had gained a lot of focus as it forms a meshwork that acts as a major structural component of biofilms [28,29,30,31]. It has been shown that

few proteins that belong to DNA-BII family of proteins can also exist extracellularly [32]. One of the proteins of the family, IHF (integration host factor) has also been shown to be present in biofilms and targeting of IHF by using anti-IHF antibodies favors clearing of biofilms [33].

HU and IHF show a very high sequence and structural similarity. HU is a non-specific DNA binder but IHF identifies a particular DNA sequence but they both bind to DNA using very similar binding modes. Although, very recently it has been reported that HU-B but not HU-A, is also a constituent of biofilms [34], we wanted to microscopically investigate the presence of HU in biofilms. By using a biofilm forming *E. coli* strain (MG1655), over-expressing fluorescently tagged HU protein, we monitored if HU is also found extracellularly and if that is so, does it bind to the cell surface?

7.2. Materials and methods

7. 2.1.Cloning of fusion proteins

hupA and *hupB* genes were amplified from *E. coli* genomic DNA by PCR. The *hupA* gene was fused to the 5' end of the RFP-encoding gene. The *hupB* gene was fused to the 5' end of the Venus-encoding gene. These fusions were done by performing splicing by overlap extension (SOE) PCR reactions using the following primers:

	Primer	**5'→3'**
1.	RFP Forward Primer 1	AGCTACTGGATCCGTGTCTAAGGGCGAAGAG
2.	RFP Reverse Primer 1	TCCTCCACCTCCGCTTCCTCCCTTGTACAGCT
3.	RFP Forward Primer 2	CGGAGGTGGAGGAAGCGGAGGCATGAACAAGA
4.	RFP Reverse Primer2	GAATACTCCCGGGTTACTTAACTGCGTCTTTCAATG
5.	Venus Forward Primer 1	GAATACTGGATCCATGGTGAGCAAGGGC
6.	Venus Reverse Primer1	CGAGCTGTACAAGGTGAATAAATCTC
7.	Venus Forward Primer 2	ATGAATAAATCTCAATTGATCG
8.	Venus Reverse Primer 2	GAATACTAAGCTTTTAGTTTACCGCGTCTTTCAGT

Amplicons were generated by performing PCR reactions using (a) Forward Primer 1 and Reverse Primer 2, and (b) Forward Primer 2 and Reverse Primer 1. Amplicons resulting from reactions (a) and (b) were spliced by using these amplicons as mega-primers and also using Forward Primer 1 and Reverse Primer 1. Using Bam HI and Hind III restriction sites, the fusion product was cloned into pQE-30 and transformed into M15[pREP4] cells and selected on kanamycin and ampicillin containing plates. Venus-HU-B was cloned into the pET-28c vector and transformed into BL21Star(DE3)[pLysS] cells and selected on kanamycin and tetracycline containing plates.

7. 2.2. Purification of fusion proteins, gel filtration chromatography and EMSA

Cells over-expressing the fusion proteins were induced and the proteins were purified from induced cells using the salt-based method for purification of HU fusions, as described in chapter 4 for HU-A or HU-B alone. Gel filtration chromatography and DNA binding studies (using EMSA) were also performed as described in chapter 4.

7.2.3. Sample preparation for fluorescence imaging

For imaging, un-induced and induced cells were collected. Cells expressing RFP-HU-A in pQE30/M15[pREP4] cells and Venus- HU-B in pET28a/BL21Star(DE3)[pLysS] were grown overnight, together with pQE30/M15[pREP4]and pET28a/BL21Star(DE3)[pLysS] controls, at 37 °C. The overnight culture was then used to inoculate 5 ml of secondary culture, which was grown at 37 °C until the O.D. at 600 nm had reached a value of 0.6. Then, 1 ml of un-induced cells was collected and the remaining 4 ml of culture were induced with 1 mM IPTG. Induced cells were allowed to grow for another 4-5 hours after which 1 ml was collected. Un-induced and induced cell samples were centrifugally sedimented into pellets, and directly resuspended in 1 ml of M9 glucose medium, washed and resuspended in the same medium. About 7 µl of these washed and resuspended cells were layered on agarose pads (made from 1.5% agarose). These agarose pads (containing cells) were placed on slides and covered with coverslips. Cells were allowed to settle on agarose pads for 5-10 minutes prior to imaging.

7.2.4. DIC and fluorescence imaging

Cells were imaged using the Delta Vision Deconvolution microscope, DV Elite (GE Healthcare), equipped with a solid state illumination and a CoolSnap HQ2 1.4 megapixel monochrome CCD camera (Photometrics). RFP-HU-A control, un-induced and induced samples of M15[pREP4] cells were illuminated with 100% transmission of the RFP wavelength output of the solid state illumination unit (531-565 nm). Similarly, Venus-HU-B control, un-induced and induced samples of BL21Star(DE3)[pLysS] cells were illuminated with 100 % transmission of the YFP wavelength output of the solid state illumination unit (505-515 nm). The imaging was done using a Plan Apochromat 100X/1.4 NA Oil DIC objective (Olympus). The exposure time was 0.150 seconds. The frame size was of 512 x 512, with 2x2 binning. False colors with intensities proportional to the recorded fluorescence intensities were used to represent RFP-HU-A images (red) and Venus-HU-B images (yellow). Intensities were quantified using ImageJ software (http://imagej.nih.gov/ij/).

7.2.5. Biofilm formation studies

Venus-HU-B fusion was cloned into pQE30 vector between the BamHI and HindIII sites using the primers; Venus BamHI Forward 5'-GAATACTGGATCCATGGTGAGCAAG GGC -3' and HU-B HindIII Reverse 5'-GAATACTAAGCTTTTAGTTTACCGCGTCTTTC AGT-3'. The pQE30 vector containing Venus-HU-B fusion was transformed into MG1655 cells by electroporation. The positive tranformants were screened by doing colony PCR and by checking the fluorescence emission from the cells. Imaging and image processing was done as previously described.

7.3. Results and discussion

7.3.1. DNA binding and dimerization properties of fusion proteins

Gel filtration chromatography was done for both the fusion proteins using a Superdex-200 column of 24 ml bed volume (GE Healthcare), described in earlier chapters. As shown in Figure 1A, the proteins eluted at elution volume corresponding to approximately 70 kDa (equal to the molecular weight of a dimer of the fusion of HU and fluorescent protein), suggesting that the tagged variants of HU are also dimeric, just like the WT protein, with dimerization presumably mediated by HU alone, at least in the case of tag-RFP). To check the DNA binding competence of the fusion proteins, EMSA assays for DNA binding competence were carried out with various DNA templates as shown in Figure 1B. The mobility shift in the EMSA gels clearly indicates that the dimeric fusion proteins are capable of DNA binding.

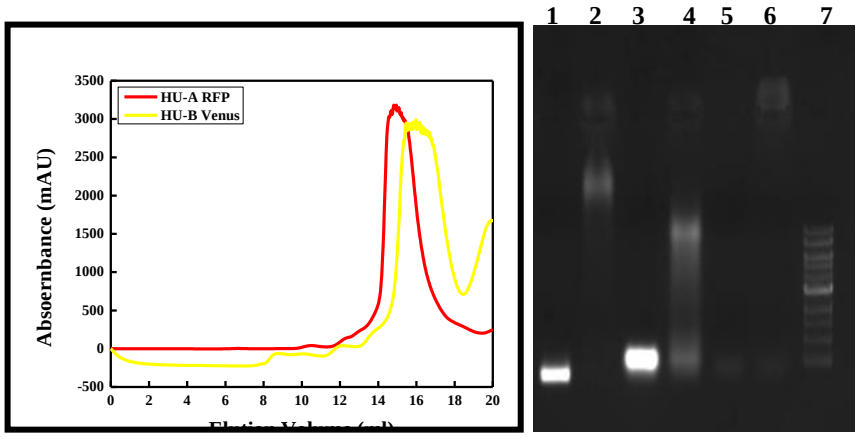


Figure 1: Panel A: Gel filtration profile of HU-A RFP and HU-B Venus on superdex-75 24 ml column. ***Panel B:*** binding of HU-A RFP to different DNA substrates; Lane 1- 270bp linear DNA control, Lane 2- binding of HU-A RFP to 270 linear DNA, Lane 3- 4WJ control, Lane 4- HU-A RFP binding to 4WJ, Lane 5- 60 bp linear DNA, Lane 6- HU-A RFP binding to 60 bp DNA and Lane 7- 100 bp DNA ladder.

7. 3.2. Leaky expression is seen in fluorescence images of un-induced cells

In the present experiments, cells expressing RFP-HU-A and Venus-HU-B were imaged using a wide-field high resolution microscope. In stark contrast to the control panels displaying results from cells containing plasmids but no gene inserts (Figures 1A and 1D for RFP-HU-A lacking cells, showing DIC images and fluorescence images, respectively, and Figures 2A and 2D for Venus-HU-B lacking cells, also showing similar DIC and fluorescence images), fluorescence signals from molecules expressed through leaky expression in some cells (occurring even under repressed conditions) are clearly detected and seen to be localized around the nucleoid (Figures 1B and 1E for RFP-HU-A gene-containing cells, showing DIC images and fluorescence images, respectively, and Figures 2B and 2E for Venus-HU-B gene-containing cells, also showing similar DIC and fluorescence images). Of course, as expected, these levels of fluorescence increase very significantly upon addition of the inducer, IPTG (Figures 1C and 1F for RFP-HU-A expression, showing DIC images and fluorescence images, respectively, and Figures 2C and 2F for Venus-HU-B expression, also showing similar DIC and fluorescence images).

7. 3.3.Quantification of levels of leaky expression

Between the panels showing fluorescence owing to leaky expression without induction for the two fusion proteins (Figures 1E and 2E), and the panels showing over-expression after induction for these proteins (Figures 1F and 2F), there are differences seen in the intensity of the fluorescence from RFP or Venus. Figure 3, shows results of quantification of the imaging data seen in the visual fields shown in Figures 1 (Figure 3A, 3C) and 2 (Figure 3B, 3D), which were obtained using ImageJ software. The figure shows (a) that the background-subtracted mean fluorescence intensity seen per un-induced cell is substantially lower than that seen after induction, and (b) also that the fold-difference seen between un-induced and induced cells is much higher for RFP-HU-A than it is for Venus-HU-B. This was done by averaging over 40 randomly selected cells in pairs of DIC and fluorescence images. At the same time, in Fig. 3C and 3D we also show scatter plots of the intensity of fluorescence seen in each of the 40 cells for the un-induced and induced cultures. The scatter plot establishes three things: (1) the highest intensity seen in any cell in the un-induced population is comparable to the lowest intensity seen in any cell in the induced population; (2) overall, cells in the induced population display much higher intensity than cells in the un-induced population; and (3) in both the un-induced and induced populations, there are large variations in the levels of expression from cell to cell.

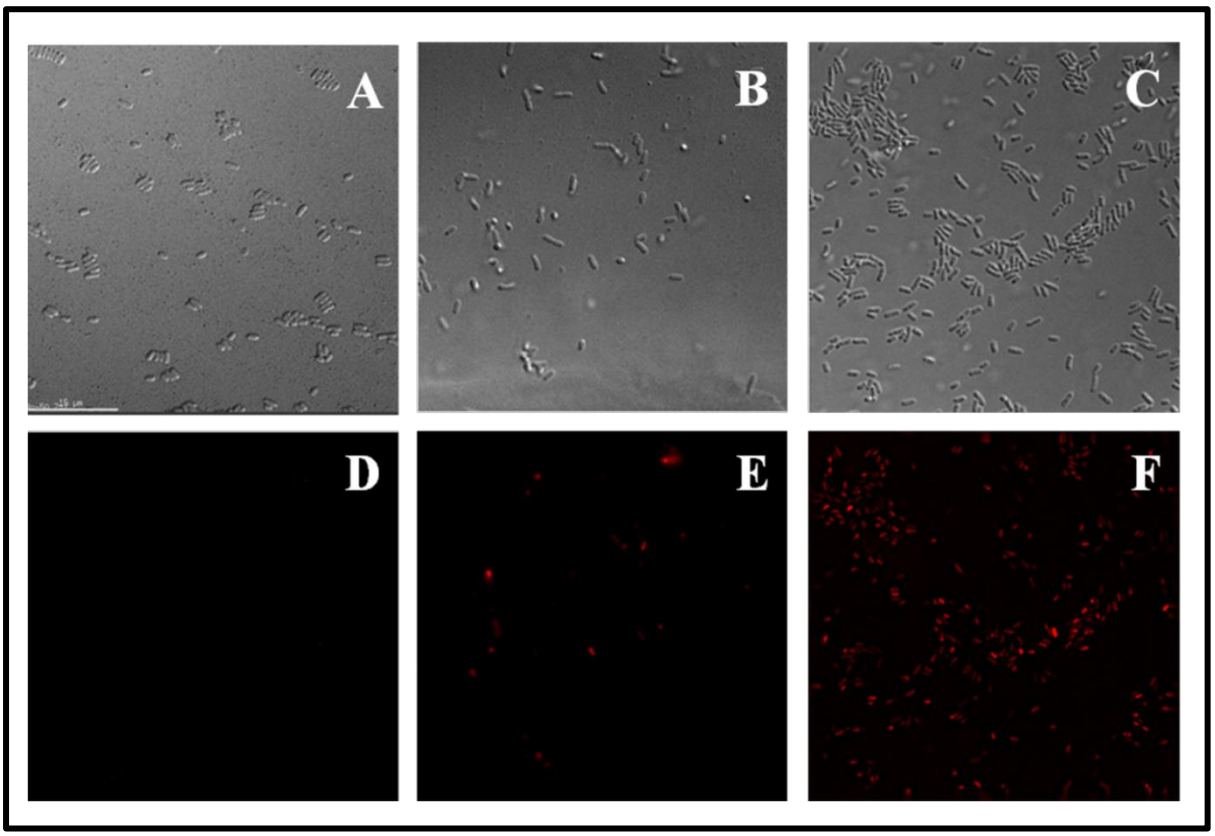


Figure 2: M15[pREP4] cells containing either the empty pQE-30 vector (lacking the insert encoding RFP-HU-A), or the pQE-30 vector bearing the insert encoding RFP-HU-A, imaged using either differential interference contrast (DIC) or fluorescence imaging using the RFP excitation wavelength, for cells that have either been subjected to IPTG induction or deprived of such induction. ***Panel A*** shows the DIC image of cells containing empty vector, whereas ***Panel D*** shows the same field of cells imaged using RFP excitation. ***Panel B*** shows the DIC image of cells containing the vector bearing the insert encoding the RFP-HU-A fusion, without any IPTG induction, whereas ***Panel E*** shows the same field of cells imaged using RFP excitation. ***Panel C*** shows the DIC image of cells containing the vector bearing the insert encoding the RFP-HU-A fusion, with IPTG induction, whereas ***Panel F*** shows the same field of cells imaged using RFP excitation.

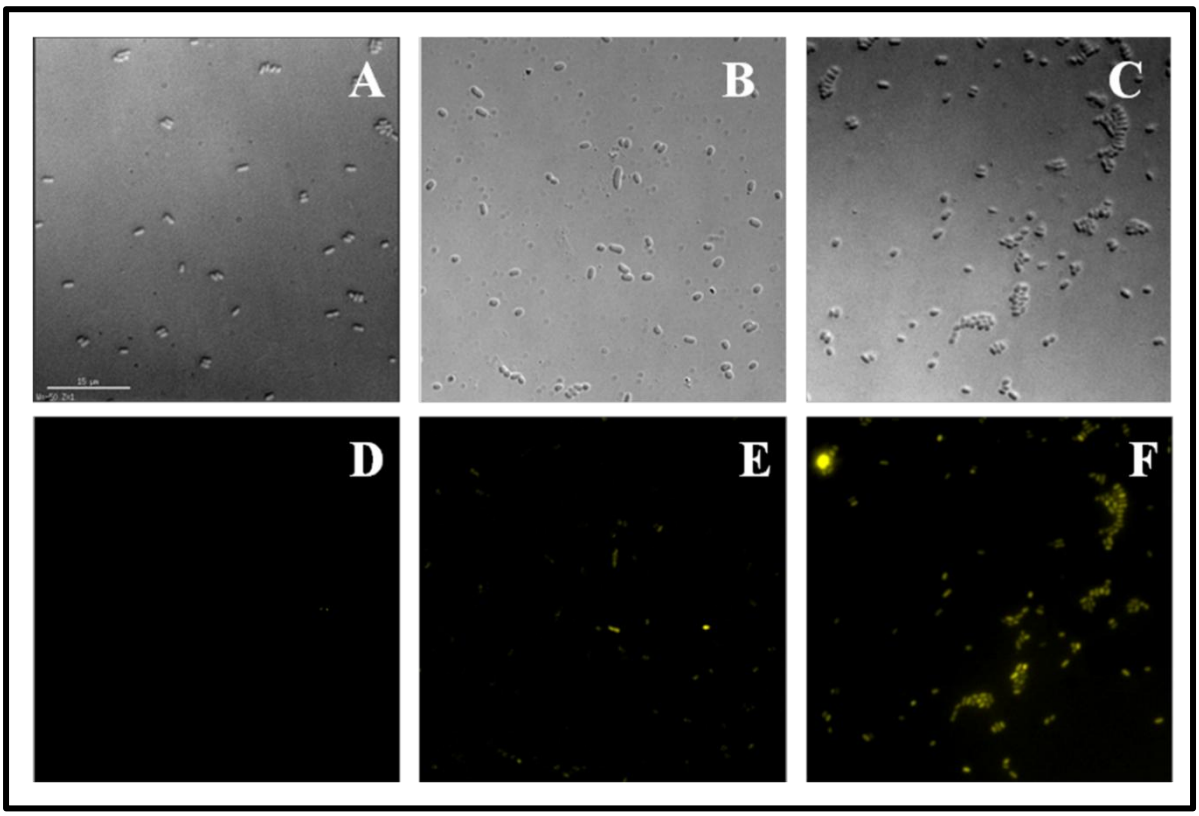


Figure 3: BL21Star(DE3)[pLysS] cells containing either the empty pET-28c vector (lacking the insert encoding Venus-HU-B), or the pET-28c vector bearing the insert encoding Venus-HU-B, imaged using either differential interference contrast (DIC) or fluorescence imaging using the RFP excitation wavelength, for cells that have either been subjected to IPTG induction or deprived of such induction. ***Panel A*** shows the DIC image of cells containing empty vector, whereas ***Panel D*** shows the same field of cells imaged using Venus excitation. ***Panel B*** shows the DIC image of cells containing the vector bearing the insert encoding the Venus-HU-B fusion, without any IPTG induction, whereas ***Panel E*** shows the same field of cells imaged using Venus excitation. ***Panel C*** shows the DIC image of cells containing the vector bearing the insert encoding the Venus-HU-B fusion, with IPTG induction, whereas ***Panel F*** shows the same field of cells imaged using Venus excitation.

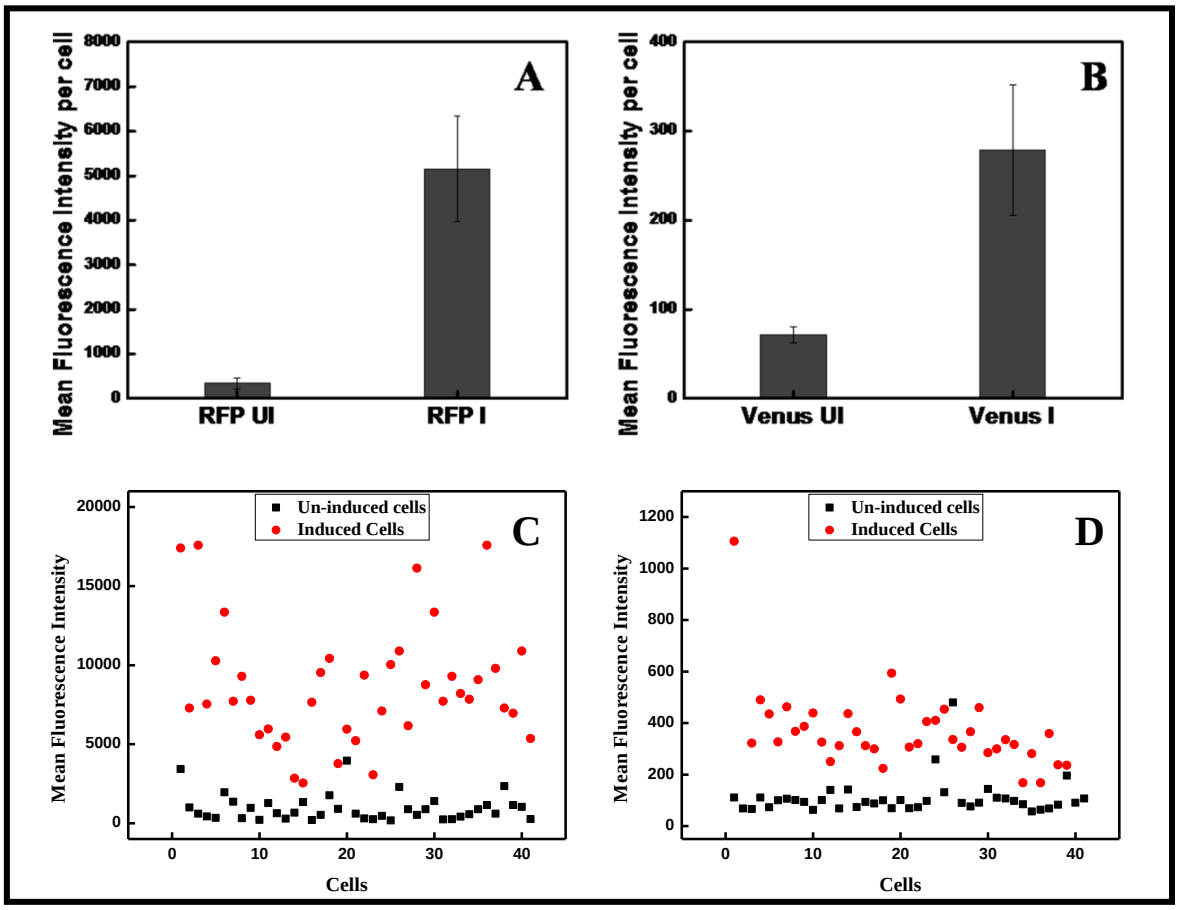


Figure 4: Mean fluorescence intensity per cell seen in the panels in Figure 1 ***(Panels E and F)*** and Figure2 ***(Panels E and F)***. 'UI' stands un-induced cells and 'I' stands for induced cells.

While it is satisfying that the leaky expression in the un-induced cells is so much lower, the fact that it is seen at all can be thought to be somewhat disconcerting, although of course the consequences of such leaky expression would depend upon the required tightness of regulation (which, in turn, could vary from expressed protein to protein, depending upon the protein's toxicity). This provides a method for the visualization and quantitation of leaky expression, while it also establishes that a certain basal level of leaky expression occurs even in the repressed state with pQE-30-M15[pREP4] (Figure 1) and pET-28c-BL21Star(DE3)[pLysS] (Figure 2). In particular, it appears that the latter system shows very insignificant fold-difference between leaky expression and induced expression; however, as already mentioned, this could owe to problems with folding of the fusion protein to the fluorescent state, or to the stability of the fluorescent fusion protein itself. Notably, a recent review indicates that the current system of choice to avoid leaky expression of toxic proteins using pET plasmids and a BL21(DE3) type of bacterial strain is one using an rha PBAD promoter system on a pET plasmid and the Lemo21(DE3) strain which is related to the BL21(DE3) strain [35]. This suggests that others have also discovered that the

pET/BL21(DE3) system has problem of leaky expression of toxic proteins, although we have not found explicit reports in the literature.

Theoretically speaking, there is reason to assume that leaky expression can occur even in tightly-regulated systems. While several scenarios facilitating such expression might be conceived, we point specifically to the 'transient de-repression' mechanism predicted by Guptasarma [36] and demonstrated by Chattoraj and co-workers [37], and to another mechanism proposed by Anthony and coworkers [38], involving read-through transcription from upstream promoters. Transient de-repression involves the temporary displacement of repressor molecule(s) from the operator sequence by the passage of the replication fork, each time a plasmid undergoes replication. For systems in which the availability of repressor is barely sufficient to keep existing operator copies repressed, the doubling of the copies of the operator sequence upon replication must be matched by the availability of more copies of repressor. This can take time if repressor molecules are not present in a significant molar excess over operator copies, allowing a small time window for expression through such transient de-repression. On the other hand, upstream read-though transcription could result from the non-availability of sufficiently effective transcriptional terminator sequences in genes present upstream of the multiple-cloning site in the vector which is kept repressed in a tightly-regulated fashion. Our results show that either through upstream read-through transcription, or transient de-repression, low levels of leaky expression are seen from systems that are otherwise tightly regulated.

The present method allows examination of cell-to-cell variations in leaky and induced expression as well as differences between populations at a very high level of sensitivity; in principle, with the right equipment, the sensititivity could be high enough to allow detection of a single molecule of the fluorescent fusion protein per cell [23], although in the present instance the sensitivity is probably of the level of hundreds of molecules per cell. In contrast, for comparison, a technique like immunoblotting would yield a highest possible sensitivity of about 1-5 picograms per population of cells [39], amounting to about 100 million molecules of a 35 kDA protein like RFP-HU-A.

7.3.3. Presence of HU-B in biofilms

To microscopically visualize the presence of HU in bacterial biofilms, we used fusion construct in which Venus, a variant of yellow fluorescent protein was fused to N-terminal of HU-B protein. The fusion construct was expressed from biofilm forming MG1655 *E. coli*

cells. DIC and fluorescent images of the Venus-HU-B over-expressing MG1655 cells were collected in X-Y-Z axis using wide-field high resolution microscope. Figure 5 A shows Z-stacks of X-Y planes of the cells and Figure 5 B shows combined XY stack of projected XZ and YZ images. The following speculations can be made from the images: 1). The Venus-HU-B over-expressing cells are elongated in shape. Elongation of cells on over-expression of HU-B alone was observed not just for MG1655, but for other strains of *E. coli* also. This elongation of cells could be because of failure of the cells to divide (data not shown here). 2). It is evident from different Z-stacks projections from top of the cell to the bottom, before fluorescence from inside the cell is observed (stack 5 of Figure 5A), owing to Venus fluorescent protein, fluorescence is observed at the surface of the cells as seen in stack 2of Figure 5A, following which, fluorescence surrounding the cells is observed as seen in stacks 3 and 4. Combined XY stack of projected XZ and YZ images in Figure 5B shows that surrounding the fluorescence from cylindrical cells, is a hollow cylinder with no fluorescence, outside which a layer of fluorescing molecules is observed. This suggests that, the over-expressed, fluorescently tagged HU-B besides being present inside the cells also coats the cells by binding to something on the surfaces of cells. It has previously been shown that a homologue of HU, hlpA form *Staphylococcus pyogenes* binds to lipotechoic acid which is a constituent of cell wall and is responsible for cell-cell adhesion [40]. Our results also indicate the HU-B over-expressed from MG1655 *E. coli* cells also binds to cell wall. *E. coli* being a gram-negative bacteria, lacks lipotechoic acid. In gram negative bacteria the observed binding of HU-B to cell surface/ cell wall could be because of binding of HU-B to some other cell wall components like lipopolysaccharides. To sum up, our data points towards binding of HU-B protein to bacterial cell wall which gives a possible physical role of HU in bacterial biofilms. We speculate the HU bound of cell wall also binds to eDNA and hence leading to stabilization of biofilms.

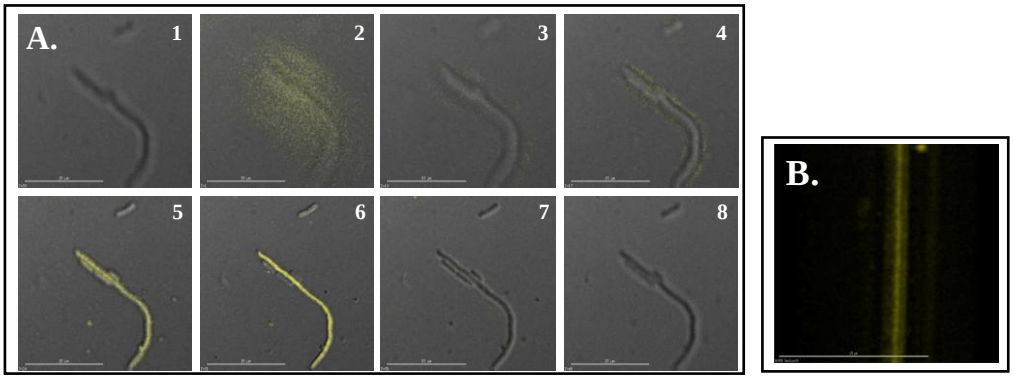


Figure 5: Panel A: Z-stacks of DIC and fluorescent images of the Venus-HU-B over-expressing MG1655 cells ***Panel B:*** Combined XY stack of projected XZ and YZ images.

7.5 References

[1] K. Terpe, Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, Appl Microbiol Biotechnol 72 (2006) 211-222.

[2] O. Khow, S. Suntrarachun, Strategies for production of active eukaryotic proteins in bacterial expression system, Asian Pac J Trop Biomed 2 (2012) 159-162.

[3] N. Broadway, Recombinant Protein Expression: Vector-Host Systems, Mater Methods 2 (2012) 123.

[4] F. Saida, Overview on the expression of toxic gene products in Escherichia coli, Curr Protoc Protein Sci Chapter 5 (2007) Unit 5 19.

[5] F. Saida, M. Uzan, B. Odaert, F. Bontems, Expression of highly toxic genes in E. coli: special strategies and genetic tools, Curr Protein Pept Sci 7 (2006) 47-56.

[6] F.W. Studier, B.A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, J Mol Biol 189 (1986) 113-130.

[7] A.C. Chang, S.N. Cohen, Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid, J Bacteriol 134 (1978) 1141-1156.

[8] J. Huang, J. Villemain, R. Padilla, R. Sousa, Mechanisms by which T7 lysozyme specifically regulates T7 RNA polymerase during different phases of transcription, J Mol Biol 293 (1999) 457-475.

[9] B.A. Moffatt, F.W. Studier, T7 lysozyme inhibits transcription by T7 RNA polymerase, Cell 49 (1987) 221-227.

[10] F.W. Studier, Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system, J Mol Biol 219 (1991) 37-44.

[11] X. Zhang, F.W. Studier, Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme, J Mol Biol 269 (1997) 10-27.

[12] V. Spehr, D. Frahm, T.F. Meyer, Improvement of the T7 expression system by the use of T7 lysozyme, Gene 257 (2000) 259-267.

[13] P. Babitzke, S.R. Kushner, The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of Escherichia coli, Proc Natl Acad Sci U S A 88 (1991) 1-5.

[14] L. Taraseviciene, A. Miczak, D. Apirion, The gene specifying RNase E (rne) and a gene affecting mRNA stability (ams) are the same gene, Mol Microbiol 5 (1991) 851-855.

[15] J.W. Dubendorff, F.W. Studier, Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor, J Mol Biol 219 (1991) 45-59.

[16] J.G. Sutcliffe, Complete nucleotide sequence of the Escherichia coli plasmid pBR322, Cold Spring Harb Symp Quant Biol 43 Pt 1 (1979) 77-90.

[17] M.R. Green, J. Sambrook, Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press New York, 2012.

[18] R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements, Nucleic acids research 25 (1997) 1203-1210.

[19] M. Chalfie, Y. Tu, G. Euskirchen, W.W. Ward, D.C. Prasher, Green fluorescent protein as a marker for gene expression, Science 263 (1994) 802-805.

[20] T. Nagai, K. Ibata, E.S. Park, M. Kubota, K. Mikoshiba, A. Miyawaki, A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications, Nature biotechnology 20 (2002) 87-90.

[21] V. McGovern, N. Higgins, R. Chiz, A. Jaworski, H-NS over-expression induces an artificial stationary phase by silencing global transcription, Biochimie 76 (1994) 1019-1029.

[22] M. Wery, C.L. Woldringh, J. Rouviere-Yaniv, HU-GFP and DAPI co-localize on the Escherichia coli nucleoid, Biochimie 83 (2001) 193-200.

[23] X.S. Xie, P.J. Choi, G.-W. Li, N.K. Lee, G. Lia, Single-molecule approach to molecular biology in living bacterial cells, Annu. Rev. Biophys. 37 (2008) 417-444.

[24] H.-C. Flemming, J. Wingender, The biofilm matrix, Nature Reviews Microbiology 8 (2010) 623-633.

[25] R. Slinger, F. Chan, W. Ferris, S.-W. Yeung, M.S. Denis, I. Gaboury, S.D. Aaron, Multiple combination antibiotic susceptibility testing of nontypeable Haemophilus influenzae biofilms, Diagnostic microbiology and infectious disease 56 (2006) 247-253.

[26] T.D. Starner, J.D. Shrout, M.R. Parsek, P.C. Appelbaum, G. Kim, Subinhibitory concentrations of azithromycin decrease nontypeable Haemophilus influenzae biofilm formation and diminish established biofilms, Antimicrobial agents and chemotherapy 52 (2008) 137-145.

[27] R. Cavaliere, J.L. Ball, L. Turnbull, C.B. Whitchurch, The biofilm matrix destabilizers, EDTA and DNaseI, enhance the susceptibility of nontypeable Hemophilus influenzae biofilms to treatment with ampicillin and ciprofloxacin, MicrobiologyOpen 3 (2014) 557-567.

[28] M. Okshevsky, R.L. Meyer, The role of extracellular DNA in the establishment, maintenance and perpetuation of bacterial biofilms, Critical reviews in microbiology (2013) 1-11.

[29] A. Nur, K. Hirota, H. Yumoto, K. Hirao, D. Liu, K. Takahashi, K. Murakami, T. Matsuo, R. Shu, Y. Miyake, Effects of extracellular DNA and DNA-binding protein on the development of a Streptococcus intermedius biofilm, Journal of applied microbiology 115 (2013) 260-270.

[30] T. Das, P.K. Sharma, H.J. Busscher, H.C. van der Mei, B.P. Krom, Role of extracellular DNA in initial bacterial adhesion and surface aggregation, Applied and environmental microbiology 76 (2010) 3405-3408.

[31] M. Harmsen, M. Lappann, S. Knøchel, S. Molin, Role of extracellular DNA during biofilm formation by Listeria monocytogenes, Applied and environmental microbiology 76 (2010) 2271-2279.

[32] D. Liu, H. Yumoto, K. Hirota, K. Murakami, K. Takahashi, K. Hirao, T. Matsuo, K. Ohkura, H. Nagamune, Y. Miyake, Histone-like DNA binding protein of Streptococcus intermedius induces the expression of pro-inflammatory cytokines in human monocytes via activation of ERK1/2 and JNK pathways, Cellular microbiology 10 (2008) 262-276.

[33] S. Goodman, K. Obergfell, J. Jurcisek, L. Novotny, J. Downey, E. Ayala, N. Tjokro, B. Li, S. Justice, L. Bakaletz, Biofilms can be dispersed by focusing the immune system on a common family of bacterial nucleoid-associated proteins, Mucosal immunology 4 (2011) 625-637.

[34] A. Devaraj, S.S. Justice, L.O. Bakaletz, S.D. Goodman, DNABII proteins play a central role in UPEC biofilm structure, Molecular microbiology (2015).

[35] G.J. Gopal, A. Kumar, Strategies for the production of recombinant protein in Escherichia coli, The protein journal 32 (2013) 419-425.

[36] P. Guptasarma, Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli?, Bioessays 17 (1995) 987-997.

[37] S. Mukhopadhyay, D.K. Chattoraj, Replication-induced transcription of an autorepressed gene: the replication initiator gene of plasmid P1, Proceedings of the National Academy of Sciences 97 (2000) 7142-7147.

[38] L.C. Anthony, H. Suzuki, M. Filutowicz, Tightly regulated vectors for the cloning and expression of toxic genes, Journal of microbiological methods 58 (2004) 243-250.

[39] B.T. Kurien, R.H. Scofield, Protein blotting: a review, Journal of immunological methods 274 (2003) 1-15.

[40] M.W. Stinson, R. McLaughlin, S.H. Choi, Z.E. Juarez, J. Barnard, Streptococcal histone-like protein: primary structure of hlpA and protein binding to lipoteichoic acid and epithelial cells, Infect Immun 66 (1998) 259-265.