Histidine near N-Extein~Intein Junction in *Mycobacterium tuberculosis(Mtu)* SufB Protein Regulates Protein Splicing

Dissertation submitted in partial fulfilment for the degree of

Master of Science in Applied Microbiology

Submitted by

Anjali Rai

School of Biotechnology (Campus 11)
Bhubaneswar, Odisha
India

Under the Supervision of
Dr. SASMITA NAYAK
ASSISTANT PROFESSOR, KIIT SCHOOL OF BIOTECHNOLOGY,
ADJUNCT ASSISTANT PROFESSOR, DEPARTMENT OF MICROBIOLOGY, KIMS,
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
DEEMED TO BE UNIVERSITY

DECLARATION

I hereby declare that the dissertation entitled **"Histidine near N-Extein~Intein Junction in *Mycobacterium tuberculosis* (*Mtu*) SufB Protein regulates Protein Splicing"** submitted by me, for the degree of Master of Science to School of Biotechnology, KIIT- Deemed to be University, is a record of *bona fide* work carried by me under the supervision and guidance of **Dr. Sasmita Nayak, Assistant Professor, School of Biotechnology, KIIT-Deemed to be University, Bhubaneswar, Orissa, India.**

Date :

Place :

Anjali Rai

Roll No.- 1662007

ABSTRACT

Protein splicing is a post translational modification of a protein, where the intervening sequence known as Inteins, catalyze their own excision from the precursor protein followed by ligation of two flanking regions(exteins)thus resulting in a functional protein. The canonical mechanism of protein splicing mainly consist of four nucleophilic displacement reactions; performed mainly by intein residues and first C-terminal extein residue. Now the studies are coming up with the role of different extein residues at intein–extein junctions that regulate the splicing reaction. That will ultimately draw a precise mechanism of splicing and its regulation in different in intein systems. *Mycobacterium tuberculosis* harbors three intein containing genes; *recA*, *dnaB* and *sufB* having. Suf B, an important member of SUF machinery as it takes part actively as sulphur acceptor, in scaffold generation and most importantly provides platform for Fe-S cluster assembly. Intein insertion in SufB adds another line of control for its functionality. Till now SufB splicing is not characterized properly. In this study we are focusing on a Histidine residue which is present at the N-extein-intein junction (-5 position with respect to intein) in *Mtu* Full length SufB protein. To identify and characterize the critical role of Histidine, it has been mutated to Alanine (H-5A). The efficiency of splicing reaction of H-5A SufB mutant has been observed at varying temperature up to 24hours. Kinetic analysis has performed to determine the rate of reaction and half-life, of H-5A. *Mtu* Full Length Suf B (WT) was taken as control and compared with H-5A mutant to decipher the role of Histidine in splicing of Suf B. Further the splicing products were confirmed by immunoblotting assay and Mass spectrometry. This study provides a preliminary idea about the critical role of Histidine present at -5 position of N-Extein- intein junction in SufB splicing reaction.

ACKNOWLEDGEMENT

Words are never enough to express the heart felt gratitude. It has been a privilege and honour to pursue my M.Sc. in School of Biotechnology, KIIT- Deemed to be University and would like to thank KSBT for granting me the opportunity to carry out the dissertation as a part of my partial fulfilment of M.Sc. degree. Any accomplishment requires the effort of many people and this work is not an exception to the rule. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project. I would like to express my deepest appreciation and sincerest gratitude to my supervisor, **Dr. Sasmita Nayak**,Assistant Professor, School Of Biotechnology, Adjunct Assistant Professor, Department Of Microbiology KIMS, KIIT-Deemed to be University, who has supported me throughout my thesis with her patience and knowledge whilst allowing me the room to work and provided me with an excellent atmosphere for carrying out this project. She has been exceptionally supportive and encouraging in her gesture of allowing me to join her research group and spending 6 months of quality time with her and her lab.

I express my sincere thanks to **Ms. Sunita Panda** and **Ms.Ananya Nanda** who helped me in practical issues beyond the textbooks and patiently helped me in carrying out my experiments. I am extremely grateful to them for being a constant critic and a strict disciplinarian and for helping me throughout in my project duration. I would like to thank **Dr. Mrutyunjay Suar**, Head of the Department and my other department faculty members for their support and help whenever required.

I would like to thank my Applied Research Labmates for cooperating with me and supporting me when ever needed and would like to thank my parents, brother, my friends and many more for their encouragement, constant support which has enable me to reach this stage of success so far. Last but not the least I must pay my homage and devotion to Almighty for giving me patience and strength to overcome the difficulties and accomplishing this endeavour.

Date:

Place : **Anjali Rai**

CONTENTS

LIST OF ABBREVIATIONS:

1) µl – Microlitre
2) mM - millimolar
3) nm - nanometer
4) CV- Column- volume
5) APS – Ammonium Persulphate
6) Asn – Asparagine
7) Cys – Cysteine
8) His – Histidine
9) IPTG–Isopropylβ-D-1-thiogalactopyranoside
10) *Mtu – Mycobacterium tuberculosis*
11)NaCl – Sodium Chloride
12) NaOH – Sodium Hydroxide
13) PAGE – Polyacrylamide Gel Electrophoresis
14) SDS – Sodium Dodecyl Sulphate
15) Ser – Serine
16) TB – Tuberculosis
17) TEMED – N,N,N',N' tetra methylene diamine
18) Thr – Threonine
19) TCEP – tris (2-carboxyethyl) phosphine.
20) ABC - Ammonium Bicarbonate.
21) CAN/ACN - Acetonitrile.
22) TFA - Trifluoroacetic acid.
23) BSA - Bovine Serum Albumin
24) TBST - Tris Buffer Saline + Tween 20.
25) HRPO - Horseradish Peroxidase.
26) Ni-NTA - Nickle Nitrilotriacetic acid.
27) IAA - Iodoacteamide

LIST OF FIGURES

LIST OF TABLES

School of Biotechnology
Kalinga Institute of Industrial Technology (KIIT)
Deemed to be University
(Established U/S 3 of UGC Act, 1956)
Bhubaneswar, Odisha, India

CERTIFICATE

This is to certify that the dissertation entitled **"Histidine near N-Extein~Intein Junction in *Mycobacterium tuberculosis* (*Mtu*) SufB regulates protein splicing"** submitted by **Ms. Anjali Rai** bearing **Roll No. - 1662007 & Registration No. - 16647551472** to School of Biotechnology, Kalinga Institute of Industrial Technology-KIIT (Deemed to be University; Declared U/S 3 Act of UGC 1956), Bhubaneswar, Odisha, India for the degree of Master of Science in Applied Microbiology is her original work, based on the results of the experiments and investigations carried out independently by her during the period from **20.12.2017** to **11.05.2018** of study under my guidance.

Further, it is also to certify that the above said work has not been previously submitted for the award of any degree, diploma, or fellowship in any Indian or foreign University.

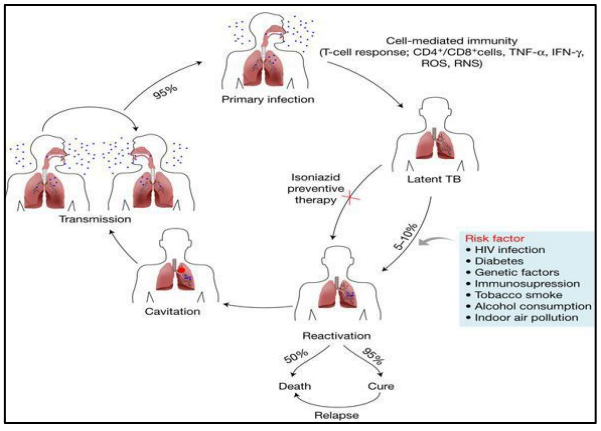
Sasmita Nayak

Dr. Sasmita Nayak
Assistant Professor
School of Biotechnology
Kalinga Institute of Industrial Technology-KIIT
(Deemed to be University; Declared U/S 3 Act of UGC 1956)
Bhubaneswar, Odisha
India

Asst. Professor
School of Biotechnology
Kalinga Institute of Industrial Technology
Deemed to be University, Bhubaneswar-24

CHAPTER 1: INTRODUCTION

Tuberculosis is one of the most debilitating disease mainly caused by *Mycobacterium tuberculosis,* which affects the human race most abominably**.** This disease is currently considered one of the most infectious diseases in the world and one of the leading causes of death worldwide. It generally affects the lungs (pulmonary TB) but can also affect other parts of the body (extra pulmonary TB). Not everyone infected with TB bacteria becomes sick. As a result, two TB-related conditions exist: latent TB infection (LTBI) and TB disease[1]. If not treated properly, TB disease can be fatal. Due to indiscriminate use of antibiotics in the past many drug resistant *Mtu* strains have evolved making its treatment more challenging. The Scientific communities are trying to overcome drug resistance, whereas *Mtu* is evolving at an alarming rate thwarting any attempts to conquer it. Tuberculosis is one of the top 10 causes of death worldwide. According to WHO, in 2015, about 10.4 million people fell ill with TB and 1.8 million people died from the disease. Over 95% deaths occur in low or middle income countries among which India is leading followed by Indonesia, Nigeria, China, Pakistan, South Africa.



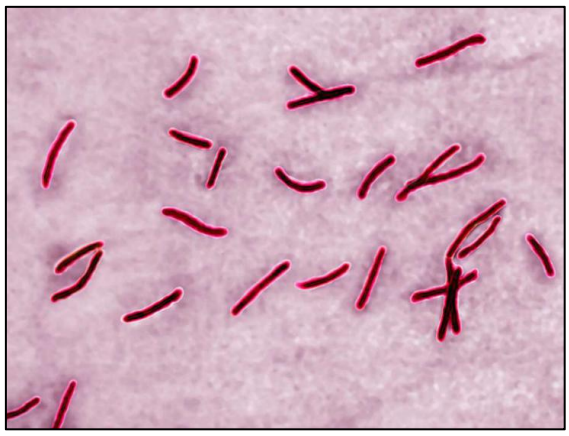


Fig 1: Life Cycle of *Mtu* and progression of TB **Fig 2 : *Mycobacterium tuberculosis***

Transmission of this deadly disease occurs mainly by aerosol and droplet. When a patient (with active Tuberculosis) coughs or sneezes, expels aerosol droplets. Each of these droplet may

transmit this disease, since the infectious dose of tuberculosis is very small. People with prolonged, frequent, or close contact with people with TB are at particularly high risk of becoming infected, with an estimated 22% infection rate. A person with active but untreated TB may infect 10-15 (or more) other people per year. Development of multiple drug resistance mycobacterial strains due to irregular medication and inappropriate treatment has been a boon to the threat of this disease. Earlier combination of drugs was used to treat tuberculosis but at present due to the emergence of multidrug resistance strains, vaccination is the only way to eliminate TB in the long term . Vaccination is helpful for developing countries, still it does not address to the needs of 2 billion patients affected by multidrug resistant TB. In such situation, the discovery of the three genes i.e., *recA, dnaB*, and *sufB* that are interrupted by inteins in *Mycobacterium tuberculosis* and their activation is dependent on intein splicing, suggests protein splicing as a novel target for anti TB drugs. These genes have different functions like DNA repair, replication and in [Fe-S] cluster formation [2].

The Suf system is an iron-sulfur cluster assembly pathway expressed in stress conditions mainly in *Escherichia coli, Cyanobacteria* etc. In case of *Mycobacterium tuberculosis* and some other Archaea, SUF machinery is the only system present for [Fe-S] cluster assembly. The Suf machinery in *E. coli* is composed of six proteins encoded by the sufABCDSE operon. The sufBCD complex is a component in the suf machinery that is responsible for the denovo [Fe-S] cluster biogenesis. The SufS cysteine desulfurase & SufE sulfur shuttle protein act together to provide sulfur for construction of nascent [Fe-S] cluster. SufA is an Fe-S carrier protein that transfers [Fe-S] clusters to target apo-proteins. SufB accepts sulfur transferred from SufE, & SufD may play a role in iron acquisition. SufC is a member of the ABC ATPase superfamily[3,4]. SufB, SufC, SufD interact with each other & three distinct states have been reported:

- SufBCD ternary complex
- SufBC sub-complex
- SufCD sub-complex.

In *Mycobacterium tuberculosis* the formation of this complex is dependent on the protein splicing of SufB, suggesting that this process is a potential new target for antituberculosis drugs. The splicing inhibitors are specific for *Mtu* and will directly interfere with survival of *Mtu* .

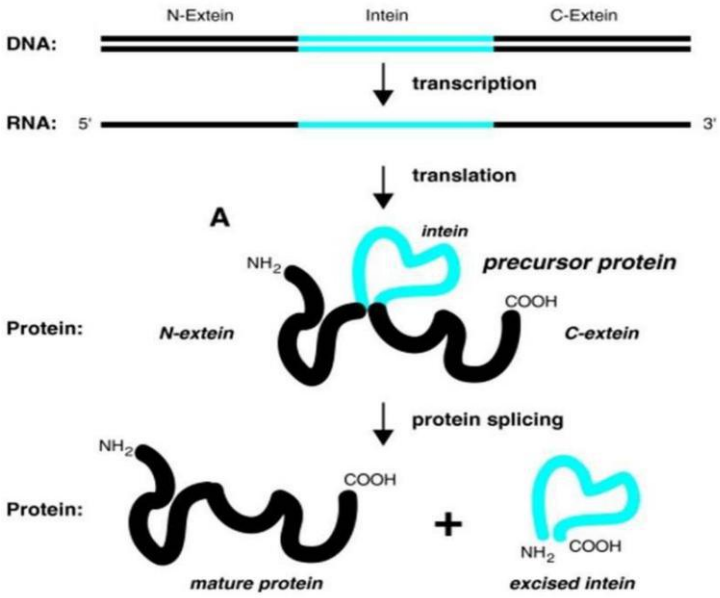


Figure 3: Protein Splicing. The intein coding sequence is transcribed into mRNA and translated into a nonfunctional protein precursor, which then undergo a self-catalytic arrangement in which the Intein is spliced out and exteins are joined to yield mature protein[5].

Protein splicing is a post translational modification where a small intervening sequence is spliced out from the precursor protein. Upon removal of this sequence the flanking regions of protein join together, hence form a functionally active protein. Splicing is a very active process, so by

inhibiting this process, protein activation can be blocked . In 1990, Stevens and colleagues found that some proteins acquire the capability of post translational modification. Protein splicing is quite similar to DNA or RNA splicing and is an intramolecular event [6].

1.1 SUF Machinery :

Suf machinery is important for iron-sulfur [Fe-S] cluster formation. The [Fe-S] clusters are metal centres consisting of elemental Fe and S in various molar ratios.[7] These [Fe-S] containing proteins are present in all living organisms and perform broad spectrum of cellular functions including respiration, central metabolism, gene regulation, RNA modification, DNA replication and repair.

Extensive studies on SUF system shows that three different systems like NIF (Nitrogen fixation), ISC (Iron - sulfur cluster), and SUF (sulfur mobilization) has been identified for the biosynthesis of [Fe-S] cluster in *E.coli*.

In E.coli, SUF system gets activated only under oxidative stress and iron starvation. SUF machinery is the only machinery for [Fe-S] cluster formation in *Mtu.*[8] SUF system is encoded by the *suf* operon consisting of 6 genes; Suf A, B, C, D, S, E. These genes encode for 6 different proteins namely, Suf A, Suf B, Suf C, Suf D, Suf E, and Suf S.[8] All of these proteins help in the [Fe-S] cluster assembly and its delivery to target proteins. Suf C is an atypical cytoplasmic ABC - ATPase, that forms a complex with Suf B and Suf D. Suf BCD complex is an essential component of Suf machinery as it acts as a scaffold protein for [Fe-S] cluster assembly and its delivery to target proteins via Suf A. Suf S is a cysteine desulfurase which mobilizes the sulfur atom from cysteine and provides it to the cluster. Suf E accelerates the activity of Suf S through BCD complex.[9]

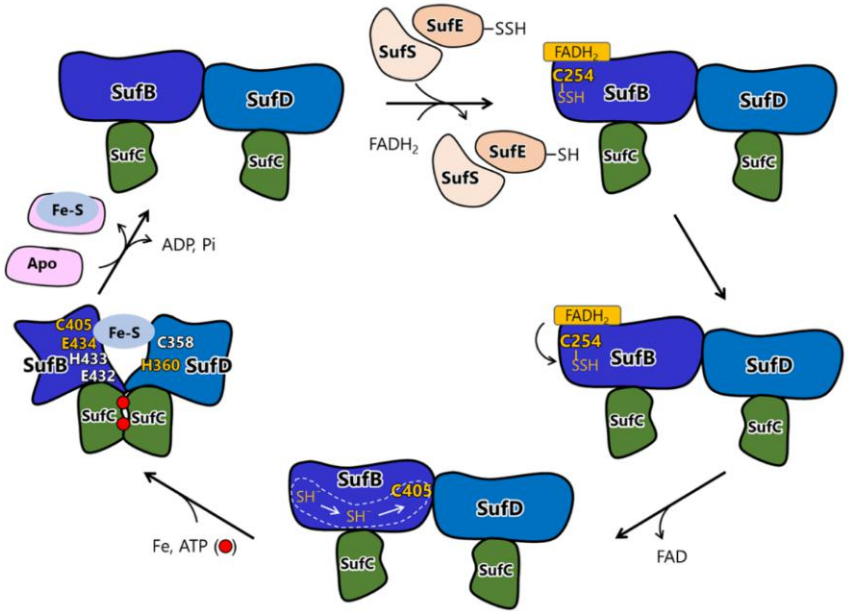


Fig 4 : Mechanism for Fe-S Cluster Biogenesis.[9]

Suf System in *Mtu* :

In case of *Mtu*, pps1 gene locus is an operon that codes for *Mtu* suf machinery with pps1 gene encoding the central SufB element. Pps1 harbors 38-39 % identity with SufB protein from *E. coli* and *Erwinia chrysanthemi.* pps1 belongs to a locus of 7 genes (ORF no. Rv 1460 - Rv 1466). Among the 7 genes, 3 are located downstream of pps1(Rv 1462 - Rv 1464) and encodes SufD, SufC & SufS ortholog proteins with respectively 24%, 48%, & 47% sequence similarity to *E. chrysanthemi* proteins, while no ORF coding for SufA or SufE was found in *Mtu.*[10]

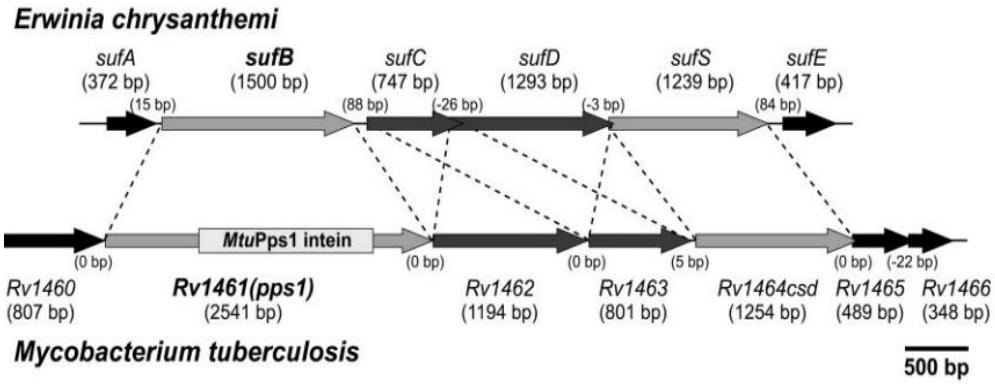


Fig 5: Alignment of Suf loci from *E. chrysanthemi* and *M. tuberculosis.*[10]

1.2 Suf B:-

Suf *B* is highly conserved protein coded by gene Rv 1461 and is interrupted by an intein sequence. It is widely distributed among various organisms. The most interesting property of Suf B intein is only found in pathogenic species of mycobacteria such *as M.tuberculosis, M.leprae and M .gastri* but is absent in non-pathogenic stains like *M.smegmatis*.[11] In other hand the exteins are similar , but intein insertion points are different in *M.tuberculosis, M.leprae and M .gastri.* Previous works show that the splicing of Suf *B* is crucial for the functionality of Suf BCD complex. Suf B provides a platform for the formation of [Fe-S] cluster assembly. Protein splicing of Suf *B* is essential for the activation of Suf *B* protein.[12]

1.3 INTEIN:

Intein is an intervening sequence that is capable of excising itself from the precursor protein forming a functional protein. After intein excision, the flanking regions left behind are called exteins. Exteins are the coding sequence of the protein and are present on both the side of intein. These exteins ligate and form a functional active protein.

Inteins are found in all the three domains of life in Eukaryotes, Eubacteria, Archeae and in some viral and phage particles. The first intein was discovered in 1988 through the sequence comparison between *Neurospora crassa* and carrot vacuolar ATPase (without intein) and the homologous gene in yeast (with intein) that was first described as a putative calcium ion transporter.

Inteins are found in conserved regions of host protein where any deletion or mutation is likely to give rise to a non –functional protein. The transcription and translation takes place along with the

host protein inteins are spliced out post translationally by a complex self-catalysed process called protein splicing in order to form a functionally active protein.

Inteins contain two types of structural domain; one being the catalytic domain and other is the endonuclease domain.[13] Catalytic domain contains all the critical residues that help in splicing out the intein from the precursor protein whereas Endonuclease domain helps in the lateral transfer of an intervening sequence to a homologous allele that lacks the intein sequence. Depending upon the site of splicing, two types of splicing domains are present, one is N-terminal splicing subdomain and other is C-terminal splicing subdomain. Each of these blocks contains conserved blocks of aminoacids. Blocks like A, N2, B, and N4 are found in N-terminal subdomain of precursor protein where as block G and F are found in C-terminal subdomain[14][15]. Although not conserved in their entirety, several positions in each motif contain highly conserved groups of similar amino acids. Most inteins begins with Cys or Ser and end in His-Asn, or in His-Gln. The first amino acid of the C-extein is an invariant Ser, Thr, or Cys, but the residue preceding the intein at the N-extein is not conserved. Inteins are classified into: large inteins and minimized inteins (mini inteins)[16] . Large inteins contains both the domains but the mini inteins lack the endonuclease domain. Inspite of lacking the endonuclease domain mini inteins are splicing active. This shows that the two domains of inteins are functionally independent.

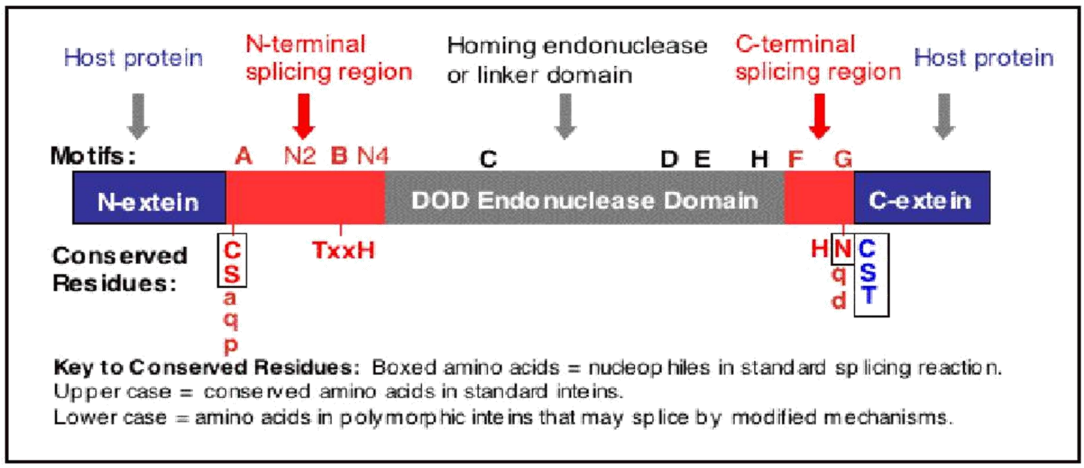


Fig 6: Schematic representation of unspliced precursor SufB host protein[17]

1.4 Mechanism Of Intein Splicing:

Protein splicing is defined as the self-excision of an intervening sequence (the INTEIN) from a protein precursor along with the ligation of the two flanking protein fragments (the EXTEINS) to form a mature active host protein. Intein-mediated protein splicing results in a native peptide bond between the ligated exteins. The standard protein splicing mechanism consists of four nucleophilic displacement reactions directed by the intein plus the first C-extein residue. The first residue of C-extein is Ser or Thr or Cys. These residues often function as nucleophiles. Inteins use the same mechanism like other enzymes to perform catalysis and are considered as single turn over enzymes with the proximal extein residues as substrates. Folding of the intein with in the precursor protein brings together the two splice junctions and the critical residues that assist catalysis.

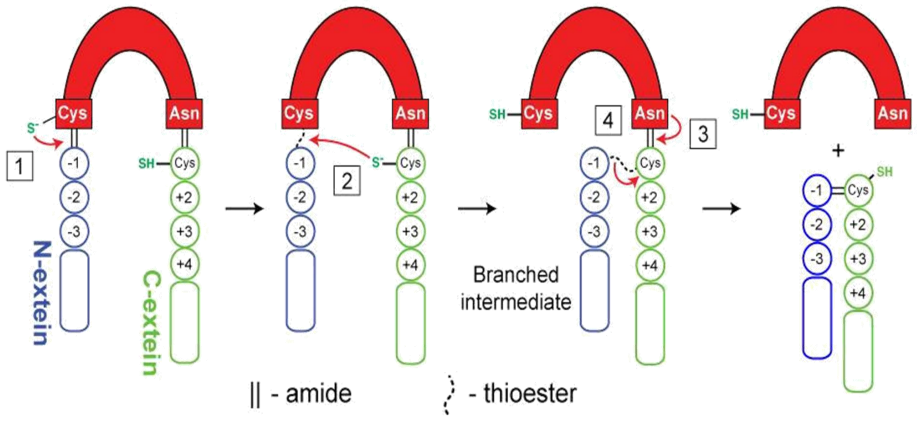


Fig 7 :- Mechanism of Protein Splicing[18].

The first step is initiated by nucleophillic attack by C1 or S1 that converts the peptide bond to a thioester bond at the N- splice junctions by an N-S or N-O acyl shift. In next step, a second nucleophillic attack by C+ 1 form a branched chain intermediate at the side chain of C+1 by trans-esterification. Next, the branched intermediate gets resolved by Cyclization of terminal Asn on intein leading to ligation of exteins and cleaving off intein. In the final step, rearrangement of the thioester bond forms a stable peptide bond in the ligated exteins[18].

In protein splicing the following four important steps are involved:

1. <u>First acyl rearrangement</u> : The protein splicing is initiated by a nucleophilic attack with a N - O shift if the first intein residue is Serine (Ser) or N - S shift if the first intein residue is Cysteine (Cys). The peptide bond present between the N extein and the N terminal of the intein gets converted to thio-ester linkage resulting in a linear intermediate.

2. <u>Trans esterification</u> : The newly formed thio-ester bond is attacked by the OH- or SH group of the first residue in the C extein (Cys, Ser, or Thr). This leads to the transfer of N extein to the first residue of the C-extein leading to the formation of branched intermediate.

3. <u>Asparagine Cyclization</u>: The last residue at the C terminus of intein is generally Asparagine (Asn). The cyclization of Asn to succinamide results in the release of the intein and the exteins are thus linked by thio-ester bonds.

4. <u>Second acyl rearrangement</u>: A spontaneous S-N or O-N acyl shift leads to the formation of an amide bond by the rearrangement of the thio-ester bond between the ligated extein segments.[15]

1.5 Why Histidine residue is important in splicing reaction?

Different Intein and Extein residues co-ordinate with each other via different interactions assisting the splicing reaction. The residues present in splicing domain are highly conserved at different conserved motifs. Likewise, Histidine residues play an important role in forming pi-pi interactions or pi-hydrogen interactions with other aromatic amino acids and this residue is more important as it is also a metal chelating residue assuming its interaction with [Fe-S] cluster assembly. Conserved Histidine residues present in B-block and F-block helps in linear thioester generation and stabilization, also helping in resolution of branched intermediate in protein splicing.[19][20]

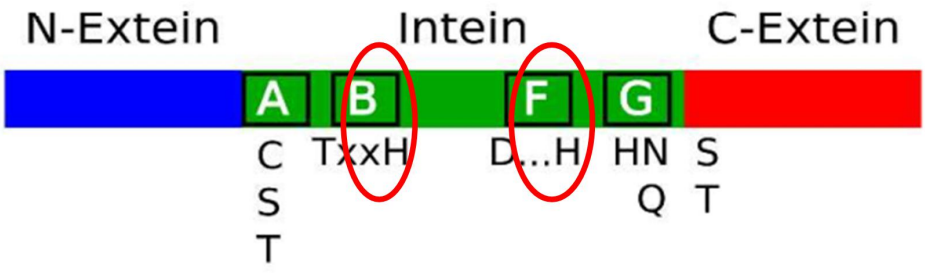


Fig 8: SufB Precursor protein showing conserved Histidine residues at B-Block and F Block in splicing domain.[19]

In *Mtu* SufB protein, another Histidine residue is present at -5 position, nearer to N-extein intein junction. It has already been shown that distal extein residues also play a vital role in protein splicing. In our study to characterize the role of Histidine at N- Extein-Intein junction in splicing reaction, it has been mutated to Alanine.

1.6 KINETIC ANALYSIS :

Kinetic analysis gives an idea about the % of splicing, rate of reaction and half-life of the reaction. It is used to compare the rate of splicing in WT and other mutants so that we can have a precise idea how a mutant can regulate the on/off pathway of splicing reaction.

CHAPTER 2: REVIEW OF LITERATURE

Inteins where initially reported in 1990 by two individual reports (Hirata *et al.*, 1990; Kane *et al.*, 1990). Both studies showed that the gene VMA1 from *Saccharomyces cerevisiae* produced a 69 KDa mature protein and a spacer protein of 50KDa. Kane.et al (1990) reported that inteins are part of mature mRNA and the inteins will produce after translation on mRNA. The final proof of this process being a post translational modification came with the development of the first *in vitro* protein splicing system (Xu *et al.*, 1993). The system showed that the intein sequence is first excised from the gene after translation by a posttranslational protein splicing process. Intein mediated protein splicing is a self-catalytic and intra-molecular process that does not require any external energy source such as ATP or other macromolecules (Kawasaki *et al.*, 1997).

Since the discovery of the first intein in *S. cerevisiae*, inteins have been also discovered in the three domains of life mainly, Eukaryote, Bacteria, and Archaea, and in viruses, and bacterial phages (Pietrokovski, 1998a; Lazarevic *et al.*, 1998). But, inteins have only been identified in unicellular organisms. To keep track of the number of intein sequences discovered a database has been created (http://tools.neb.com/inbase/) in which more than 550 intein sequences from eucarya, eubacteria, and archaea are listed (Perler,2002).

Henry Paulus (2000) reported that Intein mediated protein splicing is a self-catalysed and intra molecular process that does not require any external energy source such as ATP or other macromolecules. Due to this self catalytic properties he suggested that intein works as an enzyme. Inteins were believed to most ancient as they were found in all the three domains of life i;e Eubacteria , Eukaryotes and Archeae. Inteins are found only in unicellular organisms. Inteins are related to Hedgehog protein because they shared a common protein folding with Hedgehog C-terminal domain.[21]

Intein sequences range from 134 residues for the smallest *cis* splicing intein to more than 1000 residues (Evans *et al.*,1999; Perler,2002). Many discovered intein sequences are bifunctional because they contain an endonuclease domain (Homing Endonuclease, HEN) inserted in the intein sequence. Homing endonucleases domains are DNA cutting proteins that have specific DNA recognition sequence from 12-40 base pairs (Roberts and Macelis,1997). Inteins lacking an endonuclease domain are in some cases referred to as mini-inteins and they are up to ~200-300 amino acid residues.

Inteins have low sequence homology and only few residues are highly conserved among inteins. The most conserved residues are the first residue of the intein, a Cys or Ser, and the last residue, which is an Asn. Additionally, protein splicing is dependent on the first residue of the C-extein (the +1 residue) being a Cys, Ser, or Thr (Xu and Perler, 1996). However, these residues are not fully conserved among inteins and exceptions are found.

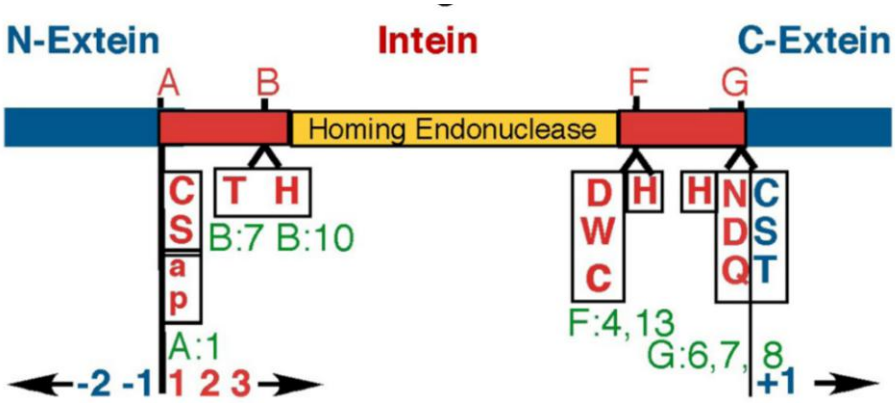


Fig 9 : Schematic representation of a precursor protein showing intein and extein domains. [15]

Inteins share some sequence similarity and sequence motifs, which contain conserved residues. The sequence motifs are referred to as block A, B, F, and G (Perler et al., 1997). The sequence block A and B are located near the N-terminus of the intein whereas the blocks F and G are located near the C-terminus.

Chong et.al (1996) examined the splicing of chimeric protein containing the intein of vacuolar ATPase subunit (VAM) of *Saccharomyces cerevisae* that involves cysteine rather than serine in reaction centre.[22]They deduced the protein splicing steps by analysing intermediates and side products that accumulated as a result of amino acid substitution. They developed the first mesophilic in vitro protein splicing system as well as strategies for modulating the rate of protein splicing and for converting the splicing reactions to an efficient protein cleavage reaction at either spliced junction.

Huet et.al (2006) investigated the role of Suf BCD complex as an essential component of the SUF machinery in [Fe-S] cluster biogenesis. They showed that the formation of Suf BCD complex was dependent on protein splicing of Suf *B* suggesting potential new target for anti TB drugs.[23]

Huet et.al (2005) established *pps1* gene as a central element of SUF system that would play an essential function for *Mtu* survival in iron limited and oxidative stressed condition.[10]

Natalya I. Topilina et.al (2015) discovered an exceptional sensitivity of *Mtu* SufB intein splicing to oxidative and nitrosative stresses when expressed in *Escherichia coli.* Experiments revealed that reactive oxygen species and reactive nitrogen species inhibit SufB extein ligation by forcing either precursor accumulation or N-terminal cleavage.

Paulus (2000) reviewed in an article that inteins can self-catalyse protein splicing reactions suggesting that intein works as an enzyme.[24]

B Perler (2005) reviewed in an article that inteins are the protein splicing elements that employ standard enzyme strategies to excise themselves from precursor proteins and ligate the surrounding exteins. He also discussed about the protein splicing mechanism; the four nucleophilic displacement reactions directed by the intein plus the first C-extein residue. Understanding intein proteolytic cleavage and ligation activities has led to the development of many novel application in fields of protein engineering, enzymology etc.[25]

Mujika et.al., (2017) showed that B-Block Histidine can act as a base activating the side chain of Cys^{1} residue and Histidine plays other two important roles which are - a) ground-state destabilization maintaining the scissile peptide bond and b) protonation of leaving amide group.[19]

Chen et.al. (2000) showed that Histidine at F-block helps in third step of splicing reaction that is, Asparagine cyclization . [20]

David W.Wood (2014) showed that after engineering a mini-intein splicing domain where the endonuclease domain of *Mtu* RecA intein was removed ,the remaining domain that ligated by several short peptides resulted in very low splicing efficiency than full-length intein. But native splicing efficiency was restored upon V67L mutation. V67L mutation lowers the global fluctuation in all modelled mini-inteins and stabilizes mini-intein construction.[26]

Qin WU(2014) Suggested that conserved residues $^{N}Arg^{50}$ and $^{C}Ser^{35}$ in DnaE split inteins are critical in the trans splicing of Npu DnaE intein.The conserved non-catalytic residues of split

inteins modulate the efficiency of protein trans-splicing by hydrogen-bond interactions with the catalytic residues at the splice junction.[27]

Francine B.Perler(2000) suggested an alternative protein splicing mechanism in KlbA family. The conserved residues Ser or Cys at the intein N-terminus and conserved intein penultimate His are absent in the KlbA intein. As N-terminal Ala cannot initiate the splicing reaction, so C-extein nucleophile attacks a peptide bond at N-terminal splice junction followed by branch resolution by Asn cyclization and an acyl rearrangement to form a native peptide bond between ligated exteins.[28]

Lana Saleh et.al (2011), for the first time investigated the kinetics splicing by *Methanococcus Jannaschii* KlbA intein. They constructed a Mja KlbA intein-mini extein precursor using intein mediated protein ligation and engineered a redox switch that initiates the splicing reaction in presence of reducing agent DTT. They suggested that formation of the branched from precursor is reversible where as the productive decay of this intermediate to form the ligated exteins is faster and occurs with a rate constant of $2.2 \times 10^{-3\ s-}$.[29]

Lixin Chen et.al (2000) for the first time investigated the role of penultimate histidine in splicing mechanism. They examined the splicing efficiency in *Methanococcus jannaschii* phosphoenolpyruvate synthase and RNA polymerase subunit A' which lack the penultimate Histidine. They suggested that loss of this Histidine inhibited but may not have blocked the splicing process.[30]

Zhenming Du (2009) suggested that protein splicing occurs by acid-base catalysis in which the ionization states of active site residues are crucial for reaction mechanism. According to him many conserved histidine have been shown to play important roles in protein splicing including

the most conserved B-Block histidine. This histidine acts as a general base to initiate the splicing with an acyl shift and then as a general acid to cause the breakdown of scissile bond.[31]

Natalta l.Topilina (2015) showed some critical residues in native exteins also regulate the intein splicing mechanism. Some environmental factors like High temperature and solution increases their efficiency.[32]

K.V.Mills et.al., (2004) have showed that the asparagine residue important in third step of splicing reaction when mutated with C-terminal Glutamine improves the rate of splicing which they have also shown in detailed kinetic analysis of splicing efficiency over varied temperatures starting from 30℃ .[33]

Mutant analysis shows the differences in the splicing and cleavage reaction and make mechanistic approach more clear where as the kinetic analysis can quantitate it.

Protein splicing reaction is a pseudo first order reaction. Mutant analysis and kinetic analysis has been done in many previous studies that emphasises the importance of residues.

CHAPTER 3: HYPOTHESIS

HYPOTHESIS:

- Metal chelating residues near N-terminal Intein-Extein Junction (H-5) may regulate Mtu SufB intein Splicing.

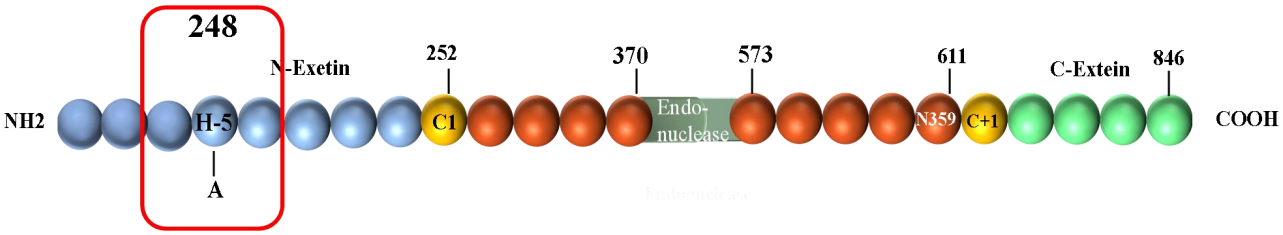


Fig 10: Schematic diagram of Mtu FL SufB Protein highlighting different conserved residues essential for SufB Splicing.

Assumptions :

- The Histidine residue present near N- Extein-Intein junction at -5 position may regulate SufB protein splicing.

OBJECTIVES :

- To check the splicing efficiency of SufB mutant protein (H-5 to alanine) relating to WT SufB at different temperatures.
- Detailed Kinetic Analysis.

CHAPTER 4: MATERIALS AND METHODS

4.1 MATERIALS :-

Table 1: List of Chemicals

Chemical Reagents	Make
Acrylamide	Himedia
Acetic acid	Merck
Ammonium persulphate	Himedia
β - Mercaptoethanol	Spectrochemicals
Bis-acrylamide	Himedia
Bovine Serum Albumin	Himedia
Bromophenol blue	SRL
Comassie Brilliant Blue R-250	Merck
Developer	Ace Chemicals
ECL Solution	Abcam
Fixer	Ace chemicals
Glycerol	Himedia
Glycine	SRL
Imidazole	SRL
IPTG	Sigma
Isopropanol	Himedia
L-Arginine	Himedia
Prestained Protein Precision Marker	Bio-Rad
SDS	Merck
Sodium phosphate buffer	Himedia

Tetramethylethylenediamine (TEMED)	SRL
TCEP	Sigma
Tris Base	Himedia
Urea	Himedia

Table 2 : List of Instruments

Instrument Name	**Company**
Centrifuge	Sigma 18K
Incubator	Labtech
Ultrasonicator	Hielscher
UV Spectrophotometer	Shimadzu UV-1800
Vertical mini SDS Gel Apparatus	Bio-Rad
Western Blot Apparatus	Bio-Rad
Water Bath	Srico
MALDI-TOF-TOF	Absci-ex

- **Bacterial Cells:**

- BL21(DE3) cells (*Escherichia coli* cells for checking the protein expression).
- DH5-ALPHA (For storage purpose).

- **Expression Vector:** pACYC Duet-1.

- **Recombinant Plasmid Constructs:**

i. pACYC Duet-1 with full length SufB (Wild Type- WT).

ii. pACYC with SufB mutant : H-5A

iii. pACYC with SufB SI (Splicing inactive double mutants).

All these plasmids were prepared by Dr. Sasmita Nayak in Belfort Lab in New York, USA.

Software - Gel Quant. NET- BiochemLabSolutions.com, GraphPad Prism.

4.2 METHODS :

4.2.1 PLASMID ISOLATION BY MINIPREP METHOD FOR LOW COPY NUMBER PLASMIDS :

QIAgen mini-prep system is a fast, simple and cost effective method for plasmid isolation when small amount of plasmid must be isolated. This procedure is based on alkaline lysis of bacterial cells followed by adsorption of DNA onto silica membrane in presence of high salt. The unique silica membrane used in QIAgen Mini-prep Kits completely replaces glass or silica slurries for plasmid minipreps.

The three basic steps involved in this procedure are as following: -

A) Preparation and clearing of bacterial lysate :

Bacteria were lysed under alkaline conditions, and the lysate were subsequently neutralized and adjusted to high-salt binding condition. After lysate clearing the samples were purified on the QIAgen mini-prep silica membrane.

B) DNA adsorption to the QIAprep membrane :

The silica membrane present in QIAgen mini-prep column strips and plates that helps in selective adsorption of plasmid DNA in high-salt buffer and elution in low salt buffer. The optimized buffers in lysis procedures, combined with the unique silica membrane, ensure that only DNA would be adsorbed, while RNA, cellular proteins and metabolites were not retained on membrane but were found in the flow- through.

C) Washing and elution of plasmid DNA :

Buffer PB was used to briefly wash off the endonucleases. The buffer PB wash step is essential for purifying low-copy plasmids in large culture volumes. Salts were efficiently removed by brief washing with PE buffer. High quality DNA was eluted from the QIAgen mini-prep column with 50-100µl of EB buffer or water. The purified DNA was ready for immediate use in a range of applications. Precipitation, concentrating, or desalting of the product is not necessary. Elution efficiency depends on pH, maximum elution efficiency is achieved between pH 7.0 and 8.5. When using water for elution the pH must be adjusted. The DNA when eluted with water should be stored at -20ºC since the DNA might degrade in the absence of buffering agent.

Standard Protocol:

Bacterial cell pellet was resuspended in 250µl of buffer P1 in which RNAase was added previously and transferred to a micro centrifuge tube. 250 µl of buffer P2 was added and tubes were gently inverted 4-6 times to mix. No cell clumps should be visible after suspension of the pellet. Buffer N3(350 µl) was added and tubes were immediately inverted gently 4-6 times. To avoid localized precipitation, the solutions were mixed gently but thoroughly, immediately after addition of buffer N3. The solution turned cloudy. The cells containing the buffers were

centrifuged at 13,000 rpm (17,900xg) for 10min in a table top micro centrifuge. Supernatant were applied to QIAgen mini-prep Spin column by decanting or pipetting and centrifuged again for 30 ~ 60 sec. After the flow through was discarded the QIAgen mini-prep spin column was washed by adding 0.75ml of buffer PE & PB & centrifuged for 30 ~ 60 sec. Flow through was discarded again and centrifuged for an additional 1 min to remove residual wash buffer. Next QIAgen mini-prep column was placed in a 1.5ml sterile micro centrifuge tube to elute the DNA. 0.5ml of EB (10mM Trish pH 8.5) or water was added to the center of each QIAgen mini-prep spin column, allowed to stand for 1 min, and centrifuged again for 1 min.

4.2.2 PREPARATION OF CHEMICAL COMPETENT CELLS:

The cells that are capable to uptake DNA and other genetic material from the outer environment are called competent cells. Most types of cells cannot take up DNA efficiently unless they have been exposed to special chemical or electrical treatments to make them competent. The standard method for making the bacteria permeable to DNA involves treatment with calcium ions. Brief exposure of cells to an electric field also allows the bacteria to take up DNA and this process is called as electroporation. In general, *E.coli* BL21 (DE3) strain is widely used for protein expression as the cells possess λDE 3 lysogen, which expresses T7 RNA polymerase under the control of lac UV promoter and are deficient of protease Lon and OmpT.

Standard Protocal for Preparation Of Chemical Competent Cells:

E. coli BL21(DE3) cells were grown overnight in 5ml of LB and next day subculture was done in 200ml LB (1:100 dilution) and incubated at 37◦C for $2^{1/2}$ hours at 200rpm (till O. D_{600} reached around 0.6), then the flask containing the cells was kept on ice for about 30minutes (to restrict cell growth). Cells were harvested at 5000rpm for 10 minutes at 4◦C. After harvesting the cells,

the supernatant was discarded and the pellets were resuspended in 20ml of 100mM $MgCl_2$ and centrifuged at 5000 rpm for 10minutes at 4◦C. Then the supernatant was discarded and the pellet was resuspended in 40ml of chilled 100mM $CaCl_2$, incubated for 20mins on ice and centrifuged at 5000rpm for 10mins at 4◦C. Finally the pellets were resuspended in 25ml of ice-cold 15% glycerol and 85mM $CaCl_2$ and centrifuged at 5000 rpm for 10minutes at 4◦C. Finally the cells were resuspended in 500µl of 15% glycerol + 85mM $CaCl_2$ and 80µl of the mixture was aliquoted in 1.5ml micro centrifuge tubes and stored at -80◦C for future use.

4.2.3 TRANSFORMATION OF BL21(DE3) CELLS:

DNA can be exchanged among bacteria by three methods: transformation, transduction and conjugation. Transformation is one of the most popular techniques of molecular genetics because it is often the best way to reintroduce experimentally altered DNA into cells. During the process of transformation, genes are transferred from one bacterium to another as ‘naked’ DNA solution. There are two major parameters involved in efficiently transforming a bacterial organism. The first is the method used to induce competence for taking up the naked DNA and the second major parameter is the genetic constitution of the host strain of the organism being transformed. Competent cells are capable of taking up DNA from their environment and expressing DNA as functional proteins. Bacteria can take up DNA artificially by using different techniques such as electroporation, heat shock, Ca^{2+} treatment of cells and protoplast uptake of DNA. Calcium chloride treatment is one of the best methods for the preparation of competent cells. Cells made permeable to DNA by calcium ion treatment will take up both single stranded and double stranded DNA. In case of electroporation, a brief and high intensity electric pulse is applied to bacterial cells to create transient pores, which facilitates the entry of exogenous molecules like DNA, RNA etc.

➢ TRANSFORMATION BY CHEMICAL METHOD :

Cells are chemically transformed by using Calcium Chloride. Calcium chloride elutes the foreign DNA into the cell and a brief heat shock opens up the pores of the cell to help in transformation.

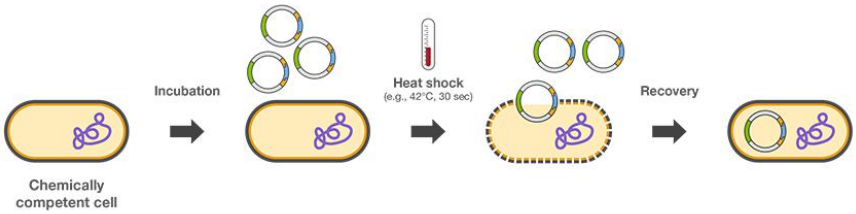


Fig 11 : Chemical Transformation

Standard Protocal :

The competent cells stored at -80°C were allowed to thaw on ice and 3µl of respective plasmid (15ng/µl conc.) was added to 80µl of competent cells and the mixture was then kept at 42◦C for 90 seconds. Immediately after that, 950 µl of LB medium was added for recovery of the cells and kept in ice for 3 minutes (cold shock). Cells were then incubated at 37°C for 2.5 hrs, while shaking at 200rpm. After incubation 450µl of the cells were spreaded on to chloramphenicol (CAM,25 µg/ml) containing LB agar plates and incubated overnight at 37°C. Next day single transformed colonies were inoculated in 5ml LB+ CAM and left to grow overnight at 37◦C.

4.2.4 PROTEIN OVEREXPRESSION:

The process in which genetic information present in gene is used to synthesize any functional gene product (like Protein) is called Gene Expression. Sometimes inducer is used to over express a protein. In our case, the transformed cells [BL21 (DE3)] contain full length SufB

gene and also mutated SufB genes. These genes need to be over expressed in order to obtain the protein products.

The competent BL21(DE3) cells contain a T-7 polymerase gene controlled under lac UV promoter. T-7 RNA polymerase is expressed upon addition of isopropyl-1-thio-β-D-galactopyranoside(IPTG) which induces a high-level protein expression from T-7 promoter driven expression vector, which in our case is pACYC duet-I containing our gene of interest (*Mtu* FL-SufB). *E. coli* BL21(DE3) strain lack both proteases: lon protease and the ompT membrane.

Standard Protocal:

Next day, subculture was done in LB broth containing 25µg/ml via 1:100 dilution and incubated at 37°C, 200rpm till the OD_{600} reaches >0.6 (nearly 2hr 30min). Then, IPTG (0.5M) was added for over expression of the desired protein and incubated for 4 hours, and the cells were harvested by centrifugation at 6000rpm for 10 min at 4◦C. Supernatant was discarded and the pellets were stored at - 80°C.

4.2.5 ULTRA SONICATION :

Sonication is a process where high intensity ultra-sonic sound wave is used to agitate a particle inside the solution. It is used to lysate the cells so that the over expressed protein come out of the cell and dissolve in the solution. Sonication is done using different amplitude with different cycle arrangements i;e half cycle or full cycle with different on and off duration depending upon the type of cell used for sonication and its density. Sonication is most commonly performed using ultrasonic bath or an ultrasonic probe. Mechanical energy from the probe initiates the formation of microscopic vapor bubbles that form momentarily and

implode, causing shock waves to radiate through a sample. Cons of sonication include heating of the sample, potential variations in yield as well as generation of free radicals that can react with other molecules. So, to avoid excessive heating sonication is usually performed in multiple short bursts while the sample is immersed in ice.

Standard Protocal:

The cell pellets were taken out of -80°C and the pellet was resuspened in 5ml of Lysis buffer maintaining the condition at 4°C . Sonication was done for 3cycles with 10sec on and 10sec off for 10 minutes at 45% amplitude.

4.2.6 Inclusion Bodies Isolation:

After ultra- sonication, the cells were centrifuged at 12000rpm 4°C for 20mins to separate the inclusion bodies (IB) from the soluble cellular proteins. The IB pellet was washed thrice to remove any soluble contaminant and was resuspended in 3.2ml of 8M urea (containing lysis buffer) to solubilize the protein.

4.2.7 PROTEIN PURIFICATION UNDER DENATURING CONDITION:

Ni-NTA Column provides a simple method for rapid screening and purification of 6x-His tagged proteins. Protein purification by Ni-NTA column is based on the principle of affinity chromatography, where Ni-NTA resins act as the stationary phase, and 6x His-tagged proteins from crude cell lysates are in the mobile phase. The high affinity of the Ni-NTA resins for 6x-His-tagged proteins or peptides is due to both the specificity of the interaction between histidine residues and immobilized nickel ions and to the strength with which these ions are held to the NTA resin. Imidazole competes with proteins for binding to metal-ion-charged chromatography

media. Equilibration buffer and sample are usually complemented with a low concentration of imidazole to reduce non-specific binding of host cell proteins. At somewhat higher concentrations, imidazole may also decrease the binding of histidine-tagged proteins. A concentration of 500mM imidazole in the elution buffer ensures complete elution of the histidine-tagged protein.

Standard Protocal:

The Ni-NTA column was washed with 5-10CV (column volume) of water followed by equilibration with 5-10CV of binding buffer. The sample was prepared in binding buffer and was slowly applied to column, avoiding the air bubbles, the flow throw sample was collected. Then the column was washed with 10-15CV of washing buffer.After that, the target protein was eluted from the column by using elution buffer. Then the elution fraction was used for quantification of protein by Bradford Assay.

4.2.8 PROTEIN ESTIMATION BY BRADFORD ASSAY :

Bradford assay is a calorimetric protein assay developed by Marion M.Bradford. It is a spectroscopic analytical procedure used to measure the concentration of protein in a solution by measuring the absorbance shift of the dye Coomasie Brilliant Blue G-250 from 465nm-595nm when binding to a protein. Both hydrophobic and ionic interactions stabilize the anionic form of the dye, causing a visible color change. The dye reagent reacts primarily with Arginine residues and less so with His, Lys, Tyr, Trp, Phe residues. The increase of absorbance at 595nm is proportional to the amount of bound dye and thus the amount of protein present in sample.

In our case we are going for Bradford in order to quantify the amount of protein we load in SDS-PAGE so that each well contain equal amount of protein. For that we prepare standard sample that contain 10µg to 50µg of BSA with 100times dilution.

Standard Sample Preparation:

Table No. 3: Bradford Sample Preparation

Reagents / Concentration	10µg	20µg	30µg	40µg	50µg
BSA	10µl	20µl	30µl	40µl	50µl
Water	990µl	980µl	970µl	960µl	950µl
Bradford	800µl	800µl	800µl	800µl	800µl

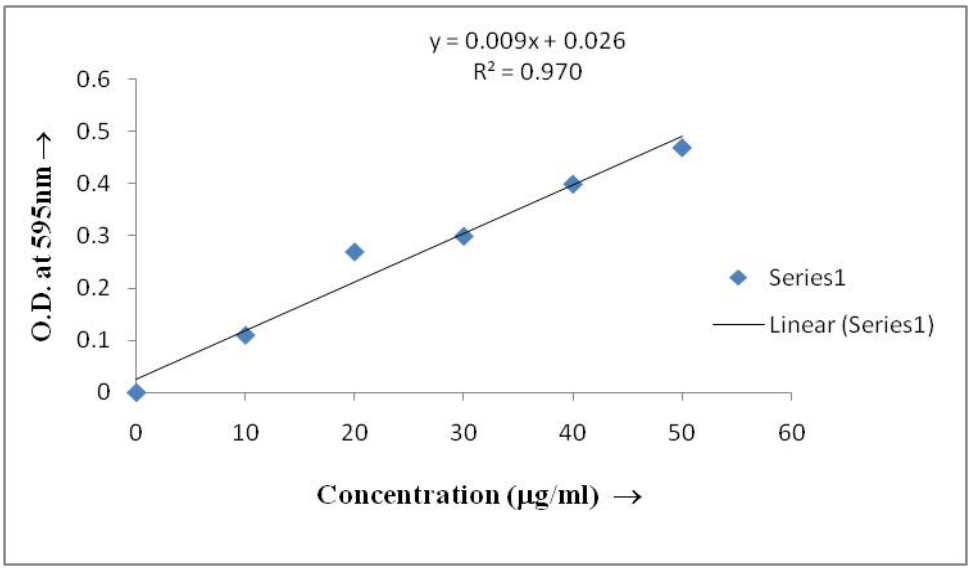


Fig 12: Standard curve of Bradford Assay

4.2.9: INVITRO REFOLDING OF PROTEIN:

After purification and quantification of protein, the estimated amount of protein(maintaining the concentration 240µg/ml) and 2mM of TCEP was added in 1ml of Renaturation buffer (pH 7.5) at different individual MCTs which was incubated at different temperatures (16˚c, 20˚c, 25˚c, 30˚c and 37˚c). Each MCTs containing samples were taken out of their respective temperatures in regular time interval (i.e. 30 mins interval upto 8hours and 1 hour interval from 9 hour upto 24 hour) and was stored in -80˚C .

4.2.10:VISUALISATION OF *Mtu* SufB SPLICING PRODUCTS VIA SDS-PAGE:

Electrophoresis is a process where charged molecules are separated on the basis of their molecular weight in response to an electric field. Polyacrylamide Gel Electrophoresis is a process used specifically to separate different components of protein depending on their molecular weight. This technique is based upon the principle that a charged molecule will migrate in an electric field towards an electrode with opposite charge. The strongly anionic detergent SDS is used in combination with a reducing agent and heat to dissociate the proteins before they are loaded onto the gel. The denatured polypeptides bind SDS and become uniformly negatively charged. Different proteins have different molecular weight and size so they are separated via molecular sieving effect based on their size and weight where lighter or smaller polypeptides migrate faster than heavier polypeptides.

Table 4: Resolving and Stacking Gel Preparation

Reagents	Stacking gel(pH6.8)	Resolving gel(pH8.8)
Distilled water	2.44ml	3.96ml
Tris-HCL	1 ml	2.5ml
Acrylamide	480 ul	3.32 ml
SDS (10%)	40ul	100ul
Ammonium persulfate(APS)	40ul	100ul
TEMED	4ul	6ul

Standard Protocal for Sample Preparation :

The loading dye was added to sample in 1:4 ratio and boiled at 95◦C for 5mins,followed by a quick spin at 10,000 rpm for 1 minute. 3µl of protein ladder was loaded to the first well of the gel and to the rest, 30µl of protein samples were loaded and the gel was run at 80V. After the gel was fully run, it was stained with staining solution (Comassie Brilliant blue-R250),overnight; followed by de-staining (also overnight).After the gel was de-stained properly, image was taken and quantified by a software: GelQuant.NET.

4.2.11: KINETIC ANALYSIS :

The current study is considered as a pseudo first order reaction; where the product of splicing is Ligated Extein and Intein. % of Splicing is calculated by using the formula % of splicing = (LE+I) / (LE+I+P). Then the % of splicing is plotted against the time(t) showing a non-linear regression ; Characteristics of a Pseudo first order kinetics. Then the curve is fitted into the formula $Y= Y_0 + (Plateau - Y_0) * (1- e^{-kt})$ in the software Graphpad Prism. After successful fitting, K(rate constant) , $t_{1/2}$ (half-life) and V_{max} are calculated.

4.2.12:CONFIRMATION OF IDENTITY OF SPLICING PRODUCTS VIA WESTERN BLOT:

The western blot (or protein immunoblot) is a widely used analytical technique used to detect specific protein in a sample of tissue homogenate or extract. It is based on the ability of protein to bind specific antibodies. It was first described by Towbin et. al. in 1979 and has since become one of the most commonly used methods in life science research. It is an analytical method wherein a protein sample is electrophoresed on an SDS-PAGE and electrophoretically

transferred to PVDF/Nitrocellulose membrane. The transferred protein is detected using specific primary or secondary labeled antibody. Antibodies bind specific sequence of amino acids, known as epitopes. Finally, a substrate that reacts with an enzyme is used to visualize the antibody/protein complex. Western blotting could detect target protein which is as low as 1ng due to high resolution of the gel electrophoresis and strong specificity and high sensitivity of the immunoassay.

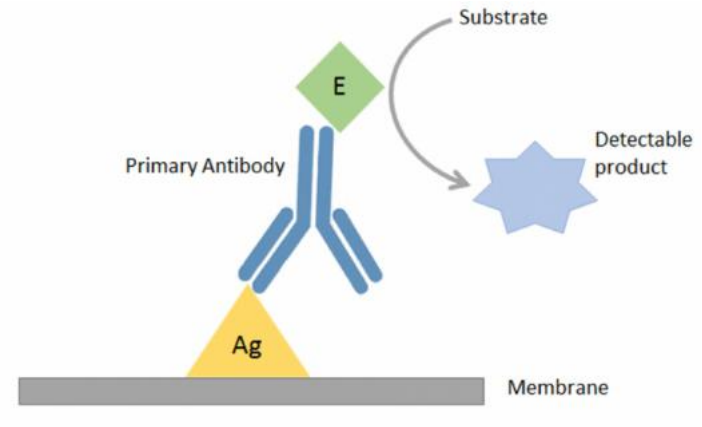


Fig 13: Detection of target protein via Western Blot

Standard Protocal :

The SDS-PAGE gel was set for Western blot in a blotting apparatus (in pre-chilled 1X Transfer buffer) and run overnight at 28V in 4◦C. Then the membrane was blocked with BSA (5% dissolved in 1X TBST) for 2hours at room temperature. Membrane was then washed thrice with 1X TBST (3 minutes each). HRPO conjugated antibody (1:2500) was added and incubated overnight at 4◦C.Next day the membrane was washed thrice with 1X TBST (5 minutes incubation each time) and then developed in the dark room.

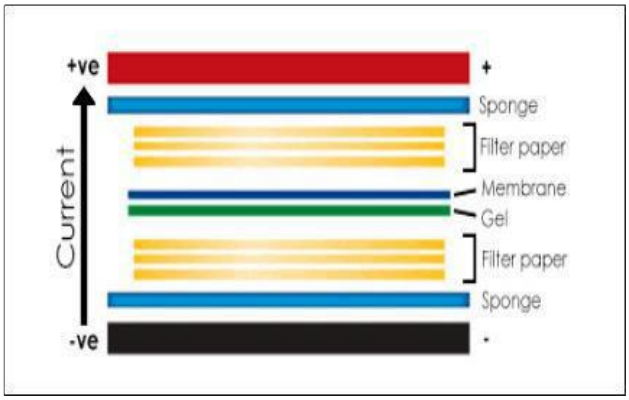


Fig 14: Schematic diagram of Western Blot Sandwich

Equal volumes of Solution A and Solution B were mixed and the membrane was incubated in this mixture. The membrane was then placed inside the developing cassette and an x-ray film was exposed to the signal (for approx. 10-15sec). The film was removed and washed in developer solution, fixed in the fixer solution and finally washed in water and thus the image was acquired.(All the steps for signal development using X-ray film was performed inside the dark room).

4.2.13 IDENTIFICATION OF PROTEIN VIA MASS SPECTROSCOPY

Mass spectrometry is a powerful analytical technique used to quantify known materials, and to identify unknown materials within a sample and to elucidate the structure and chemical properties of different molecules. The complete process involves the conversion of sample into gaseous ion with or without fragmentation which are then characterized by their mass to charge ratio (m/z) and relative abundances.

Matrix-assisted laser desorption/ionization (MALDI)- Time Of Flight(TOF) is a "soft" ionization technique that uses a laser energy absorbing matrix to create ions from large molecules with minimal fragmentation which provides microscale analysis of protein and peptides and determines molecular mass of proteins and peptides. Along with that it also identifies the expected bands of the desired proteins. The matrix(with analyte) absorbs UV/IR

light (wavelength-337nm) and heats up rapidly. A small part of matrix is evaporated along with the sample. The charged ions of various sizes are generated on the sample slide. The ions with smaller m/z value (lighter ions) and highly charged ions moves faster through the drift space until they reach the detector. Consequently the time of ion flight differs according to mass-to-charge ratio of ions. Basing upon this difference in the time of flight, peaks of various intensities are generated and thus analysed.

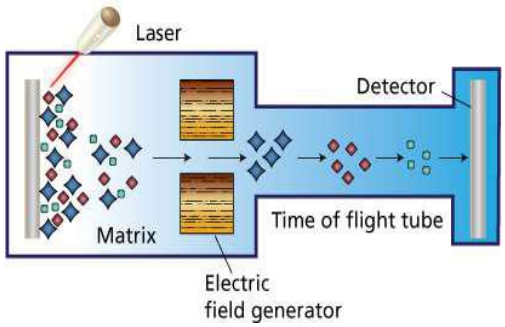


Fig 15: Schematic diagram of MALDI - TOF

Standard Protocal :

A) SAMPLE PREPARATION:

Gel slices of particular bands were rinsed with 25mM Ammonium Bicarbonate (ABC) thrice with 5 minutes incubation at each step followed by dehydration with 100µl of Solution A(1:1 mixture of CAN and ABC). Supernatant was removed and 100µl of 25mM ABC was added and incubated for 5 minutes on shaker at room temperature. The supernatant was removed again and dehydrated with 100µl of solution for 5 minutes on shaker at room temperature. The above were repeated again for one more time. Again, dehydration was done by Solvent A and ABC.(For complete dehydration Speed Vac was done for 15mins). Then, the sample was treated with trypsin for digestion. The solution was then aspirated and dispersed to elute the peptides from gel and was vortexed for 30 minutes followed by spin at 500 rpm for 1 minute

and the gel was chopped properly and the supernatant was collected carefully. Finally, extraction was done with CAN and TFA then vortexing was done and again supernatant was collected. Then, the sample was lyophilized and stored at -20℃.

B) The sample stored at -20℃ was mixed with suitable matrix and was detected and analyzed by the detector of MALDI-TOF.

CHAPTER 5: RESULTS

A) In-vitro Analysis and Comparison of H-5A SufB mutant protein and WT SufB protein splicing by resolution through 4-10% gradient SDS-PAGE at different temperatures:

1) Protein Splicing of H-5A and WT at 16˚C :

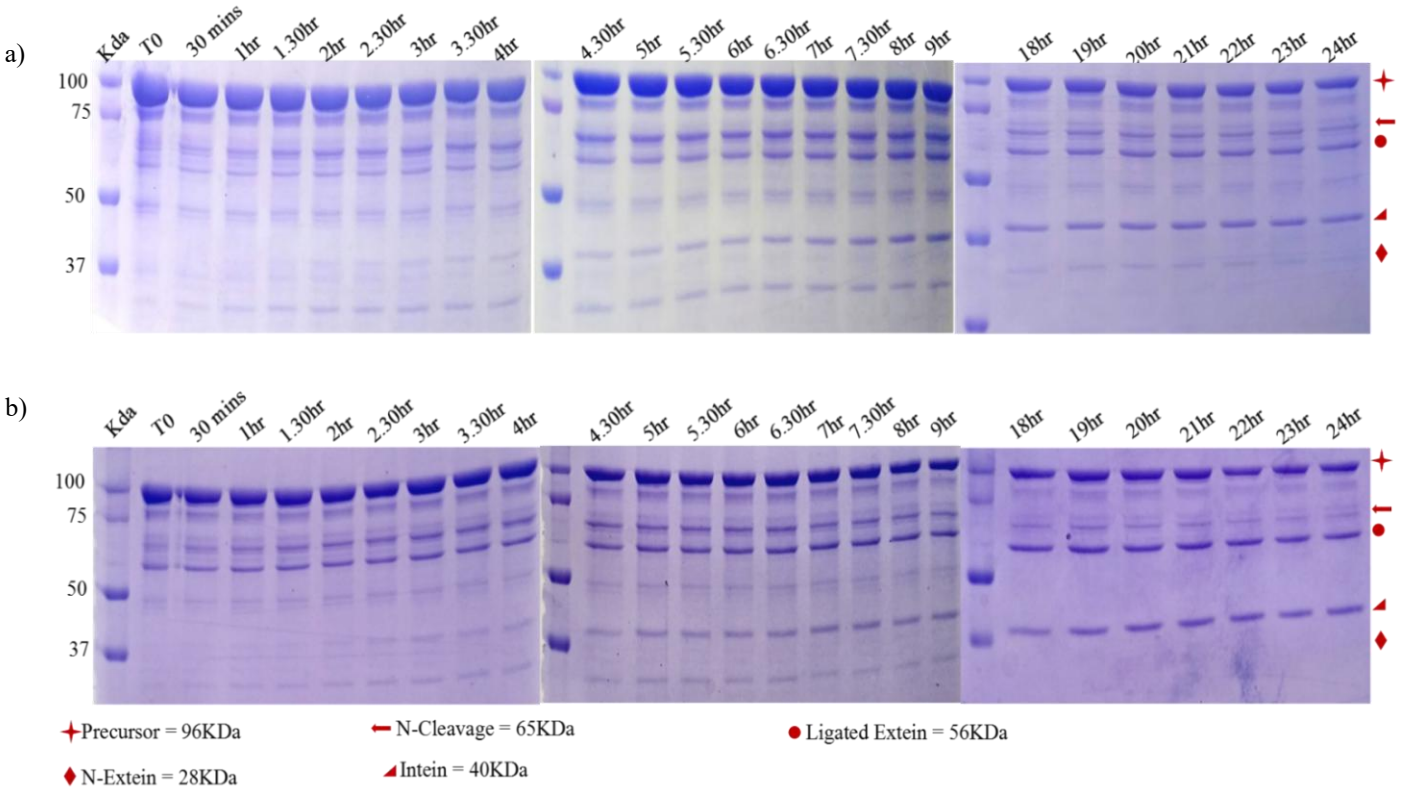


Fig 16: Effect of H-5A mutant on SufB splicing at 16˚C a) H-5A(Histidine at -5 position is mutated to Alanine) and b) WT (Wild Type) were purified and allowed to renature in *in-vitro* condition over different time periods (T_0 to 24hr). The splicing products were resolved via 4-10% SDS-PAGE. Precursor bands (96KDa) , N-Cleavage bands (65KDa) , Ligated Extein Bands (56KDa) , Intein bands (40KDa) , N-Extein bands (28KDa).

2) Protein Splicing of H-5A and WT at 20℃

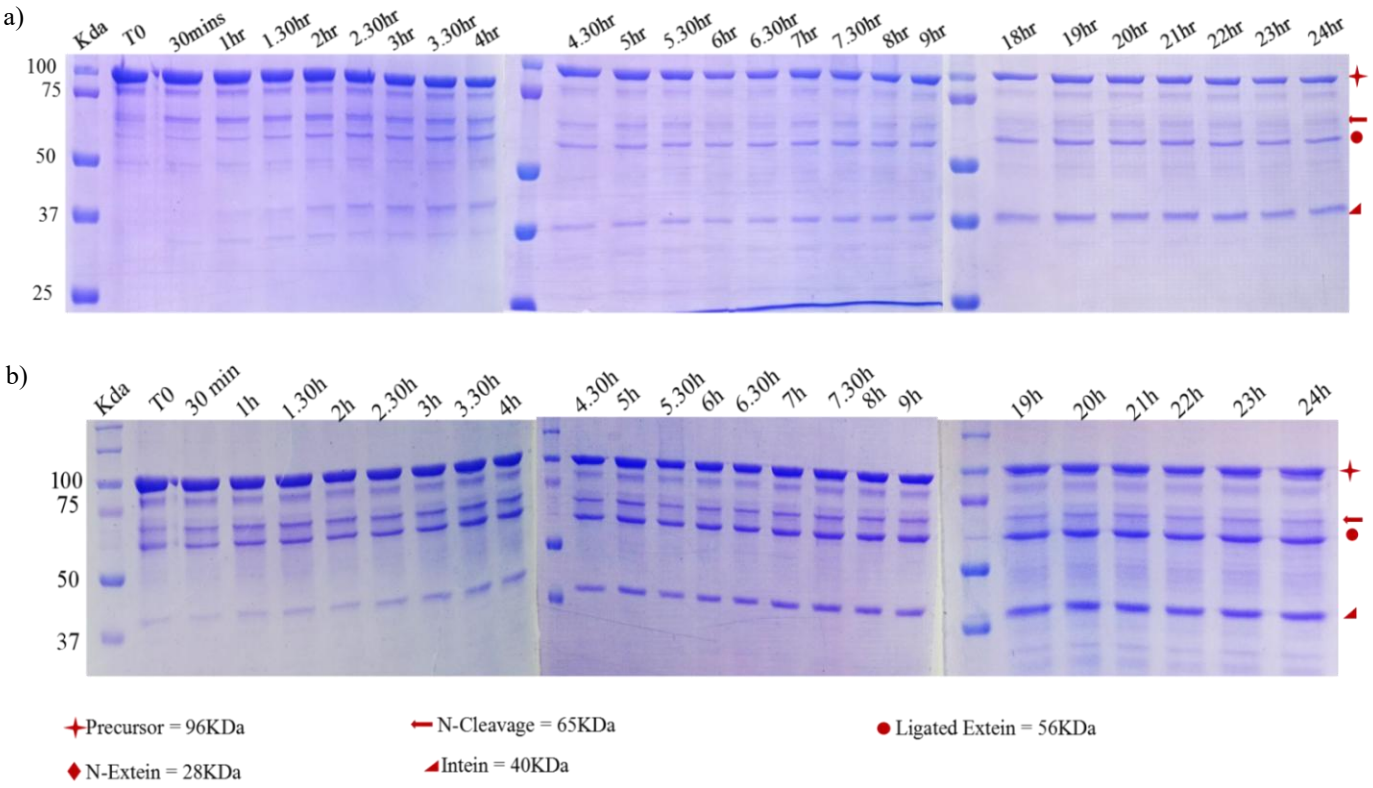


Fig 17: Effect of H-5A mutant on SufB splicing at 20˚C a) H-5A(Histidine at -5 position is mutated to Alanine) and b) WT (Wild Type) were purified and allowed to renature in *in-vitro* condition over different time periods (T_0 to 24hr). The splicing products were resolved via 4-10% SDS-PAGE. Precursor bands (96KDa) , N-Cleavage bands (65KDa) , Ligated Extein Bands (56KDa) , Intein bands (40KDa) , N-Extein bands (28KDa).

3) Protein Splicing of H-5A and WT at 25℃

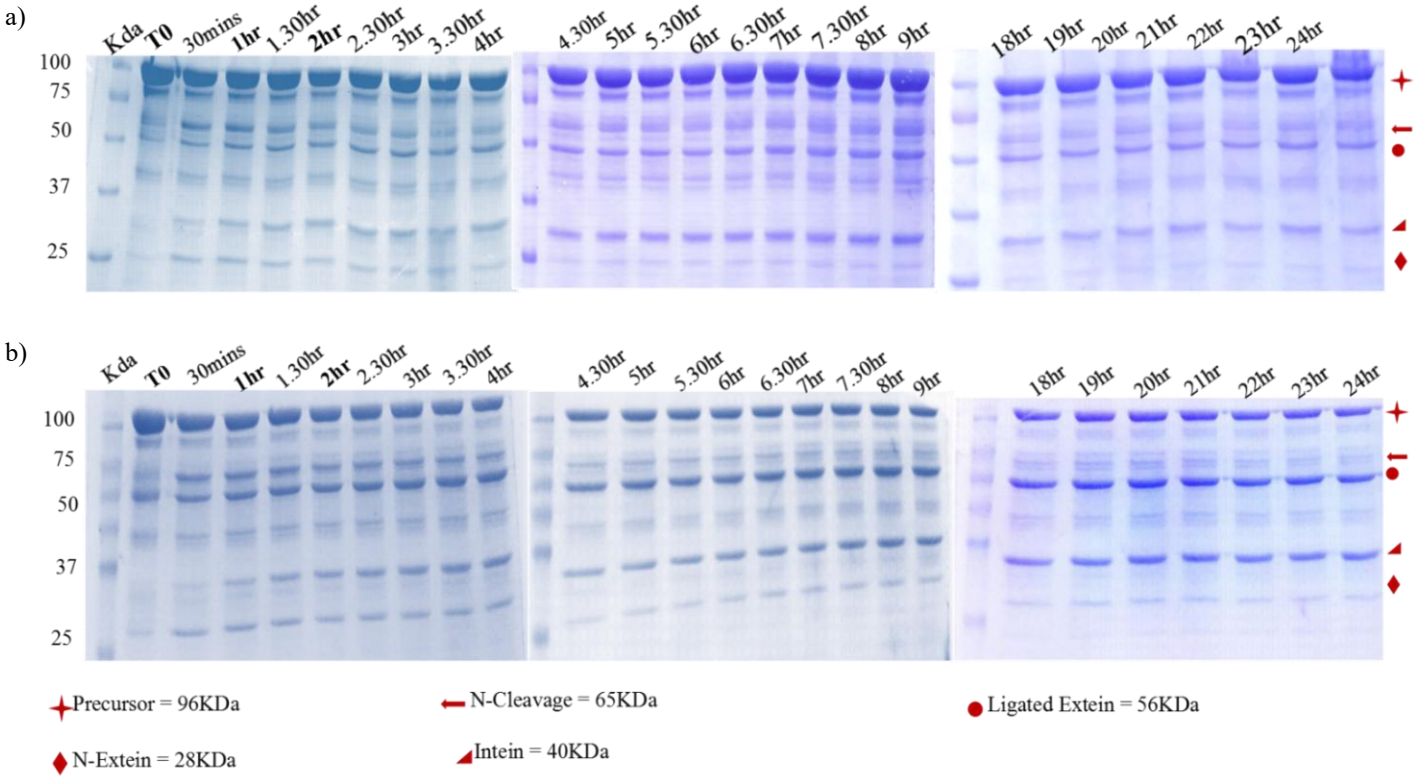


Fig 18: Effect of H-5A mutant on SufB splicing at 25˚C a) H-5A(Histidine at -5 position is mutated to Alanine) and b) WT (Wild Type) were purified and allowed to renature in *in-vitro* condition over different time periods (T_0 to 24hr). The splicing products were resolved via 4-10% SDS-PAGE. Precursor bands (96KDa) , N-Cleavage bands (65KDa) , Ligated Extein Bands (56KDa) , Intein bands (40KDa) , N-Extein bands (28KDa).

4) Protein Splicing of H-5A and WT at 30℃

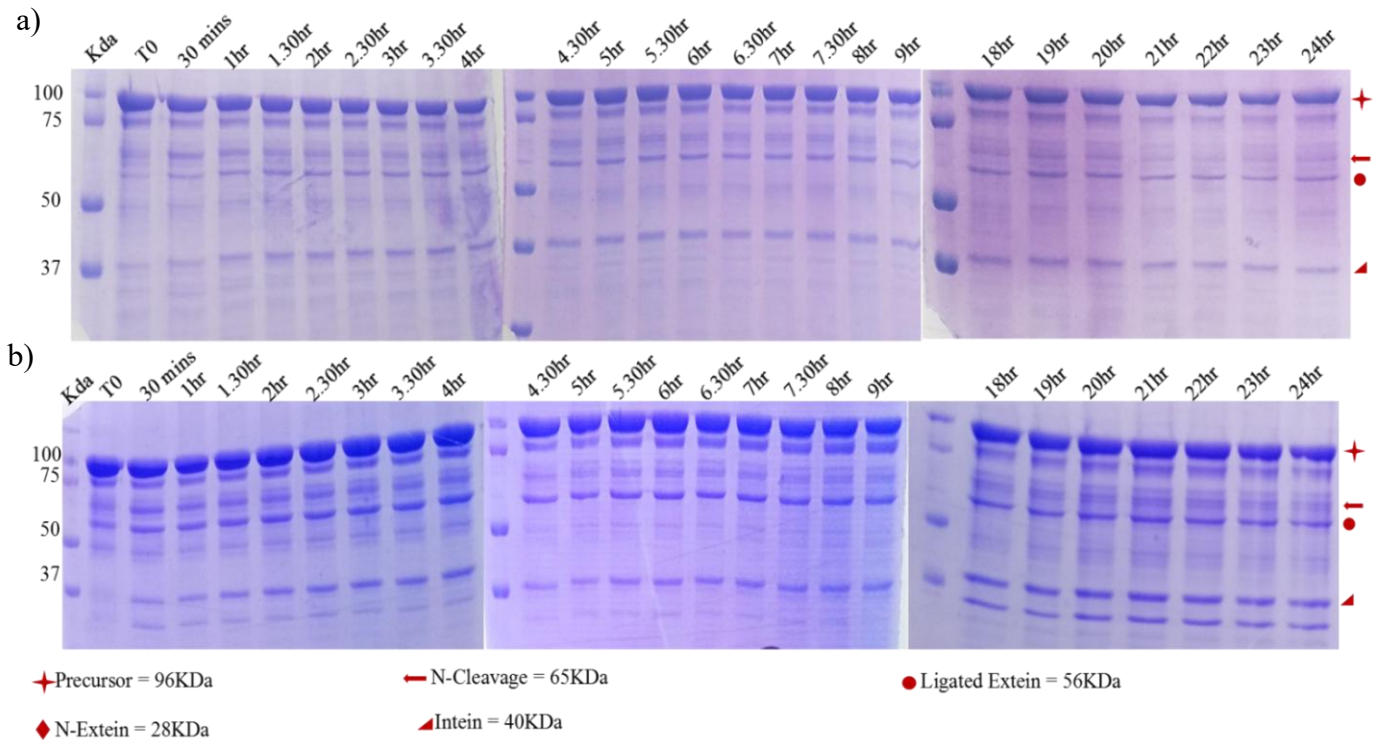


Fig 19: Effect of H-5A mutant on SufB splicing at 30˚C a) H-5A(Histidine at -5 position is mutated to Alanine) and b) WT (Wild Type) were purified and allowed to renature in *in-vitro* condition over different time periods (T_0 to 24hr). The splicing products were resolved via 4-10% SDS-PAGE. Precursor bands (96KDa) , N-Cleavage bands (65KDa) , Ligated Extein Bands (56KDa) , Intein bands (40KDa) , N-Extein bands (28KDa).

5) Protein Splicing of H-5A and WT at 37℃

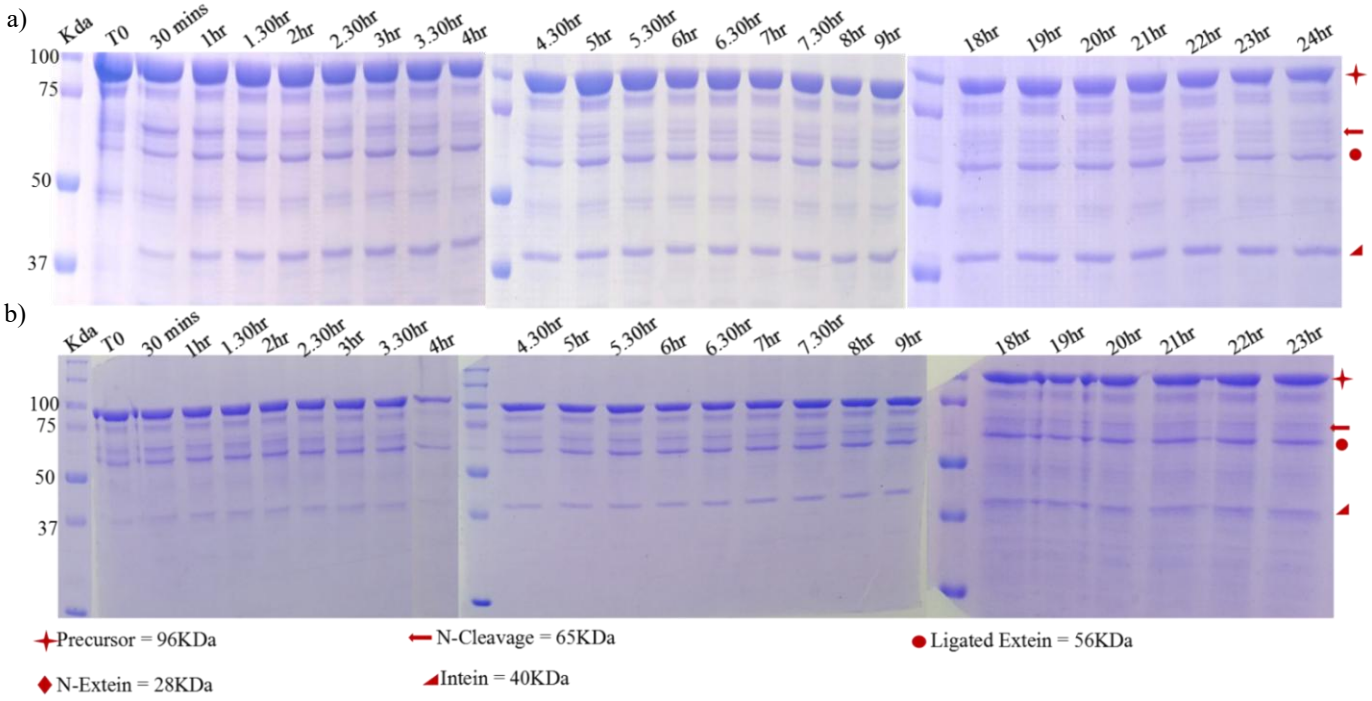


Fig 20: Effect of H-5A mutant on SufB splicing at 37˚C a) H-5A(Histidine at -5 position is mutated to Alanine) and b) WT (Wild Type) were purified and allowed to renature in *in-vitro* condition over different time periods (T_0 to 24hr). The splicing products were resolved via 4-10% SDS-PAGE. Precursor bands (96KDa) , N-Cleavage bands (65KDa) , Ligated Extein Bands (56KDa) , Intein bands (40KDa) , N-Extein bands (28KDa).

B) Confirmation of Protein Bands via Western Blot :

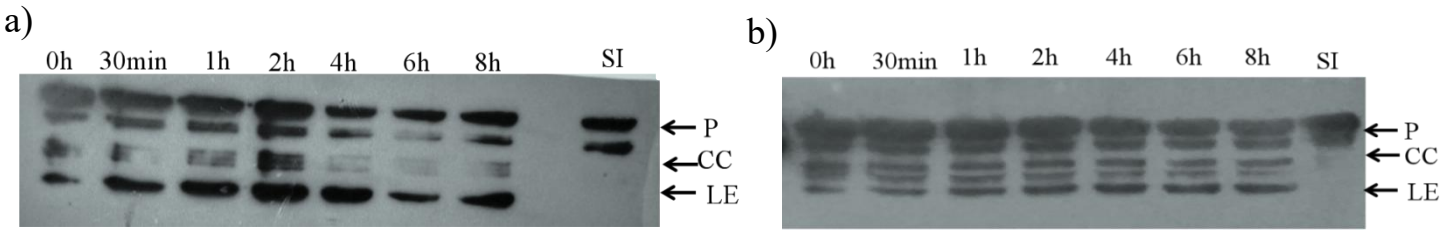


Fig 21:Confirmation of identity of splicing products via Immuno-blotting assay, where the splicing products were detected via Anti-His Monoclonal antibody. In a) WT and b)H-5A splicing products, Precursor (P) , C-terminal cleavage (CC) and Ligated Extein (LE) were blotted. SI (Splicing inactive) was considered as a negative control for splicing reaction.

C) Indentification of Protein via MALDI-TOF/TOF:

AB SCIEX\ProteinPilot Data\Results\All\ILS Proteomics2-ILS-153-27052017-Job Run-9 ncbi nr mtb all ms\ms.mrf]

Protein Summary

Spot ID	Best Protein Accession	Best Protein Mass	Best Protein Score	
A15	wti_WP_072520643.1|	40546	81	Fe-S cluster assembly protein SufB [Mycobacterium tuberculosis]
A16	Wtp_WP_072520643.1|	94852	480	Fe-S cluster assembly protein SufB [Mycobacterium tuberculosis]
A21	wtNC_WP_072520643.1|	65258	362	Fe-S cluster assembly protein SufB [Mycobacterium tuberculosis]
A22	wtLE_WP_072520643.1|	53805	203	Fe-S cluster assembly protein SufB [Mycobacterium tuberculosis]
B2	wtNE_WP_072520643.1|	27957	207	Fe-S cluster assembly protein SufB [Mycobacterium tuberculosis]

Table 5: MALDI-TOF MS/MS protein spot identification of *Mtu* FL SufB protein against the NCBI nr protein database. For each search, best protein score (significance threshold >25) and corresponding accession number, protein name and species are reported.

D) Kinetic analysis of splicing efficiency of WT Mtu SufB protein and H-5A SufB mutant protein :

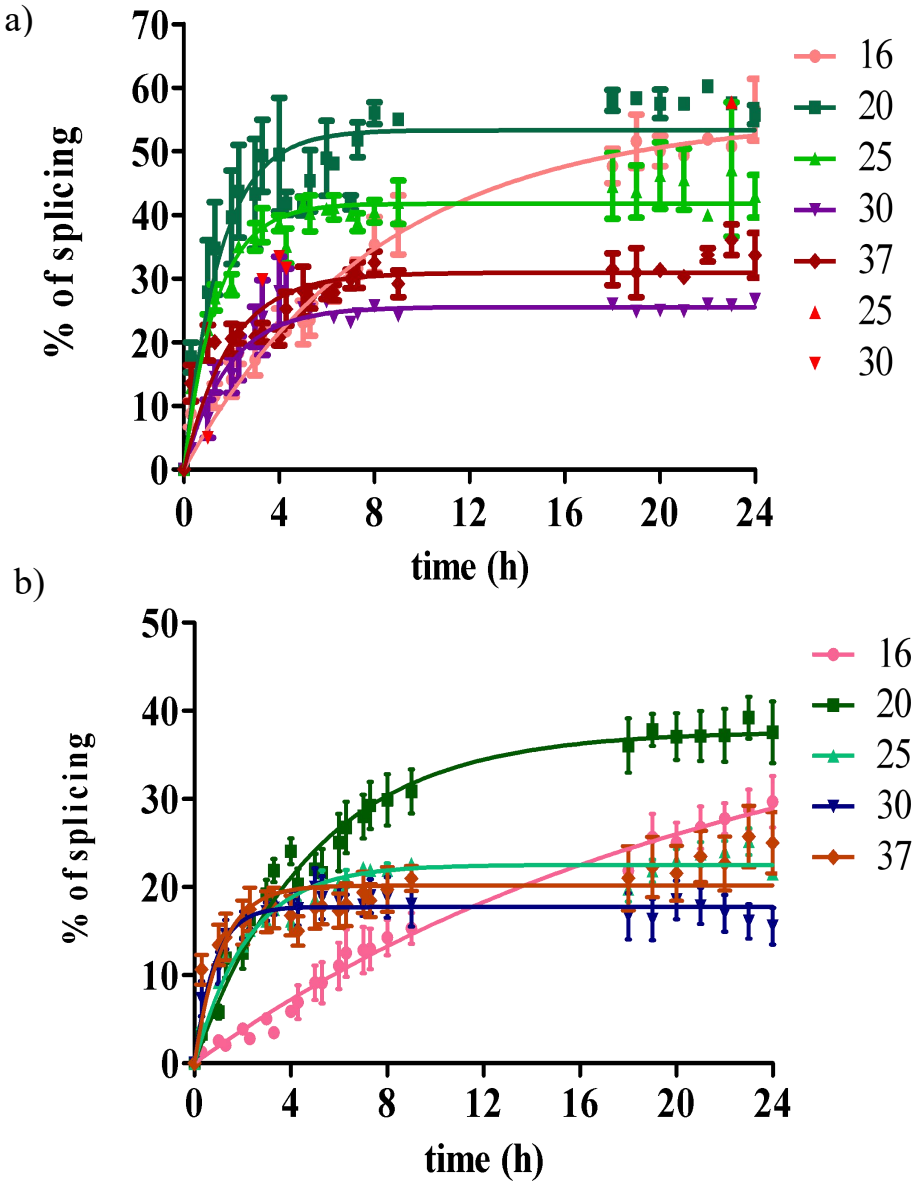


Fig 22: Kinetic Analysis of protein splicing at different temperatures(16˚C,20˚C,25˚C,30˚C and 37˚C) in a) WT and b) H-5A SufB mutant. The splicing products were calculated over different time periods and curve was fitted in a pseudo-first order reaction with equation $P = P_0 (1-e^{-kt})$. All the experiments were performed in triplicates and error bar represents SEM.

	16°C	20°C	25°C	30°C	37°C
V_{max}	55.37 ± 1.6	53.33 ± 1.3	41.79 ± 0.68	25.52 ± 0.37	30.91 ± 0.86
K	$1.2\pm0.008\times10^{-1}$ hr^{-1}	$6.5\pm0.08\times10^{-1}$ hr^{-1}	$7.2\pm0.06\times10^{-1}$ hr^{-1}	$5.4\pm0.03\times10^{-1}hr^{-1}$	$5.1\pm0.06\times10^{-1}$ hr^{-1}
$t_{1/2}$ (half-life)	5.640 hr	1.054 hr	0.9552 hr	1.281 hr	1.339 hr

Table 6: This table shows V_{max} , rate constant(K) and half-life ($t_{1/2}$) of splicing reaction of WT at different temperatures calculated via Kinetic Analysis.

	16°C	20°C	25°C	30°C	37°C
V_{max}	43.65 ± 3.527	37.68 ± 0.7275	22.51 ± 0.4184	17.75 ± 0.2640	20.18 ± 0.6937
K	$4.5\pm0.005\times10^{-2}$ hr^{-1}	$2.0\pm0.010\times10^{-1}$ hr^{-1}	$4.4\pm0.033\times10^{-1}$ hr^{-1}	$1.20\pm0.133hr^{-1}$	$9.0\pm0.195\times10^{-1}$ hr^{-1}
$t_{1/2}$ (half-life)	14.03 hr	3.617 hr	1.549 hr	0.6187 hr	0.7196 hr

Table 7: This table shows V_{max} , rate constant(K) and half-life ($t_{1/2}$) of splicing reaction of H-5A at different temperatures calculated via Kinetic Analysis.

CHAPTER 6: DISCUSSION

In *Mycobacterium tuberculosis*(*Mtu*) intein splicing is essential for the functionality of SufB protein. Intein splicing is a self-catalyzed process that doesn't require any external co-factor or energy, rather it depends on a proper folding of protein and formation of catalytic cleft. Intein contains many conserved residues that regulate the splicing mechanism. It has been previously reported that some distal residues of Intein~Extein junction also play a critical role in splicing reaction but their detailed mechanism is still unkown. In *Mtu,* SufB forms a platform for Fe-S cluster synthesis and it contains many critical residues like Histidine, cysteine that interact with the complex. In our SufB protein, there is a metal chelating histidine residue present near N-extein~intein junction. As it is already known that metal chelating residues are the targeted site for Fe-S cluster forming protein so we assumed that this Histidine may also regulate the SufB splicing. In order to depict the exact role of this histidine (248 residue), which is present at -5 position of N-extein it has been mutated to Alanine.

In-vitro splicing assay was done at different temperatures (16˚C, 20˚C, 25˚C, 30˚C and 37˚C) over a time period of 24hours with 30mins interval to check at which temperature splicing is optimum for *Mtu* FL SufB in compared to H-5A(histidine at -5 position is mutated to alanine) SufB mutant. After resolution of splicing products at different temperatures (Fig No.:-16-20), it was observed that Precursor accumulation for H-5A SufB mutant was higher in compared to WT at all temperatures. A comparable difference was also detected for the splicing products (Ligated extein and intein bands) between WT and H-5A SufB mutant protein at all temperatures. The identity of the splicing and cleavage products were confirmed via Western blot by using Anti-His antibody. Further the identity was confirmed via MALDI-TOF/TOF analysis (Fig No.:-21) by giving the best protein score, protein accession number and species name. The densitometry analysis was done followed by detailed Kinetic analysis where the values were fitted in the curve

(Fig No.:- 22) by using Pseudo first order kinetics equation $Y=Y_0+(Plateau-Y_0)(1-e^{-kt})$ where Y is the % of splicing at a specific time, Y_0 is the splicing value at T_0 , Plateau is the maximum % of splicing value after attaining the saturation point, k is the rate constant of the reaction and t is the time (in hour).

Both WT and H-5A SufB mutant follow pseudo first order kinetic reaction and there is a significant difference in % of splicing was observed at all the temperatures. From the detailed kinetic analysis (Table no.:- 6) for WT at different temperatures, it was observed that WT has maximum splicing efficiency at 20˚C as it attends highest splicing 53.33-54.6% having $t_{1/2}$ (half-life) 1.054 hr. The rate constant K- $6.5\pm0.08\times10^{-1}hr^{-1}$ was higher in compared to other temperatures. As we want to check the role of Histidine, we compared the splicing efficiency of H-5A mutant with WT. So we considered splicing reaction of H-5A at 20˚C. From kinetic analysis of H-5A SufB mutant (Table no.:-7), we found that there is approximately 3fold reduction in % of splicing with Vmax- 37.68 ± 0.72 and half-life $t_{1/2}$- 3.617 hr, which is much higher in compared to half-life of WT at 20˚C. Also the rate constant for the reaction was K- $2.0\pm0.01\times10^{-1}$ hr^{-1} which was also less than WT.

Hence we conclude that Histidine at -5 position play a vital role in splicing reaction. As this residue resides near N-Extein~Intein junction so it may regulate the first step of the splicing reaction. The role of histidine can be further be analyzed by detailed N-cleavage assay, as it may target the first step of protein splicing. Detailed structural analysis can also confirm the role of Histidine in catalytic cleft formation as well as in splicing reaction.

FUTURE GOAL:

As our assumption we observed that the efficiency of splicing reaction is decreased when Histidine at -5 position is mutated to Alanine in compared to WT. We can say that, as this residue resides near N-Extein~Intein junction so it may regulate the first step of the splicing reaction. The role of histidine can be further analyzed by detailed N-cleavage assay, as it may target the first residue(C1) of intein regulating the first step of protein splicing. Detailed structural analysis can also confirm the role of Histidine in catalytic cleft formation as well as in splicing reaction.

REFERENCES :

[1] Sushil Ahmed et.al. , Pathogenesis, Immunology, and Diagnosis of Latent Mycobacterium tuberculosis Infection, *Clinical Development Immunology.,(*2011)

[2] Comas,I.; Gagneux, S., The past and future of Tuberculosis research. *Plos pathogen.,* (2000)

[3] Huet, G., Daffé, M. & Saves, I. Identification of the Mycobacterium tuberculosis SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly: Evidence for its implication in the pathogen's survival. *J. Bacteriol.* **187,** 6137– 6146 (2005).

[4] Huet, G., Castaing, J. P., Fournier, D., Daffé, M. & Saves, I. Protein splicing of SufB is crucial for the functionality of the Mycobacterium tuberculosis SUF machinery. *J. Bacteriol.* **188,** 3412–3414 (2006).

[5] Jeans Pieters .,*Mycobacterium tuberculosis* and the Macrophage: Maintaining a Balance. *Cell Press*.,(2008).

[6] Gogarten, J.P.; Senejani,A.G,; Zhaxybayeva, O.; Olendzenski.; Inteins : Structure, Function, and Evolution. Annual Review of Microbiology.,(2002)

[7] Ayala-Castro, C.; Saini, A.; Outten, F. W., Fe-S Cluster Assembly Pathways in Bacteria. *Microbiology and Molecular Biology Reviews.*(2008).

[8] Nachin, L.; El Hassouni, M.; Loiseau, L.; Expert, D.; Barras, F., SoxR-Dependent Response to Oxidative Stress and Virulence of *Erwinia chrysanthemi*: The Key Role of SufC, an Orphan ABC ATPase. *Molecular Microbiology,* 39, 960-972. 9.(2001)

[9] EikiYuda et.al.;Mapping the key residues of SufB and SufD essential for biosynthesis of iron-sulfur clusters.,*Scientific Reports,*(2017).

[10] Huet et.al., Identification of the *Mycobacterium tuberculosis* SUF Machinery as the Exclusive Mycobacterial System of [Fe-S] Cluster Assembly: Evidence for Its Implication in the Pathogen's Survival.,*Journal of Bacteriology* 187(17):6137-46.,(2005).

[11] Huet, G.; Daffe, M.; Saves, I., Identification of the *Mycobacterium tuberculosis* SUF machinery as the Exclusive Mycobacterial System of Fe-S Cluster Assembly: Evidence for its Implication in the Pathogen's Survival. *Journal of Bacteriology.*,(2005)

[12] Huet, G.; Castaing, J. P.; Fournier, D.; Daffe, M.; Saves, I., Protein Splicing of SufB is Crucial for the Functionality of the *Mycobacterium tuberculosis* SUF Machinery., *Journal of Bacteriology (*2006)

[13] Derbyshire, V.; Wood, D. W.; Wu, W.; Dansereau, J. T.; Dalgaard, J. Z.; Belfort,Genetic Definition of a Protein-splicing Domain: Functional Mini-inteins Support Structure Predictions and a Model for Intein Evolution. *Proceedings of the National Academy of Sciences of the United States of America* .,(1997)

[14] Paulus, H., Inteins As Enzymes. *Bioorganic Chemistry* 29, 119–129.(2001)

[15] KV Mills, Protein Splicing:How inteins Escape from Precursor Proteins, Minireview, JBC Department of Biochemistry.(2014)

[16] Paulus, H., Inteins As Enzymes. *Bioorganic Chemistry*, 29, 119–129.,(2011)

[17] Perler, F. B., InBase: The Intein Database. *Nucleic Acids Research.* (2002)

[18] Callahan, B. P.; Topilina, N. I.; Stanger, M. J.; Van Roey, P.; Belfort., Structure of Catalytically Competent Intein Caught in a Redox Trap with Functional and Evolutionary Implications. Nature Structural & Molecular Biology, 18, 630-633.,(2011).

[19] Mujika et.al., Unveiling the catalytic role of B-Block Histidine in the N-S Acyl shift step of Protein Splicing.Journal of Physical chemistry,(2017).

[20] Chen et.al., Protein splicing in the absence of Penultimate Histidine

[21] Henry Paulus, (2000) Inteins as Enzymes., Boston Biomedical Research Institute, Bioorganic Chemistry 29, 119–129,(2000).

[22] Shaorong Chong, Protein Splicing Involving the *Saccharomyces cerevisiae* VMA Intein, New England Biolabs, Beverly, The Journal of Biochemistry, (1996).

[23] Gaelle Huet et.al, Protein splicing of *SufB* is crucial for functionality of *Mycobacterium Tuberculosis* Suf Machinery, Journal of Bacteriology. (May 2006)

[24] Henry Paulus (2000), Intein as Enzymes (minireview),(2000).

[25] Francine B.perler,Protein Splicing Mechanism and Applications Critical Review.(July 2005).

[26] David W.Wood, Dynamic differences between active and inactive intein,European Journal of Medical Chemistry.(May 2014).

[27] Qin WU et.al, Conserved residues that modulate protein trans-splicing of Npu DnaE split intein, Biochemical journal(April 2014).

[28] Francine B.Plerler,An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile , The EMBO journal Vol.19 N0.18 pp.5019-5026.(2000).

[29] Lana Saleh.,Branched intermediate formation is the slowest step in the protein splicing reaction of the Ala1 KlbA intein from *Methanococcus jannaschii.,*American Chemical Society.(October 2011).

[30] Lixin Chen, Francine B.Perler ,Protein splicing in the Absence of an intein Penultimate Histidine, The journal of Biochemistry,(July 2000).

[31] Zhenming Du,Highly conserved Histidine plays a dual catalytic role in protein splicing: a pKa shift mechanism, American Chemical Society., (August 2009).

[32] Natalya l.Topilina , Post-translational environmental switch of RadA activity by extein-intein interactions in protein splicing, Nucleic Acids Research, (June 2015).

[33] K.V.Mills et.al., Protein splicing of *Pyrococcus abyssi* Intein with a C-terminal Glutamine, (January 2004).

APPENDIX :

1) Buffer Compositions :

A) Lysis Buffer (100ml) [pH 7.5] :

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

Volume make upto 100ml with distilled water.

B) Renaturation Buffer(100ml) [pH 7.5] :

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

L-Arginine - 8.7gm

Volume make-up to 100ml with distilled water.

C) 8M Urea (100ml) :

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

Urea - 48gm

Volume make-up to 100ml with distilled water.

D) Binding Buffer (100ml) [pH 7.4]:

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

8M Urea - 48.04gm

Imidazole(20mM) - 0.136 gm.

Volume make-up to 100ml with distilled water.

E) Washing Buffer (100ml) :

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

8M Urea - 48.04gm

Imidazole(40mM)- 0.272 gm.

Volume make-up to 100ml with distilled water.

F) Elution Buffer (100ml) :

NaH_2PO_4 - 0.238gm

Na_2HPO_4 - 0.283gm

NaCl - 2.92gm

8M Urea - 48.04gm

Imidazole(500mM)- 3.4gm.

Volume make-up to 100ml with distilled water.

G) Running Buffer (10X) [1L] :

Glycine – 144gm

Tris base – 30.2gm

Volume make-up to 1 liter with distilled water.

For preparation of 1 liter of 1X Running buffer, mix 100ml of 10X Running buffer,10ml of 10% SDS and 890ml of distilled water.

H) Transfer Buffer (10X) [1litre]:

Glycine – 144gm

Tris base – 30.2gm

Volume make-up to 1 liter with distilled water.

For preparation of 1 liter of 1X Transfer buffer, mix 100ml of 10X Transfer

buffer,200ml of methanol and 800ml of distilled water.

2) **REAGENTS :**

a) Acrylamide-Bisacrylamide solution (100ml) :

Acrylamide – 29.2gm

Bis-acrylamide – 0.8gm

Distilled water – 60ml

Stir continuously for 30 minutes and then adjust volume to 100ml with distilled water. Filter and store at 4°C in amber colored bottle.

b) Loading dye (5X) [50ml] :

Bromophenol blue – 10mg

Tris (pH 6.8) – 2ml

SDS – 1ml

Glycerol – 5ml

Volume make-up with Tris (pH 6.8). Aliquots of 950µl are made and 50µl of β-mercaptoethanol is added to it.

c) Staining solution (1L) :

Coomassie Brilliant Blue R-250 – 1.25gm

Methanol – 500ml

Glacial acetic acid – 400ml

Distilled water – 100ml

d) De-staining solution (100ml) :

Distilled water – 50ml

Methanol – 40ml

Acetic acid – 10ml

(Mix water, methanol and acetic acid in the ratio of 5:4:1).

e) TBS (10X) :

Tris base (pH 7.6) – 24gm

NaCl – 88gm

Distilled water – 900ml

pH 7.6 is adjusted and then volume is made up to 1L.

For preparing 100ml of 1X TBST, add 10 ml of 10X TBS to 90ml of distilled water and to it add 0.1% Tween-20.

f) Bradford Reagent (1L) :

Coomassie Brilliant Blue G-250 – 50mg

Methanol – 50ml

Dissolve and to it add 100ml of 85% Ortho-phosphoric acid and 500ml of distilled water, and mix properly. Filter to remove any precipitate, and add 350ml of distilled water. Store at 4 C° in dark.

g) 10% APS :

0.1gm of APS

Autoclaved Water : 1ml.

h) Blocking Solution :

3% BSA in 1X TBST or 5% Skim Milk in 1X TBST.

i) ECl Solution :

Solution 'A' :

250mM Luminol(dissolved in DMSO) - 1ml

90mM p-Coumaric Acid Solution - 0.44ml

1M Tris-HCl (pH 8.5) - 10ml

Volume make-up to 100ml with distilled water.

Solution 'B' :

30% Hydrogen Peroxide(H_2O_2) [30%] - 64µl

1M Tris - HCl (pH 8.5) - 10ml

Volume make-up to 100ml with distilled water.

3) Media preparation:

A) Luria-Bertani broth (1L) [pH 7.0] :

Tryptone – 10gm

Yeast extract – 5gm

NaCl – 10 gm

Volume adjusted to 1L with distilled water.

B) Luria-Bertani agar (1L) [pH 7.0]

Tryptone – 10gm

Yeast extract – 5gm

NaCl – 10 gm

Agar – 15gm

Volume adjusted to 1L with distilled water