

PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF HUMAN BLOOM SYNDROME (RECQ FAMILY) HELICASE.

Submitted in partial fulfilment of the requirements for the degree of

Master of Biotechnology

By

DEBASISH ROY

Reg No:09MSB022



VIT

UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)

Vellore - 632 014, Tamil Nadu, India

June, 2011

DECLARATION

I hereby declare that the thesis entitled “PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF BLOOM SYNDROME (REC Q) HELICASE” submitted by me, for the award of the degree of *Master of Biotechnology* to VIT University is a record of bonafide work carried out by me under the supervision of Dr. Mihály Kovács.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place:

Signature of the Candidate:

CERTIFICATE

This is to certify that the thesis entitled “PURIFICATION AND BIOCHEMICAL CHARACTERIZATION OF BLOOM SYNDROME (REC Q) HELICASE” submitted by DEBASISH ROY (School of Biosciences and Technology) VIT University, for the award of the degree of *Master of Sciences in Biotechnology*, is a record of bonafide work carried out by him under my supervision, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The thesis fulfils the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Internal Guide

Divisional Leader

Internal Examiner

External Examiner

Abstract

DNA helicases are involved in the fundamental processes inside a cell, namely the replication, recombination, and repair. One of the most widely studied of the helicases are the RecQ family of helicases, which have involvement in the repair processes inside the cell by the Homologous Recombination Pathways. Mutation in these helicases thereby is associated with a number of disorders, predisposition to cancers and premature ageing. The functionality of these enzymes are dependent upon their translocation along the ssDNA. Although we still don't have the overall picture over the translocation mechanism of these helicase due to lack of structural information. The work produced in this thesis is basically carried out in a Human Isoform of the RecQ helicases. The broader aim of this project is to have crystal structure of the BLM and *E.coli* RecQ helicases in the DNA bound form. This will give us the actual picture of the translocation mechanism and help us delve deep into the involvement of these enzymes in the complex repair processes.

In this study a truncated version of the wild type BLM protein (BLM⁶⁴²⁻¹¹⁹¹, 549 amino acids, pI= 9.6, 63.5 kDa), which lacked the HRDC domain is used. Purification trials for obtaining 99% pure protein is attempted, for which about 95% purity was assessed. Translocation assays revealed the processivity, binding site size and other important translocation parameters which were in lieu with the earlier results.

The truncated version was found more active in comparison to the wild type and it retains all the necessary translocation parameters. With this study and further experiments, protein of 99% purity will be obtained and crystallization experiments can be carried out.

Key words: BLM, translocation, homologous recombination, crystallization.

Acknowledgements

First I would like to express my heartfelt thanks and gratitude towards my parents and family whose support has been an enormous strength for me throughout the period of my work. Without their constant motivation I couldn't have done this. Thank you Ma and Baba for everything.

It gives me immense pleasure and gratitude to express my heartfelt thanks to Dr. Mihály Kovács, Research Associate Professor, Dept. Of Biochemistry, ELTE, Budapest, Hungary. His consistent guidance and support for my work and many things apart from it has made me learn a lot of things not only from the science but for things away from it. I gratefully acknowledge the opportunity that he gave me to work with him and exposed me to this extremely competent international atmosphere of science. My words are not enough to describe the philosophy of life and science that I got to learn from him.

I would like to take the opportunity to acknowledge Dr. Kavitha Thirumurugan, Associate Professor, School of Biosciences and Technology, Division of Biomedical Sciences, CBMR, VIT, India. She was the first person to introduce me to the field of structural biology and in fact it was practically for her correspondence I got an opportunity to work with Dr. Kovacs. Her consistent follow ups and guidance during my work helped me learn a lot. I thank her for being a great mentor to me for helping me learn the philosophy of life and science.

It gives me immense pleasure to express my gratitude to Dr. Mate Gyimesi, PDF @ mk-lab, who was the backbone for me in my work. He was the one who taught me the minute basics of the work, right from the way I handle my pipette to everything. It reminds me how he used to ask me note down every single details of the experiments to achieve ultimate perfection in it. I learnt a lot by looking him work and understood the importance of time management in research.

I would put my words of appreciation and gratitude for Gabor Harami, who has been a friend, a brother and almost everything along with Mate, for my work. He helped me learn the art of handling the columns and other many such details. I would thank you for all that you have made me learn.

It's my pleasure to acknowledge other members of the lab, Nikolett Nagy, Zsuzsa Kocsis, Kata Sarlos, Ezer Molnar and Judit Gervai who has been wonderful people around and made my stay a wonderful experience.

I owe a special thanks to Dr. V. Raju, VC, VIT University; Dr. Seenivasan. R, Professor,SBST; Mr. Subbaji Rao, Assistant Director, IR office, VIT; Mr. James

Osborne, IR Office, VIT, and most importantly to Dr. G. Vishwanathan, Chancellor, VIT University.

And finally I would like to express my heartfelt thanks to Tanvira, who has been a consistent support for me, Saptarshi, Sreeharsha and Deba Prakash whose inevitable support is inexpressible.

Place:

Date:

Debasish Roy

Abbreviations

ADP	: Adenosine 5'- Diphosphate.
ATP	: Adenosine 5'- Triphosphate.
BLM	: Bloom's Syndrome helicase.
DTT	: Dithiothreitol
PK/LDH	: Pyruvate Kinase/ Lactate Dehydrogenase.
RecQ CT	: RecQ C Terminal.
SDS	: Sodium Dodecyl Sulphate.
SF	: Super Family
WH	: Winged Helix.
WT	: Wild Type.
NAD	: Nicotinamide Adenine Dinucleotide.
BS	: Bloom Syndrome.
RTS	: Rothmund Thompson Syndrome.
WS	: Werner's Syndrome.
PBP	: Phospahte Binding Protein
HRDC	: Human RNaseD C-terminal.
WH	: Winged Helix.
SF	: Stopped Flow
HR	: Homologous Recombination.
HP	: Heparin column.
HA	: Hydroxyapatite.

Contents

ABSTRACT.....	05
ACKNOWLEDMENT.....	06
ABBREVIATIONS.....	08
CONTENTS.....	09
Chapter1	
Introduction.....	14
Chapter2	
LiteratureReview.....	15
2.1Different Helicase Families.....	15
2.1.1. Superfamily 1.....	17
2.1.2. Superfamily 2.....	18
2.1.3. Superfamily 3.....	19
2.1.4. Superfamily 4.....	20
2.1.5. Superfamily 5.....	20
2.1.6 Superfamily 6.....	21
2.2 Properties of Helicases.....	21
2.2.1. NTP binding and hydrolysis.....	21
2.2.2. Nucleic Acid Binding.....	22
2.2.3Translocation Along Nucleic Acid Strand.....	22
2.2.4DNA Unwinding activity.....	23
2.2.5 Strand Annealing Activity.....	27
2.3. The RecQ Family of helicases.....	27
2.3.1 <i>E.coli</i> RecQ helicases.....	28
2.3.2 Yeast RecQ helicases.....	31
2.3.3 Human RecQ helicases.....	31
2.3.3.1 BLM- Bloom Syndrome Helicase.....	31
2.3.3.2 WRN-Werner Syndrome Helicase.....	33
2.3.3.3 RTS Rothmund Thompson Syndrome Helicase.....	34.
2.3.4 Model for Homologous Recombination (Role of Helicases).....	35
Chapter 3- Aims and Objectives.....	38
Chapter 4- Materials and Methods.....	40
4.1 Reagents.....	40
4.2 Nucleic Acid Substrates.....	40
4.3 Heparin.....	40

4.4 Protein Construct and Engineering.....	40
4.5 Expression and Protein Purification.....	41
4.5.1 Transformation.....	41
4.5.2 Expression.....	41
4.5.3 Purification.....	42
4.6 ATPase Assay.....	44
4.7 Transient Kinetics.....	44
Chapter 5- Results.....
5.1 Purification.....	46
5.2 Determination of Protein Concentration by Bradford method.....	50
5.3 Effect of prolonged ice storage on ATPase activity.....	51
5.4 Effect of Concentration on ATPase activity.....	52
5.5 Determination of K_{DNA} values of different DNA lengths.....	53
5.6 Determination of k_{cat} and its length dependence.....	56
5.7 Single Round Translocation Assay.....	58
Chapter 6 Discussion.....	61
6.1 BLM¹¹⁹¹ Purification.....	61
6.2 Biochemical Activities of BLM¹¹⁹¹	62
6.3 Dependence of the steady state ATP consumption on ssDNA length.....	62
6.4 Transient kinetics.....	62
Chapter 7 Conclusion.....	64
Appendix.....	65
References.....	68

List of Figures

	Page
2.1 Classification of Helicases.....	13
2.2 Catalytic cores of helicase (N and C core).....	14
2.3 (a) Bar representation of the domain structure of a typical SF1 helicase (PcrA).	
(b) N and C core (1A and 2A) insertions (1B and 2B).....	15
2.4 Space filling model highlighting the motifs.....	15
2.5 The structure of NS3 along with the motifs.....	16
2.6 The NS3:DNA complex.....	17
2.7 The SV-40 hexamer.....	17
2.8 Crystal structure of the hexameric T ₇ gp4.....	18
2.10 Models of NA binding.....	20
2.11 Inchworm model (a) Nucleotide free unbound state (Upon binding and coordination between ATP and metal ion)	
(c) ATP hydrolysis cause conformational change.....	
2.12 (a) Wedge Model (b) Torsional Model (c) Helix destabilizing Model.....	22
2.13 Schematic representation of the Human members of the RecQ helicase family.....	24
2.14 High resolution structure of the E.coli RecQ catalytic core.....	26
2.15 High resolution structure of some of the important domains of the RecQ Helicase family.....	26
2.16 (a) Chromosomes from lymphocyte of a normal person (b) Chromosomes from lymphocyte of a BS person (c) BS patient.....	28
2.17 Multiple pathways of the D loop processing by the prokaryotic and eukaryotic helicases.....	30
2.18 Werner Syndrome.....	32
4.1 pTXB3 Multi Cloning Site.....	35
5.1 Purification profile of the CH-HP Column.....	40
5.2 CM column Purification Profile.....	41
5.3 Affi-Blue Purification Profile-(1).....	42
5.4 Affi-Blue Purification Profile-(2).....	42

5.5 Hydroxyapatite Purification profile.....	43
5.6 Purification profile of the additional wash step employed before the HP elution.....	43
5.7 Concentration plot of BLM (before concentration).....	44
5.8 Concentration plot of BLM (after concentration).....	44
5.9 Concentration plot of BLM used for the Biochemical studies.....	45
5.10 Effect of Ice Storage on the ATPase activity.....	45
5.11 Effect of concentration on the ATPase activity.....	46
5.12 Titration Curves of different Oligo-dT with BLM along with their respective K_{DNA}	48
5.13 K_{DNA} vs Oligo-dT length.....	49
5.14 k_{cat} vs Oligo-dT length.....	
5.15 (a) Fluorescence traces of the Pi (0-6 μ M, bottom to top) vs time (b) Amplitude of the fits (of the traces in (a)) vs respective Pi concentrations	56
5.16 Transient kinetics of the Pi production from ATP upon mixing of the 50nM BLM ¹¹⁹¹ plus dT 7(50 μ M) & dT 15-90 (2.5 μ M) with 500 μ M of ATP plus 4mg/ml Heparin in a stopped flow monitored by MDCC-PBP fluorescence (3 μ M in all syringes.....	56

List of Tables

Table 1: Oligo-dT length and their respective K_{DNA} values.....	51
Table 2: Oligo-dT length and their respective k_{cat} values.....	54

Chapter 1

Introduction

The code of life has been inscribed in the stable and simple double helical structure, the DNA, which resides inside the cell in a highly organised manner in the chromosomes. But we need to unlock this script so as to follow the instruction that is there in it, in order to perform the fundamental processes, like the replication, recombination etc. For the unlocking of this highly stable double helix we need enzymes which can separate the duplex so as to access the information in the script. This strand separation function is catalysed by the DNA helicases.

Helicases are nucleic acid depending ATPases which are capable of unwinding the duplex DNA or RNA structures, and are considered as molecular motors as they use the free energy of the NTP hydrolysis or binding to translocate along the double or single stranded DNA (*Singleton et al 2007*). They are directional enzymes and can be functionally defined as 3'-5' or 5'-3' helicases, both of these hydrolyses ATP to couple the translocation along the given strand of DNA. The directionality of the helicases is defined by the orientation of the strand on which it translocates. RecQ family of helicases, one of the most highly studied groups, belongs to the SF2 Superfamily and is involved in double stranded DNA break repair, by Homologous Recombination pathway, interacting with many protein partners involved in this fundamental pathway. Any mutations in this important class of protein can lead to several diseases like the Blooms, Werner and Rothmund Thompson Disorders, which are generally associated with various malfunctions and predispositions to cancer.

Though extensively studied but, the molecular mechanisms, key to understand the function of these enzymes are not known completely and they are slowly being deciphered. Crystal structures of these proteins in the DNA bound form along with the biochemical analysis, will give us the precise understanding of the actual blueprint of the function of these proteins.

\

Chapter 2

Literature Review

2.1. Different helicase families

Classification of the different Helicase Families

Helicases can be classified based on the sequence homologies between helicases from various organisms. The directionality of the enzyme is determined by the type of DNA or RNA it translocates (3'-5' or 5'-3'). Originally the helicases were classified (Gorbalenya & Koonin, 1993) based on the seven conserved "signature" motifs, but later several other motifs were found specific to certain superfamilies and so presently helicases are classified into 6 superfamilies. The six superfamilies and their characteristic motifs are given below:

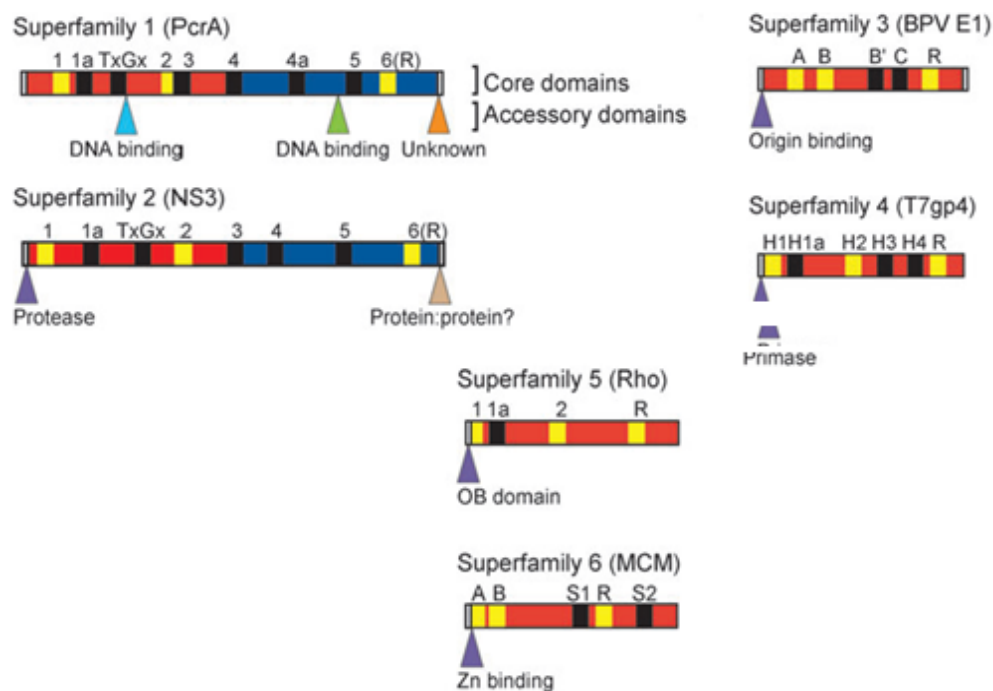


Fig. 2.1: Representation of the Gorbalenya & Koonin, 1993, classification of the helicases. The names in the parentheses are a representative of the family. Yellow represents the universal structural elements in the helicases. The position and function of the accessory domains are also shown, but these are specific to the examples given here unlike the core domains. (Ref: Singleton et.al, *Annu. Rev. Biochem.* 2007.76:23-50.)

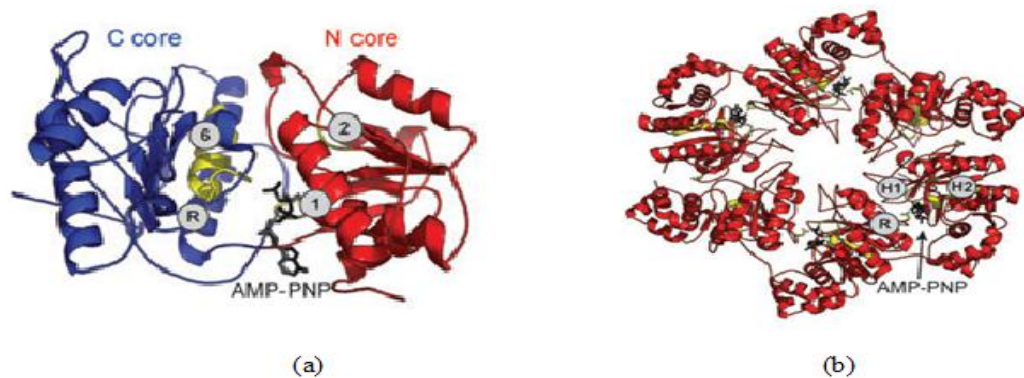


Fig 2.2: (a) SF1 and SF2 contains a monomeric core formed from the tandem repeat of a RecA like fold. The N and C terminal RecA like domains are shown. An NTP analogue (black) is bound at the interface of the two domains. (b) SF3-6 enzymes that has 6 individual RecA like folds at the core. Six nucleotide-binding pockets are present, one at each domain interface, and four are occupied with NTP analogues (*black*).

The SF1 and SF2 are the largest of the families. The motifs in the super families are a signature for the “core domains” (N- core and the C-core) that form tandem RecA-like folds, either within the same polypeptide chain or between subunits. These convert the chemical to mechanical energy by NTP binding and hydrolysis. The universal attributes of the core domains are: (1) conserved residues involved in the binding and hydrolysis of the NTP. (2) An “arginine finger” which plays a very important role in the energy coupling.

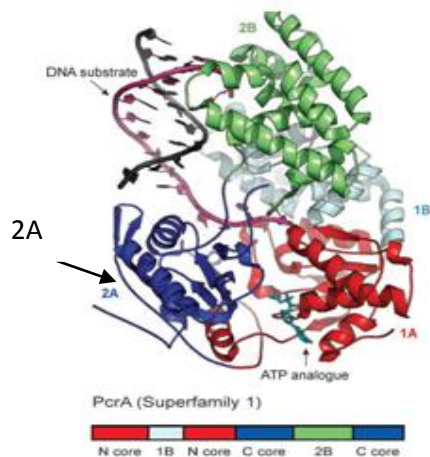


Fig 2.3: (a) Bar Representation of the domain structure of a typical SF1 Helicase(PcrA). (b) N-core and C-core (1A and 2A) and the insertions (1B and 2B).

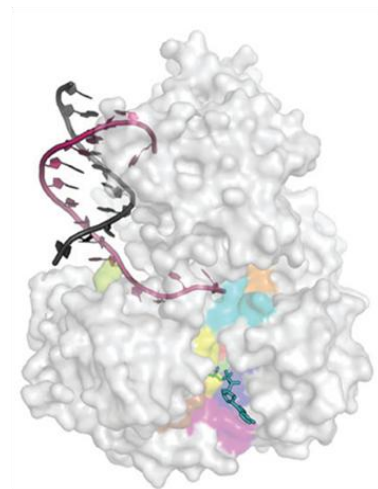


Fig 2.4: Space filling model highlighting the motifs

2.1.1. Superfamily 1(SF1):

These are one of the best characterised of the helicases. The PcrA and Rep are the best representatives of the SF1 family of helicases. Although SF1 has been further divided into 2 more families, the SF1A (3'-5') and the SF1B (5'-3') based on their directionalities in function.

- a) **SF1A:** PcrA is the best example of the SF1A family of helicases. Crystal structure of PcrA reveals that the ATP binding site is located in a cleft in between the two RecA like (N-core and C-core) folds that are lined by many of the other motifs. PcrA structures complexed with the ss-DNA undergoes conformational changes which gives us a picture of how exactly it functions (Velankar *et al*, 1999). This conformational change involves movement of the sub domain 2B, which sets up a conformation change on the sub domain 1B and 2B that interacts with the duplex DNA.

Upon ATP binding there is a change in the conformation and the cleft between the 1A and 2A (N and C-core) close around the DNA substrate. This cleft closure also causes the movement of the 1B and 2B and it forces the DNA duplex onto a negatively charged surface and it destabilizes the DNA duplex thereby causing the unwinding. The cleft closure enhances the bases in the single stranded region to

flip between the pockets of the core domains which subsequently causes the translocation of the helicases across the ss-DNA. The translocation thus follows an inchworm mechanism.

- b) **SF1B:** the RecD and the Dda are the representative of the SF1B family helicases. The current structural information of the subfamily comes from the structure of the RecD subunit of the RecBCD enzyme. But biochemical work has been deduced based on the RecBCD complex. The structure of the RecD subunit reveals that the SF1B operate as monomers (*Taylor and Smith, 1995; Nanduri et al 2002*) and they have two RecA like folds and NTP hydrolysing site in between them like the SF1A.

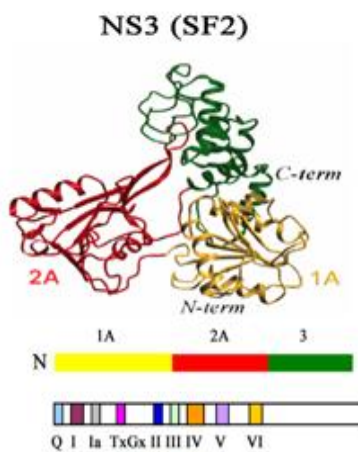


Fig 2.5: The structure of NS3 helicases belonging to SF2 Helicases along with the motifs

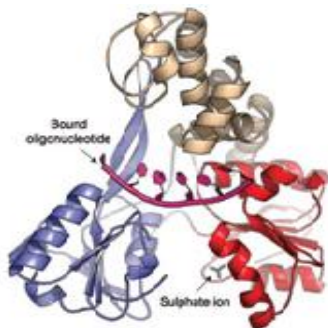


Fig 2.6: the NS3:DNA complex, domains are colored according to the panel given.

2.1.2 Superfamily 2 (SF2):

This is the largest family among all the helicase families. The helicases have been identified in a vast diversity of organisms from viruses to humans. They include the most intensively studied sub family of helicases namely the DEAD box RNA helicase (*Cordin et al 2006.*), RecQ (*Bennett and Keck 2004*) and the Snf 2 helicases . The first structural information was achieved from the crystal structure of the non structural protein NS3 of the Hepatitis C virus. NS3 can unwind the RNA and DNA duplexes but it requires a 3' overhang for this activity. It is shown that the conserved motifs of the SF1 and SF2 share similarities with one another. There are seven conserved motifs (Motif I, Ia, II-VI) in case of the SF2 helicases. The crystal structure of the NS3 reveals that the seven motifs reside in the domain 1 and domain 2 which are homologues

to the domains 1A and 2A of the SF1 family of helicases (*Korolev et al 1998, Yao et al, 2004*). Structure function studies reveal that motif I and II are involved in the NTP binding. The domain III of the SF2 connects the domain I & II unlike that of the SF1 helicases and help to couple ATP hydrolysis to a conformational change that helps it to translocate along the substrates (*Bernstein et al 2003*). The motifs IV-VI residing in the domain 2 helps in coupling ATP hydrolysis and nucleotide binding. The core domains 1 and 2 (1A and 2A) forms a groove like that of the SF1 helicases creating the nucleotide binding pocket and part of the ssDNA binding site (Fig 2.6). The core domain adopts an open conformation with the 6 bases of the ssDNA in the top surface of the core with no space for the second DNA strand to bind due to the steric hindrance from another non conserved region of the protein.



Fig. 2.7 Crystal Structure SV40 hexamer (SF3)

2.1.3. Superfamily 3 (SF3): They are quite distinct from the SF1 and the SF2 helicases, and are identified in the DNA and RNA viruses, involved in replication, origin recognition, unwinding. There are four conserved motifs in the SF3 helicases namely, A, B, B' and C (*Hickman and Dyda, 2005*). The motif A and B are Walker A and B ATPase boxes where as motif C is difference between the SF1, SF2 and SF3 is

that the former two contains two RecA like folds but the SF3 contains AAA+ core (ATPase associated with diverse cellular activities). The AAA+ domain comprised of several helix and sheets may be associated in stabilizing the hexameric and NTP binding activity. The Simian Virus 40 (SV40), the Bovine Papilloma Virus (BPV-1) E1 helicases, Adeno Associated Virus Type 2 (AAV2) are the representatives of this family. The structures of the SF3 helicases reveal that they have a six fold symmetric collar which stabilizes the hexameric ring with the AAA+ domain shown from exact radial symmetry. The central channel has an ssDNA oligonucleotide bound via interactions between the B'motif and the phosphodiester backbone. There are ADP

bound at all six sites around the ring with 3 different conformations i.e. ATP binding, ADP binding and empty. Moving round the ring we can have 2-3 sites with ADP, 2-3 sites with ATP and 2-1 site as empty. So, the position of the nucleotide binding loop correlates with the class of nucleotide binding site. ATP bound will be at the top of the spiral (5' side of the ssDNA) ADP at the intermediate and the empty at the bottom. So, during the progression, the loops each move one step downward, pushing the DNA with them.

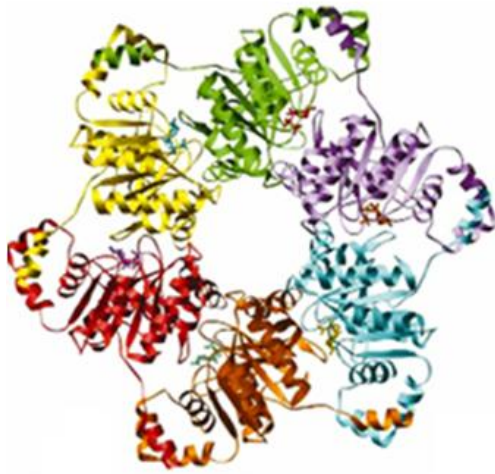


Fig 2.8 Crystal structure of T7gp4

2.1.4. Superfamily 4 (SF4):

These were first identified in the bacteria and the bacteriophages. In bacteria Helicase associates with the primase and both of them are separate polypeptides, where as in the bacteriophages both the activities are associated with the same polypeptide. All characterised SF4 helicases have 5'-3' activity.

SF4 comprises five different motifs, H1, H1a, H2, H3 and H4. The H1 and H2 motifs have a high similarity with the Walker A and B motifs but the other motifs don't have any similarity with any of the motifs in any other helicases family (Fig. 2.8). The H3 motif is a functional equivalent to the motif III of the SF1 helicases. The N terminal region of this helicase is not well conserved.

2.1.5. Superfamily 5 (SF5)

Although the Rho helicases are very closely associated with the SF4 helicases but they are placed in a separate family based on their differences in sequence. This group of helicases are responsible for the termination of transcription in bacteria, by binding to a specific sequence in the nascent RNA. The EM structures of these proteins represent open and notched forms which are in close proximity to their crystal structures.

2.1.6 Superfamily 6 (SF6)

The representative helicase of this family is the MCM helicase complex. It is thought to be the main eukaryotic replicative helicase, essential for replication initiation and elongation. Structural information is mainly derived from the archeal homologues (*Macris et al 2005*) which reveal MCM to be a homo hexameric protein. Cryo EM images shows it to have a large central channel, large enough to accommodate dsDNA.

2.2. General properties of helicases

Helicases have several biochemical properties: NTP binding, Nucleic acid Binding, NTP hydrolysis, translocation along nucleic acid strands, helix unwinding, Strand Annealing.

2.2.1. NTP Hydrolysis

Until now all the analysed helicases exhibit NTP binding and hydrolysis. Mostly they prefer to hydrolyse ATP than other NTP's. The Mg^{2+} ion which acts as a cofactor binds to the beta or gamma phosphate positions of the ATP and plays a very important role in the catalysis of the enzyme be it hydrolysis or binding. The NTPase activity of the helicases is stimulated by the binding of the nucleic acid molecule. Even though they can hydrolyse ATP independent of the DNA substrates. From the crystal structures of the various helicases various regions are identified to be very important for the functioning of the enzyme, like, there are conserved arginine fingers in the core helicase domain of the enzyme which are in close proximity to the γ phosphate of the ATP, and thus involved in the ATP hydrolysis by stabilization in the transition state of the reaction, and it also triggers the enzyme conformational change after the release of the pyrophosphate (*Matson et al 1994, Tuteja and Tuteja, 2006*). Thus it is accepted that the helicases convert the chemical energy of the hydrolysis of ATP and converts it into the mechanical force of translocation along the single or double stranded nucleic acid structures.

2.2.2 Nucleic Acid binding

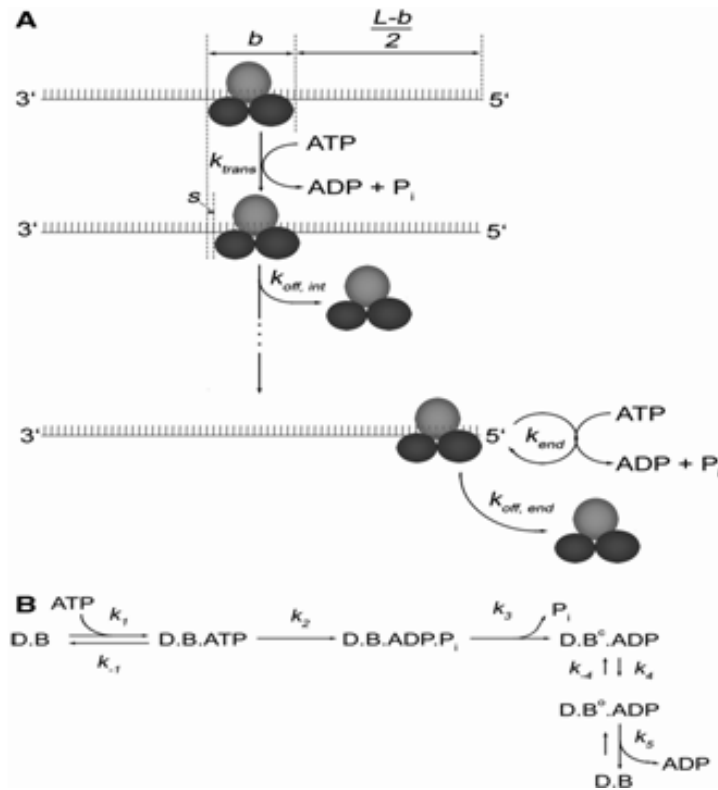
The binding of the helicases to the nucleic acids is an important step for all of its activities. Most of the helicases show strong binding to the ssDNA with a very low

K_d (dissociation constant, in nano molar range). There are conserved domains in the helicase core which are dedicated to the binding of the nucleic acid. Although the binding varies in different helicase families. Like, there is a zinc binding domain, in the SF2 helicases which has been found very crucial for the ssDNA binding (*Picha and Patel, 2004*). The zinc in the zinc motif of the helicases hydrogen bonds with some of the very crucial residues in this region namely Arg, Cys and Asp, so if these amino acid residues are mutated the binding of the DNA is readily affected (*Picha and Patel, 2004*). Most of the DNA helicases shows stronger affinity to bind to ssDNA than dsDNA.

2.2.3 Translocation along the nucleic acid strands

Translocation along the ssDNA requires the Processive movement of the helicase along the ssDNA tracts. As, DNA helicases are Processive motor enzymes they need to perform multiple enzymatic cycles in order to release end products. Till date 3 most important translocation mechanism have been explained (a) the inchworm mechanism (PcrA) (b) Brownian Motor (Thermal Ratchet) mechanism (NS3) (c) non uniform stepping model for UvrD.

Inchworm model (explained in detail below) and the Brownian model assume that hydrolysis of a single ATP molecule per kinetic step is governed by the rate limiting step of the cycle. Inchworm model describes the processive stepping of the enzyme unidirectionally, where as the Brownian model describes a diffusion driven process.



On a recently proposed model for translocation (*Gyimesi et al, 2010*) gives us key parameters of the ATP driven translocation of Bloom along the ssDNA, using fluorescence signals to directly monitor the ATP consumption also the interaction of the enzyme with DNA and ATP. According to this model, BLM translocates along ssDNA with moderate processivity, using an active inchworm mechanism, in which the rate limiting step is the structural transition between two ADP bound sites which might directly lead to the stepping along ssDNA.

2.2.4 DNA Unwinding Activity:

Two types of unwinding mechanism: active and passive:

Active: somehow disrupts the bases, usually by an aromatic residue that disrupts the bases and traps it the opened form. This can be located on a wedge domain (ex. RecG) or on RecA domains (PcrA, UvrD) or on additional domains (like RecQ1) etc. To test

that the helicase has an active unwinding mechanisms you have to measure unwinding with different types of DNA strands, so if you use a dT:dA strand or a dC:dG or with a mixed type, you should observe the same unwinding parameters. So the parameters are unaffected by base types.

Passive: The helicase waits until the thermal fluctuation disrupts the bases, than steps in and inhibits the reannealing of bases. If you measure the effect of unwinding parameters with the different base types, you should observed differences because the stability of G:C and A:T is different, G:C is stronger, so the probability of separation of these bases by thermal fluctuation is less, ergo the helicase that is waiting for thermal fluctuation to separate this bases will be slower. So when using only A:T the helicase will be faster, and when using only G:C bases the helicase will be slower.

DNA melting: The helicase when it is near to the bases enhances the probability of the separation of bases, for example by distortion of the DNA strand (so the mechanical strain will separate the DNA).

To unwind a long stretch of DNA, a helicase has to move along unidirectionally along the length of the DNA and couple the translocation activity with the strand separation. As, helicase moves along the DNA it successively disrupts the hydrogen bonds between the bases there by opening the double stranded entity only for active mechanisms. This process is linked to NTP hydrolysis and it has been estimated that the faster the helicase moves along the DNA faster is its NTP hydrolysis rates. Unwinding a helix is somewhat unstable thermodynamically, but helicase very well does that by binding with the ssDNA (*Levin and Patel, 2003, Liu et al 2000*). In addition to the ssDNA binding it has been proposed that helicases like the RepA and PcrA binds directly to the duplex DNA and destabilizes it.

There are certain common terms which explain the exact mechanism of how the helicases translocates and unwinds a duplex DNA, they are polarity, processivity, step size and mode of action. Polarity of the unwinding is defined as the direction of the helicases moving on initially bound single stranded templates with respect to the polarity of the sugar phosphate backbone. It can have both types of polarities depending on the nucleic acid it processes. If it processes DNA from 5'-3' direction

then it unwinds the duplex with 5' single strand, for e.g. RecD. If it has a 3'-5' directionality than it unwinds in the opposite way, for e.g. RecQ.

Processivity is the probability of the helicase taking a translocation or unwinding step versus the enzyme dissociating from the DNA on the same site. This is a very important parameter in the translocation characterization of the helicases as it together with the step size (number of nucleotide units travelled in a kinetic step) determines the mean run length of the enzyme.

Several models are proposed to explain the DNA unwinding mechanism based on their state of oligomerization. If the helicases are in their monomeric or dimeric state, the most proposed model is the inchworm mechanism (*Walksman et al 2000*). In this model the core helicase sub domains bind to the ssDNA and then consequently bind the NTP's, hydrolysis of which gives a driving force for the helicases to step along the DNA, thereby it displays a "halt" and "move" step in the progression, like an inchworm (Fig. 2.11).

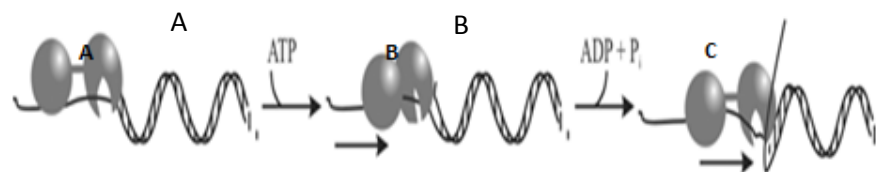


Fig. 2.11 (A) Nucleotide free unbound state. (B) Upon binding and co-ordination between ATP and metal ion, enzyme undergoes conformational change, closing a cleft between domains 1A and 2A. (C) ATP hydrolysis causes conformational change reversal and the enzyme translocates by strand separation as the bases from one strand are shuttled onto the NA binding channel during the domain movement (*Biochemical Society Transactions (2005) Volume 33, part 6, 1474-1478*).

For the dimerised helicases, each of the monomers has different affinity for the single strand or the double strand of the DNA. NTP hydrolysis causes change in the conformation of the enzyme bringing the closed monomeric forms of the enzyme

together and then they roll along the DNA, with each subunit having specific ATP hydrolysis cycle.

For the hexameric helicases and unwinding model is proposed and highly described for explaining the mechanism in which it performs the unwinding, e.g. BLM (*Karow et al,1999*). The unwinding model can further be divided into three different models.

- (a) Wedge model.
- (b) Torsional model.
- (c) Helix destabilizing model.

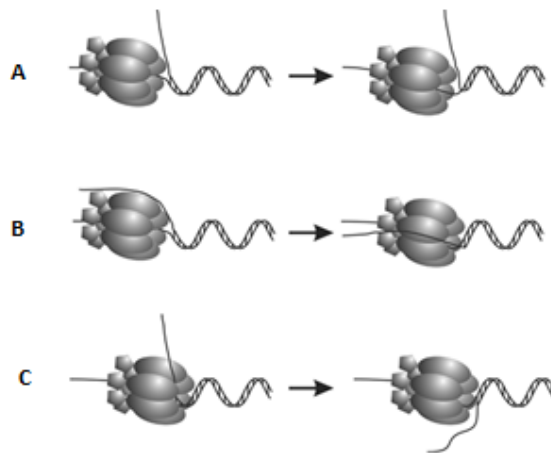


Fig 2.12 (A) Wedge Model (B) Torsional Model (C) Helix-Destabilizing Model (*Biochemical Society Transactions (2005) Volume 33, part 6, 1474-1478*).

In the Wedge model a single strand of the DNA passes through the central core of the enzyme whereas the complimentary strand passes along the outside of the ring, with no specific interactions with the dsDNA. In the Torsional model the enzyme interacts with both the included and the excluded strand. The excluded strand serves as a fulcrum upon which the hexamer rotates as it translocates down the NA lattice. In the Helix Destabilizing model is much more realistic because in this mode it has been explained that there is a specific interaction between the dsDNA and the helicases. Here the surface of the helicases is in direct contact with the dsDNA help to unwind the NA lattice.

2.2.5 Strand Annealing Activity:

Helicases are particularly known for their strand separation or binding activities, but there are evidences for the strand annealing activities as well. The RecQ helicase family have been known to have these namely the human RecQ5, RecQ4, Blooms Syndrome helicases (BLM) & Werner syndrome Helicase (WRN). It has been noted that ATP binding causes a conformational change in the helicases which promotes the strand separation or the unwinding activity, in a way is antagonistic to the strand annealing activity. So, it is basically shown in absence of ATP binding. Certain DNA binding protein also has a stimulatory activity in the strand separation and as such has an inhibitory activity on the strand annealing. As the unwinding progresses the annealing is influenced by the length of the DNA substrate, i.e. increase in length of the substrate will increase the strand annealing. So, it can be understood that the strand annealing and unwinding activity of a helicase work against each other, and in contrast it has been found that the BLM and WRN unwinding is significantly stronger than the strand annealing, although both these carry out these activities together in the strand exchange and branch migration during the DNA repair pathway and in resolving Double Holliday Junction.

2.3 The RecQ Family of Helicases

Represents one of the most highly conserved group of DNA helicases and named after the recQ^+ gene product in *E.coli* (Bachrati and Hickson, 2003; Hickson, 2003). Most of the unicellular organisms express only one family member where multicellular organisms expresses 2-7 RecQ proteins (may be more in some cases). Human contains five forms of RecQ helicases namely- RECQ1, 4 and 5, BLM and WRN along with additional splice variants at least with the RECQ5 protein. The region of primary sequence similarity is restricted to 400 amino acids in which seven structural helicase motifs characteristic of the SF2 superfamily is found. Most, but not all the members possess an extra domain named as the RQC (RecQ C- terminal), which is unique to the RecQ family members. The RQC domain is again composed of two domains namely the Winged Helix domain, which is involved in the DNA binding and the Zinc binding domain, involved in the stability and the proper folding of the RecQ helicases. One more typical domain to the RecQ helicases is the HRDC domain (Helicase and RNase D C-terminal). This is present in several helicases and RNase D

enzymes, is involved in the DNA binding, particularly in conferring some DNA substrate specificity (Berstein and Keck,2005; Wu et al 2005).

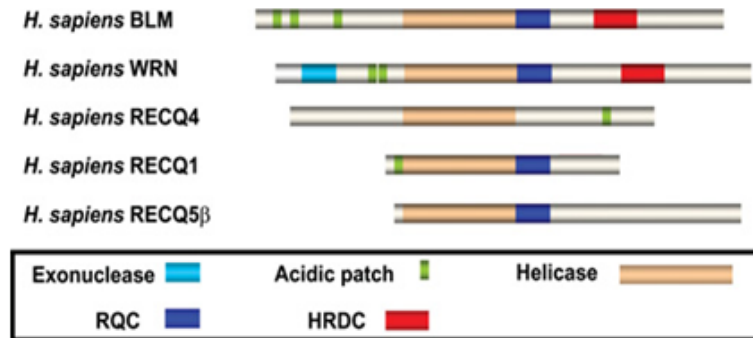


Fig 2.13 Schematic representation of the Human members of the RecQ helicase family. The colored regions indicate different conserved regions.(Biochem. J. (2006) 398, 319–337)

The RecQ family of helicases are considered of particular interest because of their specific roles in several human disorders. Mutations to these can cause several clinically defined disorders leading to severe predispositions to cancer and premature ageing, for example, BLM in Blooms Syndrome (BS), WRN in Werner’s Syndrome (WS), RTS in Rothmund- Thompson Syndrome (RTS). The multiple homologues of the RecQ family of helicases play specialised roles in the genome maintenance and are involved in the Double Stranded Break Repair, Homologous Recombination pathways.

2.3.1 *E.coli* RecQ Helicases

This is the first of the RecQ family of helicases which was discovered by Nakayama in 1984. It operates in the RecF pathway of the Homologous recombination that is involved in the repair of the stalled replication forks. This is also involved in the stopping of the illegitimate recombination which is enhanced to quite a lot in the RecQ mutants.

This protein is composed of 610 amino acids (molecular mass= 68,920 Da) (*Xu et al 2003*). It has a very wide range of DNA substrate specificity. It displays a 3'-5' polarity in DNA unwinding and DNA dependent ATP hydrolysis. *E.coli* RecQ helicase has both the abilities to bind to the ssDNA or dsDNA but it prefers the former (it prefers ss-dsDNA junctions). It binds to both the ssDNA and the dsDNA at the same site. The binding of the DNA substrate is very much dependent on the salt concentration, increasing salt concentration reduces the affinity of the DNA to bind (it refers to the role of electrostatic interactions between the amino acids and DNA).

The crystal structure of the *E.coli* RecQ as demonstrated by (*Berstein and Keck, 2003*) reveals that there are 2 additional domains which are adjacent to the helicase domain, one is the Zinc binding motif and the other is the winged helix motif (WH). The WH motif and the Zn Binding domains comprise the RecQC-Terminal domain. In addition all these there is a HRDC domain which is not included in the polypeptide for which the crystal structure was obtained. The HRDC domain structure was solved by (*Liu et al, 1999*). It has been elucidated that the Zinc binding motif has four highly conserved Cysteine residues that forms a complex with Zn^{2+} . This complex is essential for DNA binding and helicase activity, but not for ATP binding. It is also important for the protein folding.

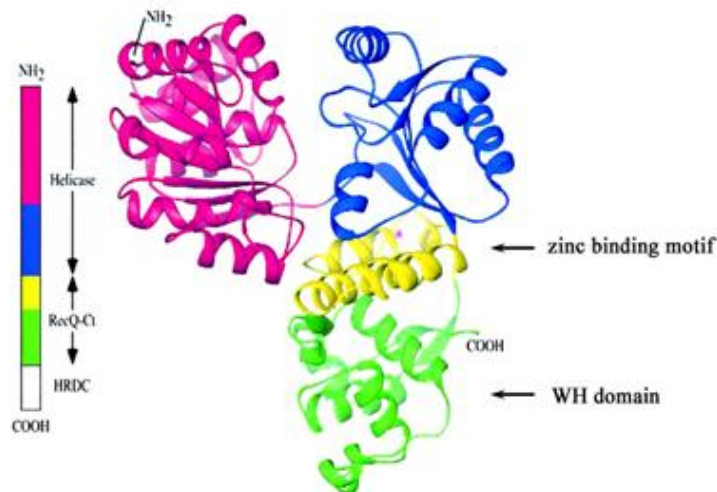


Fig 2.14 High Resolution structure of the E.coli RecQ catalytic core. (Bernstein,D.A., Zittel,M.C. and Keck,J.L. (2003) High-resolution structure of the E.coli RecQ helicase catalytic core. *EMBO J*, 22, 4910-4921.)

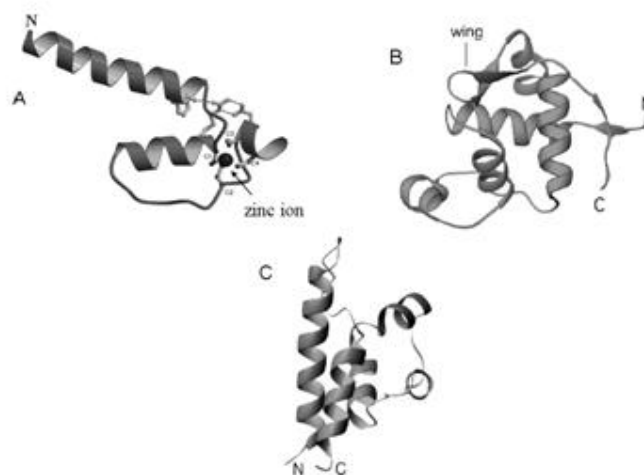


Fig 2.15. High resolution structures of some important domains of the RecQ helicases family. (A) Zinc Binding Domain (B) Winged Helix domain (C) HRDC domain.

(Liu,Z., Macias,M.J., Bottomley,M.J., Stier,G., Linge,J.P., Nilges,M., Bork,P. And Sattler,M. (1999) *Structure*, 7, 1557-1566).

RecQ plays a role in the genome stability by the following ways.

1. It plays a role in recombination by unwinding DNA at gaps or double stranded breaks, allowing RecA (recombinase) filaments formation to initiate recombination and D loop formation.

2. It dissolves Double Holliday Junctions (DHJ), in concert with Top III (*E.coli* Topoisomerase III) Also ore important is the DNA catenation resulting in inhibition or activation of HR
3. In a SOS model (SOS is a response to stress, like DNA damage. It ends in expression of many proteins), RecQ binds to a gap on the leading strand of a stalled replication fork and unwinds duplex template at 3'-5' direction, ahead of the stalled replication fork and generates SOS signalling (*Heywer, 2004*).

2.3.2 Yeast RecQ Helicase

Sgs1 is the sole RecQ helicase in the budding yeast. Its mutation cause intra and inter chromosomal hyper recombination, as well as elevated level of sister chromatid exchange (if over expressed).Sgs1 is of 1447 amino acids in length. It is an ATP dependent RecQ helicase with 3'-5' directionality. It is able to unwind DHJ and D-loops and can also disrupt DNA-RNA hybrid.

The cellular abundance of Sgs1 is cell cycle phase dependent. Its abundance is low in the M and the G1 phases, at its peak at the S phase and degrades at G2 phase.

2.3.3 Human RecQ Helicases

Till date five RecQ family members have been reported to be found in Humans, they are, RECQ1, BLM, WRN, RECQ4 and RECQ5. Mutations in these helicases have been found to be associated with various disorders as in, mutations for BLM, WRN and RECQ4 gives rise to Bloom's disorder (BS), Werner's Syndrome (WS) and Rothmund Thompson Syndrome (RTS) respectively. The WS and RTS are characterised by premature ageing where as the BS is related to be associated to predisposition to several cancers. They are considered to be tumour suppressors and they do it by maintaining the stability of the chromosome.

2.3.3.1 BLM- Bloom Syndrome Helicase

Bloom Syndrome (BS): It is an autosomal recessive genetic disorder of humans which is caused due to the mutation in the BLM gene (located on 15q26.1) (*Mc Daniel and Schultz, 1993*) and there by a non functional gene product which is unable to repair the high frequency of breaks and rearrangements in the affected persons

chromosome. This syndrome was first described by American dermatologist Dr. David Bloom in the year 1954. This was first characterised in the Ashkenazi jews. This disease is characterized by the person having short stature, and developing a skin rash shortly after the exposure to sun. Other clinical manifestations involve micrognathism, mental retardation, immune deficiency, diabetes and predispositions to cancers.

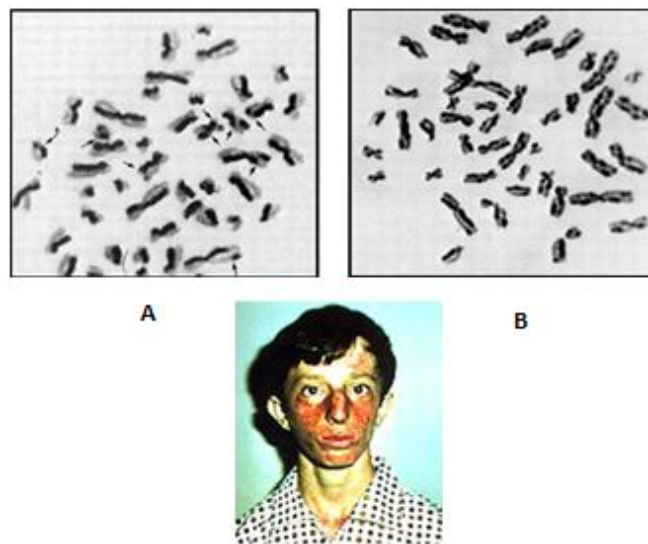


Fig 2.16 (A) Chromosomes from lymphocyte of a normal person (B) Chromosomes from the lymphocyte of a person with BS. (C) BS patient.

BLM

The full length BLM protein has 1417 amino acids in length (Molecular Mass, 159 kDa). This protein is expressed in all the tissues where there is an active cellular proliferation, including most tumour cells. BLM as a typical representative of the RecQ family of helicases is comprised of 3 conserved domain namely the Helicase core domain, RecQ-CT domain and the HRDC domain. It contains nuclear

localization signals close the C terminal. Like the other helicases BLM utilizes the energy derived from the hydrolysis of ATP for catalyzing DNA strand separation in a 3'-5' directionality (*Guo et al 2007*). It also possesses a strand annealing activity which is inhibited by the RPA and Single Stranded DNA binding protein. It can recognize and unwind a wide range of substrates of varying shapes and lengths.

It has been found that the BLM protein interacts with several other proteins to perform their tasks in the maintenance of genomic stability. Some of the most important interacting protein partners are discussed below.

hTopIII α : It is a topoisomerase, and it catalyzes transient DNA breaks in order to relieve the steric strain on the DNA due to replication. BLM and TopIII α co-localize on intermediates of DNA replication, where DNA structures generated by BLM are resolved by TopIII α . It also helps to suppress hyper recombination events.

RAD51: Recombinase, and its activity is similar to the bacterial RecA protein (both structurally and functionally) and it assists DNA double stranded breaks and HR.

It helps to promote Holliday junction branch migration. Catalyses pairing between ssDNA tail and ds DNA homologous stretch and promote strand exchange to initiate HR.

BRCA1: Role: loading of Rad51 This class of proteins acts as tumour suppressors. So, mutation in this will result in significant increase in the risk of cancers. BLM interacts with BRCA1 factor to respond to DNA damage.

2.3.3.2 Werner Syndrome Helicase (WRN)

Werner Syndrome

It is also an autosomal recessive disorder which is highly prevalent in the Japanese population. The *WRN* gene is located at chromosome 8p-12. The Werner's syndrome condition become apparently visible during the adolescence, with lack of growth spurt. Other important features includes premature ageing disorder with hair loss and greying, atherosclerosis, hypermelanosis, cerebral cortical atrophy, lymphoid depression and thymoid atrophy and Type II Diabetes.

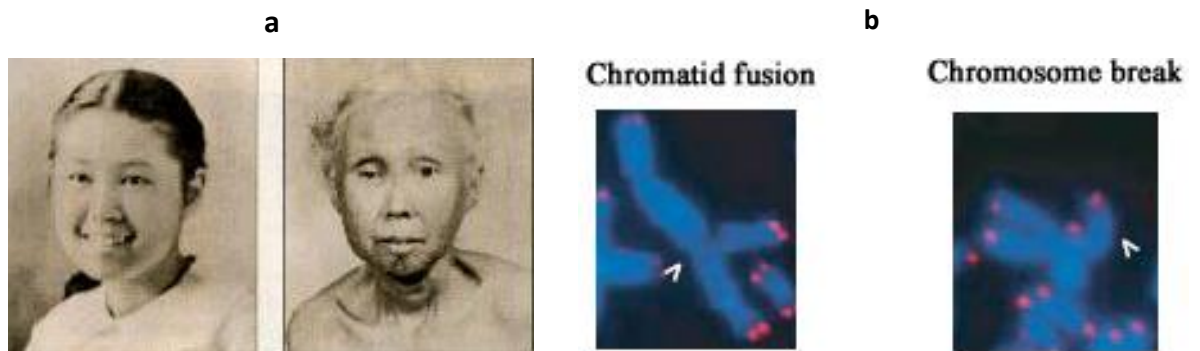


Fig. 2.18 (a) Werner syndrome patient where she looked normal as a teenager but effects of Werner syndrome were visible at 48 years, looking more old than usual.

(b)The bottom image shows chromosome breaks and fusion of WS cell, the red dots are telomeres (FTIC probed), and DNA is DAPI stained.

WRN

It is 1432 amino acid long protein. It consists of two important domains namely the central helicase core and the N-terminal exonuclease domain. There is also an unusual region of 27 amino acids between the Helicase core and the exonuclease domain. It has 3'-5' directionality and low processivity. The most important function of WRN helicase is that it can resolve quadruplex structures (G4 tetraplex DNA). It helps in the removal of DNA secondary structures or other barriers which can cause obstacles in the translocation activity. In absence of WRN, replication forks could be stalled and this would cause and this would be a great obstacle to maintain the genomic stability.

2.3.3.3 Rothmund Thompson Syndrome Helicase (RTS/RECQ4)

RT Syndrome

It is a very rare autosomal recessive disorder roughly around 200 patients reported worldwide. This was first described by Rothmund in the year 1868 where he reported patients to have unusual skin degenerations and cataracts. ⁷ **b** was further described by British dermatologist Thompson where he reported patients to have unusual skin abnormalities and skeletal malformations, Poikiloderma Congenitale. Patients suffer from post natal growth and skin abnormalities.

RTS/RECQ4

The RECQ4 protein was cloned earlier before it was known to be involved in this disorder. It lies on the p-arm of chromosome 8 at 24.3 spanning 6.5 kb and includes

21 exons (228; 235). It consists of 1208 amino acids, contains the Helicase core domain but lacks the Zinc Binding domain, WH domain and the HRDC domain. It is present in all the phases of cell cycle apparently peaks in the S and G2 phase which indicates its replicative and post-replicative roles. Like BLM its interacting protein partner is RAD51, where it cooperates with RAD51 in the HR of the DS break repair pathway. In the UV-DNA damage RECQ4 can remove the DNA lesion through interaction with nucleotide excision repair factor, XPA.

2.3.4 Model for Homologous Recombination (Role of helicases)

The double stranded breaks in the cells can be repaired by 2 major pathways. The Homologous Recombination Pathway (HR) and Non Homologous End Joining (NHEJ) Pathway. The initiation of the Homologous Recombination Pathway is the creation of a region of ssDNA that is bound by the Recombinases of RecA family (Rad51). Then RecA forms an extended helical filament with DNA lying on the centre of the longitudinal axis of the filament. Then this looks for a homologous sequence in an intact duplex, which will form the template for the repair synthesis.

RecA/Rad51 then initiates the HR by invasion of this ssDNA (filament) into the Homologous duplex, thereby created a Displacement loop, or D-loop. The next event in the HR process is the disruption of the D-loop. It can take place by two ways, in one of the ways called synthesis dependent strand annealing (SDSA), new DNA synthesis from the above mentioned strand invasion is followed by the migration of the D-loop along the DNA, without the formation of Holliday Junction. The next step here is the displacement of the invading strand and reannealing to the other broken strand (Fig 2.17 b) by complimentary base pairing.

In the other pathway the priming of the DNA synthesis region and the second arm of the breaks generates a Double Holliday Junction, following the completion of the DNA synthesis and ligation of the nicks. The Holliday Junctions are further resolved by 2 different ways. In the classical method the resolvase enzyme like RuvC creates nick in 2 of the 4 strands of the Double Holliday Junction. The second method employs the BLM helicase and its interacting partners like TopIII and RMI1.

From this model it is thus clearly evident that Helicases affect the HR pathways at several stages, which can be both effective and negative for the process. Like,

helicases creates the ssDNA tail for the strand invasion by RecA/Rad51, one of the key steps in the HR pathway. But, there are some other helicases too which are involved in the displacement of the RecA/Rad51 filaments and so can be called as anti recombinases.

But the underlying mechanisms of these are still elusive and structural studies of the proteins involved in these pathways, like the BLM can help us understand a multitude of things.

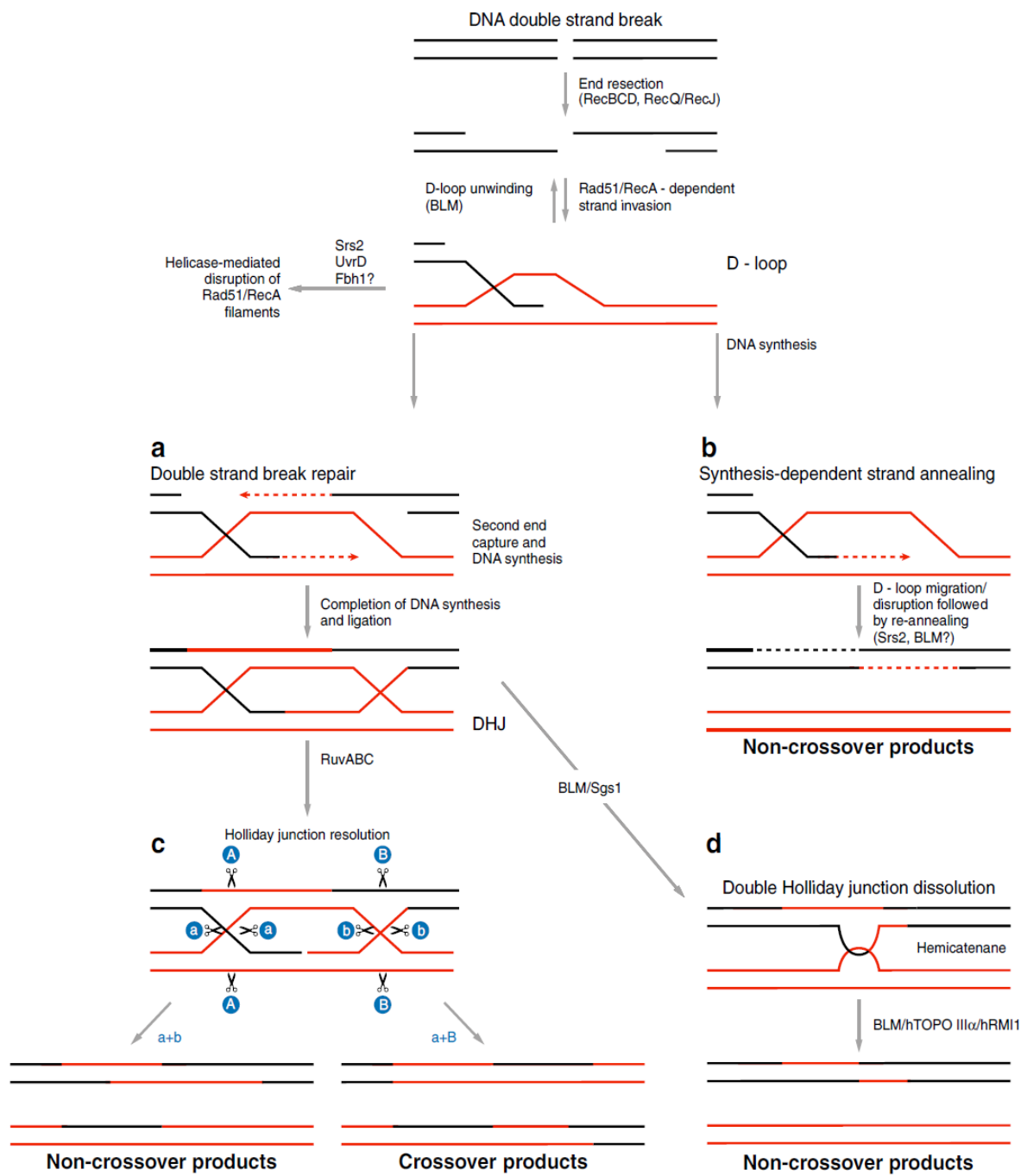


Fig 2.17 (a-d) Multiple pathways of the D loop processing in HR, by both prokaryotic and eukaryotic helicases and their interacting protein partners. (Wu, I., and Hickson, I.D. (2006) *Annu. Rev. Genet.* 2006.40:279-306)

Chapter 3

Aims and Objectives

There is only a little known about the actual unwinding mechanism of the helicases. A lot is left to be explored. There are very few crystal structures of these important family of enzymes in their DNA bound state; eg. PcrA (*Velankar et al 1999*), NS3 (*Kim et al 1998*) etc, which gives us hint about their working mechanisms in tandem with the biophysical analysis. But it is highly imperative to have a crystal structure of the *E.coli* RecQ and the BLM helicase in their DNA bound state in order to truly appreciate their unwinding mechanism.

In this work we propose to crystallize both the BLM and *E.coli* RecQ helicase in the DNA bound form, in order to unravel their mechanism, which would provide us deeper insights in order to understand their role in the Homologous Recombination pathway, key for double stranded break repair in order to maintain the genome integrity (*Singleton et al 2003*).

It is known that the helicases are involved in the Homologous Recombination Pathway (*Wu and Hickson, 2003*) mainly for the dissolution of the Double Holliday Junction, which is a key step in the HR pathway, and RecQ family of helicases are specifically involved in this pathway. One of the most important features that make this family of particular interest is their role in human disorders. Germline mutation in any of the human RecQ family genes can cause clinically defined disorders and predispositions to cancers and premature ageing. As, such understanding the mechanisms of how these proteins actually participate in the repair pathways, would enable us to develop strategies to cure them.

The broad aims of this work are:

1. Cloning and Expression of the BLM and *E.coli* RecQ. (BLM¹¹⁹¹ specific for this study, which doesn't possess the HRDC domain.)
2. Biochemical characterization.
3. Purification of the protein to 99% purity and Pre- Crystallization experiments.
4. Biochemical measurements to determine the properties of the mutants.

5. Crystallization and Structure determination.

6. Structure Analysis.

This thesis aims to purify BLM¹¹⁹¹ to crystallization grade purity and biochemically characterize its properties, before moving on to the crystallization experiments.

Chapter 4

Materials and Methods

4.1 Reagents

All the reagents that are used for the study except where not mentioned are purchased from Sigma Aldrich Co ATP from Roche Applied Science, Pi standard from Merck.

4.2 Nucleic Acid Substrates

All the oligonucleotides used for the study were purchased from VBC biotech and Invitrogen (dT(n)).

4.3 Heparin (*Sigma H3393*)

It was dissolved in sterile distilled water at 50mg/ml and then dialyzed against sterile distilled water and then against SF-50 Buffer in MWCO 3350 dialysis tubing (Serva 44183)

4.4 Protein construct and engineering

BLM⁶⁴²⁻¹¹⁹¹: The protein construct that is primarily used in this study is the BLM⁶⁴²⁻¹¹⁹¹ construct. This construct is 549 amino acids in length and retains the conserved core regions of the helicase namely the helicase core domain which is from 683-992 amino acids in the full sequence of the BLM protein. This construct doesn't contain the HRDC domain as it is supposed that it can cause structural heterogeneity and hinder crystallization.

Coding region for the BLM⁶⁴²⁻¹¹⁹¹ have been amplified by PCR from a pTXB plasmid containing the BLM⁶⁴²⁻¹¹⁹¹ construct, which was available in the lab and sub cloned into the pTXB3 vector, using the internal cloning sites of the *BglII* and the *SapI* of the multi cloning site (MCS) of the vector.

The pTXB3 has a self splicing intein coding region followed by a Chitin Binding Domain (CBD) which is specific for the affinity purification by Chitin beads.

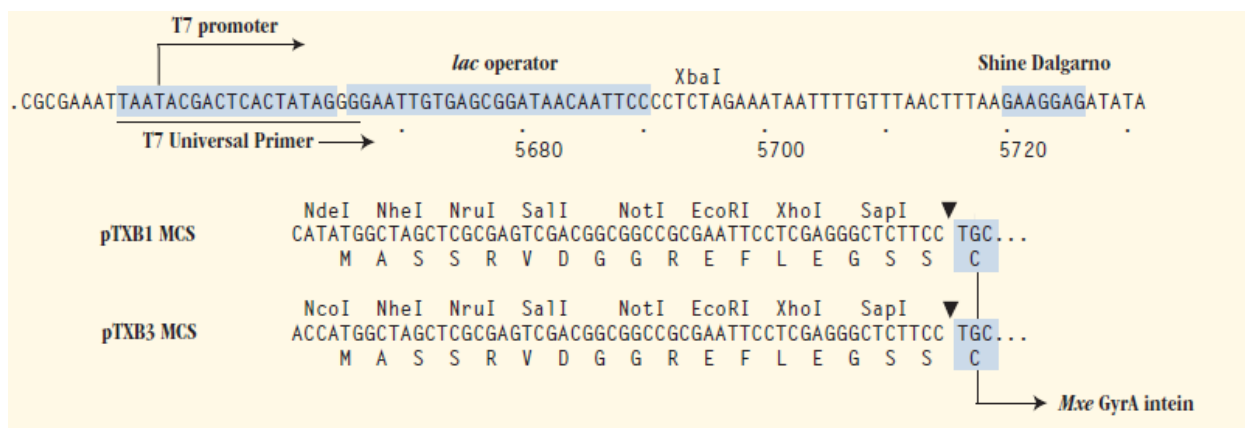


Fig 4.1. The pTXB3 Multiple Cloning Site, which shows the *SapI* and *NcoI* restriction sites.

All plasmid purification steps were done by QIAprep MiniPrep Kit (Qiagen).

4.5 Expression and Protein Purification

4.5.1 Transformation

The BLM and the *E.coli* RecQ gene constructs were cloned in the pTXB3 vector as mentioned earlier and these were transformed into the *E.coli* Rosetta (resistant for chloramphenicol) and BER cell lines respectively, which was done by the heat shock method. 250ul aliquote of competent cells (Rosetta) which have been frozen at -80°C were taken out and kept in ice till it melts down and then $1.5\mu\text{l}$ of MiniPrep plasmid, containing the desired construct mixed with it and kept in ice for 30 minutes. Then a heat shock was given to the cells by dipping them in a water bath for 1 minute which was maintained at 42°C . Then the cells were kept back in ice for nearly 5 min, $300\mu\text{l}$ of LB/2YT medium was added and shaken at 180rpm, 37°C for nearly 45 min. Then $200\mu\text{l}$ from this was plated (spread plate) onto a LB plate containing 100mg/l of Ampicillin (as pTXB3 has an Ampicillin resistant site) and 50mg/l of Chloramphenicol. The plate was kept at overnight incubation at 37°C .

4.5.2 Expression

The transformed cells will appear as small single colonies in the LB plate. A 5ml starter culture is prepared by taking 5ml LB media (Luria broth) containing exactly similar concentrations of Ampicillin (100mg/l) and Chloramphenicol (50mg/ml) and inoculating a single colony in the tubes and shake them overnight at 37°C , 180 rpm.

Mass culturing was done by using 6L, 2YT media: Peptone 16% (w/v), Yeast Extract 5% (w/v) and NaCl 5% (w/v), with Ampicillin (100mg/l) and Chloramphenicol (50mg/ml). 1.5ml of the starter culture was inoculated into each of the 1L flasks and then cells were grown by shaking them at 37⁰C, 180 rpm (New Brunswick, Innova 44R) for 3-4 hours (O.D optical density λ_{600} = 0.3). Then it is induced with 200mM/l IPTG (Isopropyl β -D-1-thiogalactopyranoside) after thorough cooling of the flasks and then incubated overnight at 18⁰C, 180 rpm. The cells were then harvested by centrifugation (Beckmann Coulter, Avanti J series, Rotor-JLA-9.1000) at 4⁰C at 5,000 rpm for 10 min. Pellet was resuspended in 240ml elution buffer (CH buffer, composition given below) and was thoroughly homogenised. 10mg/ml Lysozyme was added to the homogenate and kept at 4⁰C for 1 hour. Then the homogenate was sonicated at 70-80 duty cycle, 2 min per cycle for 7cycle and then centrifuged at 4⁰C at 18,000 rpm for 30 min. The cell lysate was used for purification by chromatography.

4.5.3 Purification

By Chitin Beads (S6651, NEB Inc): The lysate (which has the fusion protein with intein-chitin binding domain) was loaded on to the Chitin Beads, 10ml (New England, Bio Labs) which was packed in a Amersham Biosciences XK 24/40 empty column and equilibrated with 20 column volumes of CH buffer (50mM Tris HCl, pH 8.0, 0.5M NaCl, 1 mM EDTA.2H₂O, 0.1 V/V % Triton X-100, 10 V/V% Glycerol) at a flow rate of 1ml/min. Then the column is washed with 20-30 column volumes of the same buffer (CH Buffer) and the flow through was checked with Bradford reagent to analyze the accuracy of washing step. The column was then left at incubation for overnight at 4⁰C after loading 3 volumes of 50mM DTT (column cleavage).

The bound protein in the column was then eluted by CH buffer, and 15-20 1.5 ml fractions were collected and checked by Bradford Reagent (BR) (Microtitter plate- 200 μ l of BR+ 10 μ l of sample). This was then diluted to nearly 2.5 times by 50mM Tris HCl, pH 7.5.

By Heparin Column (Hi Trap HP, GE healthcare 17-0407-01): The diluted sample was then loaded on to 5ml Hi-Trap Heparin cartridge (HP), at a flow rate of 1ml/

min which was pre equilibrated with 10 column volumes HP Buffer (50mM Tris HCl, pH-7.5, 200mM NaCl, 10V/V % Glycerol, 0.1mM EDTA 2Na.2H₂O, 1mM DTT) at 1ml/min flow rate. The elution of the bound protein from the HP column was done by a linear gradient of NaCl (200mM-1M, 50ml, 1ml/min) in HP buffer (1ml fractions, BLM after ~ 27mS) This was done by ACTA prime plus TM (GE Healthcare Life sciences) purification systems. The fractions containing the protein were pooled and mixed. The BLM binds to the Heparin owing to its structural similarity to the DNA/RNA, so the protein is trapped on to the column on which heparin is immobilized.

The protein sample that is obtained after the purification by the heparin column is about 90% pure which is efficient enough for conducting the biochemical assays. But a very high grade of purity (99%) is a pre requisite for the crystallization experiments. So several other resins were used to further purify the protein fraction obtained after the HP purification. These were done individually after the HP elution.

By Carboxy Methyl (CM) Column: Pooled samples were collected and diluted 3 times with 20mM Tris HCl, pH 8.0. This was then loaded (1ml/min) on to a 5x1 ml CM Sepharose FF column (GE Healthcare, 17-5056-01) pre equilibrated with (1ml/min, 50 ml) with the CM buffer (20mM Tris HCl, pH 8.0, 0.1 mM EDTA.2Na.2H₂O,10 V/V % glycerol, 150mM NaCl,1mM DTT). The bound protein is then eluted with linear gradient of NaCl (150mM-1M, 20 ml, 1ml/min) in CM buffer (0.5ml fractions). The fractions containing the BLM were pooled and collected.

By Bio-Scale Affi Blue Column

Pooled fractions (after HP elution) were collected and diluted to 3 times with 50mM Tris HCl, pH7.5 and loaded (1ml/min) to a 5ml BioScale Mini Affi Blue Gel Cartridge (Biorad, 732-4642), pre equilibrated with AB Buffer (50mM Tris HCl, pH 7.5, 10mM NaCl). The protein is then eluted with a linear gradient of NaCl (10mM-1.4 M, 30ml, 1ml/min) in AB buffer (0.5ml fractions).

By Hydroxyapatite Column

Pooled fractions (after HP elution) were collected and dialyzed with storage buffer (50mM Tris HCl, pH 7.5, 50mM NaCl, 1mM DTT) diluted to 4 times with 50mM Tris HCl and then loaded to a 5ml Bio Scale Mini CHT 40 μ m, which is pre equilibrated with HA Buffer.

This is then eluted with a linear gradient of NaCl (50mM- 1M, 30 ml, 1min/ml) in HA buffer (0.5 ml fractions).

4.6 ATPase Assay

The ATPase activity was followed as given by (*Gyimesi et al, 2010*) as a NADH coupled assay as follows: 2% LDH/PK mixture, 136.4mM ATP, 1mM PEP, 100mM NADH. Absorbance was measured at 340nm wavelength to follow the NADH consumption by PK/LDH which in turn senses the releasing ADP. The measurement is carried out in a Shimadzu UV-2101 PC Spectrophotometer. Assay buffer: 50mM Tris (pH 7.5), 50mM NaCl, 5mM MgCl₂, 1mM DTT, 50 mg/l BSA. The temperature for all the measurements were strictly controlled at 25⁰C

4.7 Transient Kinetics

All the transient kinetics experiments were carried out using KinTek-2004 stopped flow apparatus that has a 1ms dead time at the applied 15ml/sec. Temperature was kept constant at 25⁰C with a water circulator. Pi release measurements were performed by SF-50 buffer. A Pi mop (150 μ M 7-methylguanosine, 0.1 U/ml purine nucleoside phosphorylase) was present in all the solutions. MDCC-PBP (MDCC labelled *E.coli* Phosphate binding protein, PBP binds Pi with 1:1 stoichiometry, when PBP binds Pi it undergoes a conformational change, the MDCC dye is attached to a loop that moves upon the conformational change. When PBP binds Pi the environment around MDCC changes, and it results in an increased MDCC fluorescence emission signal.) Fluorescence was excited at 436nm and excitation was followed through a 455nm cut off filter. Single round translocation experiments were performed in the presence of heparin in SF-50 buffer. Transient Pi production in different ssDNA lengths was determined which will give us an insight into the Translocation model by (*Gyimesi et al 2010*).

MDCC-PBP Fluorescence calibration

To establish the correspondence between MDCC-PBP fluorescence amplitudes and Pi concentration, a calibration step was conducted by rapid mixing of 3 μ M (Post mix concentration) MDCC-PBP (premixed with Pi mop) with varying concentration of Pi standard solution of different concentration in the stop flow. The MDCC-PBP (0-3 μ M) fluorescence was linearly dependent on different Pi concentrations.

Chapter 5

Results

5.1 Purification

By CH beads and HP column

The purification profile of BLM¹¹⁹¹ as observed from the SDS-PAGE (Nu PAGE Novex 4-12% Bis-Tris Gel, Invitrogen (WG1402BOX)) reveals that after the HP column purification we can obtain nearly 90% pure protein with 2 contaminating bands (90 and 40 KDa), which were intended to be purified by the use of CM column to obtain crystallization grade protein.

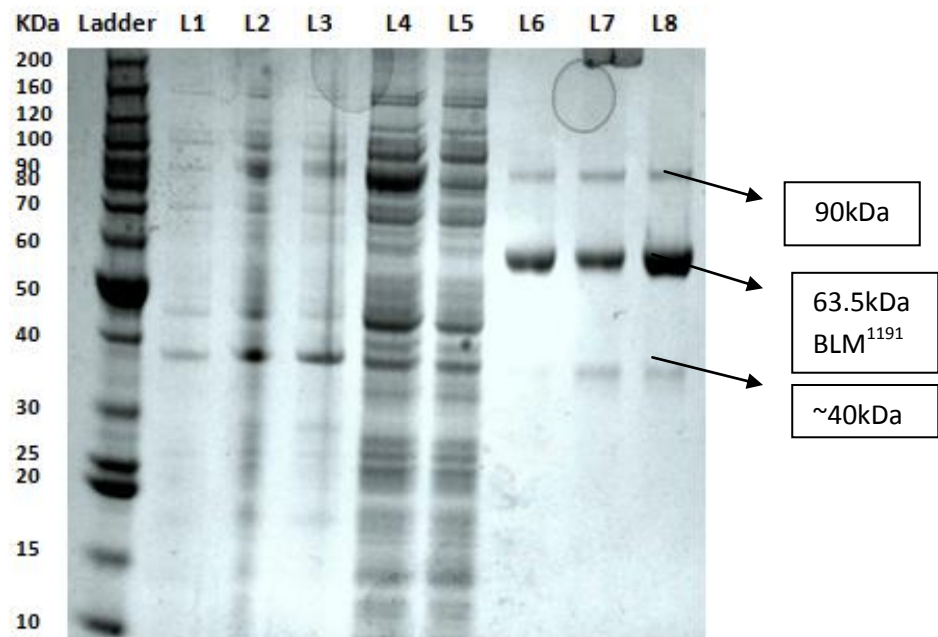


Fig 5.1: SDS-PAGE (4-12%) purification profile for CH and HP column. **Ladder:** Benchmark Unstained Protein standard (10747-012), Invitrogen. **Lanes:** 1= cells before induction, 2= cells after induction, 3=lysate pellet, 4=Lysate supernatant, 5= CH flow through, 6=CH eluate, 7= HP eluate, 8= HP eluate (after dialysis).

By CM column

The CM purification was not that efficient and it didn't remove the contaminating 90 KDa and ~40 KDa regions as these could not bind to the column and were present in the eluate along with the BLM fraction. An additional 10KDa contaminating band was observed when the protein was concentrated and then further Liq N₂ frozen.

Additional 10kDa band
observed when BLM was
concentrated and frozen

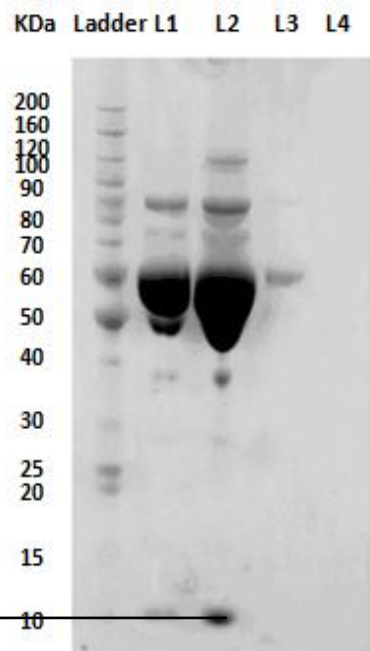


Fig 5.2: CM column Purification profile along with LiqN₂ frozen and dialyzed fractions. **Ladder:** Benchmark Unstained Protein standard (10747-012), Invitrogen. **Lanes:** 1= concentrated BLM fraction before Liq N₂ freezing, 2= after Liq N₂, 3=CM eluate 4=CM flow through

By Affi-blue Column

This column was chosen as it has multitude of binding probabilities and can act as ionic, hydrophobic, aromatic and sterically active binding site. In one of the trials with this, the protein was bound to the column but could not be eluted with Linear NaCl gradient (10mM-1.4 M). Although the bound fraction could be eluted in the column regeneration step by the use of 1.5 M NaSCN. In the second attempt a higher gradient of NaCl (500mM-3M) was planned to be used for eluting the protein from the column. Protein could not bind to the column properly, perhaps due to higher dilution as the yield after the HP elution was very low. And in addition to these there were other technical difficulties in using this column as such Hydroxyapatite column was tried.

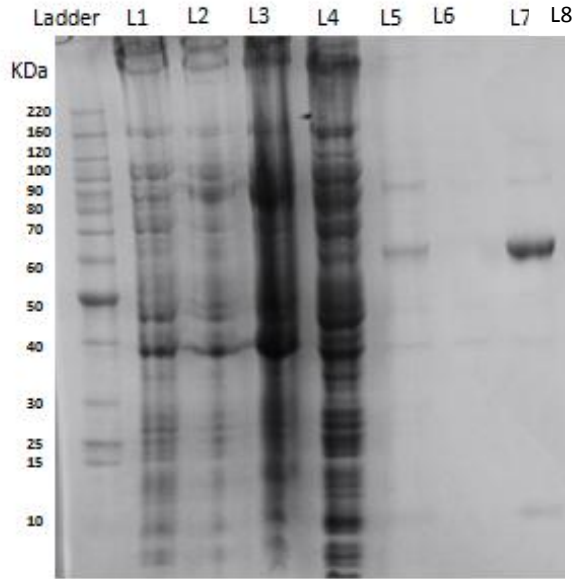


Fig 5.3: Affi-Blue Purification profile 1: Ladder: Benchmark Unstained Protein standard (10747-012), Invitrogen. Lanes: 1= cells before induction, 2= cells after induction, 3=lysate supernatant, 4=CH flow through ,5= CH eluate, 6=HP flow through, 7= HP eluate, 8= Affi Blue Flow through.

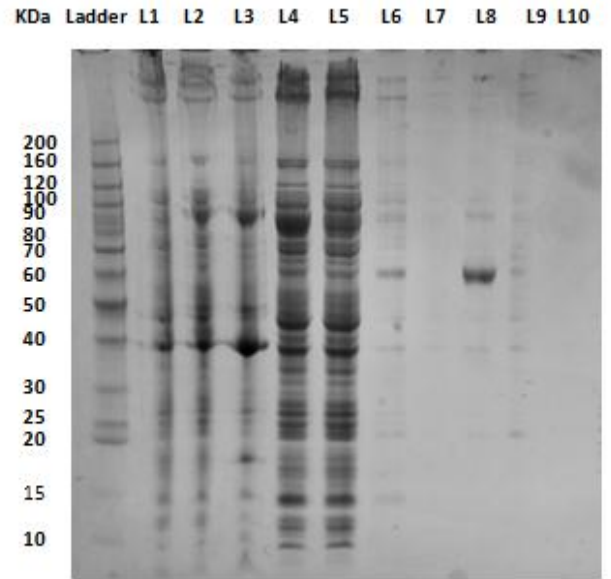


Fig 5.4: Affi-Blue Purification profile 2 Ladder: Benchmark Unstained Protein standard (10747-012), Invitrogen. **Lanes:** 1= cells before induction, 2= cells after induction, 3=lysate pellet, 4=Lysate supernatant, 5= CH flow through, 6=CH eluate, 7= HP flow through, 8= HP eluate, 9=AB flow through, 10=AB eluate.

By Hydroxyapatite Column

The Hydroxyapatite (HA) column could partially remove the contaminating 90 KDa band, although extensive washing step is required after the protein is loaded onto the column. But it cannot be eluted at a gradient that is recommended by the catalogue instead it elutes at a higher gradient concentration.

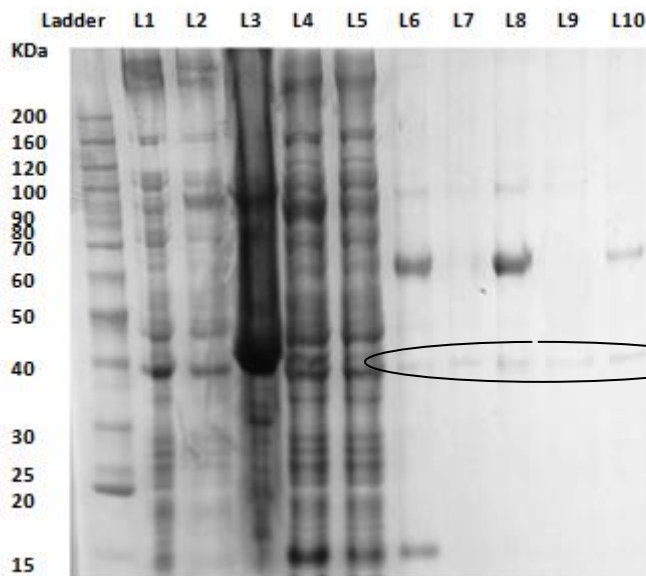


Fig 5.5: Hydroxyapatite purification profile. **Ladder:** Benchmark Unstained Protein standard (10747-012), Invitrogen. **Lanes:** 1= cells before induction, 2= cells after induction, 3=lysate pellet, 4=Lysate supernatant, 5= CH flow through, 6=CH eluate, 7= HP flow through, 8= HP eluate, 9=HA flow through, 10= HA eluate.

Contaminating 40kDa band as seen in all the flow through and elutions – CH, HP and HA column, from left to right respectively.

As, it is observed (**Fig 5.5**) that the contaminating 40KDa band is present in almost all the elutions and flow through, we speculated that an extensive washing step after the protein sample is loaded on to the HP column (15-20 Column Volumes), and also after it is loaded on to the HA column. This could successfully remove all the contaminating bands and we could possibly obtain highly homogenous BLM fractions.

As has been mentioned above an extensive washing step (by 10 column volumes) was performed after the loading the sample on the HP column and the following SDS-PAGE profile was observed.

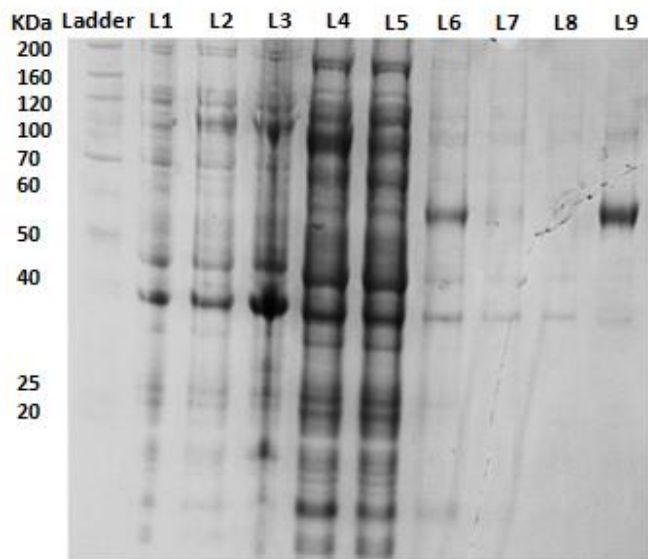


Fig 5.6: Purification profile of the additional washing step employed before HP elution. Ladder: Benchmark Unstained Protein standard (10747-012), Invitrogen. Lanes: 1= cells before induction, 2= cells after induction, 3=lysate pellet, 4=Lysate supernatant, 5= CH flow through, 6=CH eluate, 7= HP flow through, 8= HP flow (after wash) through, 9=HP eluate.

5.2 Determination of protein concentration by Bradford method.

The commonly used Bradford method is used to determine the protein concentration. This is a microtiter plate assay, where equal aliquots of Bradford reagent (200 μ l) is put in 10 consecutive wells and protein standard BSA (1.049mg/ml) is mixed on to each well in different volumes (0-5 μ l). Similarly the next row is followed by the protein sample (BLM). The total setup is covered and then absorbance is read in a micro plate reader at 595 nm.

The concentration of BLM¹¹⁹¹ used for the biochemical studies is 6.36 μ M or 0.41 mg/ml. Although concentration experiments were performed in earlier preps to analyze whether the protein precipitates upon concentrating, BLM¹¹⁹¹ doesn't precipitate upon concentrating.

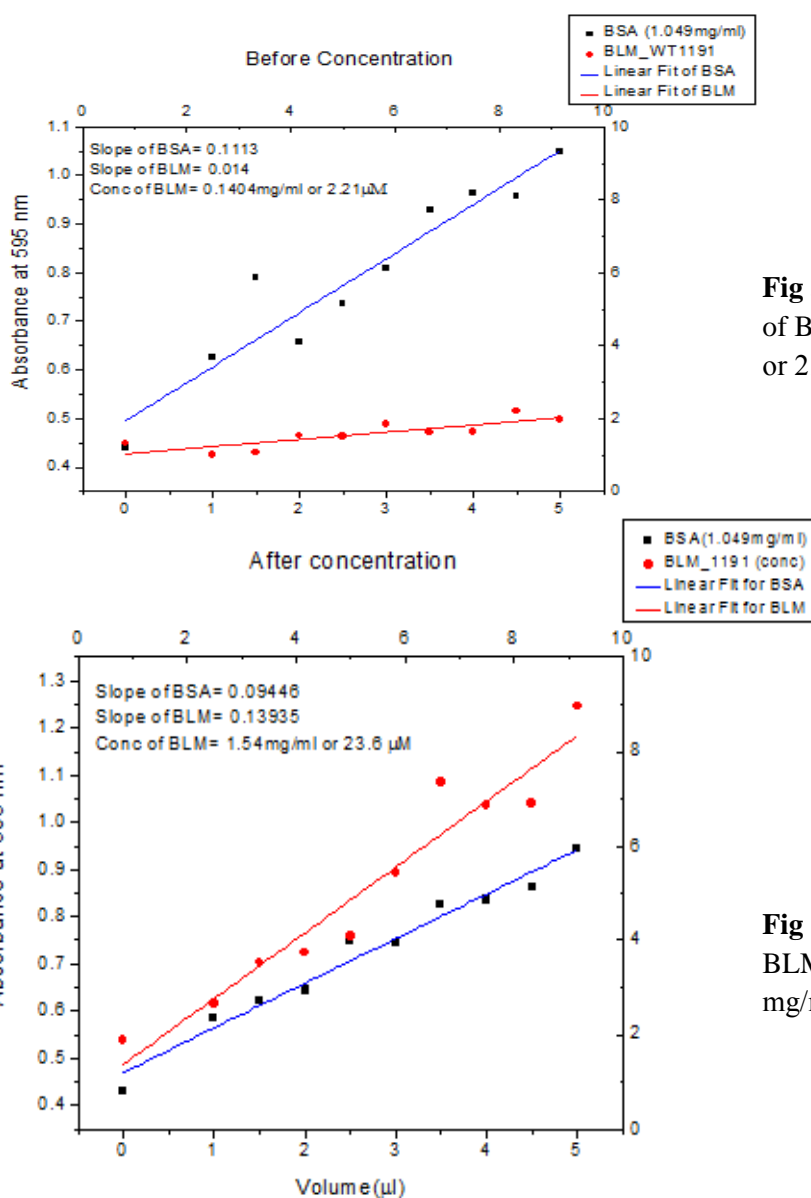


Fig 5.7: Concentration of BLM=0.1404 mg/ml or 2.21 μ M

Fig 5.8: Concentration of BLM after conc = 1.54 mg/ml or 23.6 μ M.

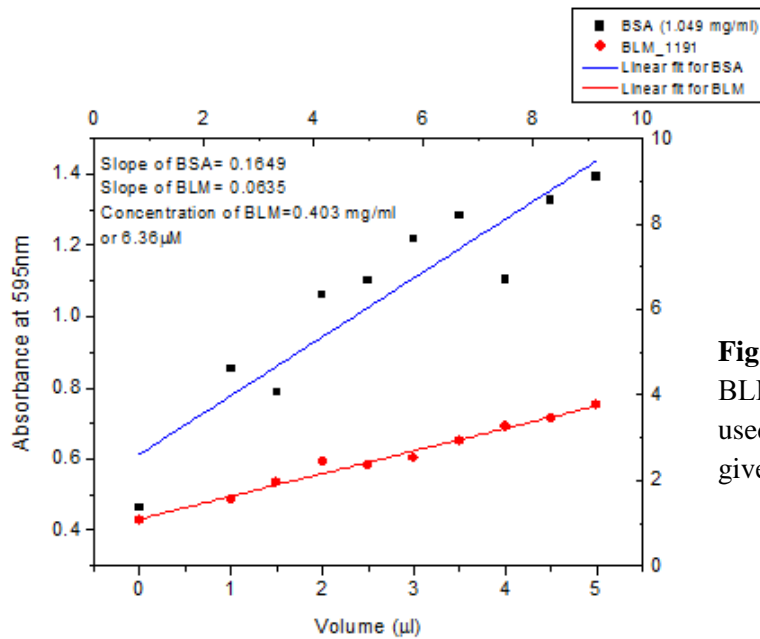


Fig 5.9 Concentration of BLM=6.36 μ M. This concentration is used for all the biochemical tests given in the text.

5.3 Effect of prolonged ice storage on the ATPase activity

Dialysed protein (BLM¹¹⁹¹) was stored in ice for 3 days to check the stability and retention of the ATPase activity. As, presumed that the protein is loses ATPases activity to some extent. The basal activity though remains almost the same the DNA activated activity is altered.

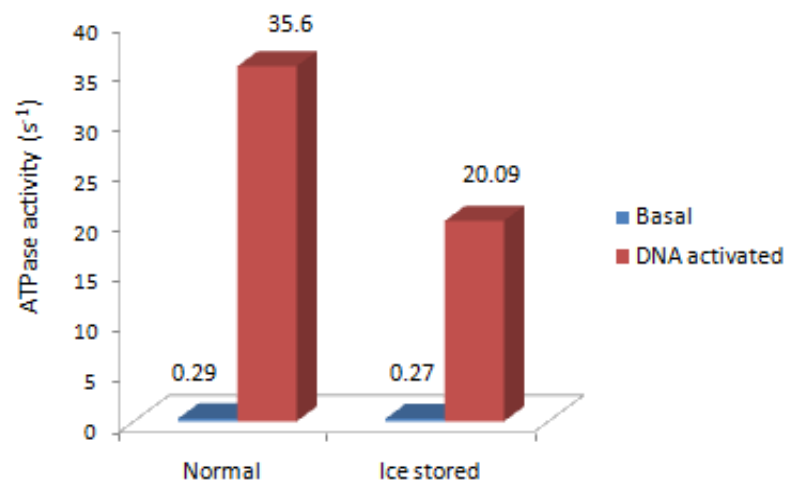


Fig 5.10 Effect of Ice storage on the ATPase activity

The ATPase activity measurement is carried out by the PK/LDH coupled assay as mentioned earlier. The basal ATPase activity of the ice stored protein is almost the same as the normal (0.27 s^{-1} , 0.29 s^{-1} respectively). The DNA activated ATPase activity is lowered from 35.6 s^{-1} to 20.09 s^{-1} in case of the ice stored one. It was found that ATPase activity remains unaltered upon Liquid Nitrogen freezing of the protein after dialysis.

5.4 Effect of Concentration on the ATPase activity

The protein was concentrated by using 2ml Millipore Amicon Centrifugal Ultrafiltration tubes (10K, EW- 2996972) by spinning for 25 minutes for a volume of $500 \mu\text{l}$ at 40°C in table top Eppendorf centrifuge (MiniSpin Plus, Rotor F-4512-11) at 13,400 rpm. The protein didn't precipitate and it was used to measure the ATPase activity by the PK/LDH coupled assay.

The ATPase activity (both basal and DNA activated) remained unaltered upon concentrating the protein upto 15.1 mg/ml or $226.3 \mu\text{M}$ (concentration data not shown in the text).

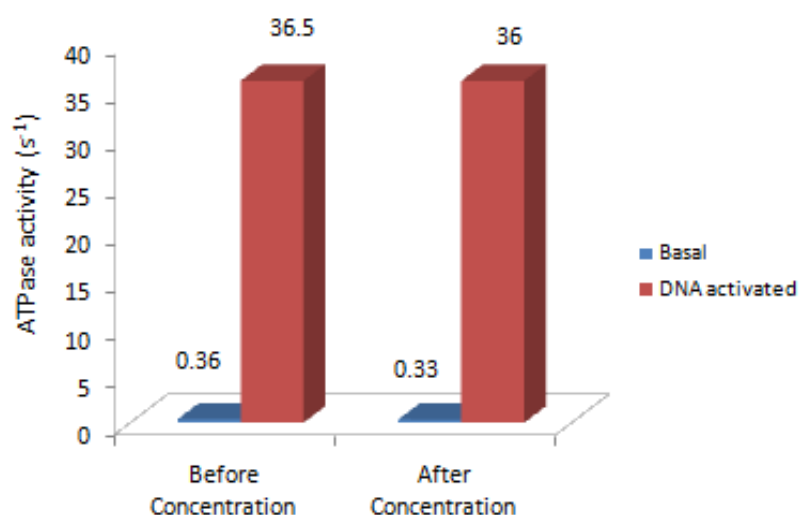


Fig 5.11 Effect of ATPase activity upon concentrating.

5.5 Determination of K_{DNA} values for different DNA lengths

K_{DNA} is the DNA concentration required for half maximal ATPase activation. In order to determine the K_{DNA} values for different lengths of DNA, titration experiments were conducted with the BLM¹¹⁹¹ by PK/LDH coupled ATPase activity measurement assay. The DNA (Oligo -dT) lengths used for the experiments are mentioned in Table 1. The data points of the titration for different concentration of each of the Oligo dT lengths were fitted to a quadratic fit following the equation (Eq-1) given below (Gyimesi et al 2010).

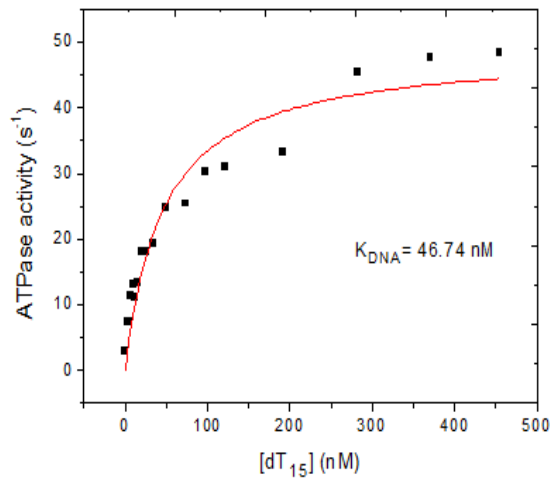
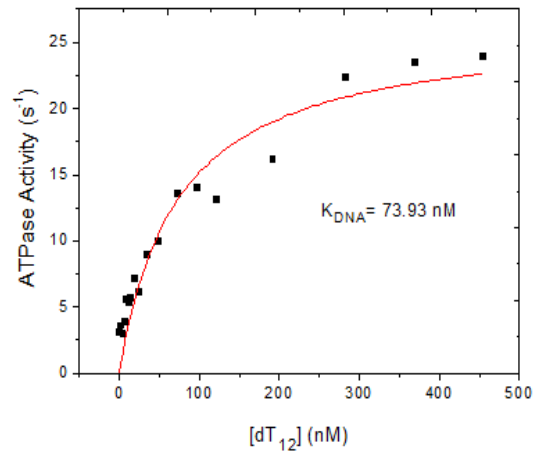
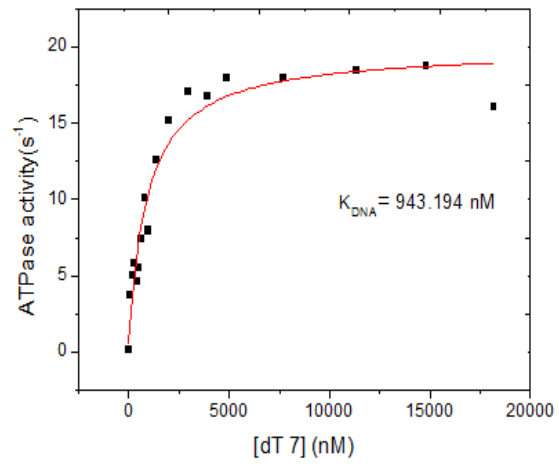
$$y = a \left(\frac{c + x + k - \sqrt{(c + x + k)^2 - 4cx}}{2c} \right) + s \quad \text{..... Eq-1}$$

[Here, a = amplitude (DNA bound fraction), c = concentration of the protein used for the assay; s = basal activity (DNA unbound fraction) and $k= K_{DNA}$].

It is observed that the K_{DNA} value decreases steeply as there is an increase in the Oligo-dT length (**Table 2**). Although on exceeding the binding site size (14 nt), the lowering of the K_{DNA} values is slightly shallow.

Table 3: Oligo-dT length and their respective K_{DNA} values.

Oligo-dT length	K_{DNA} (nM)
7	943.194
12	73.93
15	46.74
23	6.194
54	1.284



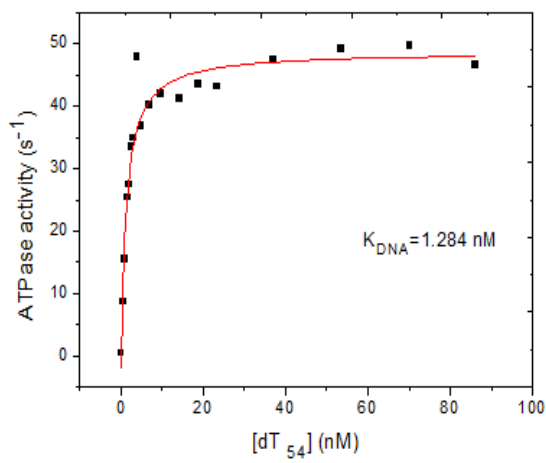
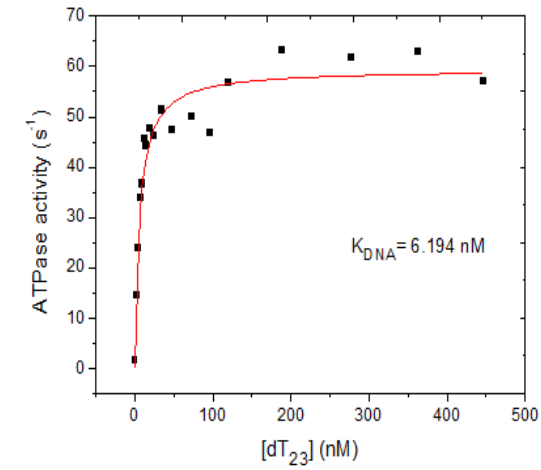


Fig 5.12 Titration curves of different Oligo-dT with BLM¹¹⁹¹ their respective K_{DNA} values as determined by fitting the data points with the Eq 1.

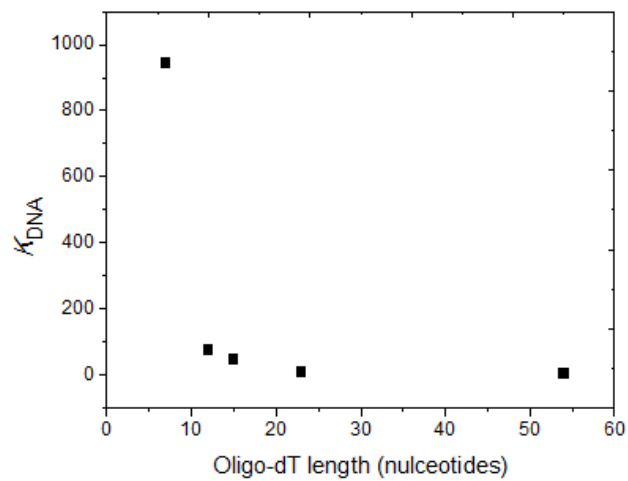


Fig 5.13 K_{DNA} vs Oligo-dT length.

The steady state ATPase activity of the DNA free BLM¹¹⁹¹ (basal ATPase activity) was determined to be 0.33 s⁻¹ as determined by the PK/LDH coupled assay. The activity was heightened to quite an extent by Oligonucleotides (for dT₅₄, k_{cat}= 35.6 s⁻¹) which is in consistence with the results of Gyimesi et al.

BLM¹¹⁹¹ (10nM) was titrated with varying lengths of DNA (oligo-dT- 7,12,15,23,54). As, seen from the K_{DNA} vs Oligo-dT length it is clearly seen that there is a steep decrease in the K_{DNA} value with increasing Oligo-dT length, until the length of the DNA substrates exceeded the binding site size of BLM¹¹⁹¹ (b=14nt). There is a little decrease beyond this range, which is determined by the binding stoichiometry of the BLM¹¹⁹¹ to the DNA substrates. All the measurements were carried out at 25⁰C.

5.6 Determination of k_{cat} and its length dependence

The k_{cat} values can be elucidated from the Eq-1(a+s, i.e the sum of the amplitude and the basal ATPase activity value), but single dose (saturating concentrations of DNA substrates) experiments were performed, where ATPase activity was measure by taking molar excess of the Oligo-dTs (dT 7= 5 μM, dT 12, 15 and 18= 1μM and dT 30-90= 200nM). The k_{cat} values with their respective lengths are given below.

Oligo-dT length (nt)	k_{cat} (s ⁻¹)
7	22.07665
12	33.25038
15	44.96453
18	49.81011
23	58.29575
30	59.22057
54	60.67609
63	62.93916
72	61.26212
79	62.65881

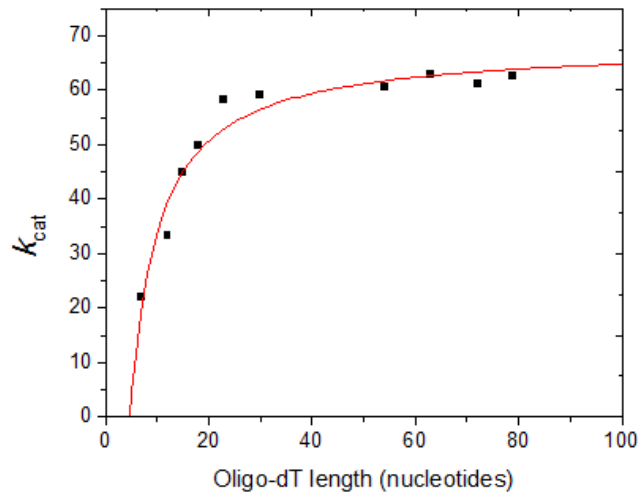


Fig 5.14 k_{cat} vs oligo dT length

The maximal DNA activated ATPase activity (k_{cat}) showed a characteristic length dependent profile. There is a steady rise in the k_{cat} values with increasing oligo-dT length and this gets stabilized at 54 nt range, with maximum value $\sim 62 \text{ s}^{-1}$. The linear increase is an indication of the the ATP consumption rate during the translocation phase, which saturates by $\sim 60 \text{ nt}$. The data points were fitted with an Eq-2:

$$k_{cat} = \left(\frac{\frac{L-b}{2s} \cdot \frac{1}{k_{trans}} + \frac{k_{end}}{k_{off,end}} \cdot \frac{1}{k_{end}}}{\frac{L-b}{2s} + \frac{k_{end}}{k_{off,end}}} \right)^{-1} \quad \text{..... Eq-2}$$

This equation is explained in the translocation model proposed by (Gyimesi et al., 2010). Here $(L-b/2s)$ is the mean number of ATPase cycles performed by BLM at a rate of k_{trans} before it finally reaches to the 5' end. There are 2 possibilities for BLM, either to dissociate from the ssDNA, $k_{off,end}$ or it can revert back for ATPase cycling, k_{end} . ($b=14, s=1$, were kept constant and rest others were allowed to float for the fit).

The k_{trans} value obtained from the fit is $68.42 \pm 16.5 \text{ s}^{-1}$

From the plot it can be seen that the k_{cat} at dT_{15} corresponds with the k_{end} value $14.12 \pm 8.023 s^{-1}$ which significantly corresponds with the earlier results with the wild type BLM indicating the binding site size ($\sim 14nt$). Following this there is steep increase in the k_{cat} values with $\sim 62 s^{-1}$ to be the maximum.

5.7 Single round Translocation Assay

After translocating till the end of the DNA substrate, BLM has a tendency to either dissociate from it or revert back for another ATPase cycle. Single translocation assays can precisely reveal the processivity parameters, key to the translocation activity. For this purpose a Heparin trap was used in which the helicase cannot revert back to bind to DNA, after dissociating from the track.

BLM was pre incubated with saturating concentrations of DNA substrates (Oligo-dT), in a Pi mop and then rapidly mixed with ATP and Heparin in the stopped flow.

50nM BLM¹¹⁹¹ plus 5 μM (for dT_7) / 2 μM (dT_{15-90}) was mixed with 500 μM ATP plus 4mg/ml Heparin in the stopped flow apparatus. The Pi production was monitored by MDCC-PBP fluorescence (Excitation wavelength= 436nm. Emission wavelength= 455 nm).

Pi Calibration of MDCC-PBP fluorescence amplitudes.

A separate calibration experiment was performed where the MDCC-PBP fluorescence was measured for different Pi (0-6 μM) concentration (Pi standard Purchased from Merck), by rapidly mixing 3 μM MDCC-PBP. The slope of this was found to be 0.1508V/ μM Pi which was used for the calculation of released Pi in the Single round translocation experiments.

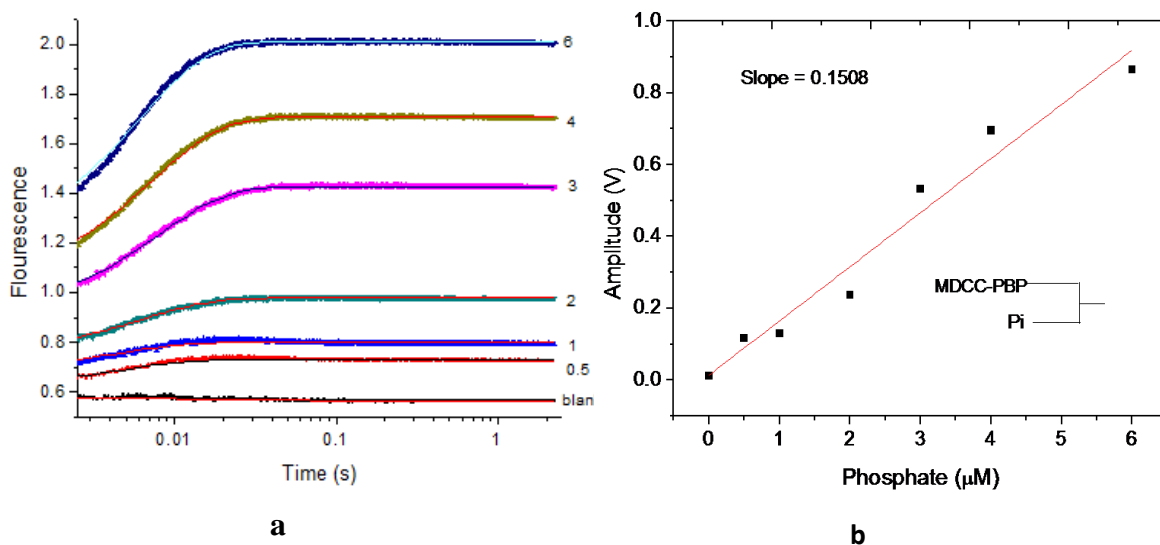


Fig. 5.15 (a) Fluorescence traces of the Pi (0-6 μM , bottom to top) vs time after rapid mixing of Pi standards with 3 μM PBP in a stopped flow system. (b) Amplitude of the fits (of the traces in (a)) vs respective Pi concentrations.

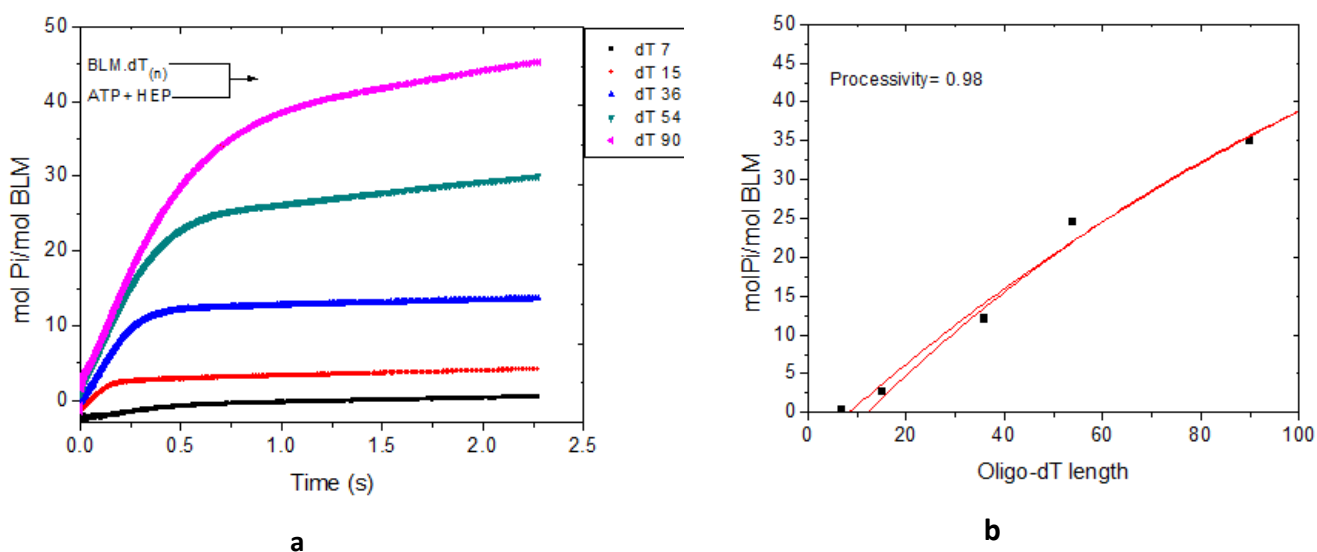


Fig 5.16 (a) Transient kinetics of the Pi production from ATP upon mixing of the 50nM BLM¹¹⁹¹ plus dT 7(50 μM) & dT 15-90 (2.5 μM) with 500 μM of ATP plus 4mg/ml Heparin in a stopped flow monitored by MDCC-PBP fluorescence (3 μM in all syringes).

(b) Oligo-dT length dependence of the Pi production.

MDCC-PBP fluorescence transients showed a multiphasic Pi release profile. Traces with substrates longer than the binding site size shows a rapid exponential burst and two distinct linear phases and the traces with lower dT length lacks the rapid linear phase. The amplitude obtained from the transient kinetic experiments were almost equivalent to the pre-steady state.

The breakpoints of Fig 5.6 (a) were plotted against the corresponding Oligo-dT lengths and fit to the Eq-3. In this the Processivity and the Coupling parameters were left to float and the binding site size was kept at 14 nt.

$$A_{\text{rand}}(L) = \frac{P}{1-P} \left(1 - \frac{1}{C(L-b)+1} \frac{1-P^{C(L-b)+1}}{1-P} \right) \quad \text{.....Eq-3}$$

Here P is the processivity value (0.98) obtained from the plot corresponds to the value given by (Gyimesi *et al.*, 2010). The coupling ratio was found to be 1.2±0.22 which is consistent with the earlier results.

Chapter 6

Discussion

6.1 BLM¹¹⁹¹ Purification

From the purification profiles that have been shown in the earlier chapter, it is evident that the BLM¹¹⁹¹ binds to the CH column efficiently, and the purification by CH column yield BLM¹¹⁹¹ with 90% purity and the remaining 10% comprises of contaminating regions of 10, 40 and 90kDa. The yield after the use of the CH column was on an average 401.4mg/l (~2.4g for 6l). Following the CH column the HP column was used to get rid of the contaminating bands (mentioned earlier). The HP column could successfully remove the contaminating 10kDa region, but the 3 others were still present. For the removal of these contaminating regions a various set of columns were tried after the CH and HP purification steps.

First of these was the CM Sepharose (CM- Carboxymethyl). As, can be seen from the Fig 5.2 the 90kDa and 40 kDa regions is faintly visible and is not efficiently eliminated by the use of CM column.

An additional step was introduced in the purification protocol as given by Gyimesi et al, right after the harvesting of the cells. After homogenisation step, the cell homogenate was left to stand for 1 hour at 4⁰C by mixing 10mg/ml Lysozyme (5ml for ~200 ml of cell homogenate) on a magnetic stirrer. This step was done for the proper lysis of the cells and to have a greater protein yield.

After the unsuccessful trial with the CM column, Affi-Blue gel column was applied for the removal of these contaminating regions. In the first trial with this column, the BLM¹¹⁹¹ fraction was bound to the column but it could not be eluted from the column by a NaCl linear gradient of 10-1.4M. So, the bound fraction was the recovered in the column regeneration step by the use of 1.5M NaSCN. In the 2nd trial we proposed to use a higher NaCl gradient of 500mM- 3.0 M, but no elution was observed as the protein didn't bind to the column (Fig 5.3), may be due to the higher dilution of the BLM¹¹⁹¹ as recovered from the earlier step. In addition to these there were some technical difficulties in the use of this column. As, such we proposed to continue with HA column.

HA Column purification could successfully eliminate the 10kDa and 90 kDa regions, although close observation of the Purification profile of the HA (Fig 5.5) column should that the contaminating 40kDa band was present in the Flow through as well as in the eluted fraction of different columns. As, such an additional washing step (with 10 column volumes of the equilibration buffer) was planned right after loading the (eluted fraction from CH column) fraction on the HP column. This washing step could remove some of the bound 40kDa region although it was not eliminated entirely.

So, in conclusion, purification by the use of HA column with a linear gradient of NaCl (500mM-1.5 M) and additionally double washing step (approximately 20-30 column volumes) before the elution of the bound fraction from both the HP as well as HA column could probably remove the contaminating 40kDa region, leading to BLM¹¹⁹¹ fraction of 99% purity eligible for Crystallization experiments.

6.2 Biochemical activities of the BLM¹¹⁹¹

As, shown by the Fig 5.10, BLM¹¹⁹¹ loses quite a considerable amount of its activity when kept for prolonged ice storage (≥ 3 days), the activity drops down to 20.9 s^{-1} from 35.6 s^{-1} indicating the loss of protein structure, but the concentration experiments showed that there is no alteration in the ATPase activity of the BLM¹¹⁹¹ even after concentration.

6.3 Dependence of the steady-state ATP consumption on ssDNA length

The steady state ATPase activity of the BLM¹¹⁹¹ (DNA free) was found to be 0.33 s^{-1} which is activated by the oligonucleotides substrates (dT₇₋₉₀). Titration experiments performed using 10nM BLM¹¹⁹¹ with varying lengths of the DNA substrates shows that the K_{DNA} value steeply decreases with the increasing length of the oligonucleotides substrates. Comparison of the k_{end} and the corresponding k_{cat} values for the dT₁₅ it gives up similar value which gives an indication of the binding site size. It also gives us other important translocation parameters, $k_{\text{trans}}=68.42 \pm 16.5$, $k_{\text{off, end}}=30.503 \pm 3.43$.

6.4 Transient kinetics

Traces with substrates longer than the binding site size consisted of a rapid exponential burst and two distinct linear phases, while in those with dT₁₅ the first,

more rapid linear phase was lacking. The exponential phase here indicates the first round of ATP hydrolysis by the DNA bound BLM¹¹⁹¹.

Chapter 7

Conclusions

The purification trials for BLM¹¹⁹¹ suggested that the with the additional optimized steps that has been included in this work to the protocol followed Gyimesi et al, could possibly yield BLM with 99% homogeneity, which would be the key to the crystallization experiments.

The protein retains all the necessary translocation parameters and also the deletion of some of the regions from the wild type protein increases the efficiency of the protein.

The further work in addition to the work produced here will include obtaining the protein of crystallization grade following the up gradation to the purification protocol made here and would focus on co-crystallization of the DNA and the BLM to obtain a DNA bound crystal structure of the BLM protein. This could help us not only to understand the basic mechanism on which these motors work but also delve deep in to the participation if these proteins in the homologous recombination pathway and involvement of these in the maintenance of the genomic stability.

Appendix

Principle of Stop Flow apparatus (KINTEK SF-2004)



The principle behind a stopped flow apparatus is simple. This apparatus uses the drive motor to rapidly fire two solutions, contained in separate drive syringes, together into a mixing device. The solutions then flow into the observation cell

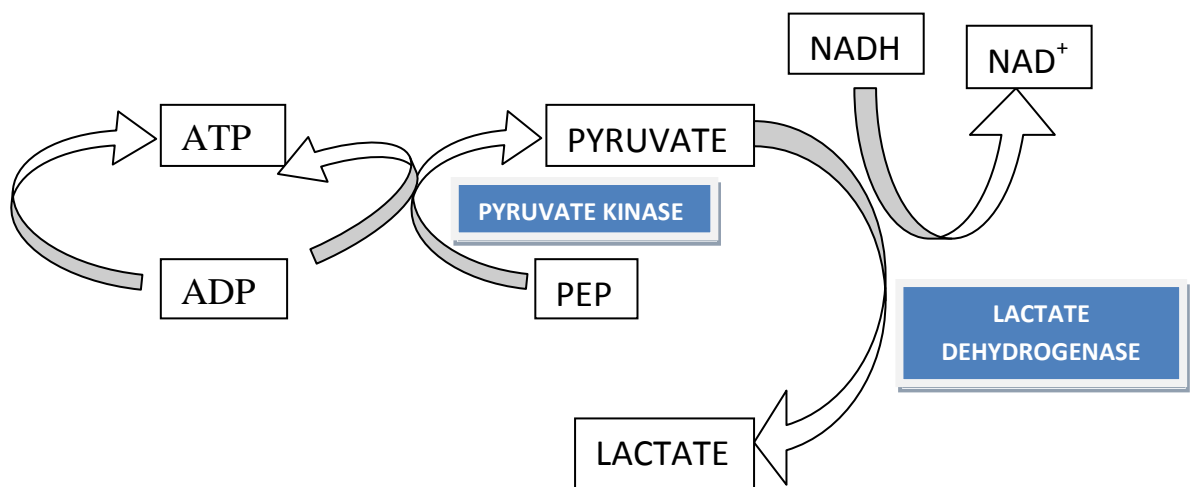
displacing the previous contents with freshly mixed reactants. A stop syringe is used to limit the volume of solution expended with each experiment and also serves to abruptly stop the flow. The flow of solution into the stop syringe causes the plunger to move back and trigger data collection. The fresh reactants in the observation cell are illuminated by a light source and the change, as a function of time, in many optical properties (Absorbance, Fluorescence, Light Scattering, Turbidity, Fluorescence Anisotropy etc.) can be measured. The measurement of these optical properties is performed by the system's detectors, which can be mounted either perpendicular or parallel to the path of incoming light depending on in which optical property we are interested.

In this more advanced system there are three drive syringes. The first two drive syringes drive two solutions together through a mixer and into a delay line. The length of the delay line along with how long it takes the solutions to pass through the delay line determines the length of the first reaction. After passing through the delay line the first two solutions (now mixed) enter a second mixer where they are mixed with a third reactant and then the mixed solutions pass into the flow cell. The reactants

displace the previous contents of the observation cell and trigger data collection as described.

The time resolution of this method is limited by the time required for the reactants to flow from the final point of mixing to the observation cell. This time is referred to as the dead time of the instrument. The SF-2004 series of Stopped-Flows have guaranteed dead times of less than 2 milliseconds. For all stopped-flows, the volume of the mixer, observation cell and flow lines all affect the volume of reactants that must be pushed to obtain a satisfactory reaction. The SF-2004 series of stopped-flows utilizes micro volume flow cells, lines and mixers.

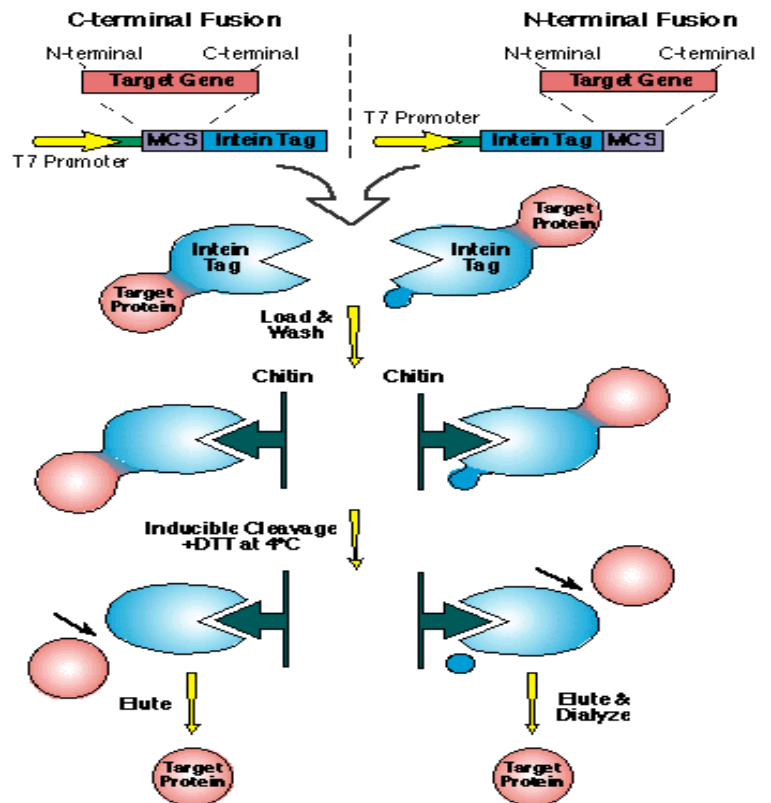
Schematic Diagram of the PK/LDH coupled assay



Principle of Protein purification by Chitin Beads

Intein Mediated Purification with an Affinity Chitin-binding Tag utilizes the inducible self-cleavage activity of engineered protein splicing elements (termed inteins) to purify recombinant proteins by a single affinity column.

Expression vectors (pTYB, pTXB) which allow fusion of a bifunctional tag, consisting of the intein and the chitin binding domain, to either the C-terminus or N-terminus of the target protein. In the presence of thiols such as DTT or β -mercaptoethanol the intein undergoes specific self-cleavage which releases the target protein from the chitin-bound intein tag.



pTXB contain a *SapI*, *NcoI*

cloning site which allows the target gene to be cloned immediately adjacent to the cleavage site of the intein tag; this results in the purification of a target protein.

References

1. Bachrati, C. Z., Hickson, I.D. (2003). RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem. J.* 374:577–606.
2. Bennett, R.J., Keck, J.L. (2004) Structure and function of RecQ DNA helicases. *Crit Rev Biochem Mol Biol*;39(2):79-97.
3. Bernstein DA, Keck JL. 2005. Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure* 13:1173–82.
4. Bernstein, D.A., Zittel, M.C. and Keck, J.L. (2003) High-resolution structure of the E.coli RecQ helicase catalytic core. *EMBO J.*, 22, 4910-4921.
5. Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M., Stillman, B. (2000) A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A.* 97(4):1530-5.
6. Cordin, A., Tanner, N.K., Doère, M., Linder, P and Banroques, R.(2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J.* 2004; 23(13): 2478–2487
7. Gorbalenya,A.E., Koonin,E.V. and Wolf,Y.I. (1990) A new superfamily of putative NTPbinding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.*, 262,145-148.21.
8. Guo, R.B., Rigolet, P., Ren, H., Zhang, B., Zhang, X.D., Dou, S.X., Wang P.Y., Gueret, M. and Xi, X.G. (2007) Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. *Nucleic Acids Res.*, 35, 6297-6310.
9. Gyimesi, M., Sarlós, K., Kovács, M. (2010) Processive translocation mechanism of the human Bloom's syndrome helicase along single-stranded DNA. *Nucleic Acids Res.* 38(13):4404-14.
10. Heyer,W.D. (2004) Damage signaling: RecQ sends an SOS to you. *Curr. Biol.*, 14, R895-R897.
11. Hickman, A.B. and Dyda, F. (2005) Binding and unwinding: SF3 viral helicases. *Curr. Opin. Struct. Biol.*, 15, 77-85.
12. Hickson, I.D. (2003). RecQ helicases: caretakers of the genome. *Nat. Rev. Cancer* 3:169–78.
13. Karow, J.K., Newman,R.H., Freemont,P.S. and Hickson,I.D. (1999) Oligomeric ring structure of the Bloom's syndrome helicase. *Curr. Biol.*, 9, 597-600.
14. Kim,J.L., Morgenstern,K.A., Griffith,J.P., Dwyer,M.D., Thomson,J.A., Murcko,M.A., Lin,C. And Caron,P.R. (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure.*, 6, 89-100.
15. Korolev S, Yao N, Lohman TM,Weber PC,Waksman G. (1998). Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicase. *Protein Sci.* 7:605–10.
16. Korolev,S., Hsieh,J., Gauss,G.H., Lohman,T.M. and Waksman,G. (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell*, 90, 635-647.

17. Levin, M. K. & Patel, S. S. (2003) in *Molecular Motors*, ed. Schliwa, M. (Wiley-VCH, Verlag GmbH, Weinheim, Germany) pp. 179–198.
18. Liu, Z., Macias, M.J., Bottomley, M.J., Stier, G., Linge, J.P., Nilges, M., Bork, P. and Sattler, M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure*, 7:1557–1566.
19. Liu, J.L., Rigolet, P., Dou, S.X., Wang, P.Y. and Xi, X.G. (2004) The zinc finger motif of *Escherichia coli* RecQ is implicated in both DNA binding and protein folding. *J. Biol. Chem.*, 279, 42794–42802.
20. Liu, Z., Macias, M.J., Bottomley, M.J., Stier, G., Linge, J.P., Nilges, M., Bork, P. And Sattler, M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure.*, 7, 1557–1566.
21. Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A. and Sung, P. (2005) Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair (Amst)*.
22. Matson, S.W., Bean, D.W. and George, J.W. (1994) DNA helicases: enzymes with essential roles in all aspects of DNA metabolism. *Bioessays*, 16, 13–22
23. McDaniel, L.D. and Schultz, R.A. (1992) Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. *Proc. Natl. Acad. Sci. U. S. A*, 89, 7968–7972.
24. Patel, S.S., and Picha, K.M., (2000) Structure and Function of Hexameric Helicases *Annu. Rev. Biochem.* 2000. 69:651–97.
25. Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, 76, 23–50
26. Tuteja, N. and Tuteja, R. (2004) Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery. *Eur. J. Biochem.*, 271, 1835–1848.
27. Velankar, S.S., Soutanas, P., Dillingham, M.S., Subramanya, H.S. and Wigley, D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, 97, 75–84.
28. Von Hippel, P.H. and Delagoutte, E. (2001) A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, 104, 177–190.
29. Waksman, G., Lanka, E. and Carazo, J.M. (2000) Helicases as nucleic acid unwinding machines. *Nat. Struct. Biol.*, 7, 20–22.
30. Wilson, S., Warr, N., Taylor, D.L. and Watts, F.Z. (1999) The role of *Schizosaccharomyces pombe* Rad32, the Mre11 homologue, and other DNA damage response proteins in non homologous end joining and telomere length maintenance. *Nucleic Acids Res.*, 27, 2655–2661.

31. Wu L, Chan KL, Ralf C, Bernstein DA, Garcia PL. (2005). The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J.* 24:2679–87.
32. Wu, I., and Hickson, I.D. (2006) DNA Helicases Required for Homologous Recombination and Repair of Damaged Replication Forks. *Annu. Rev. Genet.* 2006.40:279-306.
33. Xu, H.Q., Deprez, E., Zhang, A.H., Tauc,P., Ladjimi, M.M., Brochon, J.C., Auclair,C. and Xi,X.G. (2003) The Escherichia coli RecQ helicase functions as a monomer. *J. Biol. Chem.*, 278, 34925-34933.
34. Yao,N., Hesson,T., Cable,M., Hong,Z., Kwong,A.D., Le,H.V. and Weber,P.C. (1997) Structure of the hepatitis C virus RNA helicase domain. *Nat. Struct. Biol.*, 4, 463-467.