Mechanistic Insights into the Stabilisation of Gram Negative Bacterial Biofilms by HU And Cloning, Expression and Purification of IHF

Anupama Pavithran

MS13016

A dissertation submitted for the partial fulfilment of

BS-MS dual degree in Science

Indian Institute of Science Education and Research, Mohali

April 2018

Certificate of Examination

This is to certify that the dissertation titled ”**Mechanistic insights into the stabilization of gram negative bacterial biofilms by HU, cloning expression and purification of IHF**“ submitted by **Anupama Pavithran** (Reg. No. MS13016) for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Kausik Chattopadhyay Dr. Arunika Mukhopadhaya Prof. Purnananda Guptasarma

(Supervisor)

Dated: 20 April 2018

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Prof. 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.

Anupama Pavithran

(Candidate)

Dated: April 20, 2018

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Prof. Purnananda Guptasarma

(Supervisor)

ACKNOWLEDGEMENT

First and foremost I would like to thank my supervisor Dr. Purnananda Guptasarma for being a supportive mentor throughout the year and for all the insightful conversations shared during lab meetings that had helped me both academically and otherwise. He had always been exceptionally kind, an inspiring teacher and an exceptional thinker. I'm grateful for the excellent example he has provided as an innovative scientist with deep insights and the best teacher.

I thank the institution for providing the best research facilities and scientific environment. I express my gratitude towards all the faculties of the department for their guidances.

Extremely thankful to all the lab members who made the stay delightful even during the stressful times. Even though the stay was brief the lab had given me the best memories to cherish. A special thanks to Bhisham bhaiya for personally overseeing my work, for the creative thoughts and suggestions he had given and for the patience and support he had shown inspite of my mistakes. I thank Neeraj sir and Rajendra sir for their timely suggestions and help whenever needed, Pallavi di and Arpana ma'm for their friendly nature and support and Prince bhaiya for his cool attitude and systematic approach to every details. Thanks to Arpita di for her timely advices and for being an amazing friend, Arpita Sarkar and Archit for their remarkable sense of humour making the lab lively. Sreelakshmi, Harpreet, Angel, Kanica and Shashi for their company and bringing a smile on my face whenever overwhelmed with work.

I also owe thanks to Dr.Kanika as her PhD work had largely helped my master's thesis.

I express my love and affection to my parents and my brother who had always been the pillar of strength and had constantly motivated to go on in life. My friends in IISER were instrumental in making me what I am and had been constant source of love and criticisms. Thanks a lot for making my five year long journey worth remembering.

Finally I thank the almighty for bestowing his blessings upon me.

LIST OF FIGURES

LIST OF TABLES

NOTATIONS

EPS	-	Extra cellular polymeric substance
e-DNA	-	Extracellular DNA
DNA	-	Deoxyribonucleic acid
LPS	-	Lipopolysaccharide
MST	-	Microscale thermophoresis
Bp	-	Base pair
MQ	-	Milli Q
ORI	-	Origin of replication
LB	-	Luria bertani
NAP	-	Nucleoid associated protein
IMAC	-	Immobilisation mobility affinity chromatography
kDa	-	kilo Dalton
TAE	-	Tris-acetate-EDTA
TEMED	-	Tetramethyl ethylene diamine
µ	-	micro
IPTG	-	Isopropyl-beta-D-thio galactopyranoside
PAGE	-	Polyacrylamide gel electrophoresis

CONTENTS

ABSTRACT

Like histone and many other DNA binding proteins in eukaryotes, bacteria have NAPs (nucleoid associated proteins) which confine their chromosome in a small space called a nucleoid. NAPs (nucleoid associated proteins) help in organising the bacterial chromosome into domains for effective compaction and gene expression. Outside bacterial cells, NAPs have been implicated in biofilm formation, maintaining the structural integrity of the biofilm (being one of the major components of EPS, or extracellular polymeric substance). Despite intensive studies on their role in biofilms their architectural role has not been fully understood. Our study exposes the role of HU-the most abundant DNA binding protein in biofilms, as a "glue" between DNA and LPS, helping bacteria to be embedded in the EPS matrix, as a component of a larger ongoing study in this area. Interaction of HU with LPS was validated through flow cytometry, microscale thermophoresis, and glutaraldehyde crosslinking experiments. We further explored whether the results were applicable to other DNABII proteins as well, and this got us interested in IHF (Integration –Host Factor), a DNA binding protein in biofilms highly similar to HU sequentially and structurally. Purification of this protein in substantial quantities had been difficult. Individual proteins were unable to fold hence we co-transformed plasmids encoding both IHF proteins (A and B) into the same cell. In the study we showed a time-dependent expression of the protein(s) with maximum overexpression occurring at fifth and sixth hours. Despite being soluble and available in the supernatant, the proteins' hexahistidine tags appeared to be unavailable to bind with Ni-NTA. Subsequently, we purified the protein through cation exchange chromatography. From the studies we hypothesised the existence of a balance between the IHF proteins for their substantial expression.

Chapter 01

A BACTERIAL HISTONE-LIKE PROTEIN (HU) BINDS TO LIPOPOLYSACCHARIDE (LPS) ON BACTERIAL CELL SURFACES TO ENTRAP BACTERIA WITHIN BIOFILMS MADE OF e-DNA

1. 1. INTRODUCTION

"Proteins hold the key to the whole subject of the molecular basis of biological reactions"

-Linus Pauling

Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells. They form the molecular instruments through which genetic information is expressed. At the most basic level, proteins are a chain of smaller units called amino acids (There are 20 varied amino acids that make up a protein) determined by the nucleotide sequence in a gene. Through folding of these random coils, each protein acquires a unique characteristic and functional 3-D structure which is also their native state. Failure to achieve these states form an inactive protein causing many diseases.. To prevent protein misfolding or aggregation, cell often takes the assistance of molecular chaperones which are mostly heat shock proteins that are expressed in response to cellular stresses or high temperatures. According to the' Sequence –Structure –Function' paradigm, the amino acid sequence of a protein determines its 3-D structure that determines its function [1]. These macromolecules have four different levels of structures namely '*primary*'-linear chain of amino acid sequence in a polypeptide chain, '*secondary*' –folded polypeptide bond due to interactions between atoms of backbone; describing α-helix and β-sheet structures, '*tertiary*'-overall 3-D structure of the protein due to side-chain interactions, '*quaternary*'-protein consisting of more than one amino acid chain. The difference in properties of the proteins is due to the enormous number of ways, its monomeric units can be arranged. Based on biological function proteins are classified as below:

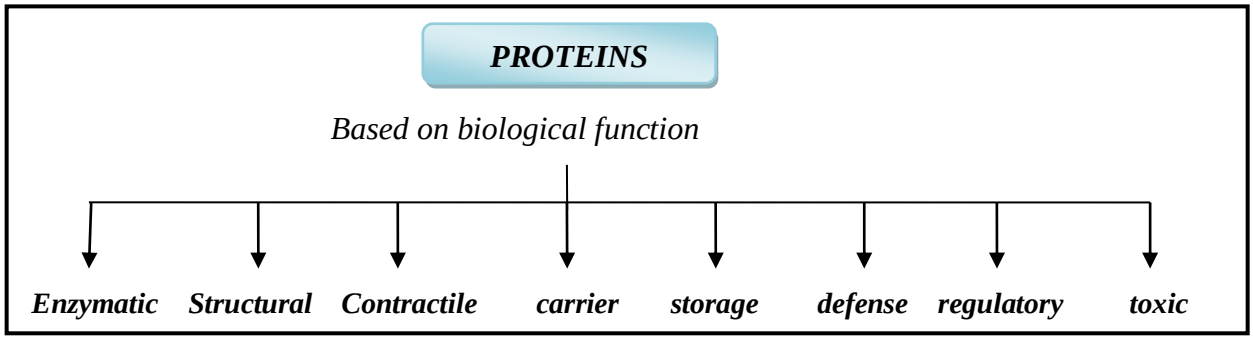


Fig:1.1 *Classification of protein based on their biological function*

Regulatory proteins basically help regulate cellular or physiological activity. The protein studied in this thesis is a regulatory protein which has a DNA binding activity and thus helping in the biosynthesis of enzymes and RNA molecules that play a role in cell division. They contain domains that is specific to DNA sites. Regulatory proteins with DNA-binding activity are termed as DNA-binding proteins.

1.1.1. Nucleoid-associated proteins

In all the organisms, the DNA has to be organised properly for storage such that they are compatible with DNA replication, chromosome segregation and gene expression. This signifies the role of certain proteins to compact the chromosomal DNA into a nucleoid that occupies about one-fourth of the intracellular volume [2]. The organisation of DNA in eukaryotes is different from that of the prokaryotic cells; the genomes of the prokaryotes are contained in single chromosomes, usually circular DNA molecules whereas the genomes of eukaryotes are composed of multiple chromosomes, each containing a linear molecule of DNA. In eukaryotes, the DNA is tightly bound to basic proteins called the' histones' that package the DNA in an organised manner in the nucleus [3]. Analogous to the histones in chromatin of the eukaryotes, the prokaryotes, have the histone-like proteins, collectively referred to as the 'nucleoid –associated proteins' owing to their cellular location that influence their gene expression. These proteins have an architectural role to maintain structural integrity and possess several DNA –binding functions such as transcription, replication, and recombination influencing their activation and repression modes. They change DNA conformation through wrapping, coiling or bending of DNA with their expression levels dependent on the growth phases, certain proteins being abundant in the logphase while others in the stationary phase. For instance, one of the DNA binding proteins, HUα is expressed mostly during the early exponential growth phase whereas HUβ is expressed during the stationary phase. HUαα homodimers are present mostly in the lag and exponential phase whereas HUαβ heterodimers expression shows up during early and late stationary phases.[2]. There are about 12 DNA binding protein species isolated from the genome of *Escherichia coli* all of which are listed in the table below[4].

Proteins or group of proteins	DNA wrapping	DNA bridging	DNA bending	Binding motif	Molecular mass	Native protomer
HU	yes		yes	A DNA structural motif in dsDNA or ssDNA with mild preference for AT –rich or curved DNA	~9kDa	heterodimer
Lrp	yes	yes		(T/C)AG(A/T/C)A(A/T) ATT(A/T)T(A/T/G)CT (A/G)	~18kDa	homodimer
MukB	yes		yes	ND	~175kDa	homodimer
Fis				A6 tracts and AT tracts	~11kDa	homodimer
H-NS				AT-rich DNA and TCGATAAATT	~15kDa	Homodimer or heterodimer
IHF	yes		yes	(A/T)ATCAANNNNTT (A/G)	~11kDa	heterodimer
Dps				ND	~19kDa	
StpA				AT-rich DNA	~15kDa	
CbpA				Curved DNA	~33kDa	
CbpB				Curved DNA	~33kDa	
EbfC				GTNAC	~11kDa	homodimer
MvaT				AT-rich DNA	~9kDa	homodimer

Table 1.1 : *all the twelve DNA binding proteins from E.coli genome*

Among the histone-like proteins Fis (factor for inversion stimulation), H-NS (Histone-like nucleoid structuring) ,HU (heat unstable) and IHF (Integration Host factor)are the most intensively studied members. [5].

1.1.2. HU (Heat Unstable protein)

Among all the DNA binding proteins HU is the most abundant and conserved protein across the bacterial species which was first isolated from *Escherichia coli* strain U93

(ribonuclease negative) and was called factor U [6]. Other than the eubacteria, it is also present in a few species archaebacteria [7], bacteriophages [8] and animal viruses [9].The major functions of this non-specific DNA binding protein include:

- Restrain DNA supercoiling
- Condense the chromosome
- Introduce negative supercoiling into circular relaxed DNA template with topoisomerase I
- Assist the action of DnaA protein in the initiation of DNA replication
- Regulate DNA replication process

Its composed of two subunits namely HU-α and HU-ß that have 70% aminoacids identity encoded by *hupA* and *hupB* respectively. The subunit composition varies over the growth period with overall HU declining during the stationary phase [5].

HU is 90 amino acids long with monomeric weight of ~10 KDa, basic and dimeric belonging to the DNABII family of DNABII binding proteins. The intracellular concentrations of this protein is around 12,000-50,000 dimers/cell making it the most abundant of the NAPs.[6]. It has been reported that *E.coli* lacking HU protein are extremely sensitive to gamma radiation and that either of the subunits is essential to restore the normal survival rate[10]. Deletion of HU from E.coli genome is not lethal unless IHF and H-NS are deleted as well. But if HU is the only NAP available, then the deletion causes lethality. Deletion mutants of HU in *E.coli* acquire compensatory mutations and cause changes in gene expression with Δ*hupAB* having the greatest effect followed by Δ*hupA* and minimal effects in Δ*hupB*. Most of the species encode only one of the HU subunits in which hupB is the most conserved [11].

1.1.3. Structural features

Dimers of HU naturally lack both tryptophan and tyrosine residues, a property that is conserved across all bacterial species.

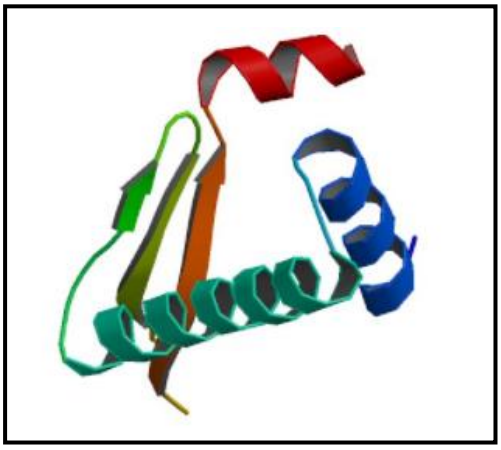


Fig 1.2: *Crystal structure E.coli HU monomer (*1MUL*)*

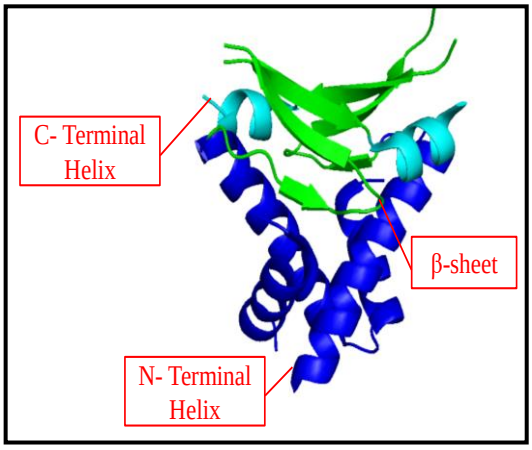


Fig 1.3*: Crystal structureof Anabena HU dimer*

Amino terminal consists of two alpha helices (blue) connected by a turn. C-terminal consists of one alpha helix (red). Rest of the structure is formed by three stranded β-sheet structure (green and orange) including a β-extension in the middle which is structured only in the presence of DNA.In case of hetrodimer α-helices of both subunits are compacted and capped by β-sheets that is extended into two β-ribbon arms[12].

Canonical and non-canonical forms of DNA binding

HU that is present at very high concentrations inside cell , induce very large flexible bends in DNA (~105- >180˚) with a minimum of 9bp; binding preferentially to distorted DNA like three- or four way junctions, nicks and overhangs. Binding to DNA happens in a structure-specific manner.

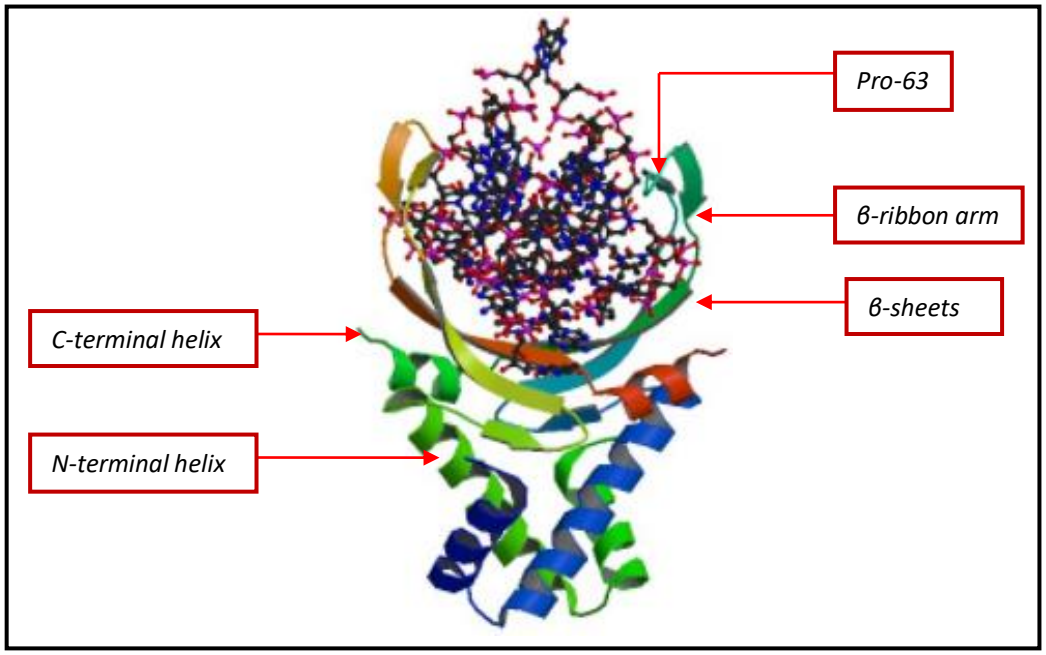


Figure 1.4: ***Anabaena HU-DNA cocrystal structure*** ***(*****1P51*****)***

In 'canonical form of DNA binding', proline residue (63rd position) at the tip of two β-strands play a critical role helping HU to interact with the minor groove of DNA. This mode of DNA binding is hypothesised to be using a lock and key mechanism as it is meant for size specific DNA sites.

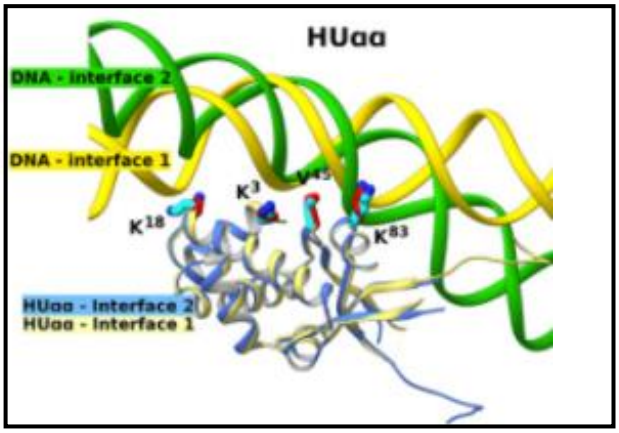


Fig 1.5 *Non-canonical DNA binding sites of HU* [12]

In 'non-canonical form of DNA binding', HU binds to DNA through its subunits side-faces carrying lysine and arginine.HU dimers are formed by inserting V^{45} into the minor groove of DNA. These interactions are universal for sequence –independent DNA binding. The HUαα-HUαα homodimers interact with DNA arginines into their minor groove with higher binding affinity [13].

1.1.4. HU-LPS interaction studies

Its estimated that 65-80% of the human chronic bacterial infections such as endocarditis, urinary tract and cystic fibrosis infections, involve biofilms hence posing a critical medical challenge. Biofilms are association of microorganisms where the cells adhere to other living organisms and non-living substrates. Primary components of biofilms are listed below:

Components	***%Total***	***Description***	***Origin***
Water	Upto 95%	Characteristics determined by dissolved determined by dissolved solutes	
Microbial cells	2-5%		
Polysaccharides	1-2%	Neutral and pol anionic polanionic ; homo and heteropolysaccharides	Extracellular
Proteins	<1-2%	enzymes	Extracellular and cell lysis
DNA and RNA	<1-2%		Cell lysis
Ions		Bound or free	

Table 1.2. *Major components of biofilms*

This is significantly based on the extracellular polymeric substance which has e-DNA as its key component. Other components include DNA binding proteins (DNABII family that include HU-A, HU-B, IHF-A, IHF-B and H-NS) and lipopolysaccharides which provide imperviousness to the antibiotics. e-DNA that contributes to the stability and development of the biofilm is released into the matrix through cell lysis of a sub-population of bacterial cells[11].

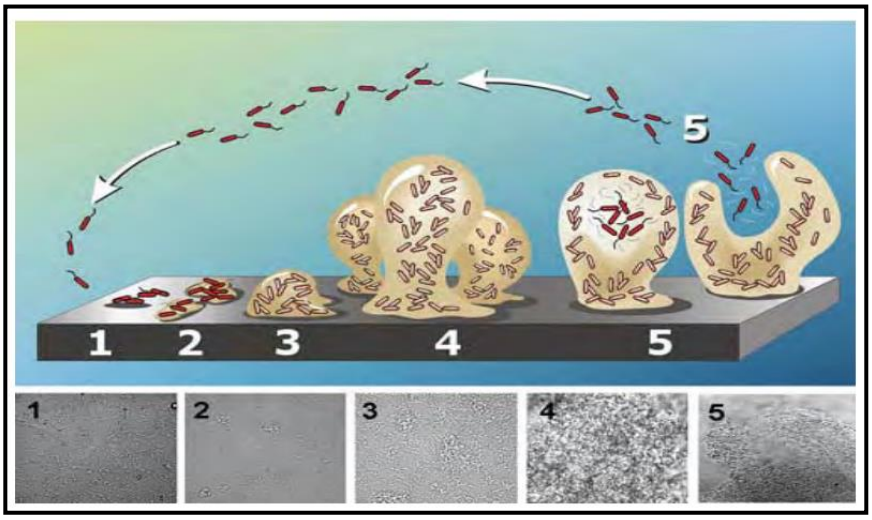


Fig 1.6 : *Sequential formation of biofilms*

The question we asked was what role does e-DNA has in biofilms. *E.coli*, being a gram negative bacteria contains LPS (lipolysaccharide), an endotoxin localised in its outer membrane. This imparts a negative charge to the cell due to the increased abundance of phosphate and carboxylate groups. Analogous to this, DNA too posses negative charge due to the sugar and phosphate backbone.

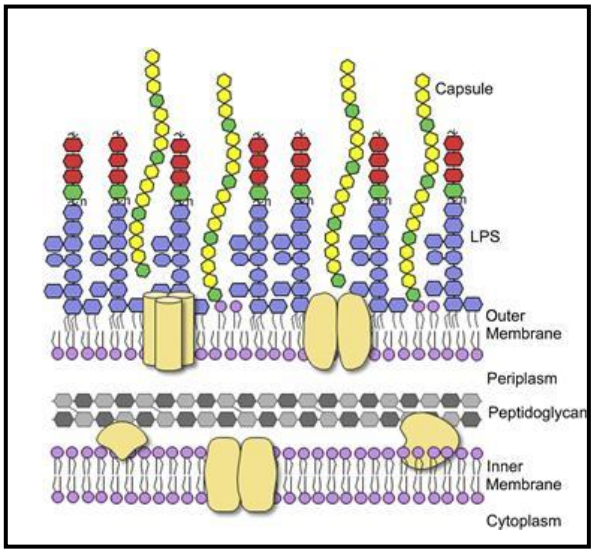


Fig 1.7 *: Structure of LPS in a gram negative bacteria*

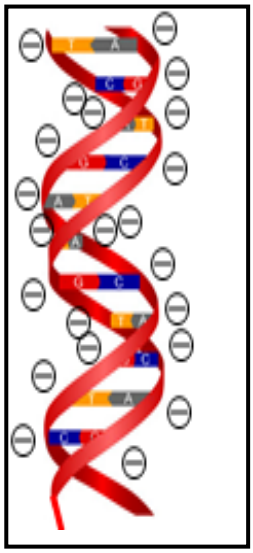


Fig 1.8. *DNA backbone showing negative charges on the surface*

With both of these structures posessing a negative charge the existence of e-DNA in the biofilms is paradoxical because if this is the case, biofilms should be actually disintegrating

through charge-charge repulsions. Hence, we hypothesised the presence of a neutraliser that could hold both the LPS on the cells and the e-DNA together in the biofilm matrix.

1.1.5.DNA binding proteins

These are proteins that modulate the functions of DNA by attaching to them.The DNABII family of proteins that include two of the NAPs; IHF and HU provide structural integrity to the biofilms by binding to DNA with high affinity. Among all the DNA binding proteins of *E.coli* , HU (Histone-like protein) is the most abundant and conserved one, hence we put forward the hypothesis that HU probably would be acting as a molecular glue between LPS and DNA. Recent research shows that HU is secreted by the bacterial cells via a Type IV secretion system that accounts for the relative abundance of the same in biofilms [13].

1.2. MATERIALS AND METHODS

1.2.1. Fluorescence Imaging

HUA-RFP bound with XL1-Blue cells and RFP control cells were imaged using Olympus corporation FV10C-03 confocal microscope in normal fluorescence microscopy mode. Cells were illuminated with RFP excitation wavelength (531-565 nm). Color intensities (red for RFP) were quantified using ImageJ software.

Sample Preparation

E.coli XL1-Blue cells grown overnight were washed thrice with 1X PBS and then resuspended in 500µl of the buffer. Two samples of 400µl each were prepared. Sample 01 had XL1-Blue cells (100µl) incubated with HUA-RFP (3µM)(300µl) and sample 02 had XL1-Blue cells (100µM) incubated with RFP (9µM) (111µl) in 1X PBS. After an incubation period of 30 minutes samples were centifuged for 2minutes at 13,200 RPM. The pellet obtained was washed twice and resuspended in 500µl of 1X PBS.

1.2.2. Microscale Thermophoresis

Principle

This is a fluorescent –based technique used to quantify biomolecular interactions. It monitors the directed movement of fluorescent molecules and also allows a precise analysis of binding with a few microlitres of sample and is independent of size or physical properties hence is advantageous over other fluorescent based methods that shows great dependance on molecular size and requirement for high sample volumes and also over non-fluorescent techniques like iosthermal titration calorimetry, surface plasmon resonance that require high immobilisation and sample consumption.

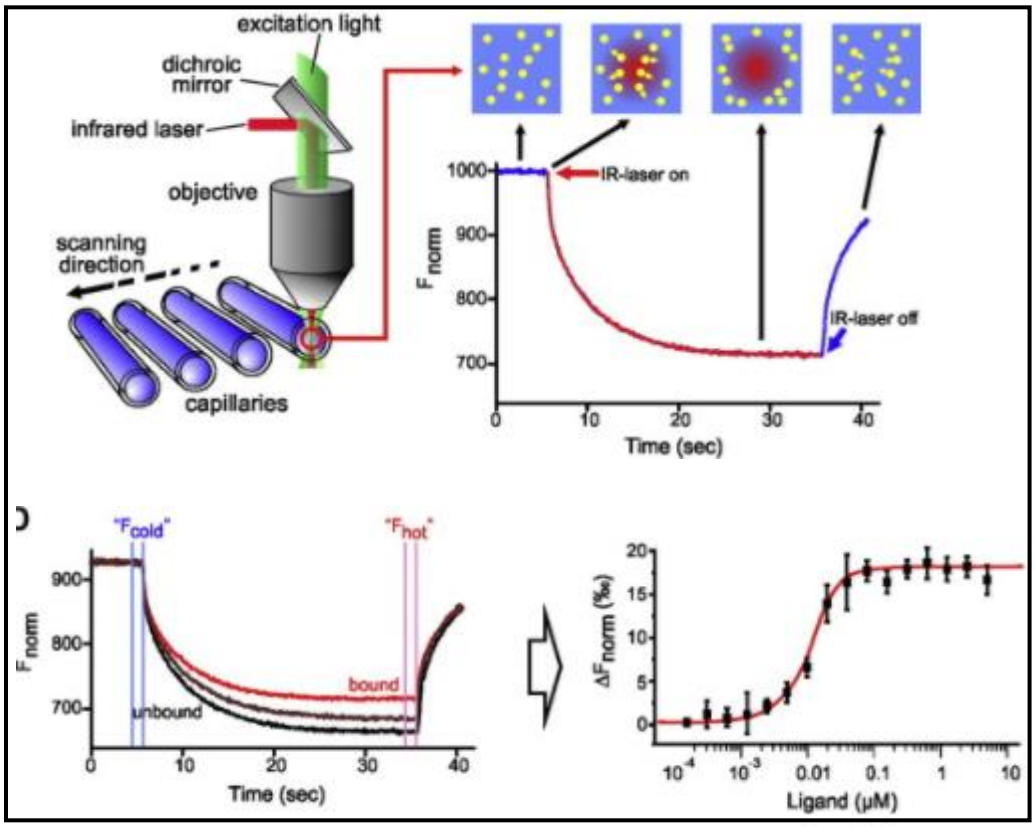


Fig 1.9 *: Working principle of Microscale thermophoresis*

Working principle of the instrument is based on 'Thermophoresis'. There are two lasers namely the IR laser and the fluorescence laser. The IR laser helps in local heating creating a microscale temperature gradients within the 16 glass capillaries that contain the samples. This

causes the fluorescently tagged molecules in the sample that were initially homogenously distributed, to undergo thermophoresis. When the laser is turned off, the molecules go back to their initial state. For each sample following observations are recorded:

- Fluorescence signal before the IR laser is 'ON'
- Temperature dependent changes in the fluorescence intensity
- Moleulcar movement under the effect of heat (thermophoresis)
- Back diffusion whren the IR laser is in 'OFF' state.

In the study,the concetration of proteins was fixed at 250nM in all sixteen capillaries while ligand (LPS) was serially diluted with 2.5mg/ml in first capillary. Experiment was done with Monolith Nanotemper NT.115. The results were analysed using MO affinity analysis software.

1.2.3. Flow Cytometry

The instrument measures different characterisitics of individual particles that flow cell by cell in a stream of fluid through a beam of laser light. Detectors detect the forward scatter light (FSC), side scatter light (SSC) and fluorescence signal. FSC measures size of healthy and dead cell whereas SSC measures granularity of the cell.

Sample Preparation

Four sample sets were prepared with a gradation of DNA concentration (1 µM, 10 µM, 20 µM and 40µM). Initially 100µl HUA tagged with RFP (1.2 mg/ml) is incubated at room temperature with 4-way junction DNA in 1X PBS (phosphate buffer saline) for 20 minutes. Futher 150µl XL-1 Blue cells (grown overnight at 37 ˚C) were added to 150 µl of the above incubated sample. After a centrifuging for 2 minutes at 13,200 RPM, the pellet was washed thrice and then resuspended in 200 µl of PBS. The sample was then acquired on a BD Accuri flow cytometer that monitors 50,000 cells for a specified time counting one cell at a time. Results were analysed using Flowjo or Cell quest pro softwares.The experiment was done on BD Acuri ™ C6.

1.3. RESULTS AND DISCUSSIONS

1.3.1. Fluorescence Imaging to see the sequestering of HU-RFP by E.coli XL1-Blue cells

The experiment was performed to see if HU can associate with the *E.coli* cells. This was a follow up experiment of a recent discovery in the lab that showed *E.coli* cells overexpressing Venus-HUB (yellow tag) had a halo of yellow fluorescence around it*(unpublished work, K.Arora,2015).* To verify whether HUB is binding to the cell surface we experimented on HUA with RFP tag to see its association with *E.coli* cells.

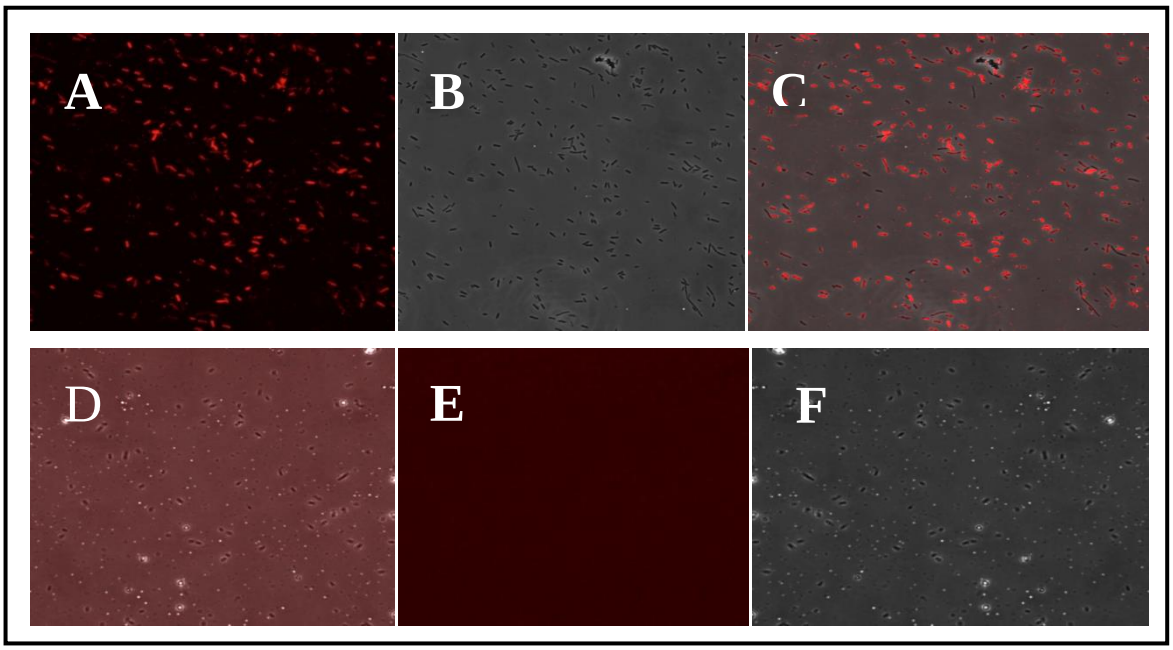


Fig 1.10: *Panel A,B and C shows fluorescence , phase contrast and merged images respectively of HUA-RFP with E.coli cells. Panel D, E and F shows fluorescence, phase contrast and merged images respectively of control RFP cells.*

1.3.2. Flow cytometry establishes clumping of bacterial cells by HU

XL1-Blue culture grown overnight was washed with PBS thrice and then resuspended in 200 µl of the same buffer. Further HUA and HUB were incubated for 10 minutes with cells in separate microcentrifuge tubes such that the concentration of proteins were very low (5 µM HUA and 10 µM) as this was very essential in flow cytometry to reduce autofluorescence. After the incubation cells were harvested, washed with PBS thrice and resuspended in 200 µl of this buffer. Samples were subjected to analysis by flow cytometry. As a control XL1-Blue cells were also prepared finally resuspended in PBS.

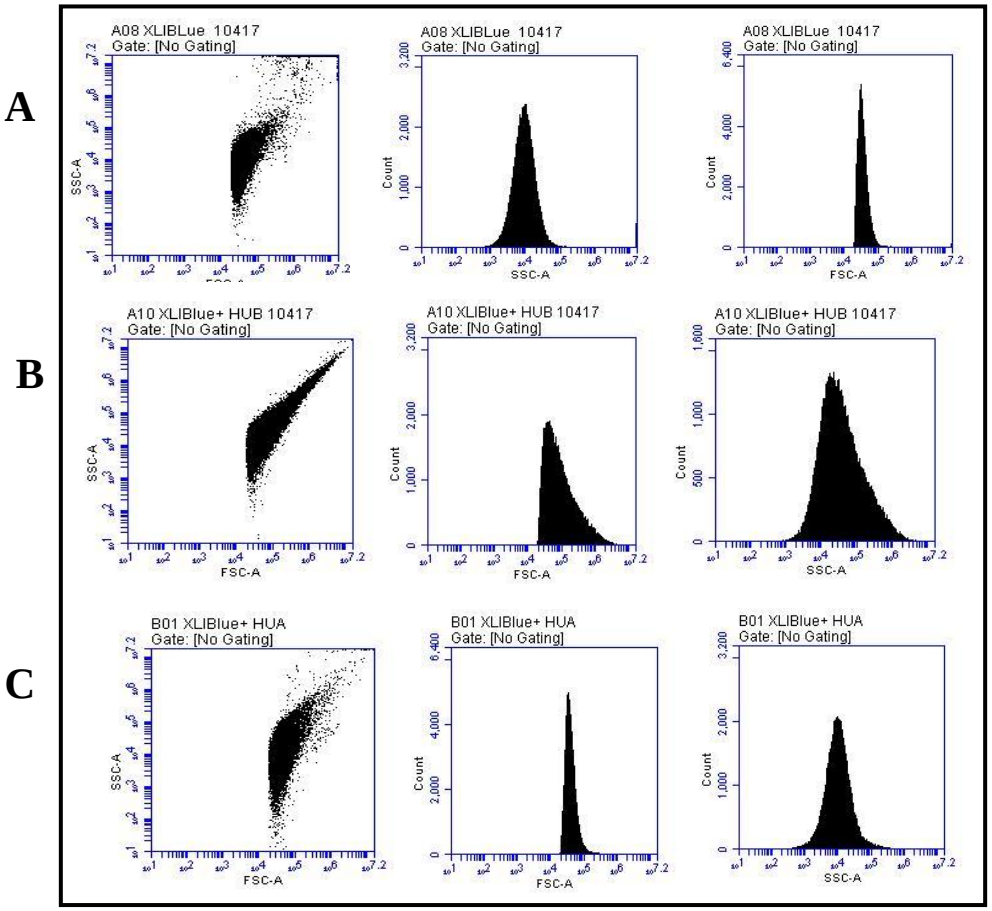


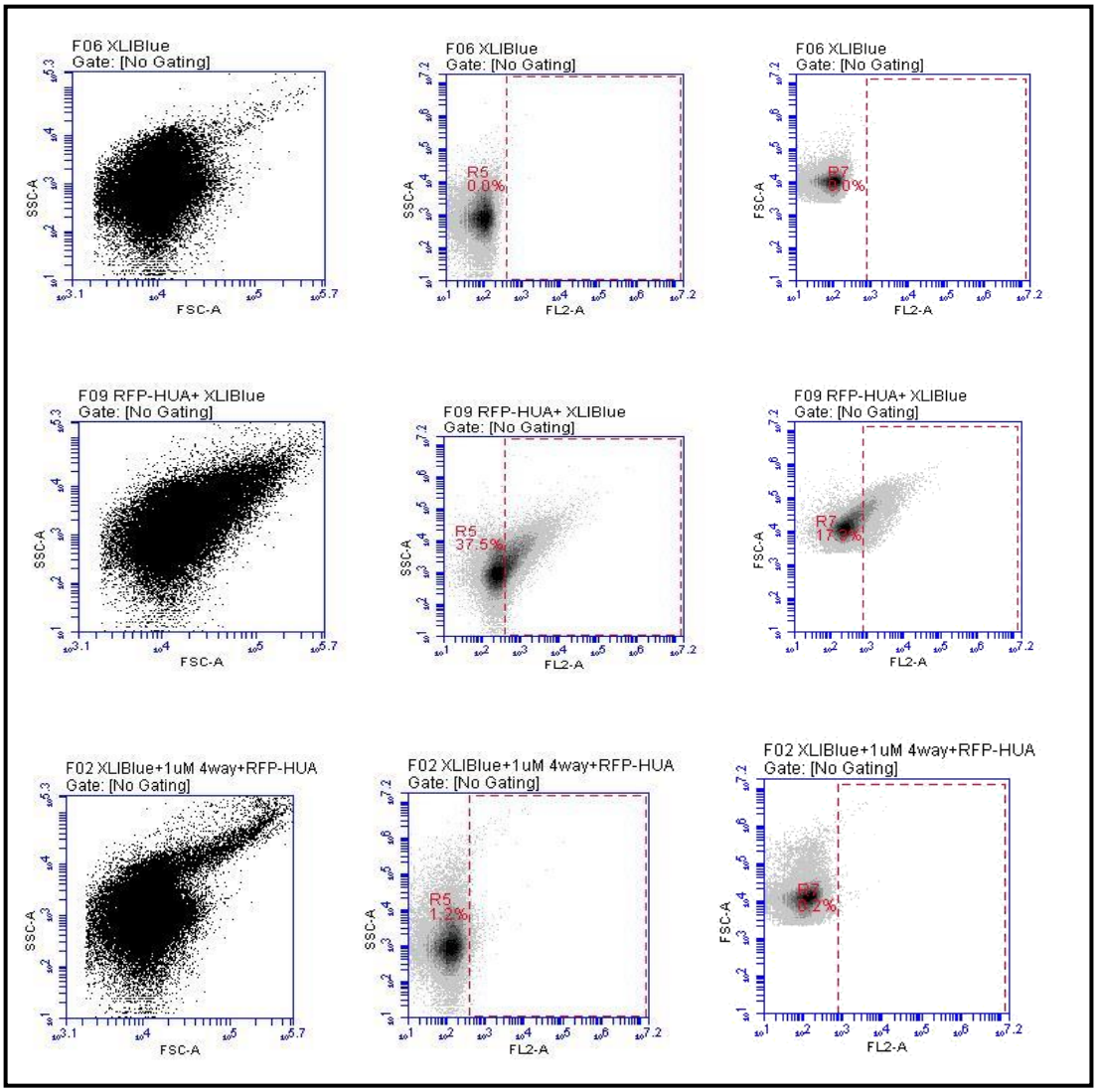
Fig 1.11: *FSC and SSC analysis in flow cytometry studies.* ***Row A*** *shows the control with just E.coli cells,* ***Row B*** *shows XL1-blue cells +HUA ,* ***Row C*** *shows XL1-Blue cells +HUB*

Flow cytometry results showed significant clumping of cells by the protein. Addition of HU increased the forward scatter and side scatter implying increase in size as well as granularity

compared to the control cells. When compared between the protein, HUB showed a better crosslinkg than HUA. Recent research from the lab shows HU-B forms and retains multimeric complexes even in urea, while HU-A only forms dimers (*Bhisham Thakur, 2017)* which provides a reasonable explanation for HUB being more potent than HUA.

Now the HU is shown to be clumping bacterial cells, the next question posed before us was if the DNA binding site of HU was essentially the same as the site through which the protein binds to the cells.

1.3.3. Flow cytometry establishes that DNA competitvely reduces HU mediated clumping



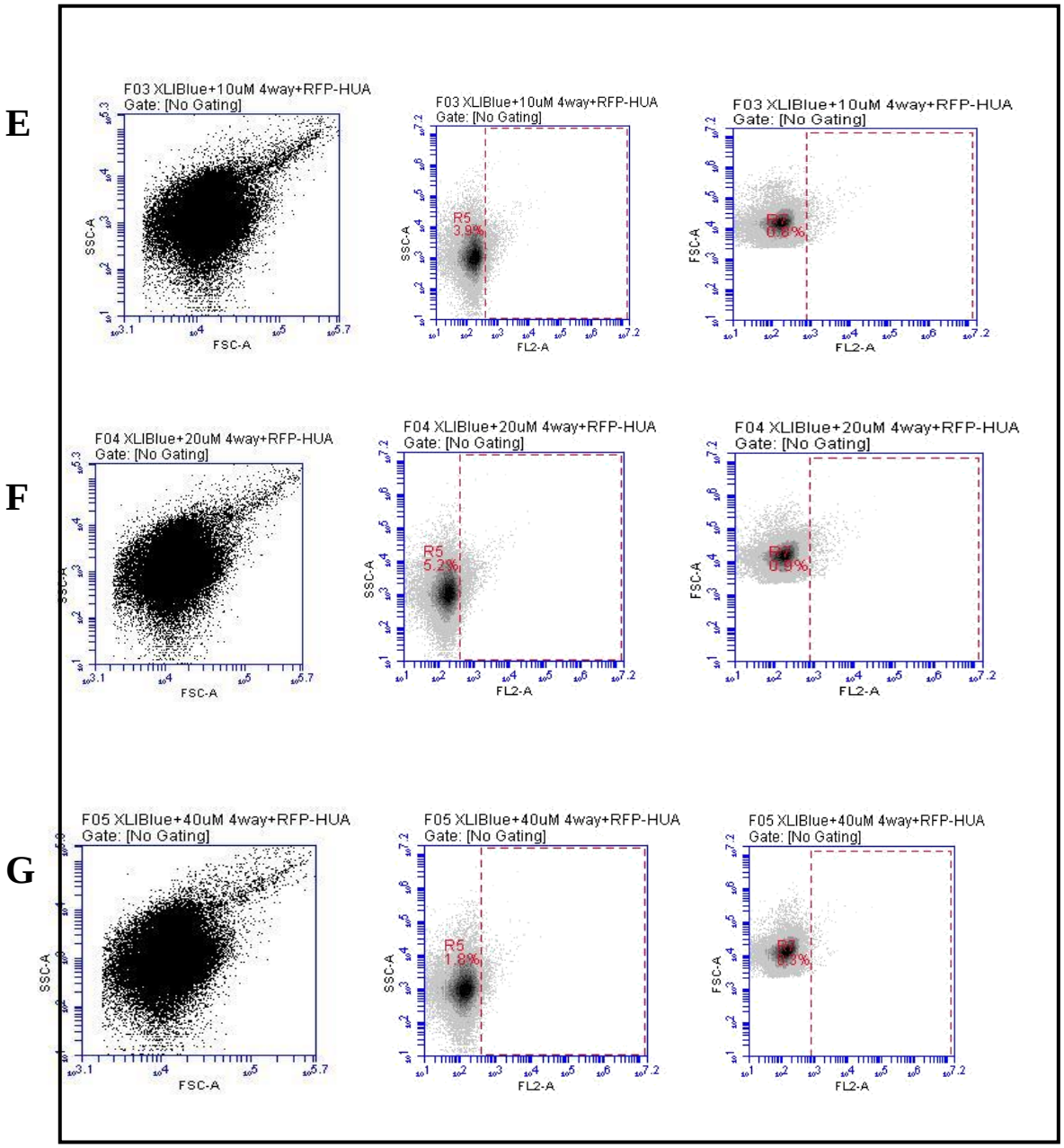


Fig 1.1 2*: FSC, SSC and fluorescence results respectively from flow cytometry in a row.* ***Row A*** *is for control cells,* ***Row B*** *is XL1-Blue cells +HUA-RFP,* ***Row C*** *is 1uM DNA+HU-RFP,* ***Row D*** *10uM DNA+HUA-RFP, ,* ***Row E*** *20uM DNA+HUA-RFP,* ***Row F*** *is 40uM DNA+HU-RFP ,DNA used was 4-way junction.*

There was an effective decrease in FSC snd SSC with increase in concentration of the four–way junction DNA. DNA competitively reduces HU-mediated clumping hence emphasising that HU binds to DNA through same sites as it bind the cells.

Further we addressed the question ‘what is the negatively charged substance in cells onto which the positively charged proteins can bind?”. Looking into the structural details of

lipopolysacchardie (LPS) ,it has sugar moeities and phosphate groups similar to DNA. Hence we put forward the hypothesis that HU might be binding to LPS of the *E.coli* cells by recognising the sugar–linked phosphate groups. Confirmatory binding experiments were performed to validate HU binding with LPS.

1.3.4. Microscale thermophoresis

The technique is handy to establish interaction between a ligand and receptor. Here the protein (receptor) is fluorescently tagged. Of the 16 capillaries the first one had just the ligand (LPS). Subsequent capillaries had ligand serially diluted each time fixing the protein concentration as 250 nM and increasing the ligand concentration. The instrument heats up the samples locally through an IR laser that provide a temperature gradient. Molecules will locally migrate away from the high temperature and come back after cooling, through molecular diffusion which in turn depend upon the hydrodynamic volume.

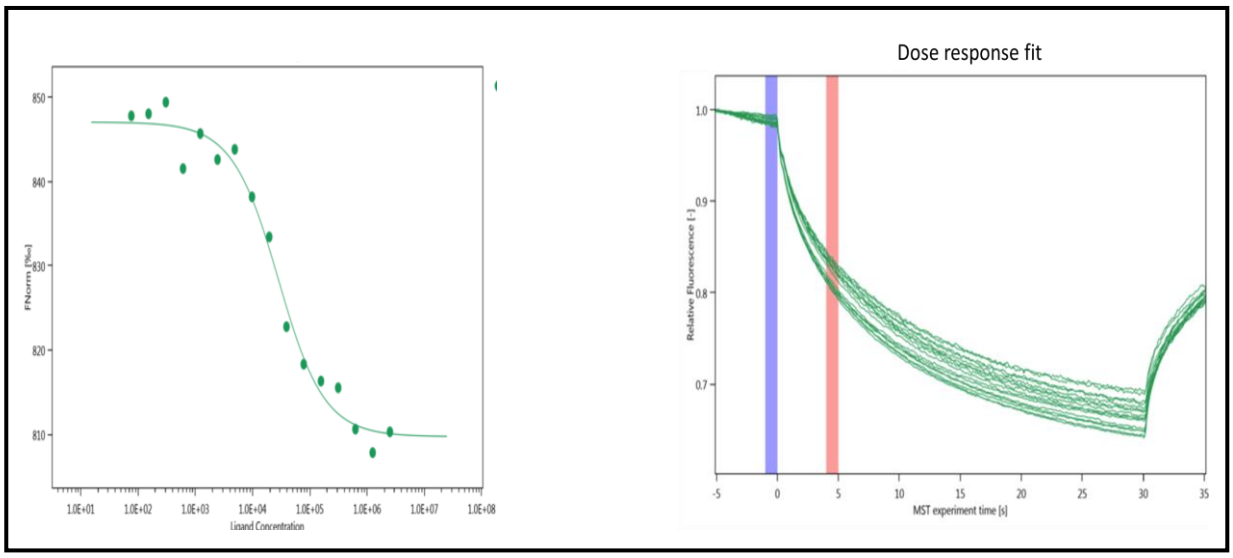


Fig 1.13: *Final fit obtained after analysing dose response fluorescence curve*

Fig **1.14**: *dose response curve showing fluorescence intensity with time*

Constant fluorescence was detected initially as the molecules were homogenously distributed at this time. Further there was a depletion in fluorescence with increase in ligand concentration owing to the increase in hydrodynamic volume which was due to the binding of

LPS with HU. Binding constants could not be estimated because LPS lacks a defined molecular weight .

1.3.5. *Glutaraldehyde crossslinking experiment*

HUA was incubated with glutaraldehyde (0.1%) in PBS for 10 minutes before adding LPS (1mg/ml). Four sets of 15µl samples were prepared of which HUA and LPS were individually prepared in PBS. Third sample had protein incubated with glutaraldehyde and the last had LPS in addition. Samples were subjected to SDS-PAGE for analysis.

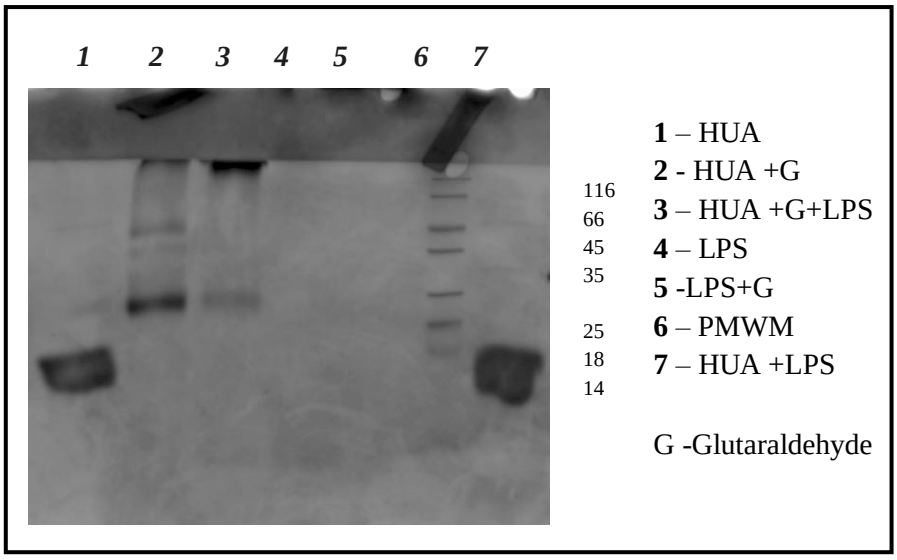


Fig 1.15 :*SDS-PAGE analysis of glutaraldehyde crosslinking experiment to confirm binding of HU with LPS.*

Results show HU by itself was monomeric but when crosslinked with glutaraldehyde showed dimeric (~20 kDa) and tetrameric (~40 kDa) populations. As expected LPS itself was not seen in SDS-PAGE either before or after the addition of glutaraldehyde. HU crosslinked with glutaraldehyde in presence of LPS formed large complexes that did not enter the gel. The results confirm binding of HU ith LPS.

1.4. CONCLUSIONS AND FUTURE DIRECTIONS

Most chronic bacterial infections involve bioflims, the major known constituents of which are extracellular DNA, DNABII proteins (HU, IHF, H-NS) along with other proteins and mucopolysaccharides. Previous researches highlight HU as a major component of biofilms [26] but, its role is still not elucidated. In our study we had put forward the hypothesis that HU bind LPS of the bacterial cells owing to its structural similarity with DNA. This hypothesis was validated through confirmatory binding experiments. The study also gave insights into binding site of HU with LPS as same as the site through which the protein binds to DNA analysed through flow cytometry. We can henceforth conclude that the existence of HU in biofilms is not by chance but have a significant role in adhering *E.coli* cells in a matrix of e-DNA. Currently studies on loop deleted mutants of HU are underway where 21 residues of the loop in canonical DNA binding were replaced with 11 residues of glycine–serine linker sequence. We also aim to mutate the lysine residues in HU to see whether this has any effect in HU binding to LPS in the non-canonical form of DNA binding. The study holds biological relevance as understanding the molecular mechanisms behind biofilm formation would help in developing therapeutic counter measures against biofilm associated chronic infections.

Chapter 02

CLONING EXPRESSION AND PURIFICATION OF IHFs

2.1. INTRODUCTION

Having known about HU being active in promoting biofilm formation, the next question addressed naturally was if this property is applicable to other members of DNABII family as well. Hence we chose IHF that possess a striking similarity to HU in terms of structure and structural fold (sequence identity between 'A' subunits of IHF and HU is 33.3% whereas between 'B' subunits, it is 34.03% which is pretty high). Even with this homology in structure and sequence, HU is sequence independent when binding to DNA while IHF requires a consensus sequence. It was proposed that IHF bind and bend DNA in the same way as HU due to their extensive sequence homology particularly in the putative DNA binding domain. Past studies show HU is not an essential protein for the bacterial species but strains lacking HU possess compensatory mutations which may possibly be due to the effect of IHF [17]. These questions signify the need to focus on IHF.

IHF is a small, basic, sequence -specific dimeric protein with significant architectural role possessing 30% aminoacids identity between the subunits.This 94-99 aminoacid residue protein binds to its 35bp cognate sequence of DNA bending it by ~160˚. IHF primarily help in lambda integration by binding to *attP*, a segment of lambda chromosome [25].

Role of IHF in structural integrity of biofilms is unknown though past researches had highlighted the significance of IHF in biofilms.In *E.coli ihfB* mutant exhibited severe biofilm defective phenotype [24].

2.1.1.Structural features

The fold of IHF is similar to HU, even the unstructured DNA binding loops become structured after binding with DNA in both the proteins.

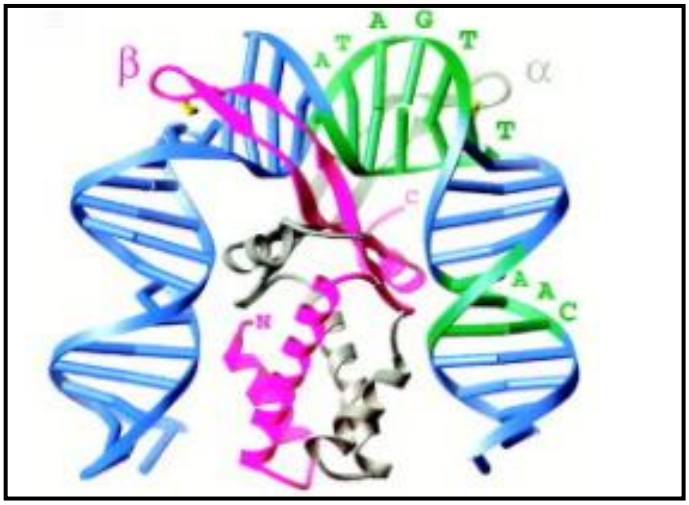


Fig 2.1: S*howing crystal structure of IHF-DNA complex. Grey depicts the a-subunit pink- β-subunit, green-consensus sequence*

IHF binds to the consensus sequence with six-base adenine tract of DNA mainly through phosphodiester bonds and the minor groove. On binding to DNA, IHF introduces two kinks where proline is present at the tip of each arm is intercalated from the minor groove, between the base pairs.

2.1.2. Comparison of HU and IHF

Differences:

- Unlike HU that is the most abundant protein in biofilms, IHF abundance is as low as 2.2 % [24].
- HU has no consensus sequence specificity but have high preferences for supercoiled DNA especially nicked, cruciformed and four-way junction DNA (distorted DNA). K_d value for undistorted DNA is higher than that for distorted DNA which is 2 nM (high affinity). In contrast IHF shows high sequence specificity WATCAANNNNTTR where W is A or T and R is a purine) binding to DNA (considerably more than 9bp) more weakly than HU with a K_d value as high as 20-30 µM (implies low affinity).[10]
- Generally, IHFs are obligate heterodimers, whereas HUs can be either heterodimers or homodimers.[13]
- HU bend DNA with 105˚-139˚ whereas IHF has a bend angle of ~160˚[22].

- The main difference between the structure of HU and IHF is the positioning of β-arms which is flexible in HU in the absence of DNA. The arm of IHFA bind DNA more efficiently than HU and IHFB due to the additional prolines in the sequence [23].

Similarities:

- 30-40% sequence identity between the proteins [9]
-

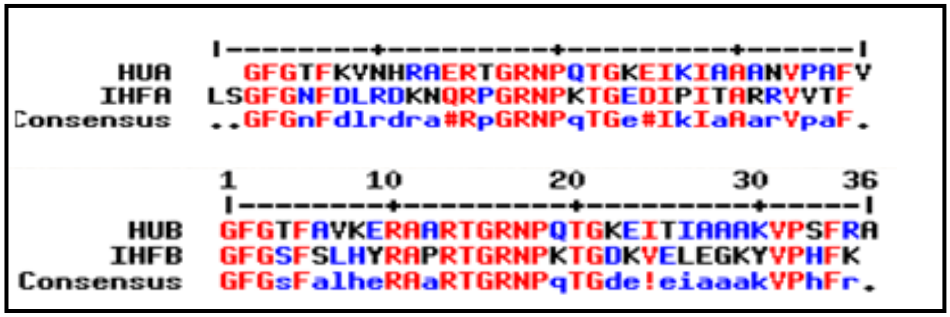


Fig 2.2 : *Sequence comparison between HU and IHF where red indicates identical residues (from Multialign)*

- Share a common fold composed of a largely α-helical 'body' with two protruding β-ribbon 'arms [7]
- Both proteins bind DNA with a pair of β-sheet arms occupying the minor groove.
-

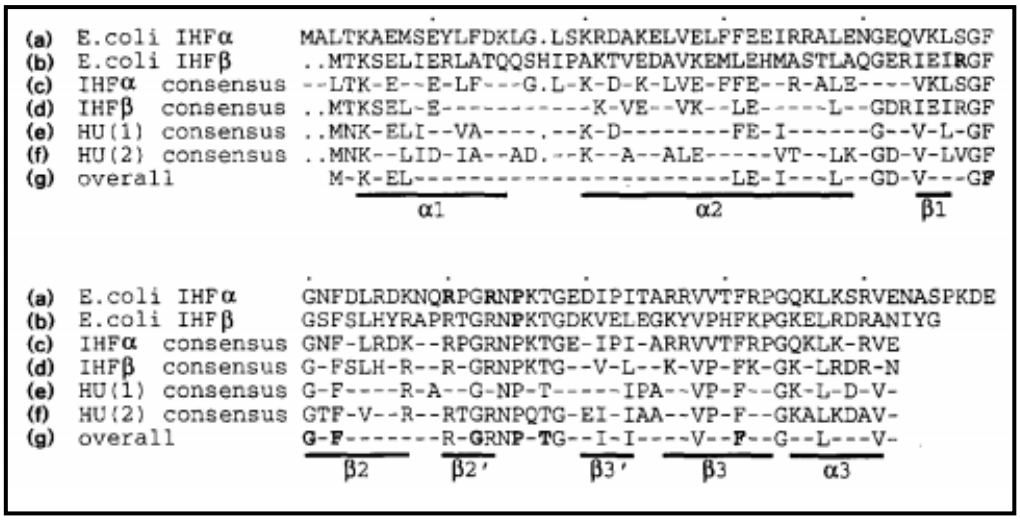


Fig 2.3: *Sequence comparison between HU and IHF family members c, d, e and f shows the highly conserved subgroups. The intercalating prolines when bound to DNA are shown in bold. Dashes are sequences that are not strongly consensus* [23].

2.2. MATERIALS

2.2.1. Bacterial strains and vectors used

Strain	Genotype	Applications.
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 form *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 pQE-30 vector
E.coli BL21star(DE3) pLyS*	F$^-$ *ompT gal dcm lon hsdS$_B$*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S) pLysS[*T7p20 ori*$_{p15A}$](Cm^R) • 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 mRNAdegradation) offering enhanced mRNA stability for protein expression	Used as expression host for all genes cloned in pET vectors.

2.2.2. Chemicals and kits

Primers were procured from Integrated DNA technology (IDT). Restriction enzymes and molecular biological reagents (polymerases and ligases) were obtained from New England Biolabs (NEB), USA. Protein molecular weight markers were purchased from Fermentas. Plasmid mini-prep kits, gel extraction kits, PCR purification kits, Ni-NTA agarose/superflow resin used in this study were obtained from Qiagen ,USA. All other fine chemicals were obtained from Sigma chemical, USA. Superdex 75 10/300 GL column was procured from GE Healthcare Life Sciences and dialysis tube from Thermo Scientific.

2.2.3. List of primers used in the study

Primer name	Length (bases)	Sequence (5' ---3')	Melting temperature (Tm)
IHFA BamHI (FP)	31	5'-ATATTAGGATCCATGGCGCTTACAAAAGCT-3'	60.4˚C
IHFA NheI (RP)	34	5'-ATATATGCTAGCTTACTCATCTTTGGGCGAAGC-3'	62.2˚C
IHFB Bam HI (FP)	31	5'-ATATTAGGATCCATGGCGCTTACAAAAGCTG-3'	60.4˚C
IHFB HindIII (RP)	33	5'-TATAATAAGCTTTTAACCGTAAATATTGGCGCG-3'	57.7˚C

2.2.4. Media

Luria Broth (LB)

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

The media was sterilised by autoclaving (15psi for minutes at 121˚C)

2.2.5. Antibiotics

Ampicillin, Kanamycin, Chloramphenicol and tetracycline used in this study were procured form Sigma chemicals, USA. For all the 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	35 mg/ml in methanol

Stock solutions of antibiotics were sterilised through 0.22 µM Millipore filter. Stocks were stored as aliquots at -20˚C.

2.2.6. Buffers used for molecular biology work

2.2.6.1. Buffer composition for the preparation of chemical competent cells

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

The solution was sterilised by passing through 0.22 µm filter followed by autoclaving and stored at 4˚C.

2.2.6.2.6X DNA gel loading buffer (In deionised water)

Bromophenol blue	0.25 %
Glycerol	30 %

2.2.6.3. 50 X 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

2.2.6.4. Ethidium Bromide stock solution (1% w/v)

Ethidium Bromide	0.1g
Deionised water	10 ml

The stock solution was stored in amber coloured vial/bottle at 4˚C till used.

2.2.7. Buffers and Solutions –SDS PAGE

2.2.7.1. Acrylamide

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

2.2.7.2. Ammonium persulphate (APS, 10%)

APS	100 mg
Deionised water	Upto 1 ml

2.2.7.3. Lower Tris (4X), pH 8.8 (pH adjusted with 6 N HCl)

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

2.2.7.4. Upper Tris (4X), pH 6.8 (pH adjusted with 6 N HCl)

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

2.2.7.5. 5X sample loading buffer

Tris.Cl (pH 6.8)	0.15 M
SDS	5 %
Glycerol	25 %
B-mercaptoethanol	12.5 %
Bromophenol blue	0.06 %
Total volume	10 L

2.2.7.6. Laemmli buffer

Tris buffer	3 g
Glycine	14.4 g
SDS	1 g
Total volume	100 ml

2.2.7.7. Gel staining solution

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

2.2.7.8. Gel destaining solution

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

2.2.7.9. SDS Gel Compostion (for 2 gels)

	Resolving gel	
	15 %	18 %
Lower Tris	2.5 ml	2.5 ml
Acrylamide	5 ml	6 ml
Water	2.4 ml	1.4 ml
APS	50 µl	50 µl
TEMED	10 µl	10 µl

	Stacking gel -5%
Upper tris	0.5 ml
Acrylamide	0.33 ml
water	1.167 ml
APS	12.5 µl
TEMED	5 µl

2.2.8. Buffers for native purification

2.2.8.1. Native Lysis Buffer

NaH_2PO_4, pH 8.0	0.05 M
NaCl	0.3 M
Imidazole	0.01 M

2.2.8.2. Naïve wash buffer

NaH_2PO_4, pH 8.0	0.05 M
NaCl	0.3 M
Imidazole	0.02 M

2.2.8.3. Native elution buffer

NaH_2PO_4, pH 8.0	0.05 M
NaCl	0.3 M
Immidazole	0.25 M

2.3. METHODS

2.3.1. Polymerase Chain Reaction

PCR reactions have the following composition:

Component	Stock concentration	Working concentration
Template	variable	1-2 ng of plasmid
Forward primer	10 mM	0.5 mM
Reverse primer	10 mM	0.5 mM
dNTPs	250 mM	250 µM
Polymerase (Taq, Vent, Deep vent, Phusion)	2 units/µl	0.02 units/µl
Buffer	10X (for Taq and vent) 5X (for Phusion)	1X

All PCR reactions were carried out in Eppendorf PCR machine with following PCR program for 30 cycles.

:

Steps	Temperature	Time
Initial denaturation	95˚C	5min (Vent/ Deep vent)
		30 sec (Phusion)
Denaturation	95˚C	30 sec (Vent/ Deep Vent)
		10 sec (Phusion)
Annealing	40 - 65˚C	30 sec
Extension	72˚C	1 min/kb (Vent/Deep Vent)
		30 sec/kb (Phusion)
Extension	72˚C	10 min

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

2.3.2. Agarose gel electrophoresis

The method is used to separate DNA fragments in an agarose matrix based on size and charge provided through an electric field. The pore sizes are dependent on the agarose concentration ; for small fragments higher percentage of agarose is used. Migration is also affected by the type of electrophoresis buffer particularly its ionic strength.

Weighed agarose is dissolved by heating in TAE buffer. Ethidium bromide was supplemented for DNA visualisation under UV. 6X gel loading buffer was added to the samples making final concentration to 1X before loading. Electrophoresis was carried out in TAE buffer at ~8V/cm or 90 volts. 100bp and 1kb ladders were used as markers to calculate the size of DNA fragments from their relative mobility.

2.3.3. Purification of DNA band(s) from agarose gel

After electrophoresis, the gels were visualised on a UV trans-illuminator and the band of interest was excised out of the gel. To elute DNA out of the gel Qiagen purification kit was used using the following protocol :

a) *Solubilisation :* Weighed gel pieces were dissolved in QG binding buffer (100µl/ 100 mg of gel pieces). This was incubated at 55˚C till completely dissolved.
b) *Binding* ; the above solution containing DNA was poured onto QIAquick spin column (provided by the manufacturer) to allow adsorption of DNA onto the gel matrix.
c) *Washing* : this was followed by washing with wash buffer PE (containing ethanol) to remove the impurities.
d) Elution : DNA was finally eluted in autoclaved distilled water and then quantitated.

2.3.4. PCR cleanup

This step was done for better yield where single band of correct size was cleaned using Sigma PCR clean up kit using the following protocol:

a) *Binding*: three volumes of binding buffer was added to the sample. The sample was loaded onto the column to allow its adsorption on the silica gel matrix.
b) *Washing* : washed with wash buffer containing ethanol to remove the impurities
c) *Elution* ; DNA was eluted in autoclaved distilled water

2.3.5. Quantification of DNA

The DNA samples were quantified by either running them on the agarose gel and visually comparing with the bands of the ladder depicting a definite amount of DNA in a definite amount of ladder loaded, or by measuring 260 nm using Nanodrop spectrophotometer. 1µl of the sample was put on the probe of the instrument after setting the baseline with buffer our sample is eluted in. Absorbance was measured at 260 nm. Purified DNA should have $O.D_{260}/0.D_{280}$ ratio between 1.8-2.0.

2.3.6. Restriction Digestion

Samples (PCR samples/plasmid) were digested using specific restriction endonucleases following the below general description:

Template	~200 ng of PCR
Fast digest bugger (10X)	1X
Restriction enzymes	Depends upon DNA amount (1ul for 1000ng)
Water	To make up the volume

Digestion was set at 37˚C for 30 minutes.

2.3.7. Ligation

Ligation reaction was set fixing the insert :vector ratio as 3:1.

Vector	50 ng
Insert	Calculated by NEB Ligation calculator
Buffer (10X for T4 DNA ligase, 2X for Quick ligase)	1X
Ligase	Depends upon amount of DNA (1ul for 100ng)
water	To make up the volume

Ligation calculator calculates the amount of insert by the following method:

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

Reaction was set for three hours at 25˚C when using T4 DNA ligase and for 10-15 minutes at 25˚C for Quick Ligase.

2.3.8. *E.coli* competent cells preparation

The cells are treated with $CaCl_2$ to help in the uptake of negatively charged DNA using the following protocol:

a) A single colony of *E.coli* was inoculated in LB media and grown overnight
b) Culture was re-inoculated into fresh LB media at a dilution of 1:100and grown to early log phase (A_{600} of 0..3-0.4)
c) The cells were chilled on ice for 15 minutes, centrifuged at 1600xg for 7 minutes at 4°C in pre-chilled centrifuge tubes. Cells were kept on ice at all subsequent steps.
d) The supernatant was discarded and cells were resuspended in 40 mL of ice cold $CaCl_2$ solution (60 mM CaCl, 10mM-PIPES, 15% glycerol) and centrifuged at 1100xg for 5 minutes at 4^0C
e) Pellet was resuspended in 20 ml ice cold $CaCl_2$ solution.
f) Cells were left on ice for 30 minutes then centrifuged at 1100xg for 5 minutes at 4°C.
g) The pellet was resuspended in 1.5 ml ice cold $CaCl_2$ solution.

h) Finally aliquots of 100 ul were made from the resuspended cells and stored at -80°C till further use.

2.3.9. Transformation

a) The chemically competent cells were thawed on ice.
b) Further, the ligation mixture was added to the cells and mixed gently. They were then incubated on ice for 15-20 minutes.
c) Heat shock was given to the cells at 42 °C for 90 seconds by incubating the samples in water bath.
d) After heat shock cells were incubated on ice for 1-2 min and then immediately 1 mL of sterile LB media was added to the cells, kept in an incubator shaker at 37 °C for 1 hour.
e) The cells were centrifuged at 5000 rpm for 5 minutes and the supernatant was discarded
f) The pelleted cells were resuspended in 100 µl of fresh media then plated on the LB agar plates with appropriate antibiotics.

The plates were incubated overnight at 37˚C and transformants were observed the next day.

2.3.10. Screening of transformants

Colonies were screened in two steps prior to sequencing

Colony PCR : colonies were picked and suspended into aliquoted mastermix in a PCR tube. Standard PCR protocol was followed and then samples were run on gel to see amplification. Clones with right size amplification were inoculated in LB media and plasmid isolated.

Restriction Digestion: Once the plasmid was isolated, it was double digested to see integration of insert into vector. After digestion plasmid was run on gel to see size of the insert. Plasmid with correct size insert was sent for sequencing.

2.3.11. Plasmid DNA isolation

Isolation of plasmid DNA was carried using QIAprep Spin miniprep kit. The procedure followed is described below:

a) Pelleting and resuspension: Bacterial cells (3-5 mL) grown overnight were pelleted by centrifugation at 13,000 rpm. The supernatant was discarded and pellet was resuspended in 250 µL of P1 buffer and mixed by pipetting.
b) Lysis: Then 250 µL of P2 buffer was added and mixed by gently inversion
c) Neutralisation: 350 µL of Neutralization buffer (Buffer N3) was then added and mixed by gentle inversion. The mixture was centrifuged at 13,000rpm for 10-15 minutes.
d) Loading: The supernatant was loaded onto the columns and spun at 13,000 rpm for 1 minute. Flow through was discarded
e) Washing: Column was washed with 750 µL of wash buffer (Buffer PE) and by centrifuging at 13,000 rpm for 1 minute. Centrifuged once more to remove residual wash buffer.
f) Elution : The plasmid DNA was eluted in deionized water in a fresh 1.5 ml microcentrifuge tube

2.3.12. Glycerol stock (15%)

1500 µl of an overnight grown culture (37˚C, 220rpm) obtained by inoculating a single colony into LB supplemented with appropriate antibiotics was mixed with 500µl of 60% glycerol (autoclaved) making final concentration upto 15%. The stock was stored at -80˚C.

2.3.13. Purification of 6x His tagged proteins

Cell pellet was resuspended in native lysis buffer at 5 mL per gram wet weight. It was further followed by sonication to lyse the cells. (if lysozyme is to be used for lysis, incubate the resuspended pellet with lysozyme-1 mg/ml for 30 mins prior to sonication). The lysate was spun at 11,000 x g for 1 hour at 4°C. The lysis supernatant or the lysate was loaded onto a pre equilibrated (equilibrated with 1X native lysis buffer) Ni-NTA column. The column's

flowthrough was collected after which the column was washed with native wash buffer. The bound protein was eluted with native elution buffer.

2.3.14. Ion Exchange Chromatography

Ion exchange chromatography was performed on AKTA purifier instrument with 1ml column having cation exchange HiTrap^{TM} SP HP resin (from GE Healthcare). The column was pre-equilibrated with phosphate buffer (buffer and the sample should be devoid of salts since the conductivity value must be less than 5µS/cm). After binding protein was eluted through a 0-100% gradient of NaCl (2M NaCl in Phosphate buffer).

2.3.15. Circular Dichroism spectroscopy

Far UV-CD spectra for protein were collected on Biologic MOS-500 CD spectrometer using a 1 mm path length quartz cuvette. Both the sample and the blank spectra were collected in the range of 200-250 nm.

2.4. RESULTS AND DISCUSSION

2.4.1. Cloning of IHF homodimers

2.4.1.1. PCR amplification

Genes encoding IHFA (*ihfA*) and IHFB (*ihfB*) were amplified from *E.coli* genome by PCR using the oligonucleotide primers listed below. PCR amplification (20µl reaction) was performed with *Taq* polymerase enzyme under conditions listed below.

template	1.0µl
ihfa forward primer	0.4µl
ihfa reverse primer	0.4µl
Taq polymerase enzyme	0.2µl
dNTP	0.4µl
Taq polymerase buffer (10x)	2.0 µl
water	15.6µl

Table 2.1*: PCR composition for ihfA gene amplification.*

template	1.0µl
ihfa forward primer	0.4µl
ihfa reverse primer	0.4µl
Taq polymerase enzyme	0.2µl
dNTP	0.4µl
Taq polymerase buffer (10x)	2.0 µl
water	15.6µl

Table 2.2 *PCR composition for ihfB gene amplification.*

Initial Denaturation	95˚C	5min
Denaturation		30sec
Annealing	56˚C	40sec
Extension	72˚C	35sec
Final extension		10min

30 cycles

Table 2.3*: PCR conditions for amplification of both the genes*

The PCR reactions were electrophoresed on agarose gel for size analysis (Expected size -300bp)

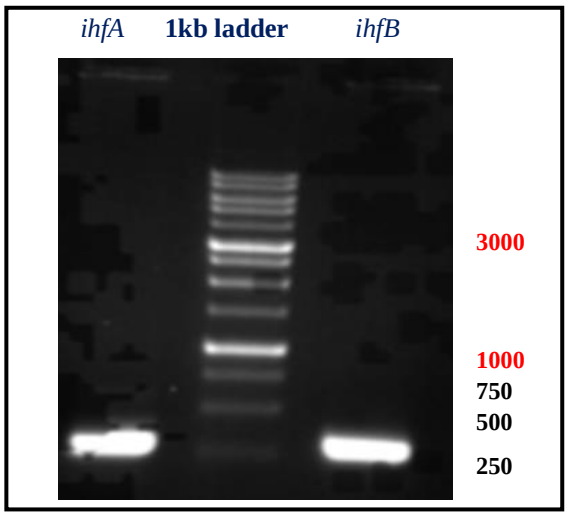


Fig 2.4 : *PCR amplification of ihfA and ihfB genes*

Under UV trans-illumination bands of interest were excised and eluted following the gel extraction protocol. DNA was quantified using a Nanodrop spectrophotometer, (*ihfA*=5.6 ng/µl,*ihfB*= 17 ng/µl)

2.4.1.2. Restriction digestion

pQE-30 is normally used to express N-terminal His-tagged proteins in *E.coli*. The clone contained in pQE-30 was grown overnight and the plasmid isolated concentration estimated: 60 ng/µl). We need our gene to be inserted in this vector between BamHI/NheI in case of *ihfA* and BamHI/HindIII in case of *ihfB*. Hence we perform the restriction digestion for both the insert as well as the vector accordingly following protocol.

template	33 µl (14 ng)
BamHI	1.5 µl
NheI	1.5 µl
10x fast digestion greenbuffer	4 µl

Table 2.4: *reaction composition for ihfA*

Template	15 µl (255 ng)
BamHI	1.0 µl
NheI	1.0 µl
10x fast digestion greenbuffer	2 µl

Table 2.5: *reaction composition for ihfB*

template	15 µl (900 ng)
BamHI	1.5 µl
NheI	1.5 µl
10x fast digestion greenbuffer	2 µl

Table 2.6 :*reaction composition for pQE-30 digested with BamHI/NheI*

Template	15 µl (900 ng)
BamHI	1.5 µl
HindIII	1.5 µl
10x fast digestion greenbuffer	2 µl

Table 2.7: *reaction composition for pQE-30 digested with BamHI/HindIII*

We are extracting our insert from the vector that is 10 times less in size to it , therefore approximately 1/10th of the DNA amount is lost after digestion. Hence we decide the elution volume accordingly such that the final concentration was not too low.

Further we did PCR purification of the insert using Sigma PCR–clean up kit to gain better yield. As for the insert we excised the band of correct size and eluted it following the standard protocol for gel extraction. Concentrations were estimated for both the inserts and vector using nanodrop spectrophotometer. (digested *ihfA* =4.4 ng/µl, digested *ihfB* =9.5 ng/µl, pQE-30 digested with BamHI/ NheI = 9 ng/µl, pQE-30 digested with BamHI/HindIII= 10.5 ng/µl).

2.4.1.3. Ligation and transformation

The digested products were ligated so as to be transformed into a host cell. For 50 ng of the vector, amount of the insert in the ligation mixture was estimated using NEB ligation calculator fixing the insert: vector ratio as 3:1.

Insert	3.5 µl
Vector	5.5 µl
T4 ligase buffer	1 µl
T4 ligase enzyme	0.3 µl

Table 2.8 : *composition for ihfA-pQE-30 ligation*

Insert	5.0 µl
Vector	2 µl
T4 ligase buffer	0.8 µl
T4 ligase enzyme	0.1 µl

Table 2.9 : *composition for ihfB-pQE-30 ligation*

Ligation mixture was transformed into XL-1 Blue cells that were chemically competent. Sufficient colonies were observed for both the transformations after an overnight incubation.

2.4.1.4. Colony PCR

To screen the positive transformants from the colonies obtained colony PCR was done. To the below mastermix 7.7ul MQ water containing the template was added.

T5 forward primer	0.5µl
T5 reverse primer	0.5µl
dNTP	0.2 µl
10x *Taq* buffer	1.0µl
Taq polymerase enzyme	0.1µl

Table 2.10*: colony PCR composition*

Initial Denaturation	95˚C	5min
Denaturation		25sec
Annealing	47˚C	35 sec
Extension	72˚C	50 sec
Final extension		10min

Table 2.11*: colony PCR settings for both the genes for 30 cycles*

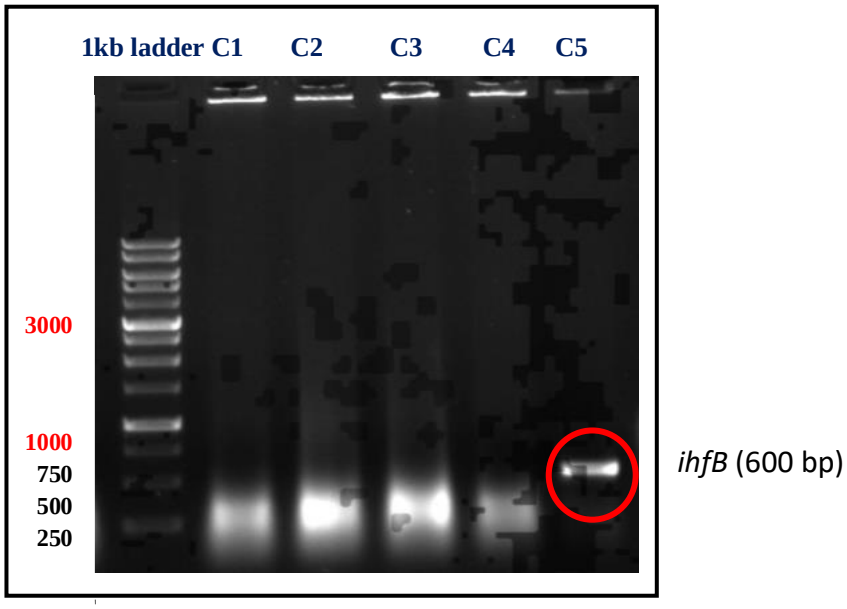


Fig 2.5 : *colony PCR of ihfB gene showing colony 5 as a positive transformant*

A positive transformant was obtained in case of the gene for IHFB. Size showed upto 600bp inclusive of the flanking regions of T5 promoter. While in case of *ihfA* none of them were positive. Colony PCR was repeated for this gene several times under different conditions of PCR settings and varied enzymes, but the results were still negative. Later we realised that T5 terminator got lost after the digestion with NheI .This was understood from the vector map for pQE-30 that showed the lambda terminator positioned before the NheI restriction site.

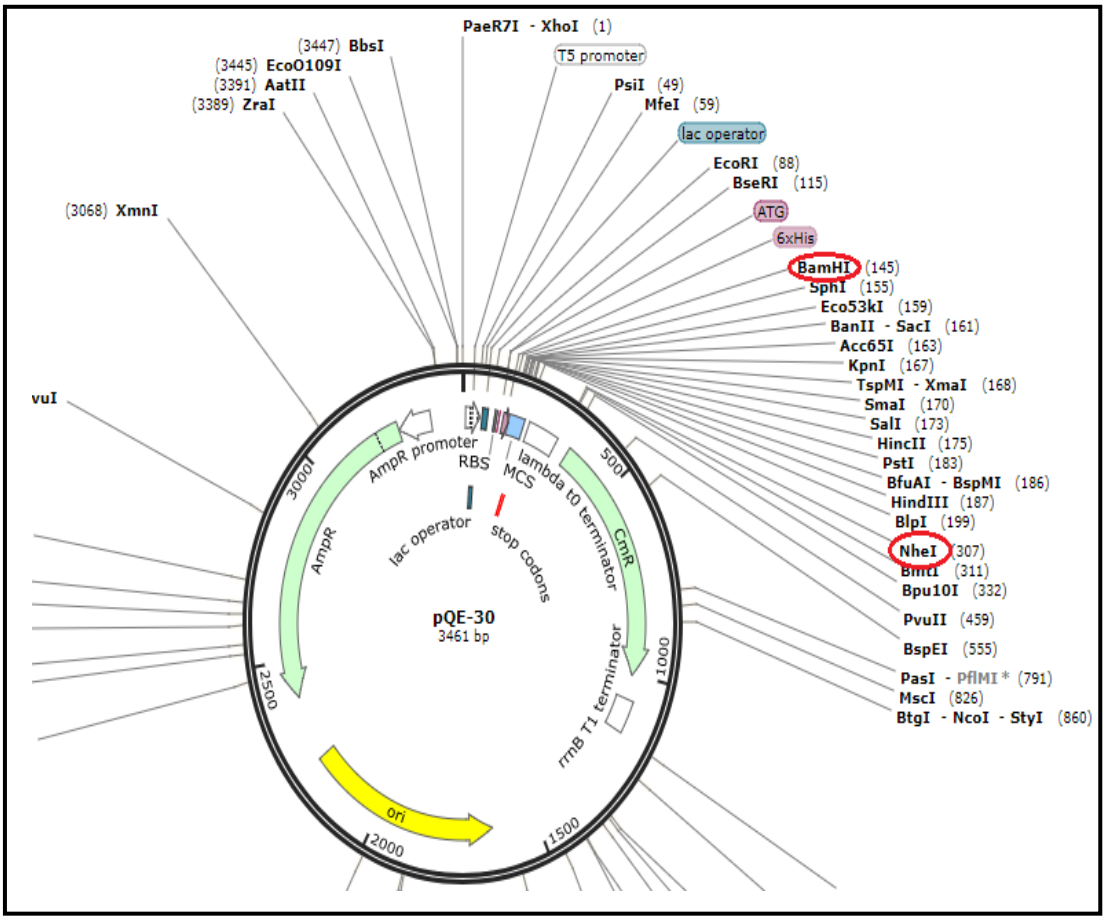


Fig 2.6: *Vector map of pQE30*

As a solution I tried amplifying the gene with T5 forward primer and gene specific reverse primer.

T5 forward primer	0.3µl
ihfA reverse primer	0.3 µl
dNTP	0.2 µl
10x *Taq* buffer	1.0 µl
Taq polymerase enzyme	0.1 µl

Table 2.12*: colony PCR composition*

Initial Denaturation	95˚C	5min
Denaturation		30 sec
Annealing	57˚C	35 sec
Extension	68˚C	40 sec
Final extension		10min

Table 2.13 *: colony PCR settings for ihfA amplification -30 cycles*

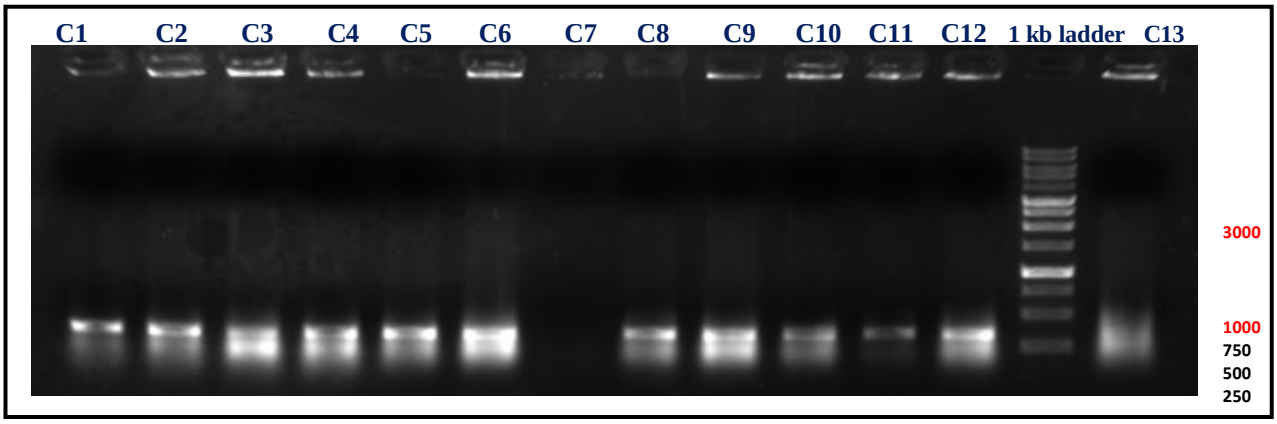


Fig 2.7*: Colony PCR for ihfA gene*

This time almost all were positive transformants. Including the initial 150bp of the promoter total gene size was 450bp. 2nd and 12th colonies were inoculated in LB media to do digestion check.

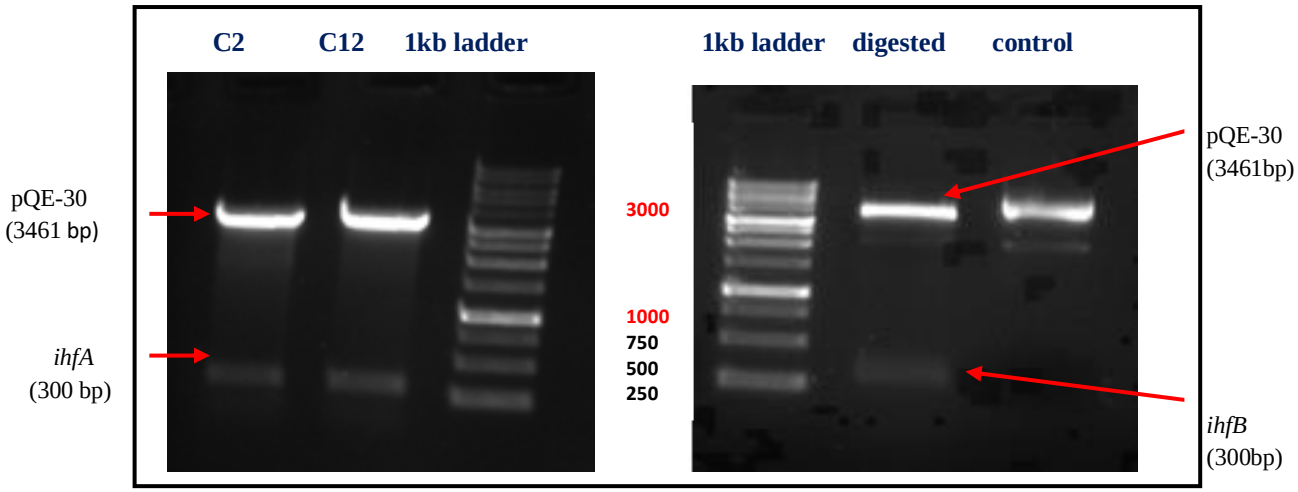


Fig 2.8: *Digestion check for ihfA with colonies 2 and 12*

Fig 2.9: *Digestion check for ihfB*

Above results confirm the insert of correct size. Clone was further confirmed through Sanger sequencing.

```
                1       10        20        30        40        50        60        70        80        90        99
                |--------+---------+---------+---------+---------+---------+---------+---------+---------+--------|
IHFAsequenced   MALTKAEMSEYLFDKLGLSKRDAKELVELFFEEIRRALENGEQVKLSGFGNFDLRDKNQRPGRNPKTGEDIPITARRVVTFRPGQKLKSRVENASPKDE
IHFA            MALTKAEMSEYLFDKLGLSKRDAKELVELFFEEIRRALENGEQVKLSGFGNFDLRDKNQRPGRNPKTGEDIPITARRVVTFRPGQKLKSRVENASPKDE
Consensus       MALTKAEMSEYLFDKLGLSKRDAKELVELFFEEIRRALENGEQVKLSGFGNFDLRDKNQRPGRNPKTGEDIPITARRVVTFRPGQKLKSRVENASPKDE
```

Fig 2.10:*IHFA sequence from NCBI aligned with sequence of the clone (ihfA in pQE-30)*

```
                1       10        20        30        40        50        60        70        80        90  94
                |--------+---------+---------+---------+---------+---------+---------+---------+---------+---|
IHFBsequenced   MTKSELIERLATQQSHIPAKTVEDAVKEMLEHMASTLAQGERIEIRGFGSFSLHYRAPRTGRNPKTGDKVELEGKYVPHFKPGKELRDRANIY
IHFB            MTKSELIERLATQQSHIPAKTVEDAVKEMLEHMASTLAQGERIEIRGFGSFSLHYRAPRTGRNPKTGDKVELEGKYVPHFKPGKELRDRANIYG
Consensus       MTKSELIERLATQQSHIPAKTVEDAVKEMLEHMASTLAQGERIEIRGFGSFSLHYRAPRTGRNPKTGDKVELEGKYVPHFKPGKELRDRANIY.
```

Fig 2.11: *ihfB sequence from NCBI aligned with sequence of the clone (ihfB in pQE-30)*

2.4.2. Protein purification and expression of the homodimers

Both the genes had 6xHistidine affinity tag at the N-terminus hence the protein was purified natively using Ni- NTA chromatography. But even after several attempts of purification, the homodimers did not show any expression in the SDS-PAGE. Then we came across a paper that shows that overexpression of the homodimers separately produces unstable and insoluble peptides emphasising the need to express both the genes conjointly [10].

2.4.3 Cotranformation of plasmids containing *ihfA* and *ihfB* into *E.coli* cells

2.4.3.1. Subcloning of *ihfB* into pET28-a

Cotransformation requires the plasmids with different origins of replication and different antibiotic selection marker, so as to be stable inside the cell. Plasmids with same ORIs that also determine the copy number, suffer incompatibity due to competition for the same machinery creating an unstable environment within the cell. pET28-a have pBR322 origin and pQE-30 have ColE1 origin hence are suitable to be cotransformed within the host [11]. The next step was to subclone (moving the gene from a parent vector to a destination vector) *ihf*B from pQE-30 into pET28-a.

2.4.3.2. Restriction digestion

Both the vector (100ng/µl) and the *ihf*B plasmid (85ng/µl) were digested with BamHI/HindIII.

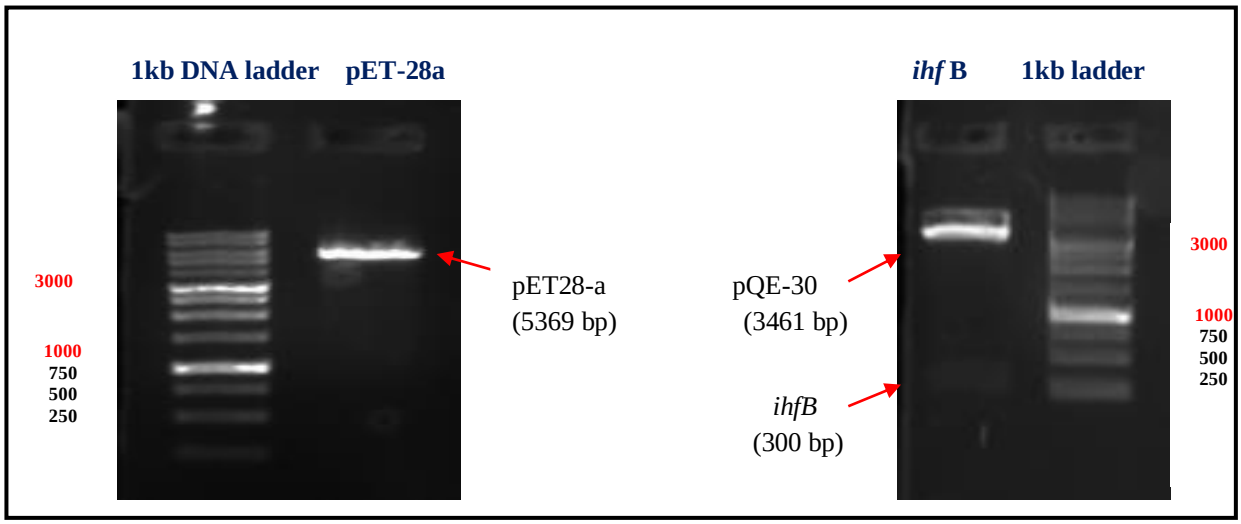


Fig 2.12 : *Digestion of pET28-a plasmid*

Fig 2.13: *Digestion of ihfB-pQE-30*

Bands of interest were excised for both the insert and vector and were further eluted following the gel extraction protocol. Concentrations were estimated at (*ihfB* =8 ng/µl, pET28-a =15 ng/µl).

2.4.3.3. Ligation and transformation

The digested products were ligated so as to be transformed into a host cell. For 50 ng of the vector, amount of the insert in the ligation mixture was estimated using NEB ligation calculator fixing the insert: vector ratio as 3:1. The 10 µl ligation mixture was transformed into XL-1 Blue cells that were chemically competent. 15 colonies were observed after overnight incubation.

2.4.3.4. Colony PCR

To screen the positive transformants, colony PCR was done with 4 colonies as template amplifying with T7 forward and T7 reverse primers expecting a product of 600 bp.

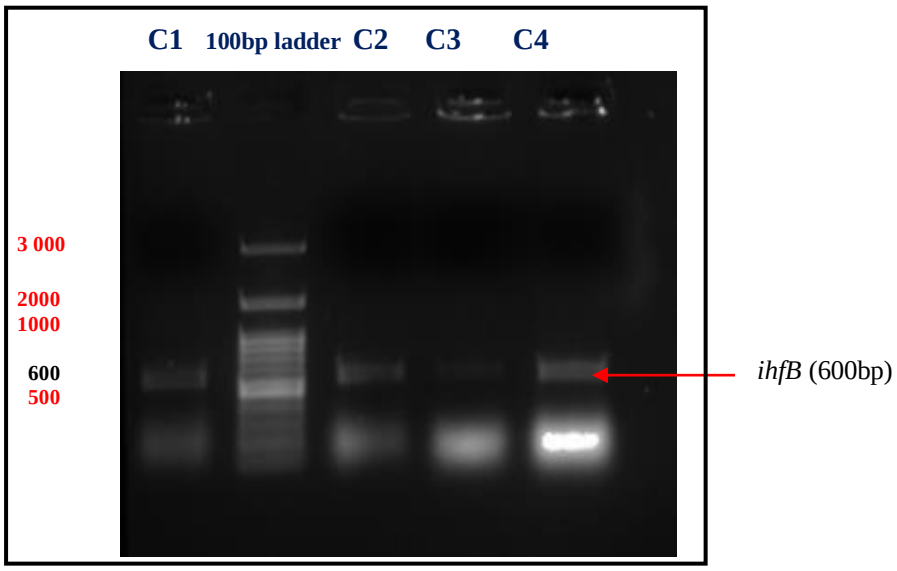


Fig 2.14 : *Colony PCR of ihfB gene*

Incoulated colony 4 in LB media overnight to do a digestion check.

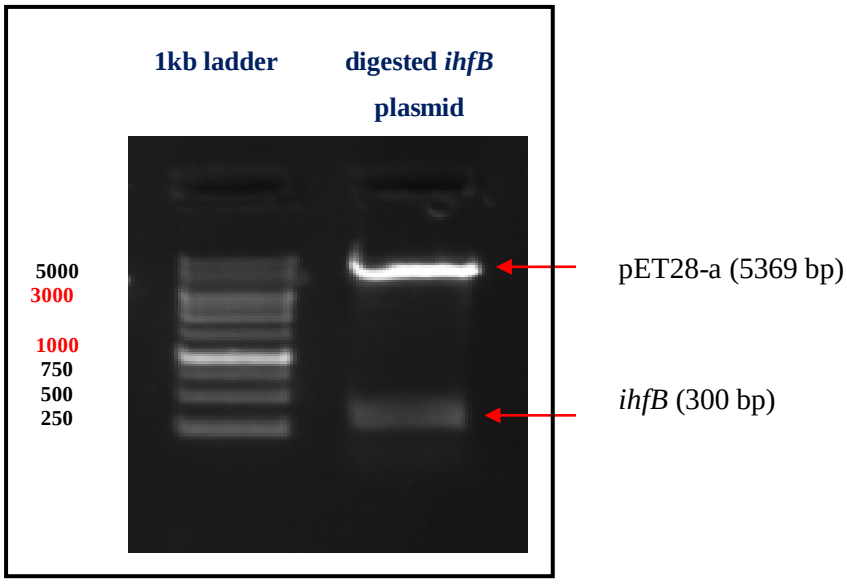


Fig 2.15 *: Digestion check of ihfB plasmid*

When the plasmid (5669 bp) was digested, it gave distinct bands of pET28-a and *ihfB* . The clone was further confirmed through Sanger sequencing.

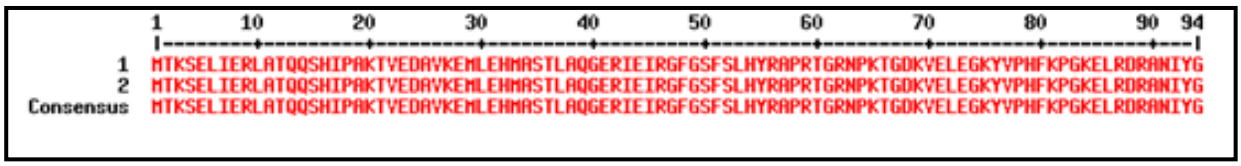


Fig 2.16: *IHFB sequence from NCBI (2) aligned with the sequence of the clone* (1)

The above clone is currently in XL1-Blue which is not an expression host. T5 promoter is native to *E.coli.*

2.4.3.5. Cotransformation.

Now the plasmids–*ihfA*/pQE-30 and *ihfB*/pET28-a were made, they need to be cotransformed into an expression host. pQE-30 has a T5 promoter that is native to *E.coli* whereas pET28-a has a T7 promoter. T7 being an external promoter cannot be expressed in XL1-Blue cloning host. Hence we chose a system such as *E.coli* BL21 STARstar(DE3)pLyS* which is T7 promoter-based expression system that can be used as an expression host for all genes cloned pET vectors It needs IPTG to induce T7 RNA polymerase. The strain will express both the vectors.

a) *pET28-a containing* BL21 Star (DE3) pLysS *cells were made chemically competent*

Usually during co-transformation cells containing one of the plasmids is made competent enough to take other plasmid. In this case *ihfB*/pET28-a was transformed into chemically competent BL21 Star (DE3) pLysS cells. Positive cells containing plasmid were made chemically competent and transformed with second plasmid, but could not get any positive clones.

b) *Both the plasmids containing IHFA and IHFB were added together during transformation*

Subsequently we tried adding both the plasmids together into the BL21 Star (DE3) pLysS competent cells *[ihfA*/pQE-30 =1µl (160ng/µl) *ihfB*/pET28-a =2µl [85 ng/µl)]. We got 6

colonies after an overnight incubation. Two colonies were inoculated in LB media with triple antibiotics selection (Ampciilin for pQE-30, Kanamycin for pET-28a and Cholaramphenicol for BL21 Star (DE3) plysS cells, kept for shaking at 37˚C at 220rpm. Secondary culture was inoculated further in 800ml culture and induced with IPTG when $O.D_{600}$ reached 0.6 and then incubated further for 5 hours. After IMAC purification under native condition, protein expression was not observed.

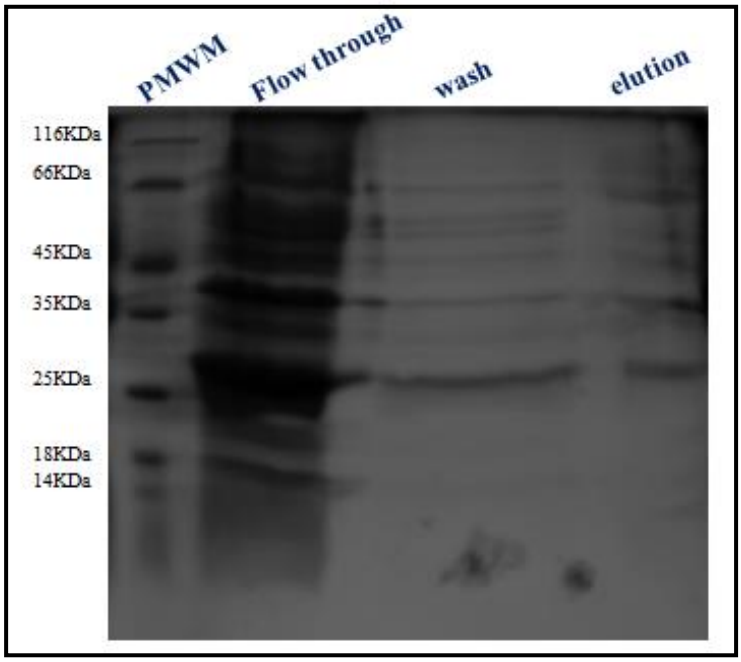


Fig 2.17:*SDS-PAGE analysis to observe the expression of proteins from the cotransformed cells.*

c) IPTG induction given from the zeroth hour

IPTG induction was necessary for the expression of *ihfB* that is in pET28-a but is not required to express IHFA (in pQE-30). Hence when we set the secondary culture IHFA starts getting produced from the first hour. Speculating IHFA is getting degraded by the time IHFB is produced we decided to provide IPTG induction from the 0^{th} hour along with the antibiotics as we do not need a tight expression. Further secondary culture is incubated for 12 hours. IMAC purification is then performed under native conditions and expression checked by SDS-PAGE.

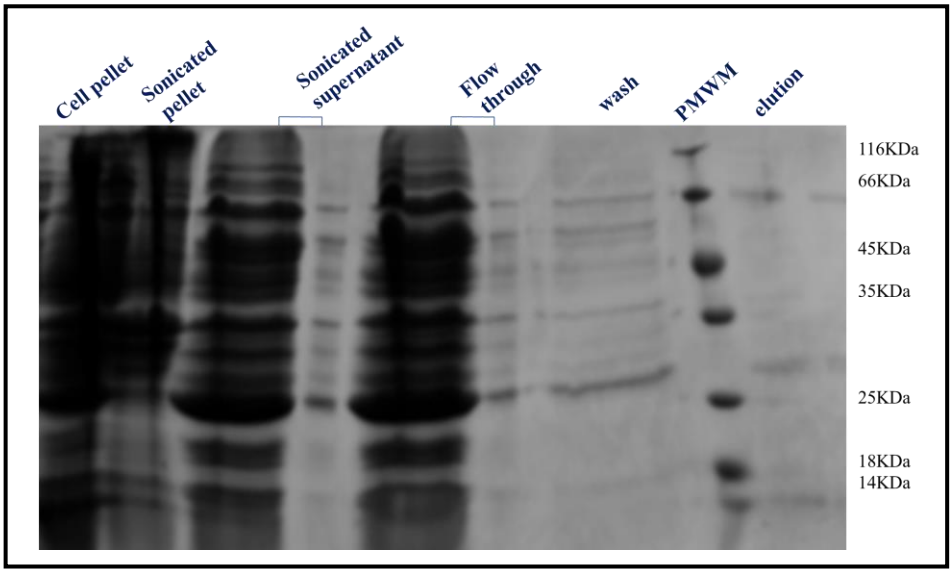


Fig 2.18 *: SDS gel showing expression of the IHF protein after inducing in the zeroth hour*

There was no expression on SDS-PAGE in the expected size range of~(12-14) kDa and but a prominent band was seen between 25-35 kDa. Speculating the band observed may be the result of anomalous mobility with protein resisting unfolding by SDS, we decided to denature the proteins by urea and guanidinium chloride (GdmCl).

d) *Flow through obtained by IMAC purification of co-transfromed cells expressing IHFA and IHFB was subjected to denaturation by urea and GdmCl.*

Flow through obtained through IMAC purification of the cotransformed culture is subjected to dentauration by urea alone (stock:10M) ranging from concentrations 3M-5M. Further a combination of 2M GdmCl+ 4M urea and 1M GdmCl +2M urea also used for denaturation. Once chemicals added samples were incubated for 45 minutes and then run in SDS-PAGE.

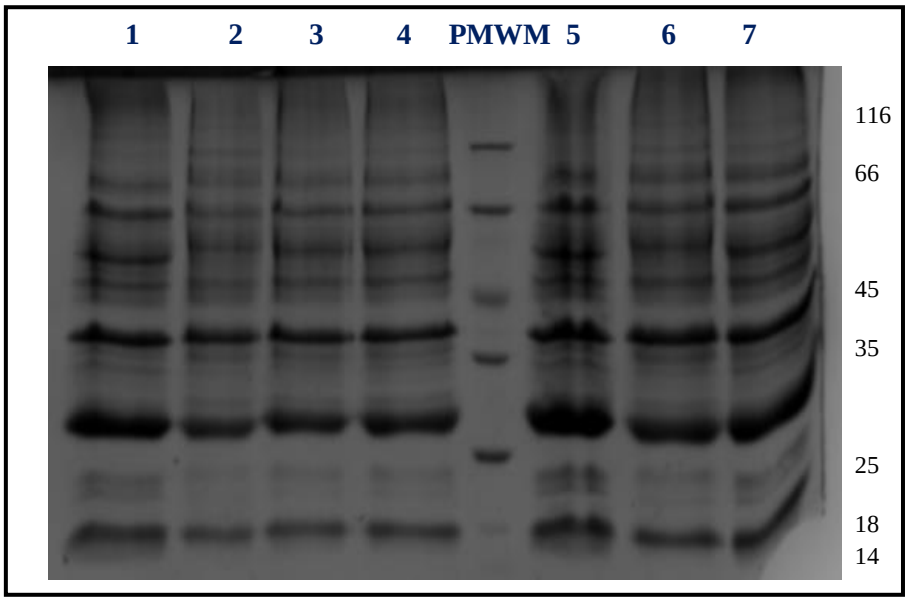


Fig 2.19: SDS-PAGE *gel showing effect of urea denaturation on IHF flow through*

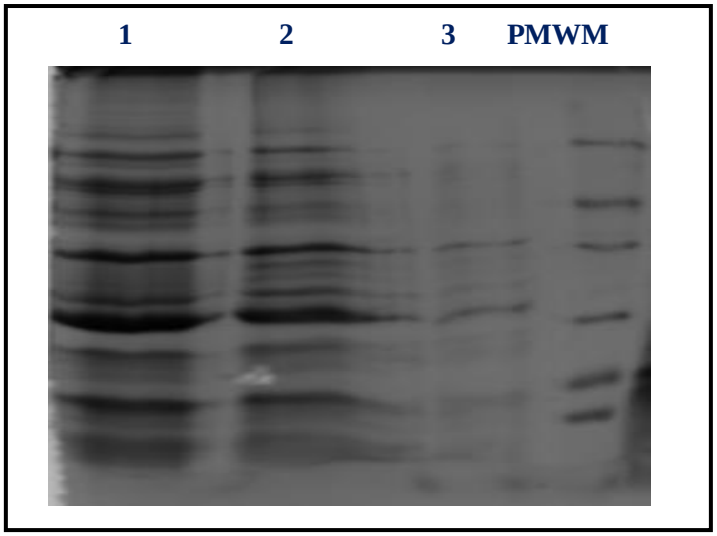


Fig 2.20: *SDS-PAGE gel showing the effect of GdmCl and urea on IHF flow through*

Even after denaturation bands were still visible in the 35 kDa range. Hence no significant results were not obtained from this experiment.

e) Time- Dependent expression of IHFA and IHFB from co-transformed cells

To see if the proteins are getting produced or getting degraded before 12 hours (time we usually incubate secondary culture) we thought of checking protein expression at regular intervals from one hour to the twelfth hour. Once the secondary culture is incubated, 1ml was aliquoted in a microcentrrifuge tube and pelleted. All the pellets from 1st -12th hour were mixed with SDS loading dye and denatured at 99˚C for 10min. after centrifugation for 5 minutes 10µl of the supernatant was run in SDS-PAGE.

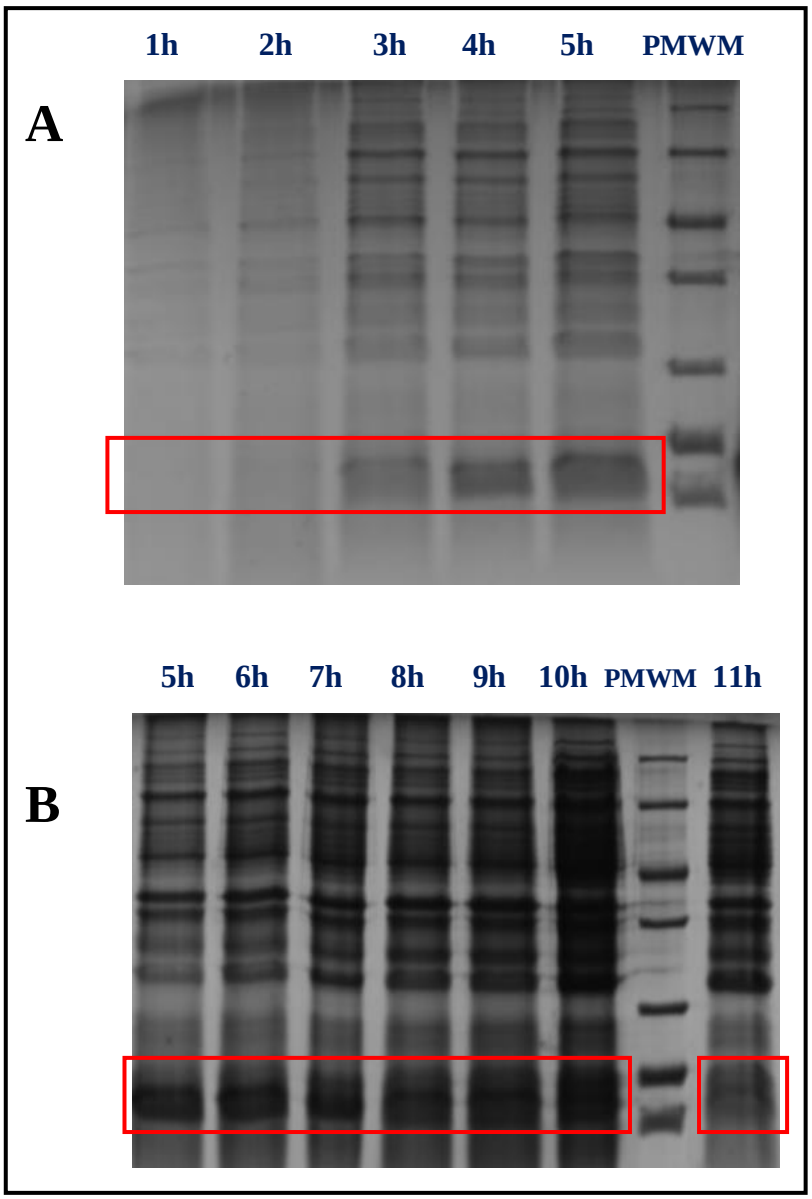


Fig 2.21:. ***Panel A*** *shows time dependent expression from 1hours to 5 hours.* ***Panel B*** *shows protein expression of co-transformed cells from 5hours to 11 hours.*

Fascinatingly a time–dependent expression was observed with expression minimal at first and second hours, subsequently increasing from the third hour and decreasing from the seventh hour. Expression was maximum at the fifth and sixth hour. Intensity of the 27 kDa band observed previously was also seen increasing with time .

f) *Purification of '6xHis-tagged' naturally expressing IHFA and IHFB by IMAC after 5hrs in the presence of lysozyme.*

Cells were harvested from 400 µl culture and resuspended in 10ml lysis buffer containing 10 mM imidazole and then incubated in 10mg lysozyme (1 mg/ml) under cold conditions for 30 minutes. Lysozyme was used this time to lyse the cells effectively and whole of the protein comes out of the genome. Cells were then lysed by sonication for a total processing time of 20 minutes giving a burst for 10 seconds followed by 10 seconds interval between each bursts. The soluble protein was further recovered in the supernatant after centrifugation for an hour at 11,000 rpm. Cell lysate was loaded onto the IMAC column pre-equilibrated with lysis buffer, purified and analysed by running on SDS-PAGE.

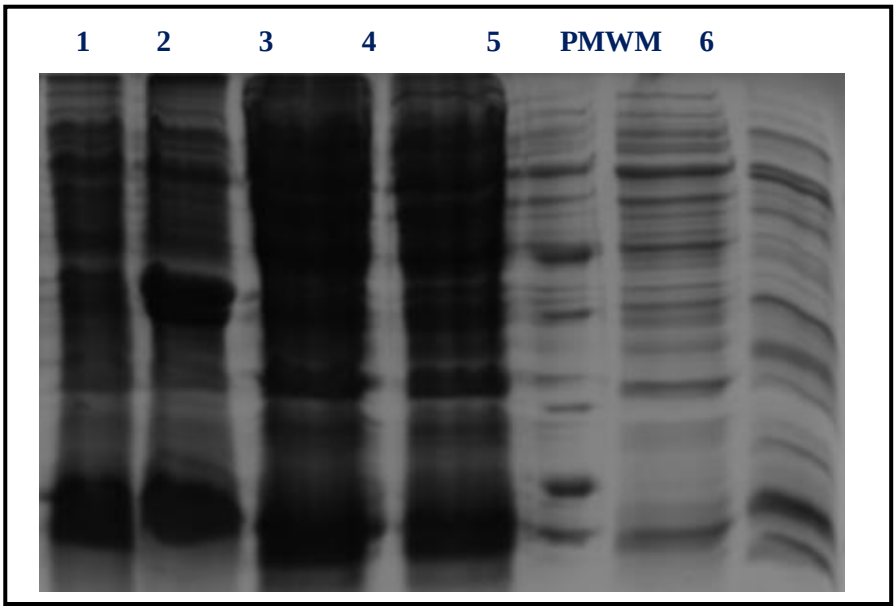


Fig 2.22 : *SDS gel showing expression of IHF after five hour of secondary growth*

Protein was visible in sonicated supernanant and flow through which confirms its solubility. No protein was observed in elute, possibly due to inaccessibility of Histine tag for binding with Ni-NTA beads. The best method to purify protein from large number of contaminants was to do ion-exchange chromatography.

g) . Ion –exchange chromatography of the whole cell lysate

Both IHFA and IHFB have pI value ~9.3. At pH-7, IHFA and IHFB have positive charges of 2.9 and 3.9 respectively (predicted by Protein Calculator v3.4). Hence cation exchange chromatography was done to purify these proteins. The 1 ml column was equilibrated with phosphate buffer (pH-7) before loading the sample. The sample was diluted 10 times with phosphate buffer since it was resupended in lysis buffer that contained 150mM NaCl which interferes with protein binding with cation exchange column due to increased conductivity (conductivity must have a value be less than 5µS/cm). Once the binding was done in the ion –exchange chromatogram, protein was eluted using a gradient of NaCl from 0-100 % (2M NaCl) in 100 column volume .

IHF do not contain tryptophan residues, hence could not absorb light of 280nm wavelength but absorb at 215nm owing to the absorbance of peptide bonds .This peak was observed corresponding to 17th fraction at around 34th volume (ml).16-21 fractions were collected for SDS-PAGE analysis.

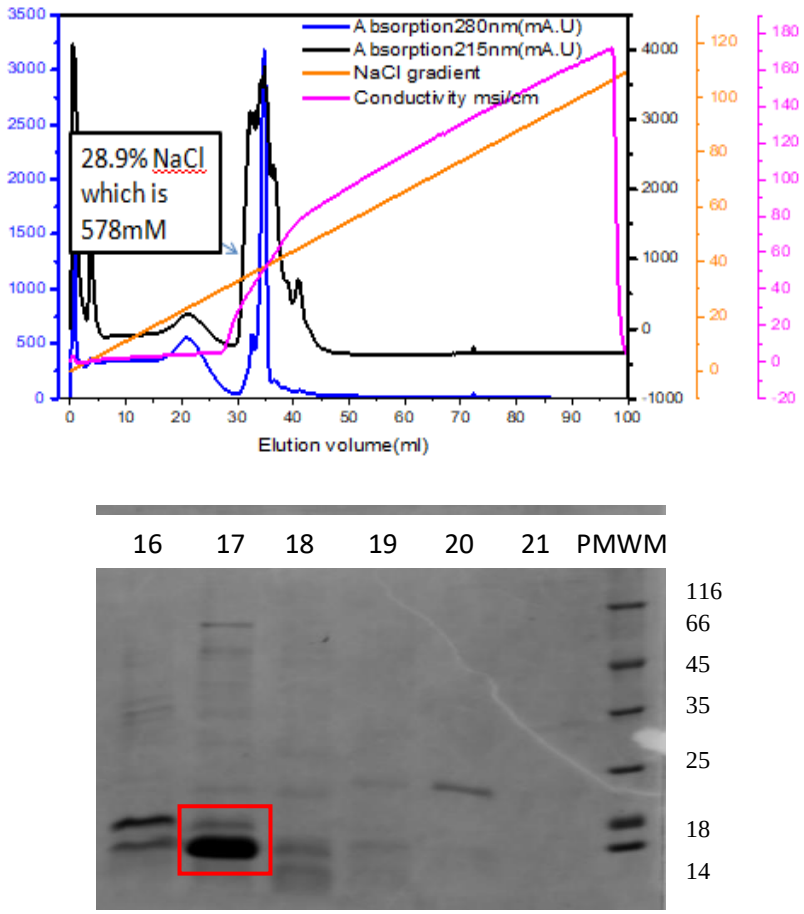


Fig 2.23: *Fractions after cation exchange chromatography*

A thick band was observed around ~10-15 kDa in the 17th fraction sample expected to be the protein of interest. The data obtained was plotted using Origin Pro software which showed protein eluted at 578mM NaCl.

h) Circular Dichroism Spectroscopy

To confirm its secondary structure, CD spectra of the 17th fraction was obtained over a range of 200-250nm.

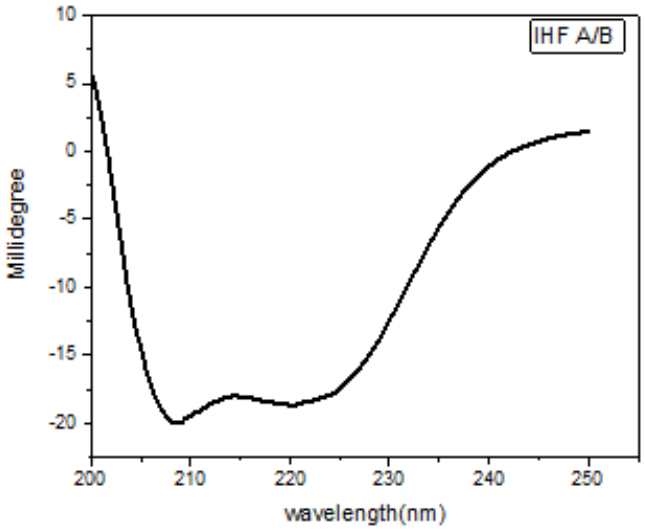


Fig 2.24*: CD spectra of the 17th fraction*

The protein was a mixture of αβ similar to HU. To do further experiment ,protein was diluted with phosphate buffer to bring its salt concentration from 578 to 154mM .. Further the sample was concentrated to its original volume by centrifuging at 4000 rpm, 2hrs using a concentrator that had a molecular weight cut off 3 kDa. The concentrated sample was subjected to SDS –PAGE analysis to obtain a clean band.

Having observed the protein getting expressed after five hours of secondary growth we thought of repeating the method for the homodimers as well. IHFA and IHFB were inoculated and grown separately for five hours adding IPTG in the zeroth hour. Cell was lysed with lysozyme. Cell pellet ,pellet after sonication and the cell lysate of both A and B subunits were subjected to analysis SDS-PAGE.

i) Purification and expression of the homodimers after 5hours of secondary growth

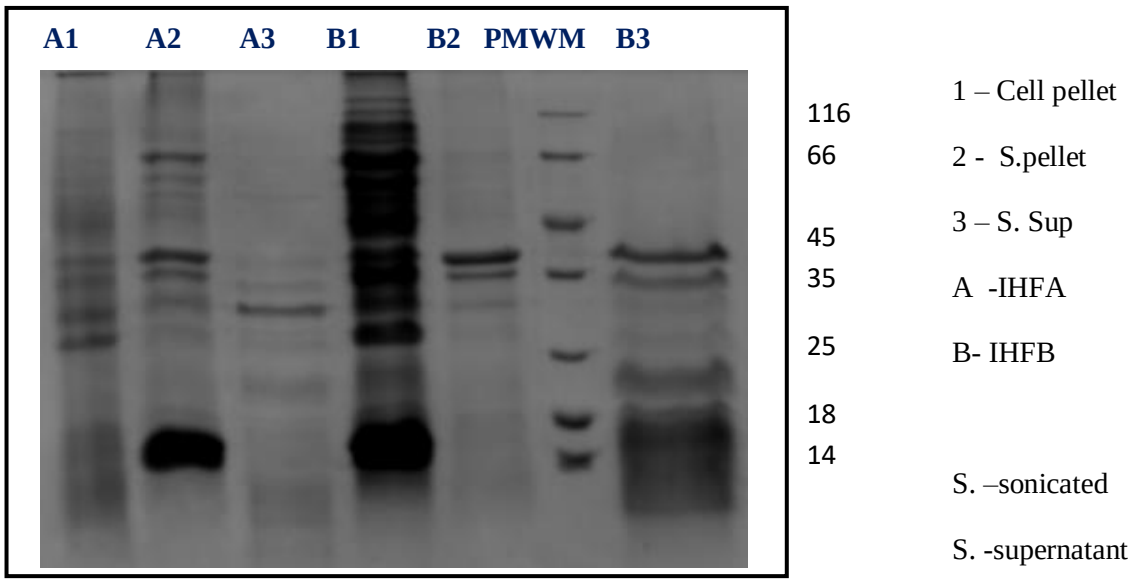


Fig*2.25: Purification of IHFA and IHFB homodimers*

In case of subunit A, high expression was observed in the sonicated pellet but not in the cell lysate. Whereas in B subunit both cell pellet and cell lysate showed expression with very thick band in the cell pellet.

As of now we are unsure whether both the subunits were expressed in the final stage or if its just one subunit. The identity of the protein needs to be confirmed through mass spectrometry. From the above results of homodimer purification and expression IHFA was observed to be insoluble as a homodimer as there was no expression in the cell lysate. It appears that there need to be a balance between certain amounts A amd B subunits so as to synthesise the protein because even though one of the subunits was showing expression (B subunit), it was very less.

2.5. CONCLUSIONS AND FUTURE DIRECTIONS

IHF is another histone-like protein of *E.coli* with high sequence homology with HU. It had been difficult to obtain substantial expression of IHF because of its low abundance unlike HU. Past researches show purification of the subunits of IHF by producing a fusion construct [20] but expression of the same naturally from *E.coli* is unexplored. Initially we tried to express occurring IHF as individual subunits (IHFA and IHFB) after 12 hours of secondary growth following native purification methods which did not succeed. We purified the protein(s) by allowing the growth of secondary after inoculation for 5 hours followed by cell lysis with lysozyme. From the cell lysate obtained after native purification ,protein was eluted by cation exchange chromatography at 28.9% NaCl (578 mM).We are unsure about the identity of the band (whether its any one subunit or both) obtained after SDS-PAGE analysis which further needs to be confirmed through MALDI. From this study we conclude on the time dependent expression shown by the protein with maximum expression at the fifth and sixth hours and the requirement of one of the plasmids containing the gene being essential for the stability and expression of the other. This was evident from the cotransformation results and by the analysis of homodimer expression. In the future we expect to characterise the protein biophysically to explore its role in biofilm formation.

REFERENCES

1. Redfern O, Dessailly B, Orengo C. Exploring the structure and function paradigm. *Current opinion in structural biology*. 2008;18(3):394-402. doi:10.1016/j.sbi.2008.05.007.
2. Dame RT. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Molecular microbiology. 2005 May 1;56(4):858-70.
3. Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. Chromosomes and Chromatin
4. Azam TA, Ishihama A. Twelve species of the nucleoid-associated protein from Escherichia coli sequence recognition specificity and DNA binding affinity. Journal of Biological Chemistry. 1999 Nov 12;274(46):33105-13.
5. Dorman CJ, Deighan P. Regulation of gene expression by histone-like proteins in bacteria. Current opinion in genetics & development. 2003 Apr 30;13(2):179-84.
6. Rouvière-Yaniv J, Kjeldgaard NO. Native Escherichia coli HU protein is a heterotypic dimer. FEBS letters. 1979 Oct 15;106(2):297-300.
7. Orfaniotou F, Tzamalis P, Thanassoulas A, Stefanidi E, Zees A, Boutou E, Vlassi M, Nounesis G, Vorgias CE. The stability of the archaeal HU histone-like DNA-binding protein from Thermoplasma volcanium. Extremophiles. 2009 Jan 1;13(1):1.
8. Geiduschek EP, Schneider GJ, Sayre MH. TF1, a bacteriophage-specific DNA-binding and DNA-bending protein. Journal of structural biology. 1990 Jul 1;104(1-3):84-90.
9. Borca MV, Irusta PM, Kutish GF, Carrillo C, Afonso CL, Burrage T, Neilan JG, Rock DL. A structural DNA binding protein of African swine fever virus with similarity to bacterial histone-like proteins. Archives of virology. 1996 Feb 1;141(2):301-13.
10. Boubrik F, Rouviere-Yaniv J. Increased sensitivity to gamma irradiation in bacteria lacking protein HU. Proceedings of the National Academy of Sciences. 1995 Apr 25;92(9):3958-62.
11. Prieto AI, Kahramanoglou C, Ali RM, Fraser GM, Seshasayee AS, Luscombe NM. Genomic analysis of DNA binding and gene regulation by homologous nucleoid-

associated proteins IHF and HU in Escherichia coli K12. Nucleic acids research. 2011 Dec 16;40(8):3524-37..

12. Swinger KK, Lemberg KM, Zhang Y, Rice PA. Flexible DNA bending in HU–DNA cocrystal structures. The EMBO Journal. 2003 Jul 15;22(14):3749-60.
13. Hammel M, Amlanjyoti D, Reyes FE, Chen JH, Parpana R, Tang HY, Larabell CA, Tainer JA, Adhya S. HU multimerization shift controls nucleoid compaction. Science advances. 2016 Jul 1;2(7):e1600650..
14. Chodavarapu S, Felczak MM, Yaniv JR, Kaguni JM. Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin. Molecular microbiology. 2008 Feb 1;67(4):781-92.
15. Rice PA, Yang SW, Mizuuchi K, Nash HA. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell. 1996 Dec 27;87(7):1295-306.
16. Swinger KK, Rice PA. IHF and HU: flexible architects of bent DNA. Current opinion in structural biology. 2004 Feb 1;14(1):28-35.
17. Bonnefoy E, Rouvière-Yaniv J. HU and IHF, two homologous histone-like proteins of Escherichia coli, form different protein-DNA complexes with short DNA fragments. The EMBO journal. 1991 Mar 1;10(3):687-96.
18. Allocati N, Masulli M, Di Ilio C, De Laurenzi V. Die for the community: an overview of programmed cell death in bacteria. Cell death & disease. 2015 Jan;6(1):e1609.
19. Devaraj A, Buzzo J, Rocco CJ, Bakaletz LO, Goodman SD. The DNABII family of proteins is comprised of the only nucleoid associated proteins required for nontypeable Haemophilus influenzae biofilm structure. MicrobiologyOpen. 2017 Dec 12.
20. Nash HA, Robertson CA, Flamm E, Weisberg RA, Miller HI. Overproduction of Escherichia coli integration host factor, a protein with nonidentical subunits. Journal of bacteriology. 1987 Sep 1;169(9):4124-7.
21. Weston A, Humphreys GO, Brown MG, Saunders JR. Simultaneous transformation of Escherichia coli by pairs of compatible and incompatible plasmid DNA molecules. Molecular and General Genetics MGG. 1979 Jan 1;172(1):113-8.
22. Grove A. Functional evolution of bacterial histone-like HU proteins. Current issues in molecular biology. 2011 Jan 1;13(1):1.
23. Rice PA. Making DNA do a U-turn: IHF and related proteins. Current opinion in structural biology. 1997 Feb 1;7(1):86-93.

24. Niba ET, Naka Y, Nagase M, Mori H, Kitakawa M. A genome-wide approach to identify the genes involved in biofilm formation in E. coli. DNA research. 2007 Jan 1;14(6):237-46.
25. Gardner JF, Nash HA. Role of Escherichia coli IHF protein in lambda site-specific recombination: a mutational analysis of binding sites. Journal of molecular biology. 1986 Sep 20;191(2):181-9.
26. Devaraj A, Justice SS, Bakaletz LO, Goodman SD. DNABII proteins play a central role in UPEC biofilm structure. Molecular microbiology. 2015 Jun 1;96(6):1119-35.