New Protein Systems for Homologous Recombination-based DNA Engineering in Bacteria

CHEN, Wenyang

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

Biochemistry

The Chinese University of Hong Kong

July 2010

Thesis Assessment Committee

Prof. CHEN Zhen Yu (Chair)

Prof. HO Wing Shing (Thesis Supervisor)

Prof. NG Tzi Bun (Committee Member)

Acknowledgements

I would like to extend my appreciation to Prof. John Wing Shing Ho and Dr. Rory Munro Watt who has been an outstanding mentor in my PhD study. Their guidance and support have been invaluable in my research. I am grateful for their comments during the preparation of the draft.

I would like to address my appreciation all the other members of my advisory committee, Prof. Chen Zhenyu and Prof. Ng Tzi Bun.

Table of contents

List of Figures

List of Tables

Abstract for thesis entitled:

New protein systems for homologous recombination-based DNA engineering in bacteria

Submitted by CHEN, Wenyang

for the degree of Doctor of Philosophy in Biochemistry

at The Chinese University of Hong Kong (July 2010)

Recombineering is a powerful tool used to manipulate or engineer DNA *in vivo*, which enables chromosomes and plasmids to be modified precisely and efficiently. It is of critical importance for research into genome and proteome function, and greatly facilitates metabolic engineering applications. The Lambda-Red (Bet and Exo) and RecET proteins constitute the most efficient bacterial recombineering systems characterized to date. However, they work only in *E. coli* or closely-related bacteria (e.g. *Salmonella* spp.), which limits their widespread application.

The Lambda-Red and RecET recombineering systems can use PCR products (double stranded DNA, dsDNA) or single stranded DNA (ssDNA, oligonucleotides) to create precise point mutations (substitutions), gene deletions and insertions in chromosomal or episomal DNA. The Exo/RecE exonuclease proteins digest dsDNA and produce long 3'-ssDNA tails. The Bet/RecT ssDNA annealing proteins bind to these 3'-ssDNA tails and promote their homologous recombination with complementary ssDNA regions on the chromosome or episome.

Novel pairs of Bet/Exo recombineering proteins were identified in the beta-proteobacterium *Laribacter hongkongensis* (LHK) and in the SXT genetic element isolated from *Vibrio cholerae*. In this research, these new recombineering proteins were functionally characterized using a variety of *in vivo* and *in vitro* techniques. The SXT-Exo and LHK-Exo proteins were both found to be alkaline exonucleases, with activities similar to those of Lambda-Exo. Both the SXT-Bet/Exo and LHK-Bet/Exo protein pairs had dsDNA recombination activity within *E. coli*.

The ssDNA recombination activities of five different Bet/RecT recombinases were directly compared using an *E. coli* reporter system. The comparison revealed that Bet protein from LHK had a higher efficiency than Lambda-Bet or RecT. Based on their predicted secondary structure, a set of rationally-designed lambda-Bet protein truncations were prepared and their biological activity was examined, to investigate structure-function relationships within this recombinase.

The results described in this thesis will be very useful in assisting the future development of novel recombineering systems that can be used for genetic engineering applications across a wide range of bacterial organisms. Such tools will greatly promote functional genomic and proteomic studies within these organisms, and may also be used for microbial engineering and biotechnological applications.

摘要

参与细菌内基于同源重组的 DNA 工程的新蛋白质系统的研究

重组工程在体内对染色体和质粒的精确有效修饰起着重要作用，因而它对于基因组功能， 蛋白质组和代谢工程的研究都起着关键作用。作为当今最有效的重组工程共功能组，Red 和 RecET 这两个来自大肠杆菌 *E.coli* 却只能在大肠杆菌或与大肠杆菌相关的细菌中发挥作用， 这大大限制了重组工程的应用。

Red 和 RecET 可以用单链 DNA 或双链 DNA 精确地在染色体上引入点突变、DNA 缺失和 DNA 插入。Exo 或 RecE 消化双链 DNA 产生 3'-单链 DNA， Bet 或 RecT 结合 3'-单链 DNA 并引导它与染色体进行同源重组。

本论文从结核弧菌的 SXT 遗传因子和 *Laribacter hongkongeis* 中开发了两个重组功能组， 其中来自 *Laribacter hongkongeis* 的重组功能组比原先的 Red 和 RecET 更有效率。本论文也深入研究了重组功能组的组成蛋白 Exo 和 Bet. 其中两个 Exo 都是碱性外切核酸酶，而 Bet 蛋白都有重组活力， 特别是 LHKBet 具有很高的活力。本论文也深入讨论了 Bet/RecT 蛋白的结构功能之间的关系， 并对它们的多聚态进行研究，发现 Bet/RecT 的多聚态非常复杂， 并且跟所在环境有莫大的关系。

本论文开发了新的高效同源重组功能体， 这对开发能够应用到更广大范围的细菌种类的重组功能体具有相当大的作用， 这也将能够大大促进基因功能组、微生物技术和药物生物学的研究

Chapter One Introduction

Part One. Recombinant DNA technology *in vitro*

Traditional recombinant DNA cloning procedures generally refer to the process in which a linear double stranded DNA (dsDNA) fragment is inserted into a plasmid (vector) which enables replication within a host (bacterial) cell. First developed by Herbert Boyer and Stanley Cohen in the early 1970s, this technology requires sequence-specific endonucleases (restriction enzymes), DNA ligase and plasmids (containing a selectable marker) that can replicate within a host (bacterial) cell [1].This powerful technology enables the DNA fragment (e.g. gene) of interest to be amplified within the host cell, enabling it to be functionally investigated, or for its encoded protein to be over-produced and studied.

In standard cloning operations, there are five steps to obtain the recombinant plasmid construct [2] : 1) the dsDNA fragment is cut by two sequence-specific endonucleases (restriction enzymes) at defined sites; 2) preparation of linearized plasmid DNA (most commonly containing an antibiotic resistance gene as a selectable marker), by cutting it with restriction enzymes that generate 'compatible' or 'sticky' DNA ends; 3) covalent ligation of the two DNA molecules using DNA ligase to produce a circular 'recombinant' plasmid molecule; 4) introduction (e.g. transformation) of the recombinant plasmid DNA into the host cells (most commonly the bacterium *Escherichia coli*), where the recombinant plasmid will replicate using the machinery provided by the host cells; 5) plating onto selective solid media, where cells that harbor the recombinant plasmid containing a selectable marker can establish colonies, and hence be identified and isolated.

The restriction enzymes and vector are most important to this traditional restriction-ligation cloning technology. Restriction enzymes can recognize and cut the dsDNA at specific DNA sequences, producing a set of smaller DNA fragments for ligation. There are three types of DNA restriction enzymes: type I enzymes cut the DNA randomly and the cutting sites can be far away (>1000bps) from the recognition site; type III enzymes cut the DNA at the sites near (~25bps) the recognition sites. Type II restriction enzymes cut the DNA at specific sites (often palindromic) that they recognize and bind to. Type II restriction enzymes are the most useful for recombinant DNA technology, and thousands of these enzymes have been isolated and characterized since the first restriction enzyme (HindII) was purified by Hamilton Smith in 1970 [3]. The cuttings of Type II restriction enzymes produce blunt ends (i.e. leave no overhanging nucleotides after cutting) or sticky ends (i.e. form a short 5'- or 3'-single stranded DNA overhang). If the restriction enzyme recognizes an N base-pair sequence, it will cut the DNA once per N^4 base pairs in theory. Consequently, the shorter the recognition sequence, the shorter the DNA products formed after cutting. However, in natural systems recognition sequences tend to occur less frequently than would be predicted. Homing endonucleases are a special class of endonucleases. Their recognition sites are longer than common type II restriction enzymes, so they can produce longer products after digestion of long DNA molecules (e.g. chromosomal DNA).

Plasmids, bacteriophages and bacterial artificial chromosomes are the most popular cloning vectors for DNA cloning with *E. coli*. Plasmids are widely found throughout the bacterial kingdom. Some are specific to a particular species, or small number of closely related species. Others, known as broad host range plasmids, can establish themselves within a diverse range of different bacterial species. Naturally found bacterial plasmids are generally covalently closed circular DNA molecules, normally 1500 to 40000 base pairs in length. They may be single copy (i.e. only one plasmid molecule present per cell); low copy (<ca. 20 plasmid molecules per cell); medium copy (ca. 20 – 100) or high copy (>100).

However, there are many commercial plasmids (vectors) available for various purposes on the market (e.g. pET vectors), many of which have high copy numbers (> 500 per cell). However, it is generally very difficult to clone long DNA fragments (>15000bp) in to a plasmid vector.

An alterative strategy for working on long DNA fragment involves the use of bacteriophages. Bacteriophages are dsDNA viruses that generally insert themselves into bacterial chromosomes (as lysogens) during a phase in their growth cycle, which is then replicated along with the hosts chromosomal DNA. Insertion points may be random or specific.

Bacteriophage λ is the most well studied one, and has been well used as a cloning vector. It has a long linear dsDNA genome (ca. 48.5 kbp), which contains 12 base-pair complementary cohesive ends that anneal to each other. After entry into the cell, these are sealed by DNA ligase (to create a cos site), forming a stable circular molecule. It has been further modified and developed for commercial use. In a typical cloning procedure, it is cleaved into three pieces, two pieces contain the minimal essential regions of DNA (required for packaging, replication and integration), and the third piece can be substituted by the inserted DNA. This method allows up to 23000 base-pairs DNA fragments to be inserted. Bacterial artificial chromosomes are designed for the cloning of very long DNA fragments, normally 100,000-300,000 base pairs. To enable uptake of large DNA molecules such as bacterial artificial chromosome, there must be some active uptake mechanism, or the cell wall of the host bacterium must be made transiently permeable e.g. by chemical treatment or electroporation.

This traditional restriction-ligation DNA recombinant technology has enabled molecular biology to evolve into a new era of genetic engineering. More and more restriction enzymes and vectors become available every year.

The advance of DNA sequencing technology provides ever-increasing amounts of genomic sequence information for bacteria, archaea and eukaryotes, which makes genomic function studies possible. Vectors such as P1 artificial chromosomes and bacterial artificial chromosomes have been developed to accommodate more than 100,000bp of foreign DNA, which greatly facilitates genomic studies within higher organisms. However, it is very difficult to clone very large DNA (megabases) via traditional restriction-ligation technology. There are several reasons for this: 1) it is very difficult (almost impossible) to find two unique recognition sites within the insert DNA and vector since the insert DNA is so long; 2) the large DNA is very fragile and it is very difficult to physically perform restriction-ligation operations. Recently developed *in vivo* procedures involving homologous recombination circumvent many of these technical difficulties, and offer numerous other advantages, providing an alternative strategy for engineering DNA molecules.

Part Two Bacterial DNA exchange and repair via *homologous recombination*

Homologous DNA recombination refers to the exchange or 'swapping' of sections of DNA between two DNA duplexes (dsDNA molecules), via regions that are identical or share high levels of homology with one another. All organisms use homologous DNA recombination to repair and/or replace sections of their chromosomes that have been damaged or improperly processed within the cell (e.g. UV damage, stalled replication forks).

There are at many proteins in *E. coli* involved in homologous recombination, such as: RecA, RecBCD, RecE, RecF, RecG, RecJ, RecN, RecO, RecQ, RecR, RecT, RuvAB, RuvC, PriA, SSB proteins, DNA polymerases, DNA topoisomerases and DNA ligase, etc. Among these, RecA and RecBCD are the best studied, and are the most important for repairing damaged DNA within the cell. (Reviewed by Kuzminov A

and Kowalczykowski SC et al)[4, 5]. Homologues of these proteins are highly conserved throughout the bacterial kingdom. RecA is the major homologous recombination-promoting protein in *E. coli*. The RecB, RecC and RecD subunits, encoded by *recB*, *recC* and *recD* genes respectively, form a protein complex known as the RecBCD enzyme, which has DNA helicase and exonuclease activities. The RecBCD complex binds to the free ends of broken dsDNA and then processively unwinds (by helicase activity) and digests (by exonuclease activity) the dsDNA as it moves along the DNA double helix, in a process that requires ATP. As RecBCD slides down the dsDNA, it encounters and interacts with *chi* sequences (5'-GCTGGTGG-3') which are scattered throughout the genome. The degradation velocity of the 3'-strand DNA is greatly reduced while 5'-strand DNA degradation is increased. This process creates a long 3'-single strand ssDNA end which is important for homologous recombination. RecA binds the 3'-single stranded DNA, forming a DNA-protein filament, then searches for its complement sequence, and pairs with it. Not less than 100 base pairs of homologous DNA region are required for efficient recombination. [6,7]

Part Three Homologous DNA recombination proteins encoded on bacteriophage lambda and the *rac* prophage of *E. coli*

During its normal life cycle, bacteriophage lambda repairs and resects multimeric chromosomal intermediates (concatamers) via homologous DNA recombination processes using the activities of three of its encoded proteins; Gam, Bet and Exo, which together are known as the Red proteins [4]. In *E. coli*, these proteins may act in concert with the RecA protein, but do not require it for full functionality. The *gam*, *bet* and *exo* genes are positioned next to each other in the P_L operon of lambda phage, and are co-transcribed and translated (Figure 1). In bacteriophage lambda, the P_L operon is under the control of cI repressor and N protein encoded upstream. A temperature sensitive mutant of the cI repressor (cI857 or cIts857) has been

developed [8]. By binding to the operator sequence, cI857 represses transcription of the PL operon at temperatures below ca. 35°C, but dissociates at temperatures above ca. 38°C, de-repressing the transcription from the strong PL promoter.

The Exo protein binds to the ends of linear dsDNA and digests it in the 5'→3' direction in a highly processive manner (4, 5, 6), producing 5'-mononucleotides and long 3'-single stranded DNA ends. Exo remains bound to the dsDNA after digestion, while the Beta protein, a single stranded DNA annealing protein, binds to the 3'-single stranded DNA region and promotes the annealing of this 3'-ssDNA end to its complementary sequence, starting the homologous recombination process. The Bet-ssDNA nucleoprotein filament cannot invades an intact DNA duplex (unlike the RecA nucleoprotein filament), but has strand displacement activity. Consequently, it can only mediate base pairing with complementary (partially) single-stranded DNA regions; i.e. near a replication fork, or at a DNA end. This is known as the single strand annealing pathway [4].

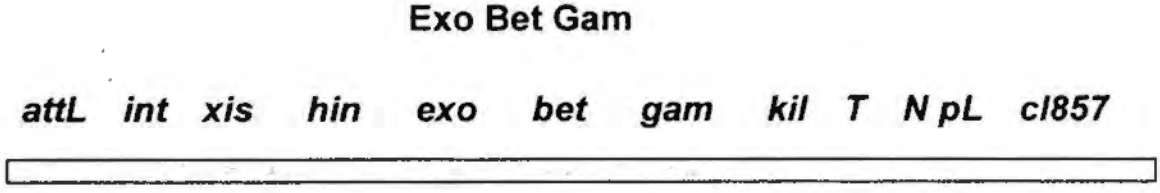


Figure1. The *pL* operon of prophage

Although there are about only 10 molecules of RecBCD in one cell, this multimeric protein complex degrades the majority of linear dsDNA molecules in *E.coli*. Similar to RecBCD, SbcCD also has nuclease activity, and degrades linear DNA and lowers the homologous recombination efficiency [9]. The Gam protein binds to SbcCD & RecBCD and inhibits their nuclease activity, thereby protecting the dsDNA from degradation and increasing the recombination activities of lambda Exo-Beta.

The RecE (ExoVIII) and RecT proteins encoded on the defective Rac prophage of *E. coli* can also mediate homologous pairing and strand exchange, analogous to the Bet

and Exo proteins from bacteriophage lambda [5,10]. In the rac prophage, the *recT* and *recE* genes are both in the same operon, with *recT* lying immediately upstream of *recE*.

Like Bet, RecT is a single-stranded annealing protein, whilst RecE is an exonuclease. Although RecE (Exonuclease VIII) has very similar activities to Lambda Exo, it is much larger (866 vs. 226aa) and shares very little sequence homology. Limited proteolysis studies have indicated that only the ca. 300 amino acids at the C-terminus are required for exonuclease activity, and for *in vivo* dsDNA recombination activities (with recT) [11, 12]. Linear dsDNA digested by RecE produces a 3'-single strand DNA ends to which RecT binds. Analogous to lambda Bet, the binding of RecT to ssDNA promotes its annealing with homologous regions of (partially) ssDNA, initiating the homologous recombination process. There are more sets of recombination function derived from various sources.

Part Four The development and application of recombineeing technologies in E. coli using lambda-Red and recET

In the late 1990, several pioneering groups were simultaneously developing technologies by which the Gam-Bet-Exo proteins of bacteriophage lambda (Lambda-Recd) and the RecE-RecT rac-encoded proteins (RecET) could be used to promote efficient homologous recombination within *E. coli*. [13, 14, 15, 16] This leads to the creation of both plasmid-based, and chromosome-integrated systems utilizing the Lambda-Red or RecET proteins to promote efficient and specific genetic exchange between introduced dsDNA molecules and the host chromosome or episomal molecules.

Of these, the most widely-used chromosomal system is *E. coli* DY380, developed by Donald Court and co-workers at the NIH [15]. This engineered strain contains a

defective lambda prophage chromosomally-integrated at the *attB* site, which encodes the gam-bet-exo genes under the tight control of the cI857 temperature sensitive repressor (Figure 1). A. Francis Stewart and co-workers at the EMBL developed two arabinose-inducible plasmids: one encoding the Lambda *bet, exo* and *gam* genes (pBAD-αβγ), and one encoding the *recE, recT* and *gam* genes (pBAD-ETγ) [16].

Largely due to the pioneering work of these groups, homologous recombination functions from bacteriophage lambda (Red) and the rac prophage (RecET) are now widely used for *in vivo* genetic engineering in *E. coli*, in a process generally referred to as recombineeing – which is short for *recombino*genic engin*eering*. This involves using PCR-prepared dsDNA molecules or synthetic ssDNA oligonucleotides to encode the desired genetic alteration. Both chromosomal and episomal DNA can be altered using recombineering in specific strains of *E. coli* (e.g. DY380, or *E. coli* DH10B housing a (pBAD-αβγ or pBAD-ETγ plasmid). This strategy can be used to create (gene) deletions, insertions, replacements or point mutations (nucleotide substitutions). To have good recombination efficiencies, the dsDNA molecules or oligonucleotides must be introduced (transformed) into the *E. coli* cells as efficiently as possible, which is generally achieved via electroporation.

For workable efficiencies, the dsDNA 'targeting molecules' must contain two regions at least 35-50bp inlength, whose sequence is identical to the regions immediately flanking the target DNA locus (section of DNA to be altered). This is most conveniently encoded on (two) long oligonucleotide primers used to PCR amplify the dsDNA targeting molecules. Usually, a selectable marker, e.g. an antibiotic resistance gene is required, and marker can be included in the dsDNA targeting molecule (dsDNA targeting cassette). Two short Frt or loxP sites may be positioned at either side of the selection marker, so that it may subsequently be removed by transiently expressing the Flp or Cre recombinase proteins, respectively[17]. In one example, recombineering and Cre-LoxP and Flp-Frt technology was used in *E. coli* to create chromosomal gene fusions to fluorescent

protein (such as GFP, YFP, RFP) genes (11). This enabled the resultant protein fluorescent fusions, which were expressed under native conditions, to be localized within the cell using fluorescent microscopy.

Recombineering can also be used for 'cloning' very large molecules e.g. biosynthetic gene clusters or large intronic eukaryotic gene regions into plasmids, cosmids, bacterial artificial chromosomes (BACs) or P1 artificial chromosome (PACs). 'Recombinogenic cloning' using 'gap-repair' is the most popular method [18,19]. The cloned DNA fragment can be up to 80 kbp in length by this method. Importantly, it is not necessary to find sequence-specific restriction sites, which is almost 'mission impossible' in an 80 kbps DNA fragment. The principle behind the gap-repair method, is that a PCR-generated dsDNA fragment is recombined *in vivo* with a linearized episome (e.g. plasmid), in a process mediated by the lambda Red or RecET proteins, to form a circular molecule. As the linearized plasmid is unable to replicate in the cell, the recombinant clones formed should result from the insertion of the dsDNA fragment into the plasmid. The molecule to be cloned is PCR amplified using oligonucleotides that have 35-50bp homology to the termini of the linearized plasmid (or BAC). In practice, the dsDNA fragment and linearized vector are co-electroporated into cells that express recombineering functions (e.g. DY380). The homologous flanking regions in the dsDNA fragment are partially degraded and 3'-ssDNA are produced. Then the 3'-ssDNA regions anneal with the homologous regions on the vector. By strand exchange, the vectors are repaired to form closed circular DNA.

Oligonucleotide-mediated recombineering may be used to introduce point mutations, repair point mutations or to create deletions on the chromosome, or on episomal DNA molecules such as plasmids or BACs. This procedure only requires the activity of the RecT or Bet protein (in addition to host DNA replication and repair proteins). Furthermore, it is generally more efficient than dsDNA recombination (up to 20% efficiencies have been reported in mismatch repair-deficient strains), and hence does

not require the use of a selectable marker [20, 21]. In Oligonucleotide-mediated recombineering, the synthetic oligonucleotide has a sequence that is complementary to that of either strand of the DNA region being targeted. The encoded mismatch or deletion must correspond to the approximate centre of the oligonucleotide, and there must be at least 30-45bp of homology on either side. Due to size and efficiency limitations, only small insertions may be encoded on synthetic oligonucleotides. However, large (tens of kbp) deletions can be created using oligonucleotide-directed recombineering.

The oligonucleotide (that directs the DNA alteration) is incorporated into episome or chromosome during replication process [20, 22]. Consistent with this 'annealing integration' model, oligonucleotides whose sequences correspond to the lagging strand in replication have higher correction efficiencies than their complements.

Since recombineering may be conveniently used to modify episomal DNA as well as for cloning large fragments of DNA, it is widely used for functional genomic studies in other organisms. In one example, the *Xenopus tropicalis Arx* gene fused with *GFP* gene was cloned into a bacterial artificial chromosome in *E. coli* via recombineering. After isolation and purification, the modified BAC was then injected into *X. laevis* embryos and was expressed in the forebrain [23]. In another example, the *C. briggsae* gene *lin-59* was cloned into bacterial artificial chromosome via recombineering, then was introduced into *Caenorhabditis elegans*, resulting in successful expression of *lin-59* [24].In mouse genomic functional studies, recombineering was used to create the *Sall4* Mutant Mouse Model, through the generation of a mutant construct housed on a bacterial artificial chromosome [25].

Part Five The biochemical and biophysical properties of the Red and RecET recombineering proteins

1.5.1. Gam protein from bacteriophage lambda

The Gam protein has a molecular weight of 16 kDa and its 3-D structure has been determined by X-ray crystallography [26]. The protein forms a dimeric structure with two protruding N-terminal helices (Figure 2). Upon binding to RecBCD, Gam

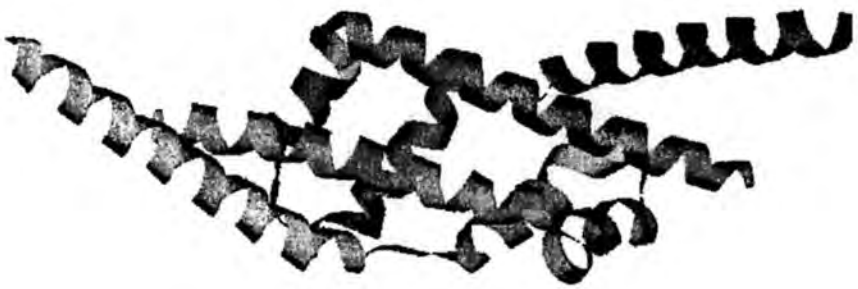


Figure 2. The structure of Gam. (modified from RCSB PDB, ID: 2UV1)

significantly inhibits nuclease and helicase activities of RecBCD [27, 28,29]. Further research has shown that the Gam protein inhibits the binding of RecBCD to dsDNA [24]. Models of inhibition have been proposed that Gam mimics either single strand DNA or dsDNA and competes for binding to RecBCD with dsDNA substrate [26].

1.5.2. Exo protein from bacteriophage lambda

Bacteriophage lambda Exo has a molecular weight of 24 kDa. Exo binds the ends of dsDNA and degrades one strand in 5'→3' direction in the presence of Mg^{2+}. It cannot initiate degradation at ssDNA breaks or ssDNA gaps [30]. The digestion is processive (Exo remains bound on the dsDNA until digestion finishes) and the rate is ca. 1000 base pairs per second [31, 32]. Exo digestion produces 5'-mononucleotides

and long 3'-single strand DNA ends. Digestion may occur at both ends of the dsDNA substrate [33, 34]. Lmabda Exo also has 3'→5' dsDNA assimilation activity; however, this activity is much slower than its 5'→3'digestion activity [35]. Lmabda Exo degradation is highly processive, which indicates that the protein topography closely matches that of the DNA substrate, and that the Exo-DNA interaction is strong enough to disfavor dissociation [36]. It is classed as an 'alkaline exonuclease', as it has optima's activity at pH values between 8 and 9.

The active form of Exo is a homotrimer that has a toroidal (ring shaped) structure (figure 3) [37, 38]. The Exo trimer has a central channel which is wide enough to

A B

Figure 3.The structure of bacteriophage lambda Exo (A: viewed from narrow diameter side B: viewed from wide diameter side Structure is modified from RCSB PDB ID: 1avq.)

accommodate dsDNA at one end (diameter ~ 30 Å), but can only accommodate ssDNA at the other end (diameter 15 Å). However, this x-ray structure did not include DNA, and so the precise nature of DNA binding and hydrolysis could only be modeled. It has been proposed that the Exo toroid acts like a bead on a 'thread' of dsDNA. It slides down the dsDNA thread while degrading one strand, leaving

3'-single stranded DNA in its walk, until reaches the end of the DNA molecule or it dissociates from the substrate dsDNA.

A divalent metal ion, ideally Mg^{2+} is required for exonuclease activity [34]. Mn^{2+} can be used to substitute Mg^{2+}, although there is a four-fold reduction in activity. The metal binding sites were identified after soaking the Exo crystal with $MnCl_2$. The divalent metal ion is chelated by the side chains of Asp119 and Glu129, and the backbone carbonyl of Leu130 [37, 38].

The ends of the substrate DNA significantly affect the dsDNA exonuclease activity of Exo. In the degradation of linear dsDNA, the relative k_{cat} for linear dsDNA substrates with different ends decreases in the order: 5'-recessed > blunt >> 5'-overhang [36]. This may be caused by several reasons: lambda Exo binds to the dsDNA after it assembles into a toroidal structure, but it cannot assemble on the end of dsDNA. The ends of linear dsDNA may interact with Exo and affect its initiation of degrading dsDNA. A 5'-overhang may hinder the binding of Exo and hence hinder the initiation of digestion sicne lambda exo digest dsDNA in a direction of 5' to 3'. When Exo binds to a 3'-overhang, it may more readily interact with the 5'-phosphate and start the degradation [36]. Research has shown that the velocity for Exo digests 5'-phosphorylated linear dsDNA (5'-P dsDNA) is more than 2-fold faster than for dephosphorylated linear dsDNA (5'-OH dsDNA), which means that the recognition of 5'-phosphate plays an important role in Exo digestion. Further research demonstrated that Arginine-28 (R28) is necessary for the 5'-phosphate group recognition at the dsDNA end, which affects the Exo catalytic efficiency. Although Exo can bind to 5'-OH dsDNA and form a relatively inert complex, the digestion is limited [39].

1.5.3. RecE protein from the *Rac* prophage of *Escherichia coli K-12*

The RecE protein (exonuclease VIII) is a linear dsDNA-dependent exonuclease. It has been reported that the active form of RecE is a trimer or tetramer [40]. The RecE protein has 866 amino acid residues corresponding to a molecular weight of 98 kDa. However, the apparent MW is about 140 kDa as determined by SDS-PAGE [41], which has been attributed to anomalous migration behavior [42, 43].

Similar to lambda Exo, RecE degrades linear dsDNA in the 5'→3' direction, creating 3'- single stranded DNA ends. It also has little activity in degrading linear single strand DNA, but has no activity on closed circular dsDNA, or dsDNA nicking [41]. However, unlike lambda Exo, RecE degrades 3'-overhang, 5'-overhang and blunt-ended linear dsDNA equally well. Genetic and molecular analysis have shown that the only the C-terminal domain is essential for RecE activity [43]. Further structure-function research shows that the truncated protein that only has C-terminal 302 amino acid residues have the same level activity as full length RecE. This C-terminal domain is reasonably resistant to the digestion of trypsin and subtilisin, which indicates that the domain is well packed [44]. The function of the N-terminal domain (1-564 amino acid residues) remains unknown.

The x-ray crystal structure of the catalytically-competent C-terminal domain of RecE has recently been solved [45]. The active part of the RecE protein crystallizes as a toroidal tetramer, which forms a tapered central channel analogous to that of the lambda Exo. A related alkaline exonuclease protein from phage SPP1 (Chu, G34.1P) has also been studied *in vitro* [46,47].A number of other Exo-like proteins have been identified in other phages, or bacterial prophages, alongside partnering Bet or RecT-like proteins. Some of these recombinase-exonuclease pairs have been shown to mediate homologous recombination processes analogous to Lambda-Red and RecET (Table 1) [48].

1.5.4. Bet protein from bacteriophage lambda

Bet protein binds to single strand DNA annealing [49] and its molecular weight is 25.8 kDa. The binding of Bet to single strand DNA covers more than 35 nucleotides [50] and protects the ssDNA from nuclease degradation [51,52]. In homologous

Table1.Recombineering with different recombinase–exonuclease pairs

Resource	Recombinase	Exonuclease
E.coli	Bet	Exo
E.coli	RecT	RecE
B. subtilis	GP35	GP34.1
L. monocytogenes	Orf48	Orf47
L. pneumophila	OrfC	OrfB
P. luminescens	plu2935	plu2936

recombination, Bet directs the annealing of ssDNA to its complementary sequence in a Mg^{2+}-dependent process, and has strand displacement activity. Unlike RecA, Bet cannot promote the invasion of an intact DNA duplex [53]. Bet cannot bind dsDNA, but remains bound to the dsDNA product after the annealing process.

In the presence of Mg^{2+} ions, Bet spontaneously assembles into a large ring-like structure, with a diameter of about 145 Å, and a central hole of 35 Å, which is estimated to comprise 12 subunits. In the presence of oligonucleotides, the Bet protein binds these oligonucleotides and forms larger rings with diameters of about 185 Å with central holes of 75 Å in diameter. When long ssDNA molecules are available, Bet protein forms large rings and left-handed helical filaments. The large rings are about 210 Å in diameter and the central hole is about 100 Å in diameter. The left-handed helical filaments diameters are 200 Å, similar to those of large rings [54]

A varity of *in vitro* experiments have indicated that residues 1-130 of the Bet protein form a 'core' structural domain. This domain is resistant to proteolytic digestion by subtilisin in the absence of DNA, and consequently it was proposed to be the ssDNA binding domain. Upon binding to ssDNA, residues131-177 have enhanced resistance to subtilisin digestion. An ssDNA binding model was proposed that the N-terminal domain (residues 1-131) and central domain bind ssDNA together, while the disordered C-terminal has no obvious role in DNA binding [55].

1.5.5. RecT protein from *Rac* prophage of *Escherichia coli K-12*

RecT is a single strand DNA annealing protein of 269 amino acid residues, its molecular mass is 30 kDa. Although it shares low homology at the amino acid level with lambda Bet[56], RecT is functionally equivalent to the Bet protein, and forms a specific binding interaction with the RecE exonuclease[12]. RecT only has dsDNA recombination activity in partnership with RecE, and does not associate with, nor function with, the lambda Exo exonuclease protein. RecT binds to the 3'-ssDNA ends [57] that produced by RecE degradation, and protects at least 13 nucleotides [58], and promotes its paring to complementary ssDNA regions. RecT has strand displacement activity, and unlike lambda Bet, it has been shown to mediate strand invasion to a certain extant. Unlike Beta protein, RecT also binds dsDNA with an affinity that is higher than that for ssDNA binding [59].

1.5.6. Single strand DNA binding protein (SSB) from *Escherichia coli*

The SSB protein from *Escherichia coli* is a non-specific ssDNA binding protein of 177 amino acid residues, with a molecular mass of 18.8 kDa. It was first purified by Bruce Alberts in the late 1960s [60, 61]. The protein is homotetramer in the absence of DNA, and the tetramern structure is very stable across a wide range of salt concentrations. The structure of SSB tetramers was determined by Raghunathan et al in 1997(Figure 4)[63]. Two SSB subunits form a dimer, and one tetramer consists of

two dimmers; with an interface formed between the two dimers. One of the interfaces stabilizes the tetramer and the other contributes the formation of a superhelical structure that may be involved in the binding to single strand DNA [63].

The N-terminal domain of SSB is involved in ssDNA binding, while the C-terminal domain facilitates tetramerization [64]. The SSB tetramer binds to ssDNA and covers a fragment of DNA whose length changes from 35 to 65 nucletotides, depending upon conditions, e.g. temperature, pH and salt concentrations [65]. This suggests two binding modes: in the first mode, the ssDNA interacts with only two SSB subunit in each tetramer, binding to 35 nucleotides; in the second mode, it interacts with all four subunits in each tetramer, binding to 65 nucleotides [66,67]. These two modes can interconvert, with the transition determined by environmental factors and the protein/ssDNA ratio [68, 69]. At higher salt concentrations SSB tends to adopt the second mode, whilst it tends to adopt the first mode at lower salt concentration [70, 71]. SSB binds to DNA in a cooperative mood. In the first binding mode, the cooperativity is unlimited, while in the second binding mode the cooperativity is limited [70]. In DNA replication, SSB stimulates helicase activity [72]. It binds to ssDNA and prevents ssDNA from annealing to each other and inhibits nuclease digestion [73]. It also promotes the binding of DNA polymerase [74] and enhances its fidelity [75, 76] and processivity [77, 78].

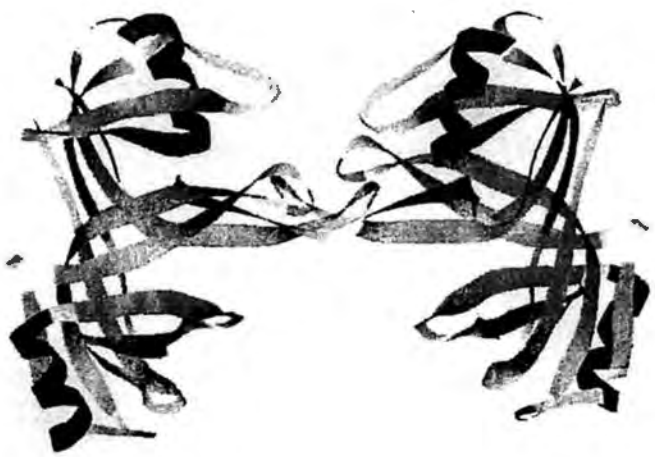


Figure 4.The structure of SSb from *E.coli.*
(Structure is modified from RCSB PDB ID: 1kaw)

As SSB is an ssDNA binding protein, it competes for ssDNA binding with RecA in homologous recombination. With the help of the RecBCD protein, RecA can displace SSB bound to ssDNA [79]. Unbound SSB can stimulate DNA strand exchange that is mediated by RecA. The stimulatory effect can take place in both of presynaptic and postsynaptic phases. SSB binding removes the secondary structure of DNA in the presynaptic phase, and prevents the ssDNA strands separated by the RecBCD helicase from re-annealing to each other [80].

Part six Introduction for SXT element

Waldor MK reviewed SXT research recently. Integrating conjugative elements (ICE) are transposon-like genetic elements that can self-transmit amongst prokaryotes. As such, they are important mediators of horizontal gene transfer [81]. This is particularly important as many encoded antibiotic and heavy-metal resistance genes. Although ICEs have some features similar to plasmids and phages, they cannot replicate by themselves. However, they insert themselves into the host chromosome (like bacteriophage lysogens) and replicate with the host chromosome, using the host-encoded replicative machinery. Since the first ICE, Tn916, was identified in *Enterococcus Faecalis* in 1980, many ICEs have subsequently been identified in Rhizobiaceae (α-proteobacteria) Burkholderiaceae (β-proteobacteria) and in a large number of γ-proteobacteria [82, 73].). SXT is an ICE found in *Vibrio cholerae*, the causative agent of the severe diarrheal cholera, and is responsive for the spread of antibiotic resistance in this species [84, 85]. It carries resistance to sulfamethoxazole (Su), trimethoprim (Tm), chloramphenicol (Cm), and streptomycin (Sm). A pair of toxin-antitoxin (mosA and mosT) is found to promote the maintenance of SXT in *Vibrio cholerae* [86].

El Tor O1 *V. cholerae* has an ICE, ICE*Vch*Ind1, which is similar to SXT from *V. cholerae* O139, but the two ICEs are not identical. Comparative DNA sequence

analysis has indicated that they arose from the same precursor [87]. pJY1, a genetic element isolated from *V. cholerae* O1 in 1977, carring Su, Cm, and Sm resistance, may be the immediate precursor of SXT [88]. Not all SXT-related ICEs in *V. cholerae* O139 have antibiotics. ICE*Vch*HKo1, an ICE isolated from *Vibrio cholerae* O139, has no antibiotic resistance genes [89]. Over the past few decades, SXT-related ICEs have been spread in Asia, in various bacterial species and genera [90, 91]. Although it was initially isolated from the recently-emerged O139 serogroup, it is now present in virtually all clinical *V. cholerae* isolates from the Indian sub-continent [92]. SXT-related ICEs also have been isolated in Mexico from an environmental *V. cholerae* isolate [93], and in Spain from the Vibrio-related fish pathogen *Photobacterium damselae* subsp. *Piscicida* [94]. The original (historical) host of the SXT ICE remains to be determined.

Although SXT-related ICEs are often associated with *Vibrio cholerae,* it may not be the major host. A genetic element derived from a South African isolate of *Providencia rettgeri*, R391, is functionally and genetically related to SXT. R391 mediates kanamycin resistance and mercury resistance [95]. The SXT and R391 ICEs both integrate into the host chromosome at the same site, the 5'-end of *prfC*. Furthermore, the *int* gene, essential for the integration of both SXT and R391, is almost identical in both ICEs [86]. Genomic comparisons have shown that R391 and SXT are closely related [97, 98, 99].

The transfer of SXT between bacterial host cells requires three steps: 1) excision from the host chromosome and circularization; 2) conjugative transfer to the recipient cell; 3) integration into the chromosome of the new host. Numerous proteins regulate the transfer process. Although *prfC* encodes release factor 3 (RF3), it can only express active protein after the integration of SXT, which provides a promoter for *prfC* as well as a novel 5'-coding sequence that is fused to the prfC protein-encoding gene, which activates it [87].The mechanism of SXT excision and integration are RecA-independent, and are thought to be similar to those of

bacteriophages. The integration site (*attP*) on SXT is a 17bp sequence that is nearly identical to the integration site (*attB)* on chromosome, which facilitates the homologous recombination [87].

SXT is excised from chromosome at the *attL* and *attR* sites via homologous recombination, and forms closed circular dsDNA. SXT encodes various Tra proteins that form the conjugative apparatus (pilus), which transfers the circularized SXT molecule to the recipient cell. The genes encoding the Tra proteins are arranged in four clusters on the SXT genome, which span more than 25kbps. Genes encoding proteins for initiation and DNA process are located in the first cluster; genes encoding cell paring and conjugative function proteins are located in the other three clusters. Genetic analysis of Tra genes from different ICEs shows that they are conserved in SXT-related ICEs [98, 99]. The recipient cell cannot accommodate redundant ICEs, and the redundant ICEs are excluded by eexS protein, which is mediated by TraG protein [99].

The regulation of SXT transfer protein expression is somewhat similar to those of bacteriophage lambda and is controlled by SetR. In the basal state, SetR binds to the operon between *setR* and *s086* (Figure 5) and represses the expression of of SetC and SetD proteins. After RecA is activated by DNA damaging agents, activated RecA (RecA*) can stimulate the autoproteolysis of SetR and release the P_L promoter. So SetC and SetD proteins are expressed. SetC and SetD activate the expression of *int* and *tra* genes and trigger the SXT transfers [100]. All SXT-realted ICEs carry *setR*, *setC* and *setD* genes, which suggests they share the same regulation mechanism.

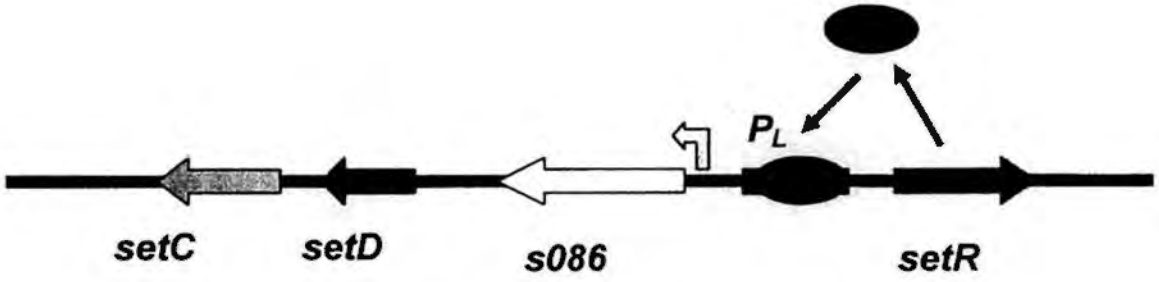


Figure 5. SXT regulation gene expression

SXT-related ICEs share a similar backbone and have a conserved core that encodes proteins for excision, integration and transfer [98, 99]. Besides the conserved core, there are also highly variable genes, such as the antibiotic resistance genes, which have considerable variation in their locations and sizes in SXT, ICE*Vch*Ind1 and R391.There are also four 'hotspots' in the backbone, where diversified insertion genes of SXT-related ICEs genes are located. This indicates that their insertion does not disrupt the overall stability, showing that ICEs are rather flexible [98].

The genomic organization and open reading frames of SXT have been revealed [98]. There are homologues of bacteriophage *lambda exo* (*s066*), *bet* (*s065*) genes, and an *E. coli ssb*-like gene (*s064*) in the SXT element. These three genes are possible to compose a homologous recombination function, they are located in the same open reading frame (Figure 6) and have been cloned into the pJB1 plasmid by the Waldor group at Harvard medical School. The *s066* (SXT *exo*) 1017bp gene sequence is

ATGAAGGTTATCGACCTATCACAACGTACTCCTGCATGGCACCAGTGGCGCATTGCAGGGGTTACG
GCATCTGAAGCCCCAATTATTATGGGGCGTTCACCCTACAAAACACCTTGGCGATTATGGGCAGAA
AAAACTGGATTCGTATTACCGGAAGACCTGTCGAATAATCCTAATGTACTTCGCGGTATAAGGTTGG
AGCCTCAAGCAAGGCGAGCATTTGAGAATGCGCATAATGACTTTCTTCTGCCGTTATGTGCAGAAGC
CGATCATAACGCAATCTTTCGAGCCAGCTTTGATGGCATCAACGATGCGGGCGAGCCCGTTGAACT
GAAATGTCCTTGCCAGTCAGTTTTTGAGGATGTGCAAGCTCACCGAGAACAAAGCGAGGCGTACCA
GTTGTATTGGGTGCAAGTACAGCATCAAATACTGGTCGCCAATAGCACGCGAGGTTGGTTGGTTTTC
TATTTTGAGGATCAACTGATTGAGTTTGAAATACAACGAGACGCGGCGTTCTTAACTGAGTTGCAAG
AAACAGCGCTTCAGTTTTGGGAGTTAGTACAGACCAAAAAAGAACCGTCAAAATGCCCTGAGCAAG
ATTGTTTTGTTCCCAAGGGTGAAGCCCAATACCGTTGGACATCGCTGTCTCGACAGTATTGCTCAGC
ACATGCCGAAGTGGTCCGACTGGAAAATCACATTAAATCTTTGAAAGAGGAAATGCGAGACGCTCA
GTCAAAATTGGTCGCCATGATGGGTAACTACGCTCATGCCGACTATGCTGGGGTCAAACTCAGTCGC
TACATGATGGCGGGCACGGTGGACTATAAGCAATTGGCCACCGATAAATTAGGCGAGCTGGATGAA
CAGGTTTTAGCCGCTTACCGAAAAGCGCCACAAGAGCGGTTGCGTATCAGCACCAATAAGCCAGAG
CAGCCCGTTGAAACACCAATCAAAATCAGCCTTGAGCAAGAGAACTTGGTTCTGCCAGGTGACTCG
CCGAGCTCATTTTATTTTTAA

Figure 6.The operon of SXT *s064*, *s065* and *s066*

SXT Exo has 338amino acid residues, 38.843 kDa, and its sequence is,

KVIDLSQRTPAWHQWRIAGVTASEAPIIMGRSPYKTPWRLWAEKTGFVLPEDLSNNPNVLRGIRLEPQA
RRAFENAHNDFLLPLCAEADHNAIFRASFDGINDAGEPVELKCPCQSVFEDVQAHREQSEAYQLYWVQ
VQHQILVANSTRGWLVFYFEDQLIEFEIQRDAAFLTELQETALQFWELVQTKKEPSKCPEQDCFVPKGEA
QYRWTSLSRQYCSAHAEVVRLENHIKSLKEEMRDAQSKLVAMMGNYAHADYAGVKLSRYMMAGTV
DYKQLATDKLGELDEQVLAAYRKAPQERLRISTNKPEQPVETPIKISLEQENLVLPGDSPSSFYF

The SXT Exo amino acid sequence has been aligned with bacteriophage lambda Exo (Figure 7). The alignment shows that the similarity of SXT Exo to lambda Exo, which indicates SXT exo is possibly an alkaline exonuclease.

The *s065* (SXT bet) gene has 819bp, and its sequence is as following,

ATGGAAAAACCAAAGCTAATCCAACGCTTTGCTGAGCGCTTTAGTGTCGATCCAAACAAACTGTTC
GATACCCTAAAAGCAACAGCATTTAAGCAACGTGACGGTAGTGCACCGACCAATGAGCAGATGATG
GCGCTCTTGGTGGTTGCAGATCAGTACGGCTTGAACCCTTTCACCAAAGAGATTTTTGCGTTCCCTG
ATAAGCAAGCTGGAATTATTCCAGTGGTAGGTGTCGATGGATGGTCTCGCATCATCAATCAACACG
ACCAGTTTGATGGCATGGAGTTTAAGACTTCAGAAAACAAAGTCTCCCTGGATGGCGCGAAAGAAT
GCCCGGAATGGATGGAATGCATTATCTACCGGCGCGACCGTTCGCACCCAGTCAAAATCACTGAAT
ACCTGGATGAAGTCTATCGACCGCCTTTTGAGGGTAACGGAAAAAATGGCCCTTACCGTGTAGATG
GTCCATGGCAGACGCACACTAAGCGAATGCTAAGACATAAATCCATGATCCAGTGTTCCCGCATTG
CGTTTGGCTTTGTGGGAATTTTCGATCAAGACGAAGCGGAGCGAATTATCGAAGGCCAAGCAACAC
ACATTGTTGAGCCATCGGTGATTCCACCCGAGCAAGTTGATGATCGAACCCGAGGGCTTGTTTACAA
GCTTATCGAGCGGGCGGAAGCTTCAAACGCATGGAATAGTGCATTGGAATACGCCAATGAACATTT
TCAAGGTGTTGAACTGACGTTTGCGAAACAAGAAATATTTAATGCACAGCAACAAGCAGCCAAAGC
GCTCACACAGCCTTTAGCTTCTTAG

SXT Bet has 272 amino acid residues, 31.041 kDa, and its sequence is:

MEKPKLIQRFAERFSVDPNKLFDTLKATAFKQRDGSAPTNEQMMALLVVADQYGLNPFTKEIFAFPDKQ
AGIIPVVGVDGWSRIINQHDQFDGMEFKTSENKVSLDGAKECPEWMECIIYRRDRSHPVKITEYLDEVYR
PPFEGNGKNGPYRVDGPWQTHTKRMLRHKSMIQCSRIAFGFVGIFDQDEAERIIEGQATHIVEPSVIPPEQ
VDDRTRGLVYKLIERAEASNAWNSALEYANEHFQGVELTFAKQEIFNAQQQAAKALTQPLAS

SXT Bet has been aligned with bacteriophage lambda Bet (Figure 8). The alignment shows the similarity of SXT Bet to bacteriophage lambda Bet, which suggests that SXT Bet is highly likley to have activities similar to lambda Bet.

The *s064* (SXT ssb) gene has 420bp, and its sequence is,

ATGAAAAACCAAGTAACACTCATAGGCTATGTTGGCTCTGAGCCAGAGACGCGAGCCTATCCATCA
GGTGATTTAGTGACCAGCATTTCACTGGCCACTTCTGAGAAATGGCGCGACCGTCAATCCAATGAGC
TCAAAGAGCATACGGAATGGCATCGGGTCGTTTTTCGAGATCGTGGTGGATTAAAGTTAGGGCTCA
GGGCAAAAGATTTAATCCAAAAAGGAGCGAAGCTTTTTGTTCAAGGGCCTCAGCGCACGCGCTCAT
GGGAGAAAGATGGCATTAAGCATCGATTGACCGAAGTGGACGCGGACGAGTTTCTGCTTCTTGATA
ATGTGAACAAAGCATCTGAGCCATCAGCGGCGGATGATGCAGGCTCCCAAACTAATTGGGCACAAA
CTTATCCTGAACCAGATTTTTAA

SXT SSB has 139 acid residue sequences, 15.742 kDa, and its sequence is:

MKNQVTLIGYVGSEPETRAYPSGDLVTSISLATSEKWRDRQSNELKEHTEWHRVVFRDRGGLKLGLRA
KDLIQKGAKLFVQGPQRTRSWEKDGIKHRLTEVDADEFLLLDNVNKASEPSAADDAGSQTNWAQTYPE
PDF

SXT SSB is aligned with *E.coli* SSB protein (Figure 9) and shows the similarity of *E.coli* SSB and SXT SSB, which suggests that SXT SSB is possible to be *E.coli* SSB-like protein.

Figure 7. The alignment of SXT Exo with bacteriophage lambda Exo using T-coffee

Figure 8. The alignment of SXT Bet with bacteriophage lambda Bet using T-coffee

Figure 9. The alignment of SXT SSB with *E.coli* SSB protein using T-coffee

Part Seven Introduction to *Laribacter hongkongesis*

Laribacter hongkongesis was first isolated in Hong Kong, from blood and empyema (pleral pus) samples taken from a patient with alcoholic liver cirrhosis[101]. The bacteria are seagull/spiral rod shape and are gram-negative, anaerobic and non-sporulating[101]. *Laribacter hongkongesis* is susceptible to imipenem and ciprofloxacin, but is resistant to ampicillin and tetracycline. It can grow on sheep blood agar and MacConkey agar at temperatures between 25 and 42°C. Its chromosome is about 3169 kbps in size, with a G+C content of 62.35 %[102]. Phylogenic analysis shows it is a member of *Neisseriaceae* family of the β-subclass of proteobacteria[101]. *Laribacter hongkongesis* has also been isolated from Chinese tiger frog[103], fish[104,105] and freshwater reservoirs in Hong Kong and in Hangzhou, China[106]. It was proposed to be a potential cause of traveler's diarrhea[107].

In order to survive in the different temperatures and environments, *Laribacter hongkongesis* regulates different expression of proteins such as transporters (especially the multidrug efflux pumps and heavy metal transporter), urease, bile salts efflux pump, adhesin, catalase, proteins involved in chemotaxis, etc[102]. Most striking is the different expression of two isozymes of N-acetyl-L-glutamate kinase — NAGK-20 and NAGK-37 at different temperatures. NAGK-20 and NAGK-37 are two isozymes of NAGK. The NAGK-20 has high expression at 20°C and low expression at 37°C while NAGK-37 has high expression at 37°C and low expression at 20°C[95], although they have the same function. Phylogenic study shows that they are more closely related to each other than to NAGK proteins in other bacteria. The two NAGKs topology in phylogenic tree is similar to 16sRNA, which suggests that the two isozyme genes evolved after *Laribacter hongkongesis* had separated from closely related bacteria, and was subjected to different evolutionary pressures[102].

The complete genome of *Laribacter hongkongesis* was recently reported [102].There are homologues of the bacteriophage lambda *exo* (LHK *exo*) and bet (LHK *bet*) genes, as well as a homologue of the *ssb* gene from *E.coli* (LHK *ssb*). These three genes are located next to each other, within the same operon (Figure 10), which suggests that they may have homologous recombineering functions.

Figure 10. The operon of LHK *ssb*, *bet* and *exo*

The LHK *exo* gene has 718bp, and its sequence is:

ATGACCATGGAACAACGCACCGAAGAATGGTTTGCCGCCCGCCTCGGCAAGGTGACCGCCAGCCGT
GTGGCCGATGTGATGACAAAAACCAAGTCCGGCTATGCAGCCAGCCGGCAAAACTACATGGCCGA.
GCTGATCTGCCAGCGCCTGACGGGAACACAGGAAATCCGGTTCAGCAATGCCGCCATGCAGCGCGG
AACCGAGCTGGAACCCCATGCCCGCGCCCGCTACATCATCGAAACCGGCGAGATTGTTACCGAGGT
CGGACTGATTGATCACCCGACCATTGCCGGATTCGGAGCCAGCCCGGACGGACTGGTCGGCGACAC
CGGCCTGATCGAGATCAAGTGCCCGAATACCTGGACGCACATTGAAACCATCAAGACCGGCAAGCC
GAAACCGGAATACATCAAGCAGATGCAGACGCAGATGGCCTGCACAGGCCGTCAGTGGTGCGACTT
TGTCAGCTACGACGACCGACTTCCGGACGACATGCAGTATTTCTGCACCCGCATCGAGCGCGACGA
CGCGCTGATTGCCGAGATCGAAACCGAGGTTTCCGCATTTCTGGCCGAACTGGAAGCCGAAATCGA
ATACCTGAAACGAAAGGCCGCCTAA

LHK Exo has 205 amino acid residues, 23.155kDa, and its sequence is:

MTMEQRTEEWFAARLGKVTASRVADVMTKTKSGYAASRQNYMAELICQRLTGTQEIRFSNAAMQRGT
ELEPHARARYIIETGEIVTEVGLIDHPTIAGFGASPDGLVGDTGLIEIKCPNTWTHIETIKTGKPKPEYIKQM
QTQMACTGRQWCDFVSYDDRLPDDMQYFCTRIERDDALIAEIETEVSAFLAELEAEIEYLKRKAA

LHK Exo is aligned with bcateriophage lambda Exo (Figure 11), which suggests that LHK Exo is very likely to be an alkaline exonuclease similar to lambda Exo.

Figure 11. The alignment of LHK Exo with bacteriophage lambda Exo using T-coffee

LHK *bet* gene has 789bp, and its sequence is:

ATGTCCACCGCACTCACCATGCTGACCAGCAAGCTCGCCGCAAAATTCGACATGGGCGACGGCACC
GAGCTGGTCTCAACCCTCAAGCAGACGGCATTCAAGGGTCAGGTGACTGACGCACAGATGACCGCG
CTGCTGATCGTCGCCAACCAGTACGGGCTGAACCCGTGGACCAAGGAAATCTACGCCTTCCCTGAC
AAAAACAACGGCATCGTGCCGGTGGTCGGCGTGGATGGCTGGAGCCGCATCATCAACGGCGACCCG
AACTTCGACGGCATGGAGTTTGAACAGAACGCCGAGAGCTGCACTTGCCGCATCTTCCGCAAGGAT
CGTGGCCGCCCGATCAGCGTGACCGAGTGGATGGACGAGTGCAAACGCGAGACAGGACCGTGGAA
ATCGCACCCCAAGCGCATGCTGCGCCACAAGGCCATGATCCAGTGCGCCCGCCTCGCATTCGGCTTT
GCCGGCATCTACGACCAGGACGAAGCCGAGCGTATTGCCGAAAAGGACGTGACGCCGCCCAAACA
GGCGTACGAGAAGGTCACCCAAGGCGACACCAGCAATCCGCGCCGGAATGAACTGCTGGCCAAGG
CCAACGAGATTGCCGAAACCGGCGACGTGGACGGCCTACGCGAATATTTCCGCTCGCTCGGCAAGG
AGGATCGTCTGCTGATTGGCACCGACGAAATGAGCCGGATGGGTGAGCAGGCTGCCAGGAACGGC
GACCTCAAAGAGGAGCAGGCCACCAATGGCAGGATCATCGAAGGCGAGCTGGTGGAAGGAGAATG
A

LHK Bet has 262 amino acid residues, 29.3319kDa, and its sequence is:

MSTALTMLTSKLAAKFDMGDGTELVSTLKQTAFKGQVTDAQMTALLIVANQYGLNPWTKEIYAFPDK
NNGIVPVVGVDGWSRIINGDPNFDGMEFEQNAESCTCRIFRKDRGRPISVTEWMDECKRETGPWKSHPK
RMLRHKAMIQCARLAFGFAGIYDQDEAERIAEKDVTPPKQAYEKVTQGDTSNPRRNELLAKANEIAETG
DVDGLREYFRSLGKEDRLLIGTDEMSRMGEQAARNGDLKEEQATNGRIIEGELVEGE

LHK Bet is aligned with lambda Bet (Figure 12). This indicates that LHK Bet is highly possibly a lambda Bet-like protein and mediates ssDNA annealing.

LHK *ssb* gene has 489bp, and its sequence is:

ATGGCATCCGTCAACAAAGTCATCCTCGTCGGCAACCTCGGGCGTGACCCGGAAGTGCGCTACATG
CCGAACGGTGAGGCCGTCTGCACTTTCAGTATCGCCACCACCGACAGCTGGAAGGACAAGAACGGC
CAGAAGCAGGAGCGCACCGAATGGCACAACATCATCCTTTATCGGCGGCTGGCCGAGATTGCCGGC
GAGTATTTGAAAAAAGGCCGCCCGGTCTACATCGAGGGCCGCATCCAGTCGCGCAAATACACCGGC
AAGGACGGTGTAGAACGCACGGCCTTCGAGATATTGGCAACCGAGCTGCAAATGCTGGGTGGCCGG
GTGGAATCGGGCGGAAGCACCCGGACAGAGCGTAGCGAGCCACCGCCACCACCGCGCCGGCAGGA
GCCGAAGCCGGCCAGCAACTTCGACGACATGGACGACGACATCCCGTTCGCGCCGCTCGGCCTGCA
ATACCGGTTTGGCCTGCACTGCATTTGA

Figure 12. The alignment LHK Bet with bacteriophage lambda Bet using T-coffee

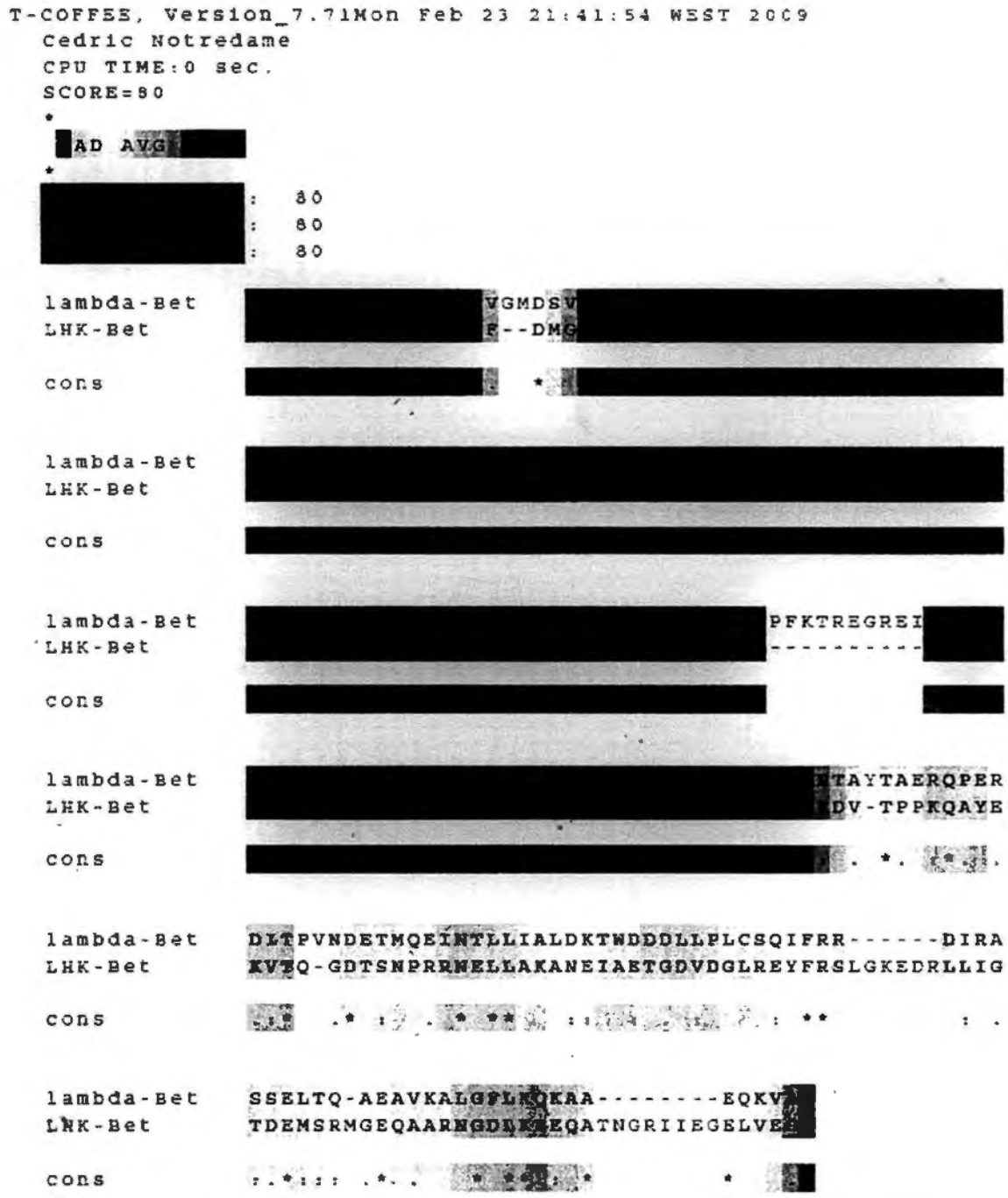


Figure 13. The alignment of LHK SSB with *E.coli* SSB using T-coffee.

LHK SSb has 162 amino acid residues, 18.341kDa, and its sequence is:

MASVNKVILVGNLGRDPEVRYMPNGEAVCTFSIATTDSWKDKNGQKQERTEWHNIILYRRLAEIAGEYL
KKGRPVYIEGRIQSRKYTGKDGVERTAFEILATELQMLGGRVESGGSTRTERSEPPPPPRRQEPKPASNFD
DMDDDIPFAPLGLQYRFGLHCI

LHK SSB has been aligned with *E. coli* SSB (figure 13), which suggests that LHK SSB is highly likley to be a single stand DNA binding protein with properties similar to those of *E.coli* SSB.

Chapter Two Materials and methods

Materials

Chemicals were purchased from Sigma and Fluka. Primers were purchased from Tech Dragon Company (Hong Kong). Bench protein ladder, 1Kb DNA ladder, and Picogreen were purchased from Invitrogen. Restriction enzymes are purchased from NEB. Centrifuges were purchased from Beckman. AKTApurifier, Superdex 200 HR10/30 column, Superdex 75 10/300 GL, Hitrap chelating HP column, and HiTrap Desalting column were purchased from GE Company. pBADαβγ, pBADETγ, pDH plasmid, Bet truncation plasmids DY380 Δ*bet* strain were kindly provided by Huang JDgroup in Biochemistry Department in the University of Hong Kong

Methods

Part One: Plasmids cloning

2.1.1. Exo, Bet and Ssb protein expression plasmids

SXT exo cloning

SXT-exo gene was amplified via polymerase chain reaction (Expand polymerase, Roche) using the primers Sexofor (TATACATATGAAGGTTATCGACCTATCAC) and SexoRevX (TTAACTCGAGTTAAAAATAAAATGAGCTCGGCGA), and pJB1 plasmid as template. PCR product was digested with NdeI / XhoI (underlined), and cloned into pET28a, to create an N-terminal hexahistidine fusion protein.Plasmid construct was named pEA1-1. pJB1 is a plasmid containing the the putative *ssb*, *bet* and *exo* genes open reading operon from SXT. Its sequence is unpublished, and has

not been deposited in the databanks. The plasmid was constructed by John Beaber, a PhD student working for Matthew Waldor at the New England Medical Center, Tufts University, Boston MA, USA.

SXT bet cloning

SXT-bet gene was amplified via PCR (Expand polymerase, Roche) using the primers SXTNdeI and SXTXhoI, and pJB1 plasmid as template. PCR product was digested with NdeI / XhoI, and cloned into pET28a, to create an N-terminal hexahistidine fusion protein.

SXTNdeI TTAACATATGGAAAAACCAAAGCTAATCCAA
SXTXhoI TATACTCGAGCTAAGAAGCTAAAGGCTGTGTGAG

SXT ssb cloning

SXT-ssb gene was amplified by PCR (Expand polymerase, Roche) using the primers SSBfor1 and SSBrevX, using the plasmid pJB1 as template. PCR product was cloned into pET2832a via EcoRI / XhoI, to create a N-terminal hexahistidine fusion protein.

SSBrevX TTAACTCGAGTTAAAAATCTGGTTCAGGATAAGTTTG
SSBfor1 TATAGAATTCACCATGGGAAACCAAGTAACACTCATAGGC

LHK Bet cloning

The putative *ssb*, *bet* and *exo* genes are located next to each other in open reading operon (Figure 10). The LHk *bet* gene was amplified via PCR using the primers LHKbetfor1 and LHKbetfor1, and LHK genomic DNA as template. PCR product was digested with BamHI / XhoI, and cloned into pET28a, to create an N-terminal hexahistidine fusion protein.

LHKbetfor1: 5'-ATATGGATCCCATATGTCCACCGCACTCACCAT-3'

LHKbetrev1 5'-TTAACTCGAGTTAAAGCTTTTCTCCTTCCACCAGCTCG-3'

LHK Exo cloning

The LHk *exo* gene was amplified via PCR using the primers LHKbetfor1 and LHKbetfor1, and LHK genomic DNA as template. PCR product was digested with BamHI / XhoI, and cloned into pET28a, to create an N-terminal hexahistidine fusion protein.

LHKexofor1 5'-ATATGGATCCATGACCATGGAACAACGCAC-3'

LHKexorev2 5'-TTAACTCGAGTTAAAGATTGGGCGGCCTTTCGTTTCAG-3'

2.1.2. Recombineerng systems plasmids

LHKbet-exo-pBAD (LHKBX, Figure 14): The pBADαβγ is lambda *bet-exo-gam* gene cloned in pBAD plasmid. LHK *bet-exo* gene was amplified from LHK genomic DNA using LHK primers and cloned in pBADαβγ to substitute the lambda *bet-exo* gene. However, the stop codon of *exo* is located within Lambda *gam* gene and RBS of *gam* gene was truncated, so the Gam protein will not express in protein expression via IPTG induction.

LHKbet-exo-ssb-pBAD (LHKBXS, Figure 14) were prepared in a method similar to LHKbet-exo-pBAD preparation. The lambda *bet-exo* gene in pBADαβγ was substituted by *bet-exo-ssb* gene, the stop codon of LHK *ssb* gene is located within lambda *gam* gene and RBS of *gam* gene was truncated, so the Gam protein will not express in protein expression via IPTG induction.

pBex4b1, pBX2B, pGX2B, pEXS1A, pαKX2A and pBAD-ETγ (Figure 14) were kindly provided by Huang JD group in Department of Biochemistry in HKU. pBex4b1 and pBX2B are modified pBADαβγ plasmids in which the lambda *bet-exo* gene was substituted by SXT *bet-exo* gene and SXT *ssb-bet-exo* respectively. The stop codon of *exo* gene is located within Lambda *gam* gene and RBS of *gam* gene was truncated, so the Gam protein will not express in protein expression via IPTG induction.

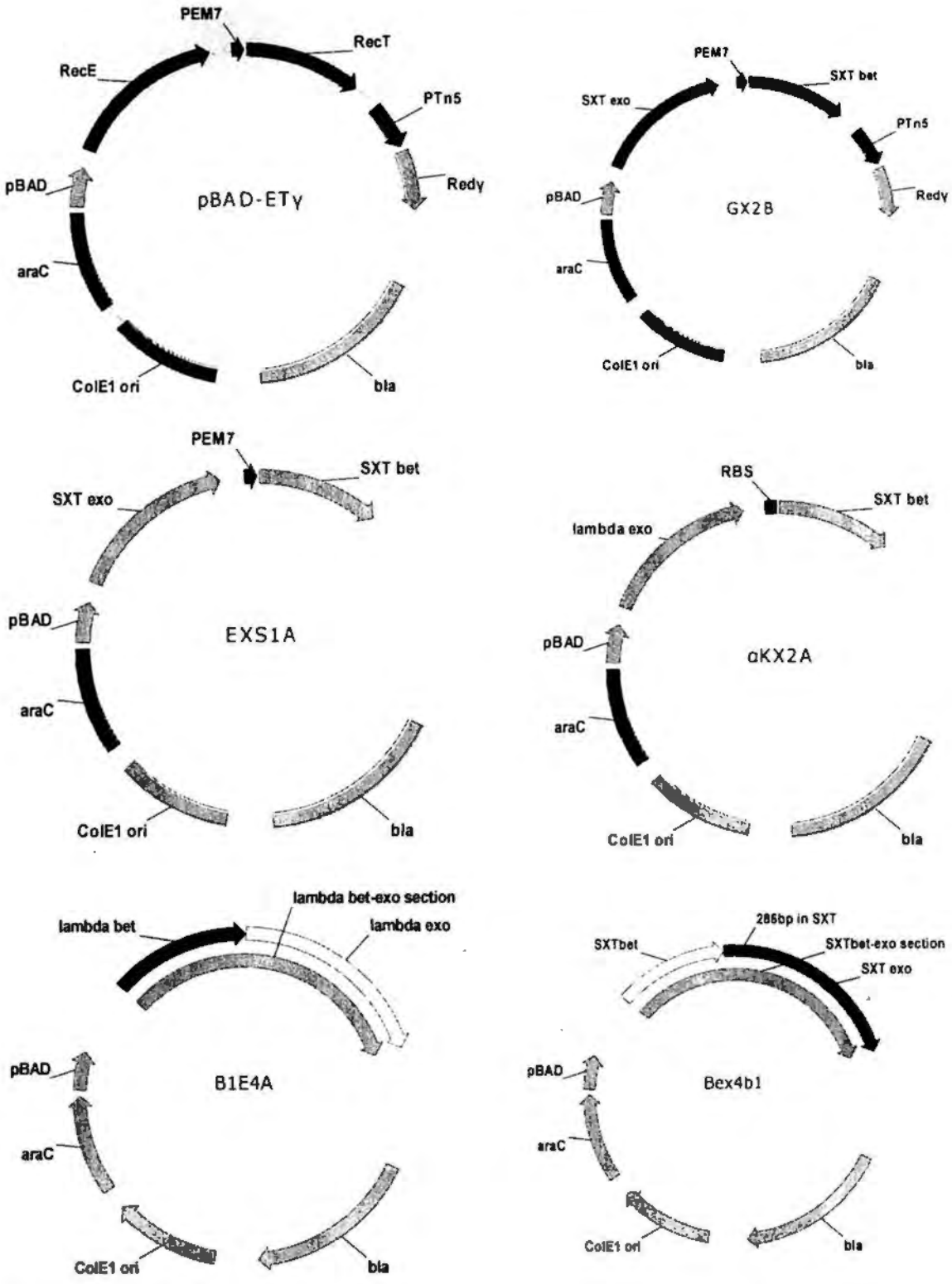
pGX2B is the modified pBADαβγ plasmid in which the lambda *bet-exo* gene was substituted by SXT *bet-exo* gene. However, the stop codon of *exo* is not located within Lambda *gam* gene, so the Gam protein will express in protein expression via IPTG induction.

EXS1A is the modified plasmid in which the lambda *bet-exo* gene was substituted by SXT *bet-exo* gene. However, EM7 promoter and RBS from pEM7/Bsd were subcloned immediately following to the end of SXT *bet* gene, to substitute the original long DNA sequence between the *bet* and *exo* gene. The SXT *exo* stop codon is located within Lambda *gam* gene and RBS of *gam* gene was truncated, so the Gam protein is interrupted and will not express in protein expression via IPTG induction.

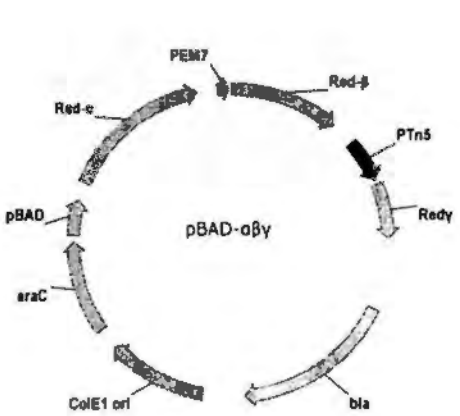
αKX2A is the modified plasmid in which the lambda *bet-exo* gene was substituted by lambda *exo*-SXT *bet* hybrid function, and the SXT-*bet* is located within Lambda *gam* gene and RBS of *gam* gene was truncated, so the Gam protein is interrupted and will not express in protein expression induction.

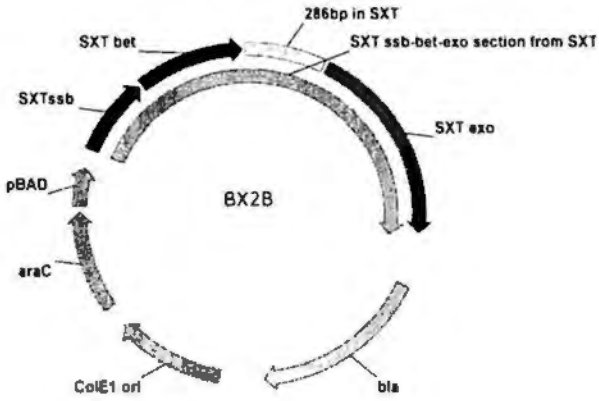
pBAD-ETγ is is the modified pBADαβγ plasmid in which the lambda *bet-exo* gene was substituted by Rac prophage *recE-recT* gene. However, the stop codon of *recT* is not located within Lambda *gam* gene, so the Gam protein will express in protein expression via IPTG induction.

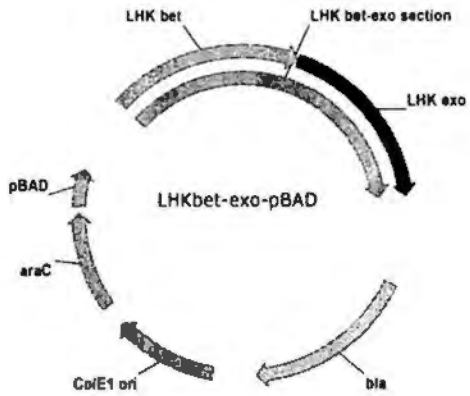
Figure 14. The recombineering plasmids constructed.

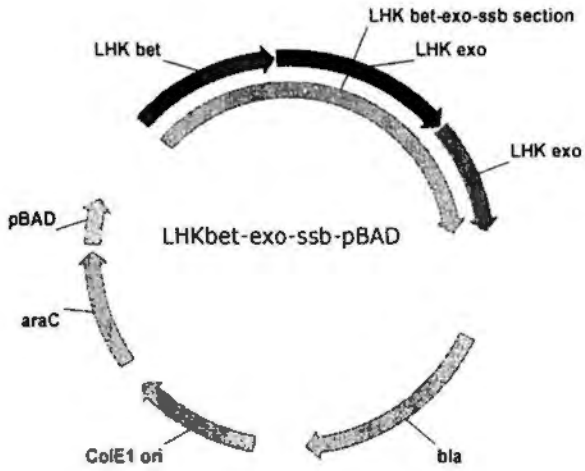


(to be continued)









(continue to Figure 14)

2.1.3. *bet* cloning in pBAD and PDH vectors

Lambda-*bet*-pBAD (Figure 15), and *RecT*-pBAD, SA-*bet*-pBAD, Lambda-*bet*-pDH, and *RecT*-pDH and SA-*bet*-pDH were kindly provided by Huang JD group in the Department of Biochemistry in HKU. SXT-*bet*-pBAD (, Lambda-*bet*-pBAD, and *RecT*-pBAD, SA-*bet*-pBAD are the modified pBADαβγ plasmids in which lambda *bet-exo* gene was substituted by lambda-*bet*, *RecT*, and *S. aureus bet* gene (as lambda-*bet*-pBAD), respectively. However, the stop codons of each *bet* genes are located within *gam* gene and the RBS of *gam* gene is truncated, so Gam protein will not express in the protein expression via IPTG induction. Various lambda Bet truncations, RecT trcuncations and SXT bet truncations in pBAD are also prepared.

PDH vector was constructed in Huang JDgroup, it has P_L promoter from lambda phage under the control of *cI857* repressor, so the protein gene expresses at heat-shock at 42 °C. Lambda-*bet*-pDH (Figure 15), SXT-bet-pDH, *RecT*-pDH and SA-bet-pDH are plasmids in which lambda *bet*, SXT *bet*, RecT and *S. aureus bet* cloned in pDH (as lambda-*bet*-pDH), respectively. LHK *bet* and SXT *bet* were subcloned from LHK-*bet*-pET28a and SXT-*bet*-pET28a, respectively, to pDH vector. LHK bet with N-terminal hexahistidine-tag in pDH (LHK-His-bet-pDH) and in pBAD (LHK-His-bet-pBAD) were also prepared to do activity assay.

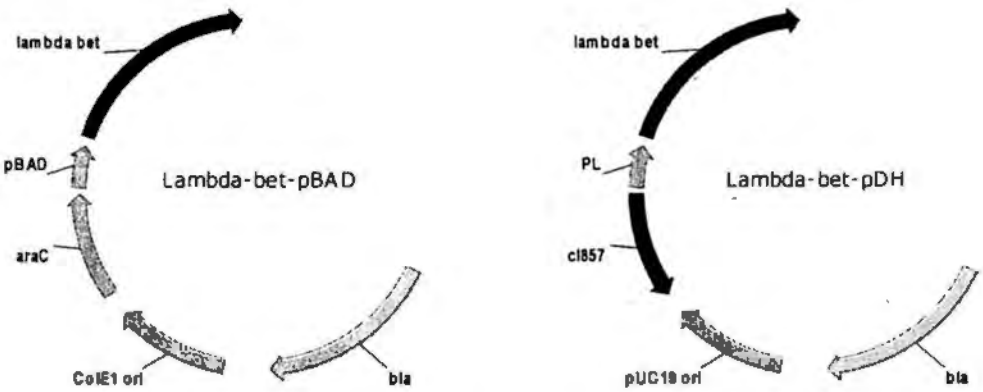


Figure 15. *bet* genes in pBAD and pDH vectors

Part Two Protein expression, purification and SDS-PAGE

2.2.1. SXT Exo protein expression and purification

SXT-exo-pET28a plasmid was transformed into BL21(DE3) with pLysS, and the colony was inoculated into 30ml L Luria-Bertani (LB) medium (containing 100μg/ml kanamycin and 34μg/ml chloramphenicol) and shaked at 250 rpm for overnight at 37°C. 20ml of overnight culture was expanded into 2L LB medium and shaked at 37 °C till OD_{600}~0.6. Protein expression was induced by 0.1mM IPTG and 1mM $MgCl_2$ at room temperature for overnight. Cells were harvested by centraifugation at 4 °C for 15 minutes. Cell pellet was resuspended in 5 volumes Bind buffer (25mM Tris-HCl, pH7.4, 500mM NaCl, 25mM imidazole) and sonicated.

The lysate was centrifuged at 15,000g for 30minutes and supernatant was applied to 1ml Hitrap chelating HP column (GE) after filtration with 0.45μM filter (Millipore). After 5 column volume wash with Bind buffer, the column was eluted with gradiently increasing concentration of imidazole in Bind buffer. Protein in fractions was checked by Bradford assay and SDS-PAGE. The fractions containing SXT Exo were pooled together and applied to 5ml Hitrap desalting column. SXT Exo eluted by Stock buffer (25mM Tris-HCl, pH7.4, 50mM NaCl), and then the protein sample was ready for further biochemical and biophysical assay.

2.2.2. LHK Exo protein expression and purification for biochemical assay

LHK-exo-pET28a plasmid was transformed into BL21(DE3), and and the colony was inoculated into 30ml L Luria-Bertani (LB) medium (containing 100μg/ml kanamycin) and shaked at 250 rpm for overnight at 37°C. 20ml of overnight culture was expanded into 2L LB medium and shaked at 37 °C till OD_{600}~0.6. Protein expression was induced by 0.5mM IPTG at 32 °C for 6 hours. Cells were harvested by centraifugation at 4 °C for 15 minutes. Cell pellet was resuspended in 5 volumes Bind buffer (25mM Tris-HCl, pH7.4, 500mM NaCl, 25mM imidazole) and sonicated. The lysate was centrifuged at 15,000g for 30minutes and supernatant was applied to 1ml Hitrap chelating HP column (GE) after filtered with 0.45μM filter (Millipore). After 5 column volume wash with Bind buffer, the column was eluted with gradient increasing concentration of imidazole in Bind buffer. Protein in fractions was checked by Bradford assay and SDS-PAGE. The fractions containing LHK Exo were pooled together and diluted with Start buffer (25mM Tris-HCl, pH7.4, 50mM NaCl, 5mM imidazole). The LHK Exo was further purified using 6ml Resource Q column (GE) with an Elute buffer (25mM Tris-HCl, pH7.4, 1M NaCl, 5mM imidazole), and then the protein sample was ready for biochemical and biophysical assay.

2.2.3. Protein expression and purification

LHK-bet-pET28a , SXT-bet-pET28a and SXT-ssb-pET28a plasmid were transformed into BL21(DE3), repectively. Colony was inoculated into 20ml LB (containing 100μg/ml kanamycin) and shaked at 250rpm for overnight at 37 °C. 10ml overnight culture was expanded into 1L LB medium and shaked at 37 °C till OD_{600}~0.6. 0.5mM IPTG was added to the culture to induce protein expression at 32 °C for 6 hours. Cells were harvested by centraifugation at 4 °C for 15 minutes. Cell pellet was resuspended in 5 volumes Bind buffer (25mM Tris-HCl, pH7.4, 500mM NaCl, 25mM imidazole) and sonicated. The lysate was centrifuged at 15,000g for 30minutes and supernatant was applied to 1ml Hitrap chelating HP column (GE) after filtered with 0.45μM filter (Millipore). After 5 column volume wash with Bind buffer, the column was eluted with gradient increasing concentration of imidazole in Bind buffer. Protein in fractions was checked by Bradford assay and SDS-PAGE. The fractions containing purified protein were pooled together and diluted with Start buffer (25mM sodium phosphate buffer, pH7.4, 50mM NaCl, 5mM imidazole). The protein were further purified using 6ml Resource Q column (GE) with an Elute buffer (25mM T sodium phosphate buffer, pH7.4, 1M NaCl, 5mM imidazole), and then the protein samples were ready biophysical and biochemiscal assay.

Part Three Preparation of linear dsDNA pUC18/PstI

To prepare linear dsDNA substrate, pUC18 was prepared using Qiagen Miniprep Kit. The plasmid was completely digested with restriction enzymes (PstI, BamHI, SspI). The samples of digestion products were analyzed on 1% agarose and stained with ethidium bromide. pUC18/Pst I was aliquoted and dephosphorylate using alkaline phosphatase (calf intestinal, NEB). The DNA was purified using Spin column (Qiagen) and eluted with 10mM Tris-HCl, pH8.0 buffer. The efficiency of DNA dephosphorylation was checked by using T4 ligase to do ligation of the

dephosphorylate and un-dephosphorylate pUC18/PstI. The ligation products were transformed in DH10B competent cells and the ligation (dephosphorylation efficiency) was compared to analyze to dephosphorylation efficiency.

Part Four Picogreen assay of dsDNA digestion by exonuclease

An aliquot of 25μL of 250-fold diluted Picogreen was added to the reaction mixtures and incubated in dark for 5minutes, and then the fluorescence was measured with an excitation 485nm and an emission 525nm. The final reaction rate was calculated by measuring the difference of the fluorescence intensity between the start conditions and the final reaction conditions. The positive control is one half of linear dsDNA substrate in reaction mixture that was boiled and then immediately chilled in ice-cold water. The negative control is a reaction mixture in which the active enzyme was substituted by the same amount of inactivated enzyme. All the nuclease assays were repeated at least 6 times.

Part Five Biochemical assay of LHK Exo

2.5.1. Optimal Mg^{2+} concentration of LHK Exo determination

Standard nuclease activity assay of LHK Exo is performed in 25mM Tris-HCl, pH8.0, 50mM NaCl, 5mM $MgCl_2$ at room temperature (22°C), the reaction volume is 70μL, the reaction was allowed to continue for 5 minutes.

The enzyme reaction with a total volume of 70μl containing 7.23ng LHK Exo (0.272 pmol monomer) and various concentrations of $MgCl_2$ in 50mM Tris-HCl pH8.0, 50mM NaCl buffer was initiated by adding 33.3ng (0.020pmol) pUC18/PstI, the reaction was allowed to continue for 5 minutes at room temperature (22°C)and then

quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.5.2. Optimal pH of LHK Exo determination

The enzyme reaction with a total volume of 70µl containing 7.23ng LHk Exo (0.272 pmol monomer) and 7.5mM $MgCl_2$ in 50mM Bis-Tris propanol, 50mM NaCl buffer at various pHs was initiated by adding 33.3ng (0.020pmol) pUC18/PstI, the reaction was allowed to continue for 5 minutes at room temperature (22°C) and then quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.5.3. Optimal temperature of LHK Exo determination

The enzyme reaction with a total volume of 70µl containing 7.23ng LHK Exo (0.272 pmol monomer) and 7.5mM $MgCl_2$ in 50mM Tris-HCl pH8.0, 50mM NaCl buffer was initiated by adding 33.3ng (0.020pmol) pUC18/PstI. The reaction was allowed to continue for 1 minute at various temperatures and then quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.5.4. Digestion velocity of LHK Exo determination

The enzyme reaction with a total volume of 350µL containing various amount (1.09, 1.36, 2.77, 5.45, 136.25pmol monomer, respectively) of LHK Exo and 2.5mM $MgCl_2$ in 50mM Tris-HCl pH8.0, 50mM NaCl buffer was initiated with substrate 166.5ng (0.1pmol) pUC18/PstI at room temperature (22°C). So, there are 33.3ng dsDNA and various amount (0.218, 0.272,0.554, 1.108, 27.25pmol) of LHK Exo in every 70µL reaction mixture. After the reactions were initiated, 70µL reaction mixture was withdrawn at different time points and quenched with 20mM EDTA. The extent of digestion was measured using Picogreen fluorescence assay. The

digestion time courses were plotted and the amount of enzyme to saturate the DNA ends was determined. Then more digestion time courses were plotted and the slopes were figured out to calculate the digestion velocity.

2.5.5. Fraction activity test

After the LHK Exo protein was purified, it was frozen with liquid nitrogen and stocked in -80°C. The enzyme fraction activity was test after it melted on ice. The enzyme reaction with a total volume of 70μl containing various amount of newly melted LHK Exo (1.088, 0.544, 0.272, 0.136,0.068, 0.034, 0.017, 0.0085pmol monomer) and 5mM $MgCl_2$ in 50mM Tris-HCl pH8.0, 50mM NaCl buffer was initiated by adding 33.3ng (0.020pmol) pUC18/PstI, the reaction was allowed to continue for 10 minutes at room temperature and then quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay. The percentage of DNA digestd vs amount of enzyme was plotted and the fraction activity was analyzed.

2.5.6. Substrate preference test of LHK Exo

The enzyme reaction of a 420μL total volume containing 7.23ng LHK Exo (0.272 pmol monomer) of LHK Exo and 2.5mM $MgCl_2$ in 50mM Tris-HCl pH8.0, 50mM NaCl buffer was initiated by adding 199.8ng (0.1pmol) of different linear dsDNA substrates (pUC18/PstI, dephosphated pUC18/PstI, pUC18/BamHI and pUC18/SspI) at room temperature (22°C). After the reactions were initiated, 70μL reaction mixture was withdrawn at different time points and quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay. The digestion time courses were plotted and the substrate specificities were compared.

2.5.7. Metal ion dependence test for LHK Exo

The 20μL reaction mixtures were in 25mM Tris-HCl pH7.4, 50mM NaCl buffer, with or without 5mM of various metal ions, containing 180ng pET28a/NdeI linear dsDNA and 1.45μg LHK Exo protein. The reaction mixtures were quenched with 20mM EDTA after incubation at 37°C for 20 minutes, and then the reaction mixtures were run on 1% agarose gel and stained with ethidium bromide.

2.5.8. Digestion time course of LHK Exo on gel

A 200μL reaction mixture in 25mM Tris-HCl pH7.4, 50mM NaCl, 5mM $MgCl_2$ buffers, containing 1.8μg pET28a/NdeI linear dsDNA and 14.5μg LHK Exo protein was incubated at room temperature for 1minute, 2minutes, 3minutes, 4minutes, 5minutes, 6minutes, 7minutes, 15minutes, then 20μL of reaction mixture was withdrawn and quench with 20mM EDTA, 1% SDS. Then the digestions were analyzed by running the mixture on 1% agarose gel. The gel was stained with ethidium bromide.

Part six Biochemical assay of SXT Exo

2.6.1. Optimal temperature determination

The standard reaction buffer (25mM Tris-HCl, pH7.4, 50mM NaCl, 0.5mM $MnCl_2$) was incubated with 2μl 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) in water bath for 5 minutes, then 2μL 0.12mg/ml of SXT Exo (6pmol monomer) was added to start the reactions. The total reaction volume is 50μL, and reactions were allowed to continue for 1 minute. The reaction was quenched with 20mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.6.2. Optimal Mn^{2+} ion concentration determination

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μL 0.12mg/ml (6pmol monomer) were added to the reaction buffer (25mM Tris-HCl, pH7.4, 50mM NaCl) containing various concentrations of $MnCl_2$ to start the reactions and the final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 20mM EDTA. The digestion was measured by Picogreen fluorescence assay.

2.6.3. Optimal Mg^{2+} ion concentration determination

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μL 0.12mg/ml (6pmol monomer) were added to the 25mM Tris-HCl, pH7.4, 50mM NaCl reaction buffer containing various concentration of $MgCl_2$ to start reactions and the final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 75mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.6.4. Optimal pH of SXT Exo determination

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μ 0.12mg/ml (6pmol monomer) were added to the 25mM Tris-HCl, pH7.4, 50mM NaCl 0.5mM $MnCl_2$ reaction buffer at different pHs to start the reactions and the final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 75mM EDTA. The digestion was measured using Picogreen fluorescence assay.

2.6.5. Substrate preference test

12μL 2.5ng/μL linear dsDNA (pUC18/PstI, de-phosphated L pUC18/PstI, pUC18/BamHI or pUC18/SspI) and 12μ 0.12mg/ml (36pmol monomer) were added to 25mM Tris-HCl, pH7.4, 50mM NaCl, 0.5mM $MnCl_2$ reaction buffer to start the reactions at 37°C and the final reaction volume is 300μL. 50μL reaction mixture was withdrawn and quenched the reactions with 20mM EDTA at different time points. The digestion was analyzed by fluorescence assay using Picogreen fluorescent assay.

2.6.6. Heparin inhibition assay

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μ 0.12mg/ml (6pmol monomer) were added to 25mM Tris-HCl, pH7.4, 50mM NaCl, 0.5mM $MnCl_2$ reaction buffer containing various concentrations of heparin to start the reaction and the final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 20mM EDTA. The digestion was analyzed by fluorescence assay using Picogreen as described above. The activities of SXT Exo in the presence of heparin were compared and normalized with the activity of the SXT Exo without heparin.

2.6.7. Salt concentration effects assay

To start the SXT Exo digestion, 2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μ 0.12mg/ml (6pmol monomer) were added to 25mM Tris-HCl, pH7.4, 50mM NaCl, 0.5mM $MnCl_2$ reaction buffer containing and various concentrations of NaCl, KCl and CaCl2, respectively. The final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 20mM EDTA. The digestion was analyzed by fluorescence assay using Picogreen. The activities of SXT Exo in the presence of different salts were compared and normalized with the activity of SXT Exo in the standard condition.

2.6.8. PO_4^{3-} inhibition assay

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μ 0.12mg/ml (6pmol monomer) were added to the reaction buffers (sodium phosphate buffer or Tris-HCl buffer at various concentrations at pH7.4) containing 50mM NaCl and 0.5mM $MnCl_2$ to start reactions and the final reaction volume is 50μL. The reactions were allowed to continue for 30 minutes at 37°C, and then quenched with 20mM EDTA. The digestion was analyzed by fluorescence assay using Picogreen fluorescence assay. The activities were compared and normalized with the activity of SXT Exo in the standard condition.

2.6.9. SO_4^{2-} inhibition assay

2μL 2.5ng/μL pUC18/PstI linear dsDNA (0.003pmol) and 2μ 0.12mg/ml (6pmol monomer) were added to 25mM Tris-HCl, pH7.4, 50mM NaCl, 0.5mM $MnCl_2$ reaction buffers containing various concentration of Na_2SO_4 or K_2SO_4. The final reaction volume is 50μL. The reactions were allowed to continue for 30minutes at 37°C, and then quenched the 20mM EDTA . The digestion was analyzed by fluorescence assay using Picogreen fluorescence assay. The activities were compared and normalized with the activity of SXT Exo in the standard condition.

2.6.10. Metal ion dependence test for SXT Exo

The 20μL reaction mixtures comtain 184ng pET28a/NdeI linear dsDNA and 0.6μg SXT Exo protein in 5mM Tris-HCl pH7.4, 50mM NaCl buffer, with or without various metal ions. The reaction mixtures were quenched by 20mM EDTA after incubation at 37°C for 30 minutes, and then the reaction mixtures were analyzed on 1% agarose gel and stained with ethidium bromide. To check whether the SXT Exo has dsDNA exonuclease activity or not in the presence of high concentration of

$MgCl_2$, a 20μL reaction mixture of 25mM Tris-HCl, pH7.4, 50mM NaCl, 20mM $MgCl_2$ containing 184ng pET28a/NdeI linear dsDNA and 0.6μg SXT Exo protein was allowed to incubated at 37°C for 30 minutes as above and check the digestion by 1% agarose gel.

2.6.11. Digestion time course of SXT Exo on agarose gel

A 200μL reaction mixture is in 25mM Tris-HCl pH7.4, 50mM NaCl, 1mM $MnCl_2$ buffers, containing 1.84μg pET28a/NdeI linear dsDNA and 6μg SXT Exo protein, was incubated at room temperature. 20μL of reaction mixture was withdrawn at 0 seconds, 30 seconds, 60 seconds, 2 minutes, 5 minutes, 10 minutes, 20 minutes, 40 minutes, and 80 minutes and quenched with 20mM EDTA, 0.4%protease K, 1% SDS. Then the digestions were checked by running the mixture in 1% agarose gel. The gel was stained with ethidium bromide.

2.6.12. Dephosphorylated pUC18/PstI digested by SXT Exo

The 20μL reaction systems were in 25mM Tris-HCl, pH7.4, 50mM NaCl, 1mM $MnCl_2$ buffer, containing 50μg linear pUC18/PstI dsDNA (dephosphorylated or non-dephosphorylated) and 0.6μg SXT Exo protein. The reaction mixtures were quenched by 20mM EDTA after incubation at 37°C for 30 minutes, and then the reaction mixtures were run on 1% agarose gel and stained with ethidium bromide

2.6.13. Digestion velocity of SXT Exo determination

A 100μL reaction mixture in 25mM Tris-HCl, pH7.4, 50mM NaCl, and 0.5 mM $MnCl_2$ buffer contains 0.15pmol linear pUC18/PstI dsDNA and various amounts (15, 30, 60, 120. 240pmol monomer) of SXT Exo. The reaction is performed in ice water bath and 20μL reaction mixture was withdrawn and quenched with 20mM EDTA at different time points. The digestion was measured by Picogreen fluorescence assay.

Part Seven Biophysical study of LHK Exo, LHK bet, SXT exo, SXT Bet and SXT SSB proteins

2.7.1. Cross-linking

After purified using a 5ml Hitrap chelating column, LHK Exo, LHK bet, SXT Bet and SXT SSB were further purified using a 6ml Resource Q column with Start buffer (25mM sodium phosphate buffer, pH7.4, 50mM NaCl) and Elute buffer (25mM sodium phosphate buffer, pH7.4, 1M NaCl), and then the proteins are used to do cross-linking using DMS and gluteraldehyde at various concentrations. The cross-linking products were checked by SDS-PAGE and gel was stained with Coommassie Blue R-250.

2.7.2. Gel filtration

SXT exo and LHK Exo multimerity determination

100μL of purified SXT Exo, lambda Exo and LHK Exo proteins were applied to Tricorn Superdex 200HR 10/300GL column on an AKTA-FPLC (GE Healthcare) in the buffer of 25mM Tris-HCl, pH7.4, 150mM NaCl, 1mM EDTA, 5mM imidazole. The column was pre-calibrated with protein standards: ferritin (440kDa); aldolase (158kDa); thyroglobulin (67kDa); ovalbumin (43kDa); hymotrypsinogen A (25 kDa) and ribonuclease A (13.7kDa). The flow rate is 0.4ml/min at 4°C.

LHK Bet, SXT Bet and SXT SSB multimerity determination

100μL Bet and SSb proteins were loaded to Superdex200 HR200 10/30 column in different buffers, the flow rate is 0.4ml/min at 4°C. Ferritin (440kdal), aldolane

(158kDal), conalbumin (75 kDal) and ovalbumin (43kDa) were used to prepare the standard curve.

Part Eight, LHK Exo -LHK Bet interaction

2.8.1. Interaction of LHK Exo and LHK Bet proteins *in vitro* study by pull down assay

LHK-*exo* was subcloned into pGex-4T1 vector. The LHK-exo-pGex-4T1 and LHK-beta-pET28a(+) proteins expression were induced for about 5 hours when the $OD_{600}$of the culture reach 0.6-0.7 with 0.5mM IPTG at 32°C in 100ml LB culture, and then the cell pellets were collected. The cell pellets were resuspended in 25ml Bind buffer (25mM Tris-HCl, pH7.4, 150mM NaCl, 1mM EDTA, 10% glycerol) and sonicated using Bradson sonifier 300 with 30% output for 2 minutes with 2 seconds on and 5 seconds off in ice-water mixture bath. The lysates were ultracentrifuged for 45 minutes and the supernatant was collected and filtered with 0.45μM filter (Millipore). The LHKexo-pGex-4T1 supernatant was loaded slowly (0.2-0.3ml/minute) in 4°C refrigerator to 1ml GSTrap FF column first, then LHKbet-pET28a(+) supernatant was loaded to the same 1ml GSTrap FF column slowly. The column was washed thoroughly till no protein detected using Bradfrod assay with about 6ml Bind buffer, and finally eluted the column with 10mM reduced GSH in Bind buffer. 1ml eluate was collected in each fraction.

The controls were: a, LHK-exo-pGex-4T1/BL21(DE3) extract was loaded to 1ml GSTrap FF column, and eluted with 10mM reduced GSH in Bind buffer after thorough wash with Bind buffer; b, LHKbet-pET28a(+)/BL21(DE3) extract was loaded to 1ml GSTrap FF column, and eluted with 10mM reduced GSH in Bind buffer after thorough wash with Bind buffer, c, pGex-4T1/BL21(DE3) extract was loaded to 1ml GSTrap FF column, and eluted with 10mM GSH in Bind buffer after

thorough wash, pGex-4T1/BL21(DE3) extract loaded to 1ml GSTrap FF column, then LHKbet-pET28a(+) extract was loaded to the same column, after 5CV bind buffer wash, the column was eluted with 10mM GSH in Bind buffer after thorough wash with Bind buffer.

2.8.2. Interaction of LHK Exo and LHK Bet proteins *in vivo* study by pull down assay

LHK *bet-exo* gene was cloned in pET32bvector via HindIII and NdeI, which makes Exo protein have His tag in the C-terminal while Bet has no tag.

The LHKbet-exo-pET32b was introduced in BL21(DE3) and protein expressions were induced by 0.5mM IPTG at 32 for 6 hours. Cell pellet from 400 ml was collected, and sonicated in 30 ml Bind buffer: 25mM Tris-HCL pH7.4, 150mM NaCl, 25mM imidazole and 10% glycerol in ice-water bath using Bradson sonifier 300 with 20% output for 8 minutes (2.5 seconds on and 17.5seconds off). The lysate was subjected to ultracentrifugation at 15,000g for 30minutes at 4°C, and supernatant was loaded to 1mL HItrap chelating column (GE) after filtered with 0.45μM filter (Millipore). The column was washed thoroughly with more than 6ml Bind buffer. Finally, the column was eluted with the Elution buffer: 25mM Tris-HCL pH7.4, 500mM NaCl, 500mL imidazole. 1ml eluate was collected in each fraction. The protein in the fractions was checked by running 15% SDS-PAGE.

The negative control was LHK-exo-pET28a expressed in BL21(DE3) which is subjected to the elution exactly the same as LHKbet-exo-pET32b above.

Part Nine Homologous recombination assay

2.9.1. dsDNA recombination

One set of two primers: ECgalKF1 (GTTTGCGCGCAGTCAGCGATATCCATTTTCGCGAATCCGGAGTGTAAGA ATAAAAATAGGCGTATCACGAG) and ECgalKR1 (TTCATATTGTTCAGCGACAGCTTGCTGTACGGCAGGCACCAGCTCTTCCG TAGTGAACCTCTTCGAGGGAC),were designed to amplify the Cm^r gene, however, at the 5'-end of the forward primer (ECgalKF1), there are about extra 50 bases which are homologous to the upstream of *galK* gene (underlined) in the *E.coli* DH10B strain. At 5'-end of reverse primer (ECgalKR1), there are also extra 50 bases which are homologous the downstream of the *galK* gene (underlined). The Cm^r is amplified and PCR product was purified for homologous recombination assay (figure 16).

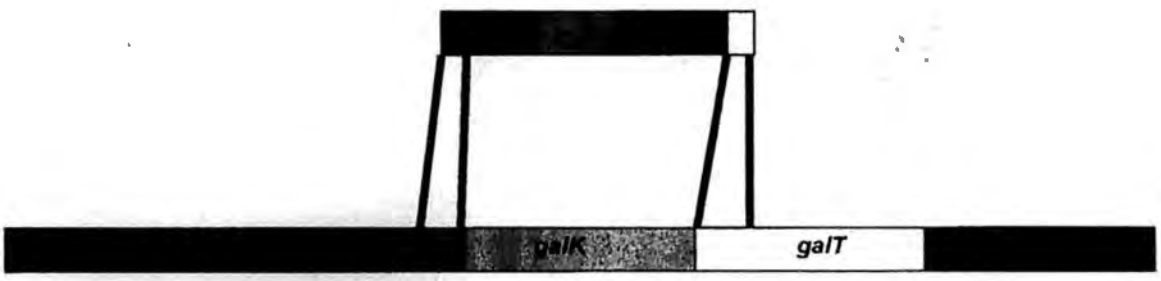


Figure 16.Homologous recombination mediated by recombineering functions

The recombination plasmids (pBex4b1, pBX2B, pGX2B, pEXS1A, pαKX2A and pBAD-ETγ, LHK-bet-exo-pBAD and LHK-bet-exo-ssb-pBAD) were transformed into DH10B, and the colonies were inoculated for overnight shaking at 37°C. The overnight culture was 1:100 transferred to 50ml LB in 250ml flasks and shake at 250rpm at 37°C for 1 hour. Then 0.1% of arabinose was added to induce protein expression for 1 hour at 250rpm at 37°C. Finally the cell culture was used to prepare competent cell that was resuspended in 190μL water. 1.18μg *galK* PCR product was

electroporated into 90μL competent cell. Cell was shaken at 250rpm at 37°C for 1.5hours before plating on LB agar media plates. The efficiency is calculated by dividing the number of recombinant cell by the number of all the survive cell in the electroporation.

2.9.2. Point mutation repair using ssDNA

DY380 is a DH10B-derived strain prepared by Dr. DL Court group at NCI-Frederick. In the chromosome of DY380 bacteria, there is a defective bacteriophage lambda prophage and the expression of recombination proteins (Bet, Exo, Gam) is controlled by repressor *cI857*, which is temperature sensitive. DY380 was further modified by Huang JD group in the Department of Biochemistry in HKU. The *bet* gene of lambda prophage in the bacteria chromosome was knockout and fragments of gene Kan^sCm^r were integrated into the chromosome respectively. The Kan^s in Kan^sCm^r is Kan^r gene with a point mutation, so the bacteria are *chloramphenicol* resistant and kanamycin sensitive. Single stand oligonucleotides was designed and synthesized to repair the point mutations respectively. For the both flanks of point mutations, there are 45 bases in the designed single stranded oligonucleotide which are homologous to them (figure 17).

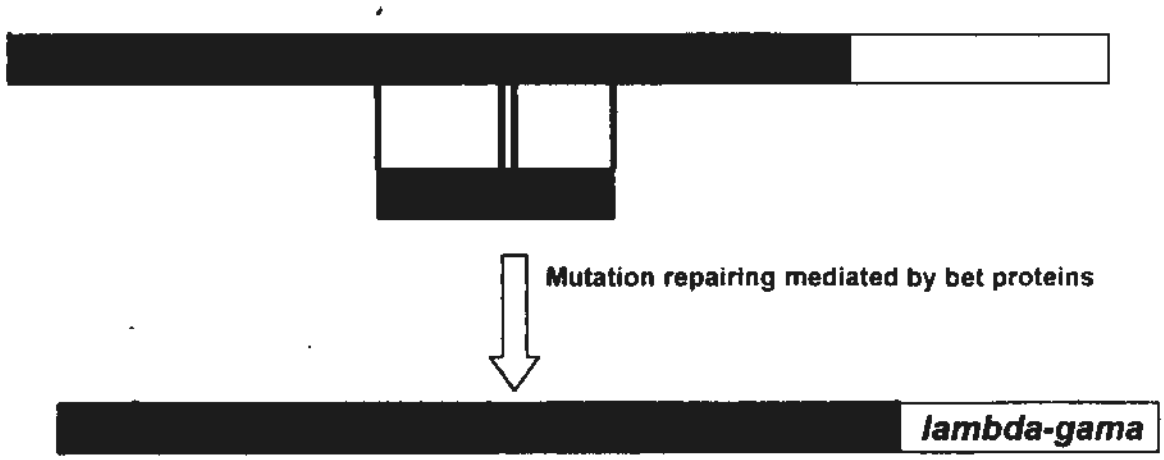


Figure 17. Point mutation repairing mediated by Bet proteins

pBAD plasmids cloned full length *bet* genes or truncated bet genes were transformed into new stain of DY380Δbet *Kan*^{r}*Cm*^{s} or DY380Δbet *Kan*^{s}*Cm*^{r} and the colony was inoculated for overnight shaking at 32°C. The overnight culture was 1:100 transferred to two sets of 50ml LB in 250 flask and shake at 250rpm at 32°C for 1.25 hour. Then 0.1% of arabinose was added to one set to induce protein expression and the shaking was allowed to continue for 1.5 hours at 250rpm at 32°C. Also, 0.1% of arabinose was added to another set and the culture were shaken for 1.25 hours, and then the flasks were shaken in in 42 °C water bath for 15 minutes before used to prepare competent cell. Competent cell was resuspended in 190 μL water. 100ng of single strand oligonucleotide was electroporated into the 90 μL competent cell. 1ml LB was added to the electroporation product and then shaken at 250rpm at 32°C for 2 hours before plating on LB agar media plates. The efficiency is calculated by dividing the number of recombinant cell by the number of all the survive cell in the electroporation in transformation.

pDH vectors containing full length *bet* gene or truncated bet gene were transformed into DY380Δbet *Kan*^{r}*Cm*^{s} or DY380Δbet *Kan*^{s}*Cm*^{r} and the colony was inoculated for overnight shaking at 32°C. The overnight culture was 1:100 transferred to 50 ml LB in 250 ml and shaken at 250 rpm at 32°C for 2.5 hour. Then the culture was shaken in 42 °C water bath for 15 minutes before used to prepare competent cell. The competent cell was resuspended in 190 μL water. 100ng of single strand oligonucleotide was electroporated into 90 μL competent cell. 1ml LB was added to the electroporation product and then shaken at 250 rpm at 32°C for 2 hours before plating on LB agar media plates. The repairing efficiency is calculated by dividing the number of recombinant cell by the number of the all cell surviving in the electroporation.

LHKHis-bet-pDH and LHKHis-bet-pBAD were transformed into DY380Δbet *Kan*^{s}*Cm*^{r} respectively. The colonies were inoculated for overnight shaking at 32°C at 250rpm. The overnight culture was transferred to 500ml LB at the ratio of 1:100 and

shaken at 32°C at 250 rpm, and then the protein expression was induced as procedure described above. 50ml culture was aliquoted and used to do competent cell preparation, ssDNA elcetroporation were performed as procedure described above. The 450ml culture left was used to purify His-tagged protein using 0.2ml Ni-NTA resin. The protein was eluted by 0.5ml elution buffer, and subjected to SDS-PAGE to check protein expression.

2.9.3. Deletion of pLysS using ssDNA

pLysS plasmid has Cm^r resistant gene, but the Cm^r gene is in the middle of Tet^r gene, and the Tet^r gene was inactivated, so the pLysS is Cm^rTet^s. If the Cm^r was deleted, the Tet^r gene will be repaired and be active again, and the modified pLysS will be Cm^sTet^r. (figure 18)

Two single strand oligonucleotides were designed and synthesized to delete to Cm^r gene in Tet^r gene. In the DNA upstream and downstream of Cm^r, the two designed single strand oligonucleotides both have homologous 40 bases which are homologous to flanks DNA of Cm^r, respectively. One single strand oligo nucleotides follows the direction of plasmid replication; the other one is reverse to the plasmid replication direction.

Lambda-*bet*-pDH / DH10B overnight culture was 1:100 transfer to 50ml LB in 250ml flask and shaken at 32 for 2. 5 hours, then the culture was shaken at 42°C water bath for 15minutes and used to prepare competent cells as procedure above. 1.24ng pLysS plasmid was electroporated to 90μL competent cell with various amount of single strand oligonucleotide, and then the electroporation product recovered at 1ml LB and shake 32°C for 2hours. The electroporation product was plate in LB agar plate. The deletion efficiency was calculated by dividing the number of cells survives in LB plate with 10μg/ml by the number of cell survives in LB plate with 34 μg/ml Cm. Then a curve of deletion efficiency vs ratio of pLysS/ single

strand oligonucleotide was plotted.

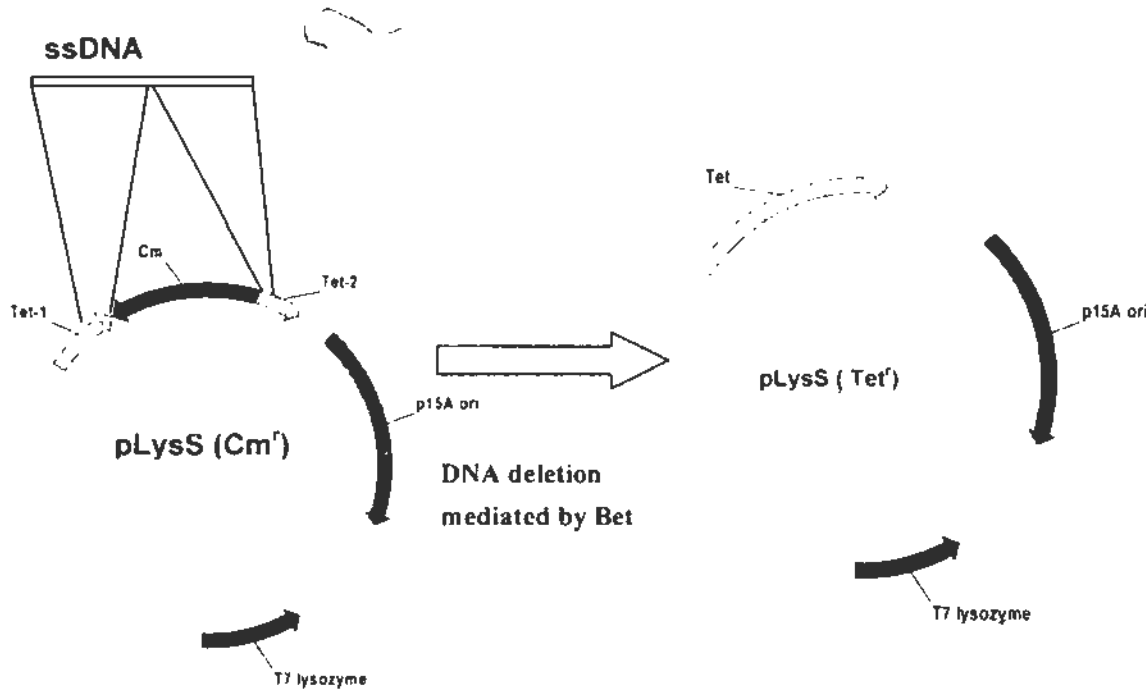


Figure 18. DNA fragment deletion mediated by Bet proteins

LHK-bet-pDH, lambda-bet-pDH, SXT-bet-pDH, SA-bet-pDH and pDH plasmid (as control, figure 62) were transformed into DH10B, and the colony was inoculated in LB for overnight incubation at 32°C. The overnight culture was 1:100 transferred to 50ml LB in 250ml flask and shaken for 2.5hours at 32°C, and then the culture was shaken at 42°C water bath for 15minutes and then used to prepare competent cells as procedure above. 1.24ng pLysS plasmid and 100ng of single strand oligonucleotide were electroporated to the competent cells together, and then electroporation product was recovered at 1ml LB and shaken 32°C for 2hours. The electroporation product was plate in LB agar plate. The deletion efficiency was calculated by dividing the number of cells survives in LB plate with 10μg/ml by the number of cell survives in LB plate with 34 μg/ml Cm.

LHK-bet-pBAD, lambda-bet-pBAD, SXT-bet-pBAD, SA-bet-pBAD and pBAD plasmids were introduced to DH10B, and the overnight culture was 1:100 transferred to 2 sets of 50ml LB in 250ml flasks and shaken for 1.25hours at 32 °C. Then 0.1%

arabinose was added to one set culture to induce protein expression for 1.5 hours, then cell culture is ready for competent cell peparation. Also, 0.1% of arabinose was added to another set and the culture were shaken for 1.25 hours, and then the flasks were shaken at 42 ℃water bath for 15 minutes before used to prepare competent cell. After preparation of competent cell as procedures decribed above, 1.24ng pLysS and 100ng of single strand oligonucleotide were electroporated to the competent cells. The electroporation product was recovered at 1ml LB and shake 32°C for 2 hours. Then the electroporation product was plate in LB agar plate. The deletion efficiency was calculated by dividing the number of cells survives in LB plate with 10μg/ml by the number of cell survives in LB plate with 34 μg/ml Cm.

Chapter Three Result

Part One, Protein expression and purification and pUC18/PstI dephosphorylation

3.1.1. SXT protein expression

Although SXT Exo had high expression level in whole cell (lane 3, Figure 19), the soluble expression was very low About 0.12 mg protein was obtained from 2L culture, which contains about 5g cell pellet. The purified SXT Exo protein was more than 95% purity after eluted from 1ml Hitrap chelating HP column as shown in figure 19 (lane 5). SXT Bet and SXT SSB protein had good soluble expressions. In the purification on 5ml Hitrap chelating HP column, the two proteins bound to the column tightly and the proteins were almost 100% pure after elution. After ion exchange purification on 6ml Resource Q column, no contaminant was found on SDS-PAGE stained with Comassie brilliant Blue R-250 (Figure 19).

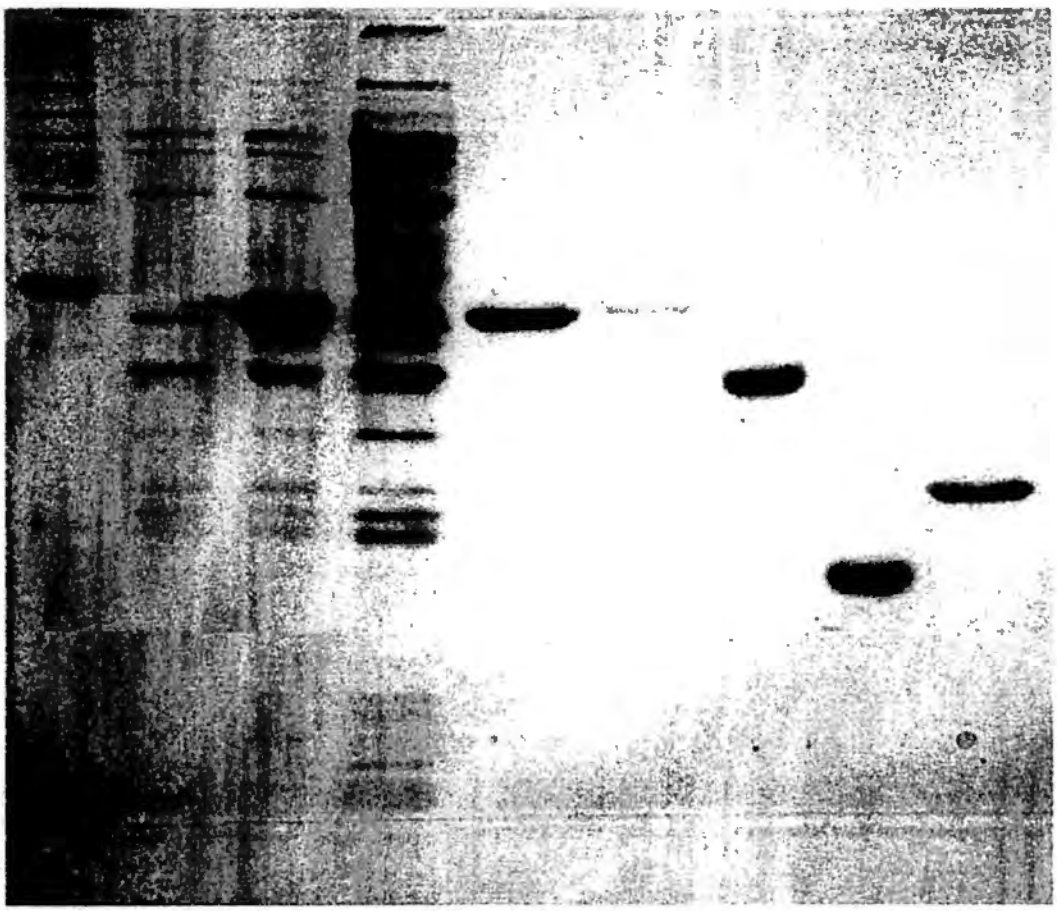
3.1.2. LHK protein expressions and purifications

LHK Exo and LHK Bet had high level of soluble expressions. The two proteins bound the Hitrap chelating HP column so tightly that they were almost 100% pure after elution. After ion exchange purification on 6ml Resource Q column, no contaminant was found on SDS-PAGE stained with Comassie brilliant Blue R-250 (Figure 20). LHK Bet also had good solublility, as the Bet proteins from SXT and bacteriophage lambda.

3.1.3. Linear dsDNA preparation and pUC18/PstI dephosphorylation

The pUC18 plasmid was cut completely by restriction enzymes, as shown in agarose gel (Figure 21). In the Picogreen assay, there was a linear relationship between fluorescence reading and the dsDNA within a range from 0.2875 to 28.75ng, which facilitated the nuclease assay very well (Figure 22).

Figure 19. The purification of SXT Exo and SXT bet, and SXT SSB and lambda Exo



1 2 3 4 5 6 7 8 9

Lane 1, BenchMark protein ladder (Invitrogen)

Lane 2, SXT-Exo-pET28a/BL21(DE3) whole cell without protein expression induction

Lane 3, SXT-Exo-pET28a/BL21(DE3) whole cell after protein expression induction

Lane 4, SXT-Exo-pET28a/BL21(DE3) supernatant after sonication and centrifuation

Lane 5, purified SXT Exo using 1ml Hitrap Chelating HP column

Lane 6, SXT Exo after desalting

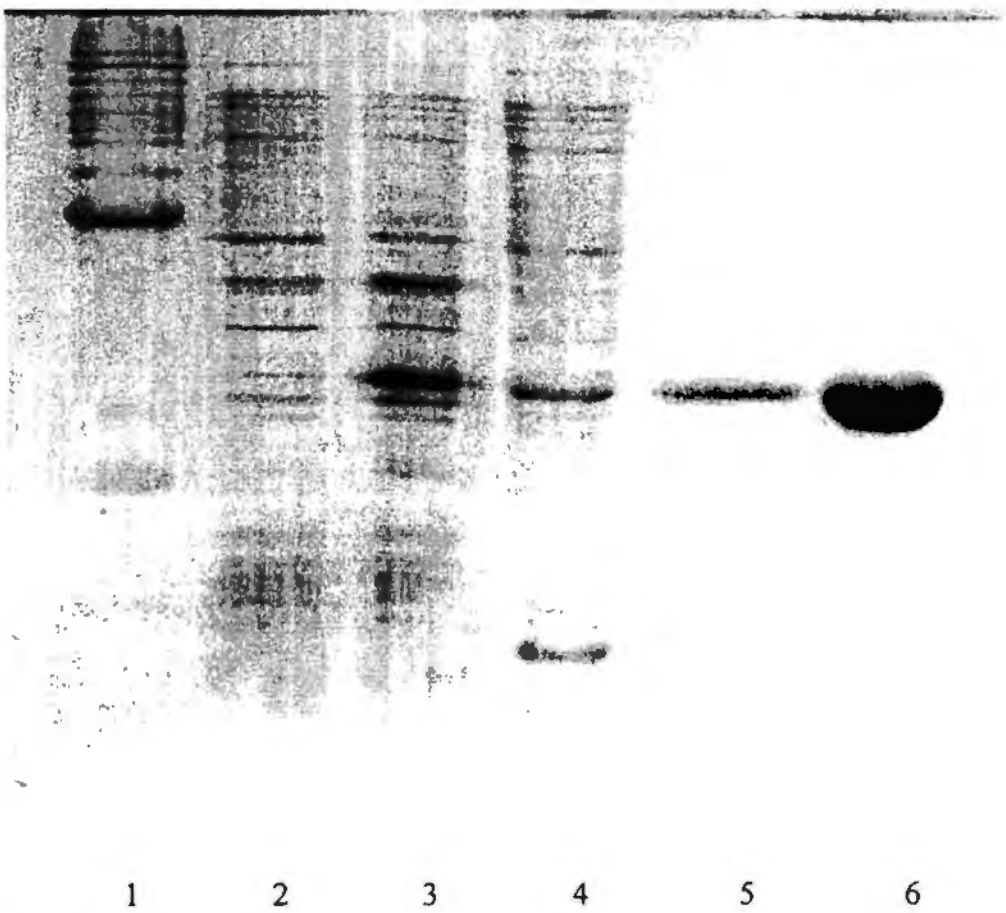
Lane 7 SXT bet protein after purification using Resource Q

Lane 8, SXT SSB protein after purification using Resource Q

Lane 9, lambda Exo protein after purification using Resource Q

Linear pUC18 DNA cannot self-ligate and form a closed DNA if the 5'- PO_4^{3-} is removed. The dephosphorylated pUC18/PstI had much less colonies than the un-dephosphorylated pUC18/PstI, there was more than 95% of linear pUC18/PstI dephophorylated as shown on Figure 23.

Figure 20. LHK protein expression and purification



Lane 1, BenchMark protein ladder

Lane 2, LHK-exo-pET28a/BL21(DE3) whole cell before protein expression induction

Lane 3, LHK-exo-pET28a/BL21(DE3) whole cell after protein expression induction

Lane 4, LHK-Exo-pET28a/BL21(DE3) supernatant after sonication and centrifugation

Lane 5, purified LHK Exo using 5mlHiTrap chelating column

Lane 6, LHK Exo protein after purification using Resource Q

Figure 21. Preparation of linear pUC18 substrate

Lane 1, 1kb+ DNA ladder (Invitrogen)

Lane 2, pUC18/PstI

Lane 3, pUC18/BamHI

Lane 4, pUC18/SspI

Lane 5, dephosphorylated pUC18/PstI

Figure 22. The pUC18/PstI dsDNA fluorescence reading

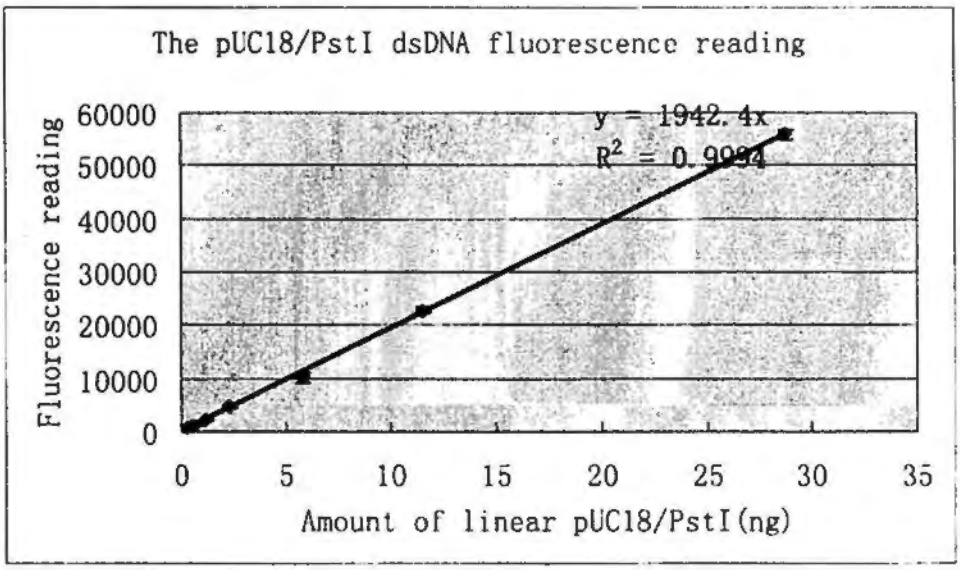
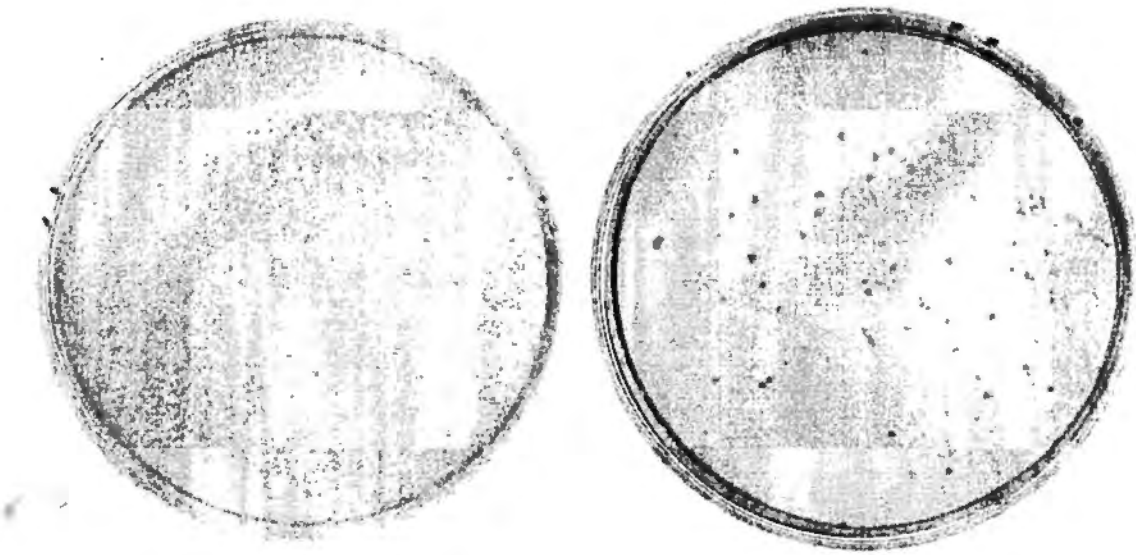


Figure 23. The percentage of pUC18/PstI dephosphorylation



A B

A, after pUC18/PstI linear dsDNA self-ligation, the ligation product was transform to DH10B competent cells

B, after dephosphated pUC18/PstI linear dsDNA self-ligation, the ligation product was transform to DH10B competent cells

Part Two, Biochemical assay of LHK Exo

As shown in Figure 24, 25, 26, LHK Exo had optimal $MgCl_2$ concentration of 7.5mM, optimal temperature of 42 °C and optimal pH 8.0. In the reaction mixture, 0.272pmol LHK Exo saturated the ends of linear dsDNA substrate (0.020pmol pUC18/PstI), more protein than 0.272pmmol could not further increase the digestion rate (Figure 27), and all the ends of linear dsDNA were saturated by LHK Exo. Calculated from slopes of the curves in the Figure 28, the rate for LHK Exo digested pUC18/PstI was 7.29±1.44 nucleotides per second at room temperature of 22℃, in 50mM Tris-HCl pH 8.0, 50 mM NaCl, 2.5 mM $MgCl_2$ buffer. In the fraction activity assay (Figure 29), 0.2742 pmol of LHK Exo monomer was enough to saturate the ends of 0.02 pmol linear dsDNA. So 43.8% of the all protein was still active after one freeze and thaw.

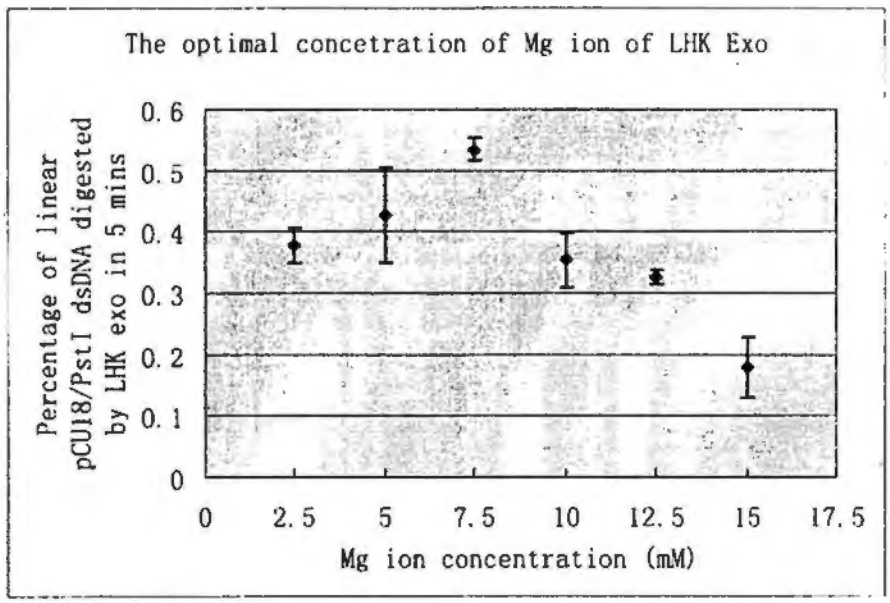


Figure 24. The optimal concentrationof Mg^{2+} ion of LHK Exo

Dephosphorylated pUC18/PstI was not a prefered substrate for LHK Exo, only a small portion was digested (Figure 30). The digestions of pUC18 by PstI, BamHI and SspI generated 4-base 3'-overhang dsDNA, 4-base 5'-overhang dsDNA, and blunt

end dsDNA, respectively. The initial degrading rate of 3'-overhang linear dsDNA (pUC18/PstI) digestion by LHK Exo was faster than 5'-overhang (pUC18/BamHI),

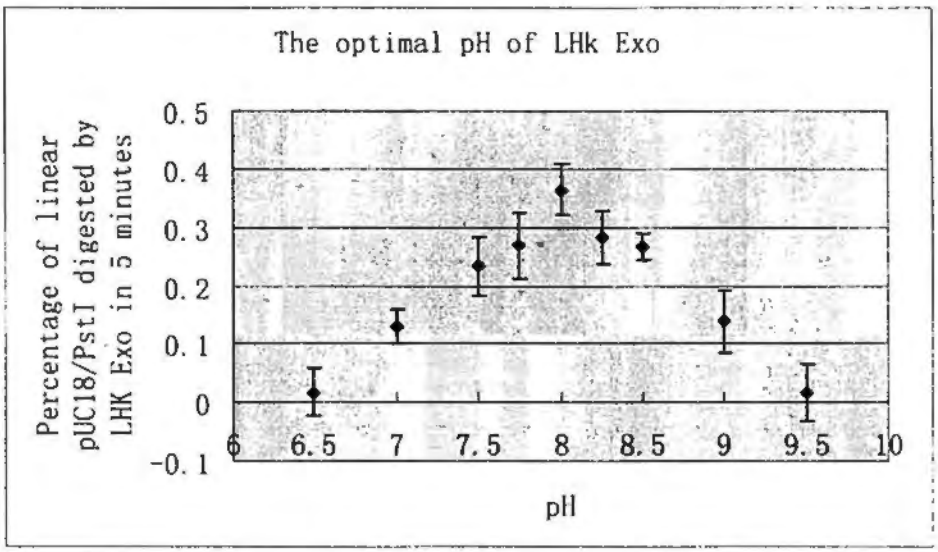


Figure 25. The optimal pH of LHK Exo

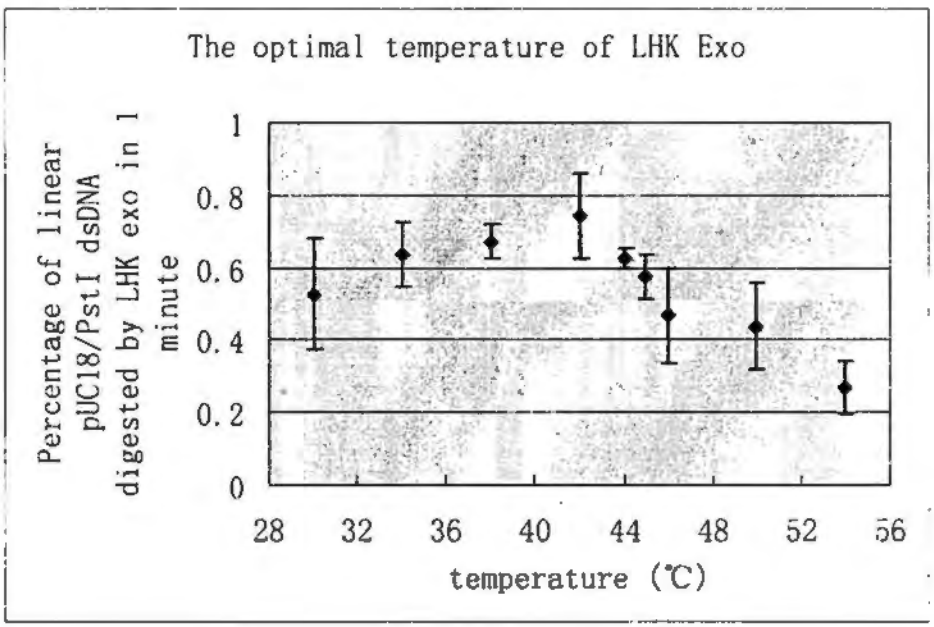


Figure 26. The optimal temperature of LHK Exo

and similar to blunt end dsDNA (pUC18/SspI) (Figure 30). LHK Exo did not cleave linear dsDNA from the middle, it digested substrate from the ends gradually, and the size of the DNA decreased gradually but relative unchanged as shown in Figure 31. LHK Exo activity required the presence of Mg^{2+} of Mn^{2+} (Figure 32).

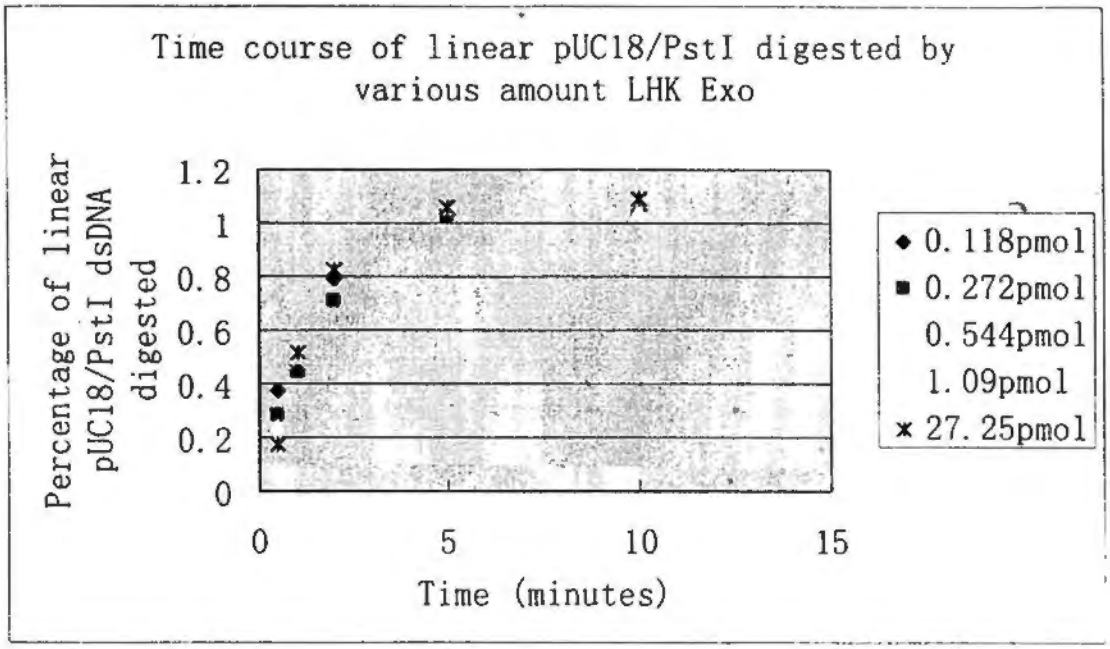


Figure 27. Timer course of linear pUC18/PstI digested by various amount LHK Exo

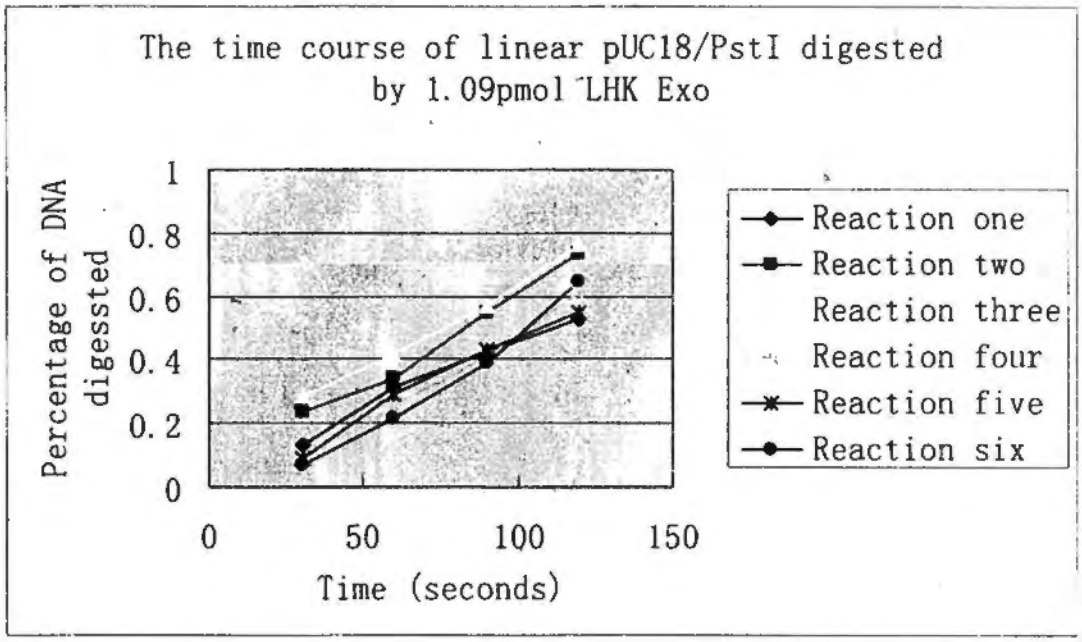


Figure 28. The time course of linear pUC18/PstI digested by LHK Exo

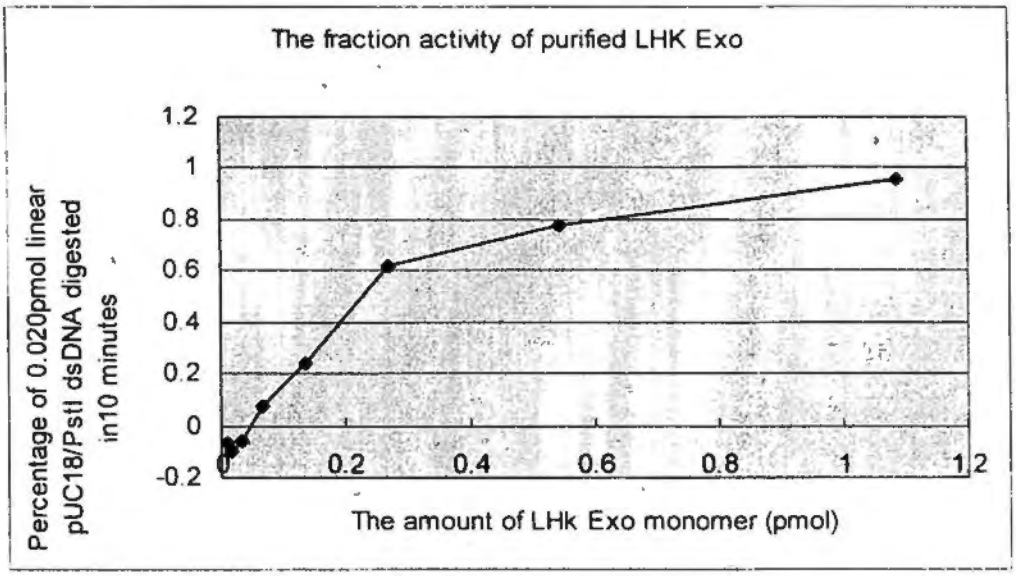


Figure 29. The fraction activity of LHK Exo

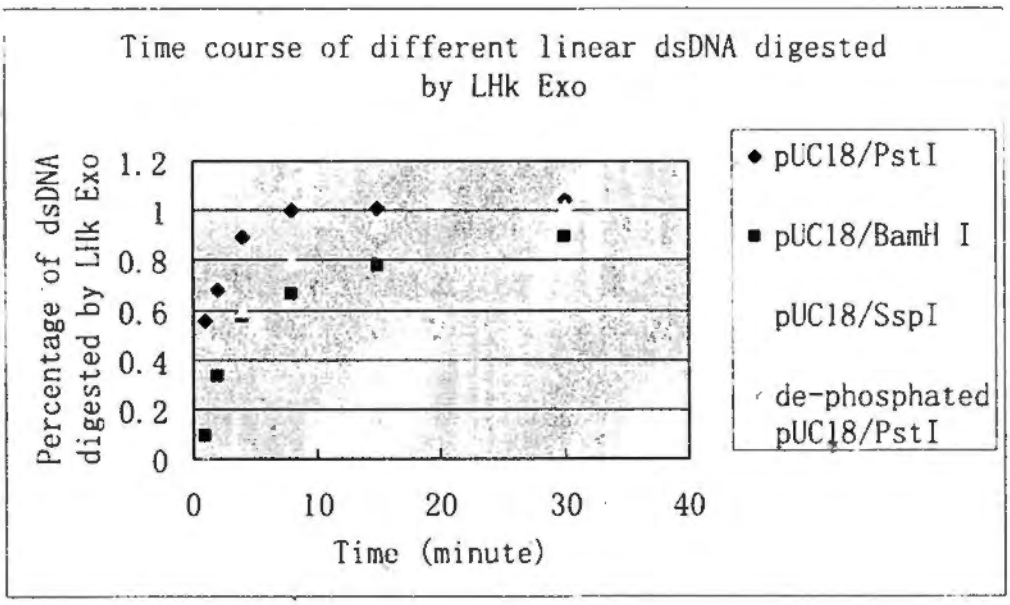
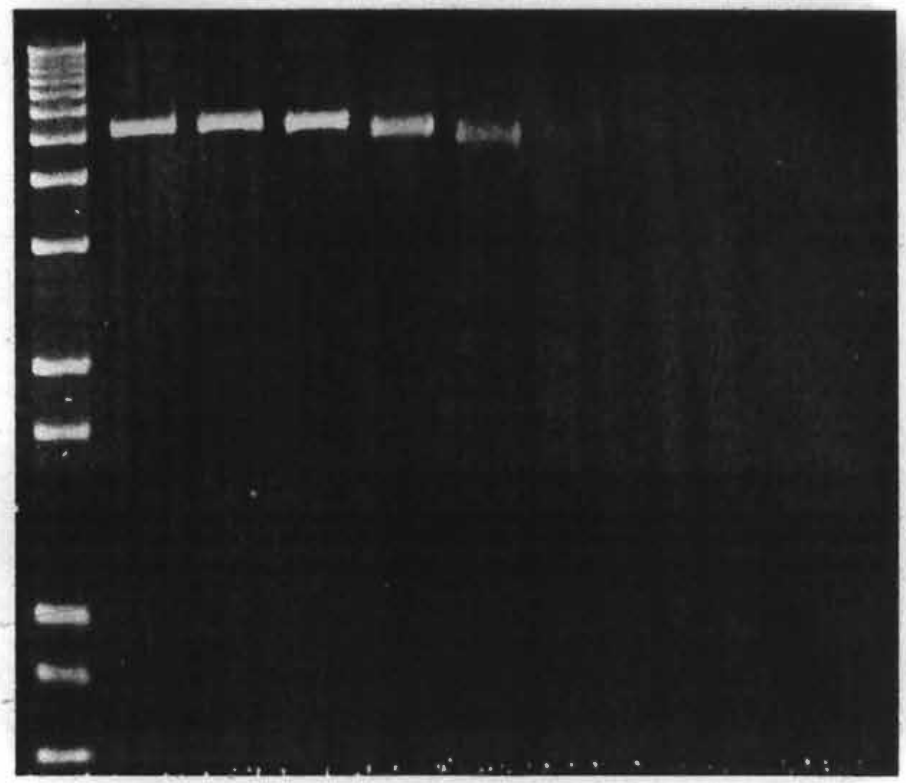


Figure 30. Time course study of diverse linear dsDNA digested by LHK Exo



1kb+ ladder 0 1 2 3 4 5 6 7 15min

Figure 31. The digestion time course of pET28a/NdeI by LHK Exo

Figure 32. The metal ion dependence of LHK Exo.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Lane 1, 1Kb(+) DNA ladder (Invitrogen); Lane 2, pET28a/NdeI DNA + Mg^{2+}; Lane 3, pET28a/NdeI DNA + LHK Exo + Mg^{2+}; Lane 4, pET28a/NdeI DNA + Mn^{2+}; Lane 5, pET28a/NdeI DNA + LHK Exo + Mn^{2}; Lane 6, pET28a/NdeI DNA + Fe^{2+}; Lane 7, pET28a/NdeI DNA + LHK Exo + Fe^{2+}; Lane 8, pET28a/NdeI DNA + Cu^{2+}; Lane 9, pET28a/NdeI DNA + LHK Exo + Cu^{2+}; Lane 10, pET28a/NdeI DNA + Co^{2+}; Lane 11, pET28a/NdeI DNA + LHK Exo + Co^{2+}; Lane 12, pET28a/NdeI DNA + Ca^{2+}; Lane 13, pET28a/NdeI DNA + LHK Exo + Ca^{2+}; Lane 14, pET28a/NdeI DNA + Ni^{2+}; Lane 15, pET28a/NdeI DNA + LHK Exo + Ni^{2+}; Lane 16, pET28a/NdeI DNA + Zn^{2+}; Lane 17, pET28a/NdeI DNA + LHK Exo + Zn^{2+};Lane 18, pET28a/NdeI DNA + LHK Exo only; Lane 19, pET28a/NdeI DNA only

Part Three Biochemical assay of SXT Exo

SXT Exo activity was depedent on the presence of Mg^{2+} or Mn^{2+} ions (Figure 33, 34).SXT Exo had optimal temperature of 41°C, optimal Mn^{2+} concentration of 3mM, optimal Mg^{2+} of 30mM, optimal pH of 8 (Figure 35-38). The initial rate of 3'-overhang linear dsDNA digestion by SXT Exo was faster than 5'-overhang substrate, and the initial rate of 5'-overhang linear dsDNA digestion by SXT Exo was faster than blant end dsDNA (Figure 39), only a small portion of dephosphorylated linear pUC18/PstI dsDNA was digested by SXT Exo. However, all dephosphorylated linear pUC18/PstI dsDNA could be digested if more SXT Exo was applied (Figure 40).

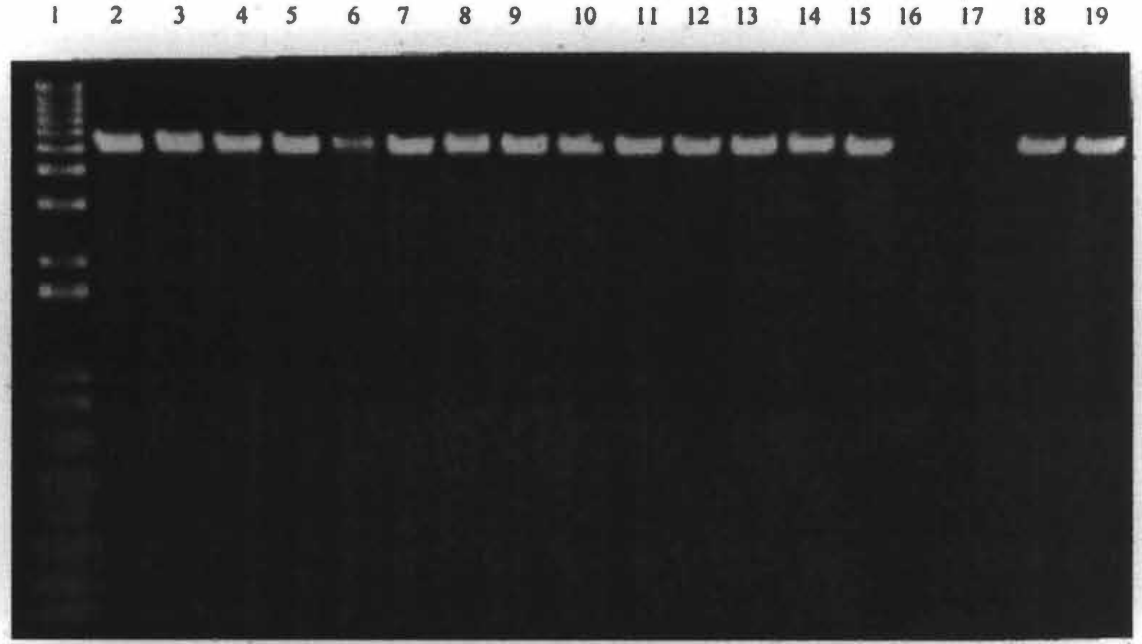
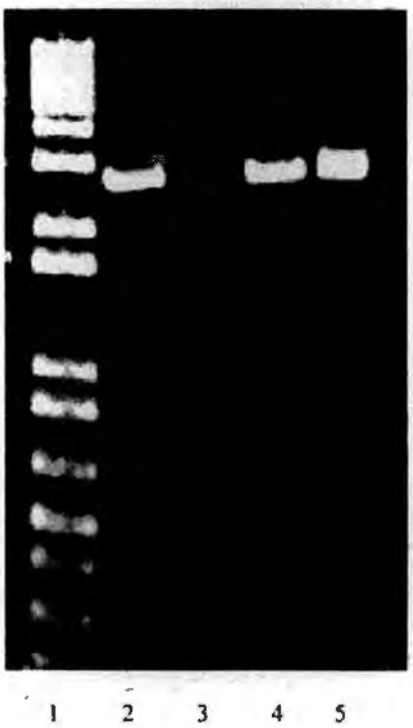


Figure 33. The metal ion dependence of SXT Exo

Lane 1, 1Kb(+) DNA ladder (Invitrogen); Lane 2, pET28a/NdeI DNA only; Lane 3, pET28a/NdeI DNA + SXT Exo only; Lane 4, pET28a/NdeI DNA + SXT Exo +Mg^{2+}; Lane 5, pET28a/NdeI DNA + Mg^{2}; Lane 6, pET28a/NdeI DNA + SXT Exo + Mn^{2+}; Lane 7, pET28a/NdeI DNA + Mn^{2+}; Lane 8, pET28a/NdeI DNA + SXT Exo + Ni^{2+}; Lane 9, pET28a/NdeI DNA + Ni^{2+}; Lane 10, pET28a/NdeI DNA + SXT Exo + Cu^{2+}; Lane 11, pET28a/NdeI DNA + Cu^{2+}; Lane 12, pET28a/NdeI DNA + SXT Exo + Ca^{2+}; Lane 13, pET28a/NdeI DNA + Ca^{2+}; Lane 14, pET28a/NdeI DNA + SXT Exo + Co^{2+}; Lane 15, pET28a/NdeI DNA + Co^{2+}; Lane 16, pET28a/NdeI DNA + SXT

Exo + Fe^{2+}; Lane 17, pET28a/NdeI DNA + Fe^{2+}, Lane 18, pET28a/NdeI DNA + SXT Exo + Zn^{2+}; Lane 19, pET28a/NdeI DNA + Zn^{2+}



Lane 1, 1Kb+ DNA marker (Invitrogen)
Lane 2, pET28a/NdeI only
Lane 3, pET28a/NdeI + 20mM $MgCl_2$ + SXT Exo
Lane 4, pET28a/NdeI + 20mM $MgCl_2$
Lane 5, pET28a/NdeI + SXT Exo

Figure 34. SXT Exo activity dependent on Mg^{2+} ion.

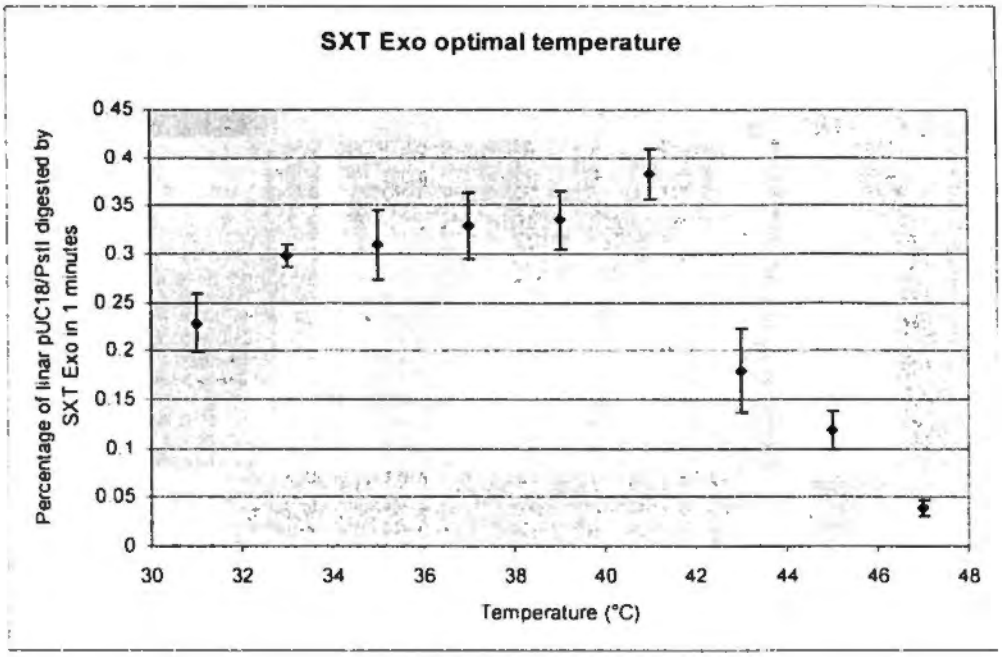


Figure 35. The optimal temperature of SXT Exo

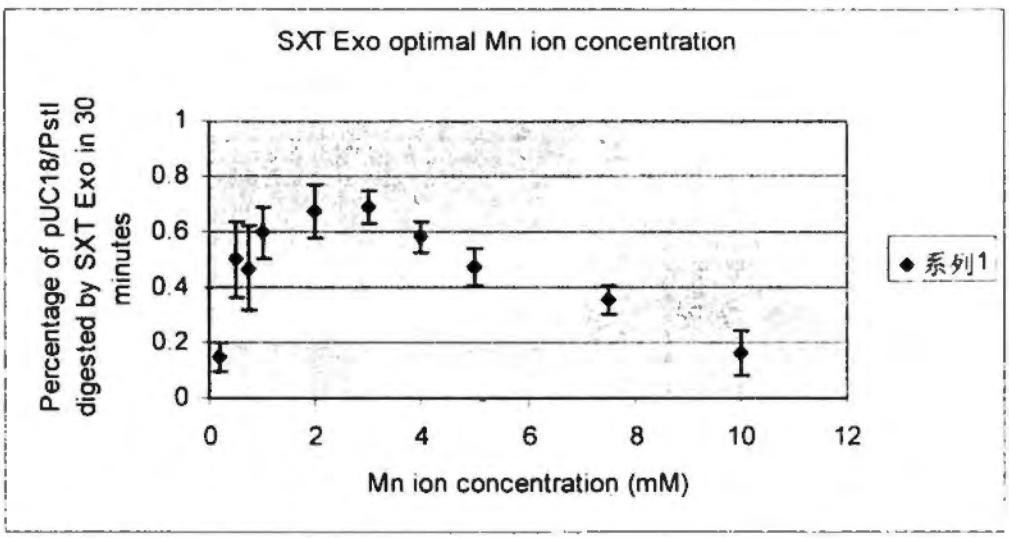


Figure 36. The optimal Mn^{2+} ion concentration of SXT Exo

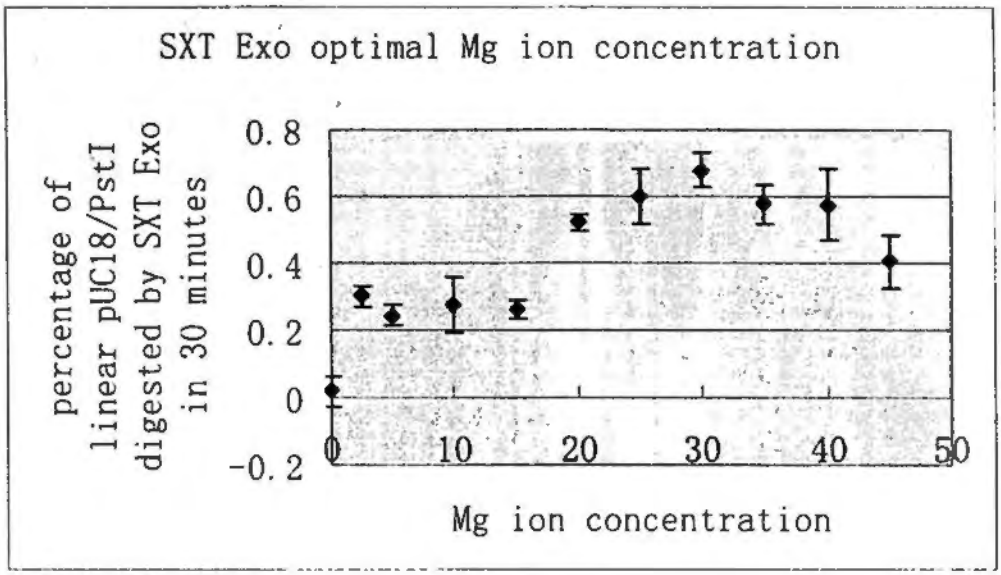


Figure 37. The optimal Mg ion concentration of SXT Exo

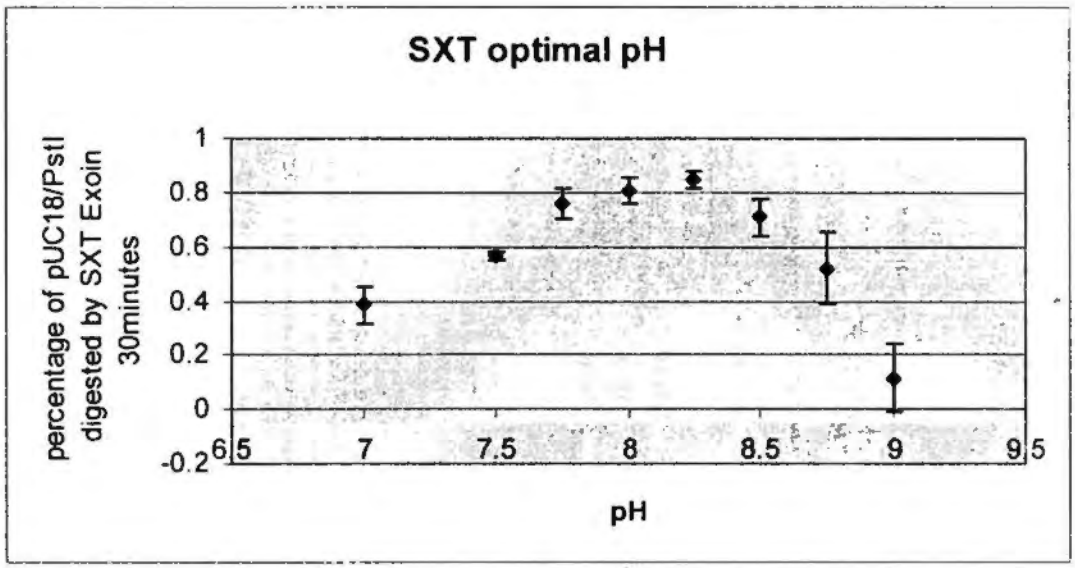


Figure 38. The optimal pH of SXT Exo

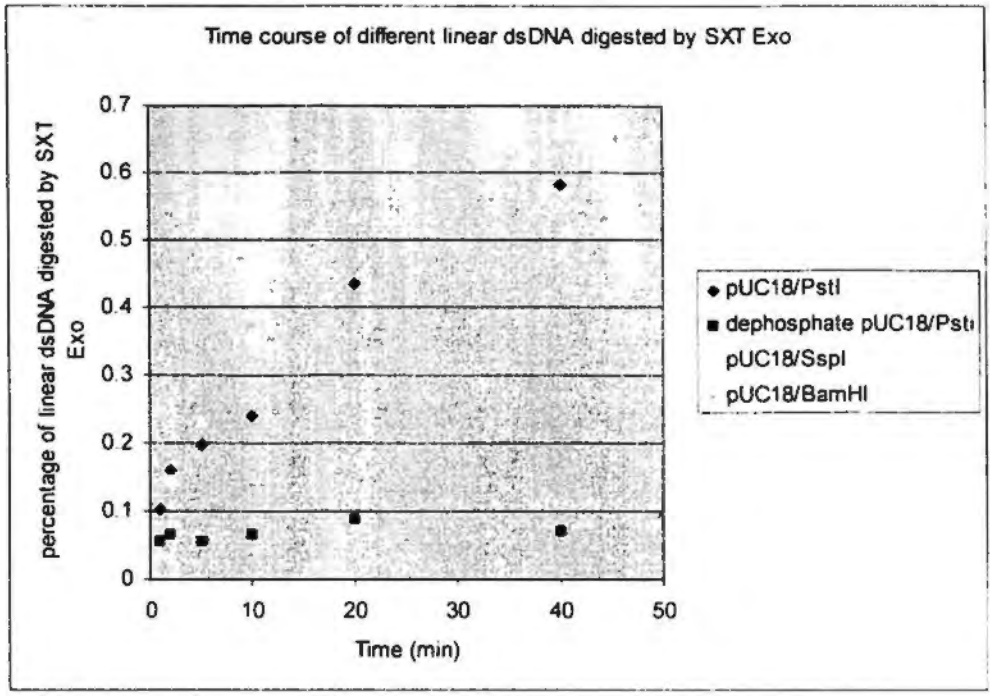
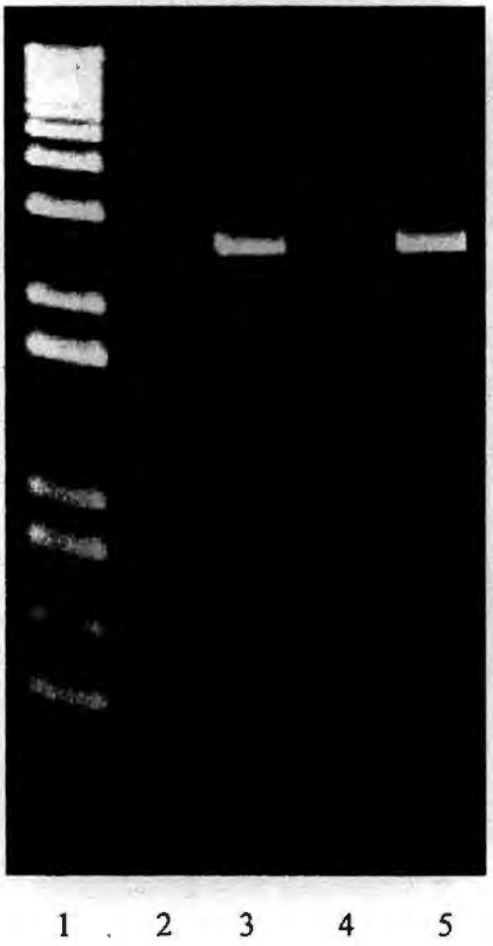


Figure 39. Time course study of different linear dsDNA digested by SXT Exo



Lane 1, 1Kb+ DNA marker (Invitrogen)

Lane 2, dephosphorylated pUC18/PstI + 1m $MnCl_2$ + SXT Exo

Lane 3, dephosphorylated pUC18/PstI + 1m $MnCl_2$

Lane 4, pUC18/PstI + 1m $MnCl_2$+ SXT Exo

Lane 5, pUC18/PstI + 1m $MnCl_2$

Figure40. dephosphorylated pUC18/PstI digested by SXT Exo

SXT Exo did not cleave the dsDNA from middle (Figure 41); it digested dsDNA from the ends gradually. Heparin inhibited the activity (Figure 42) of SXT Exo, the IC_{50} was about 0.05 mg/ml (~2.78μM). Salt concentrations affected SXT Exo activity (figure 43). Even at low salt concentrations (<1.3mM NaCl), the activity of SXT was similar to the activity at standard salt concentration (25mM Tris-HCl and 50mM NaCl). However, higher NaCl concentrations (>200mM) inhibited the Exo activity, Exo protein almost had not activity when KCl concentration was not less than 300mM, while the $CaCl_2$ inhibited the activity much, even at low concentrations (<100mM). Noticeably, at the low salt concentration (50mM -100mM), SXT Exo activity in buffer with KCl had a little bit higher activity than that in buffer with NaCl (Figure 42).

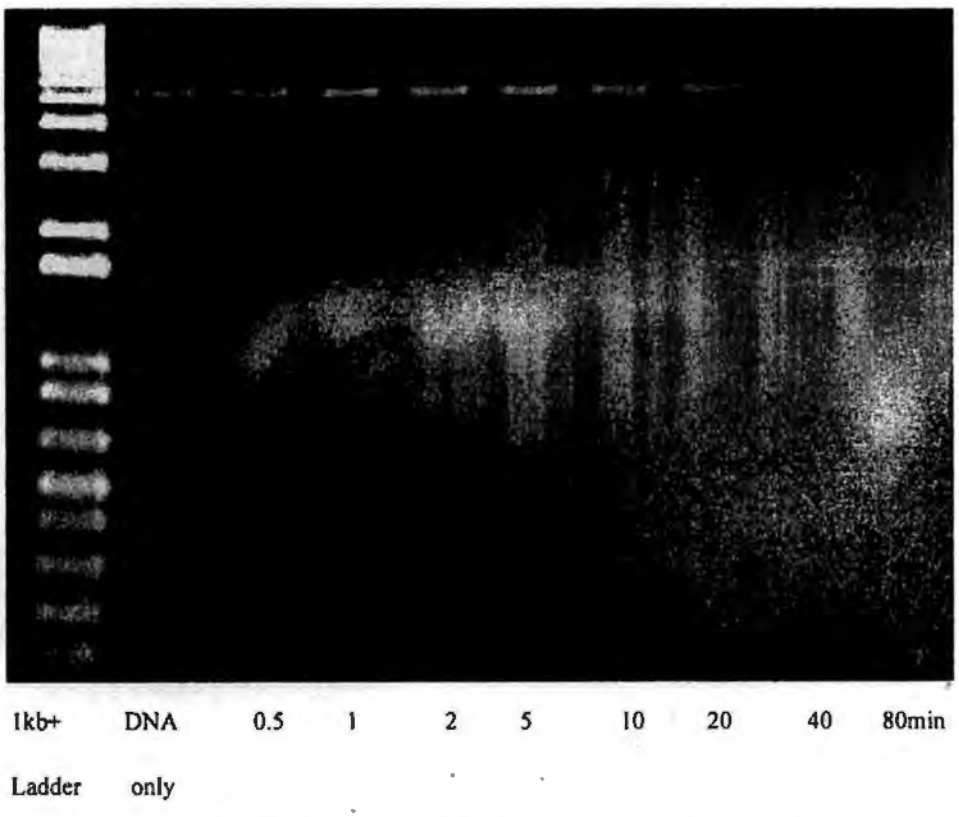


Figure 41. Time course for pET28a/NdeI digested by SXT Exo

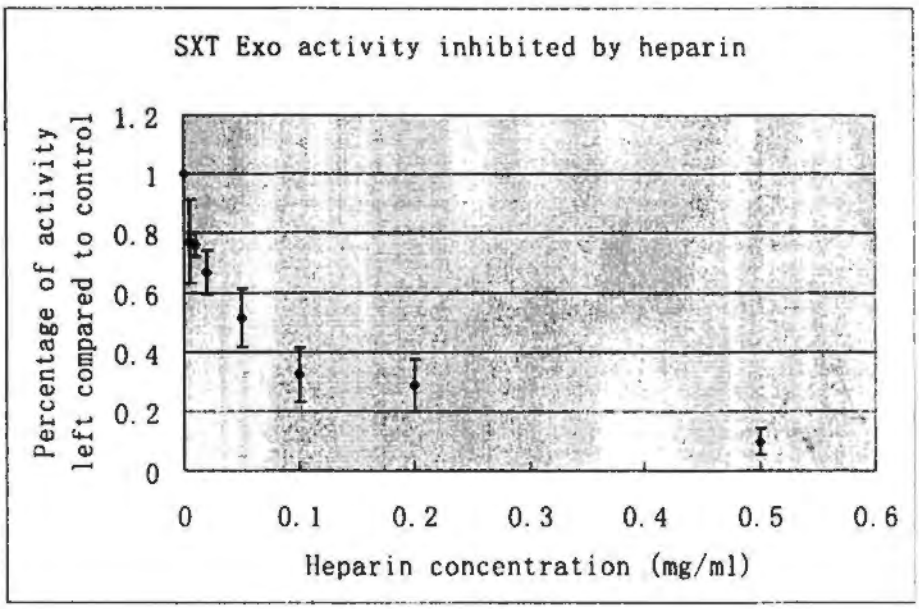


Figure 42. SXT Exo activity inhibited by heparin

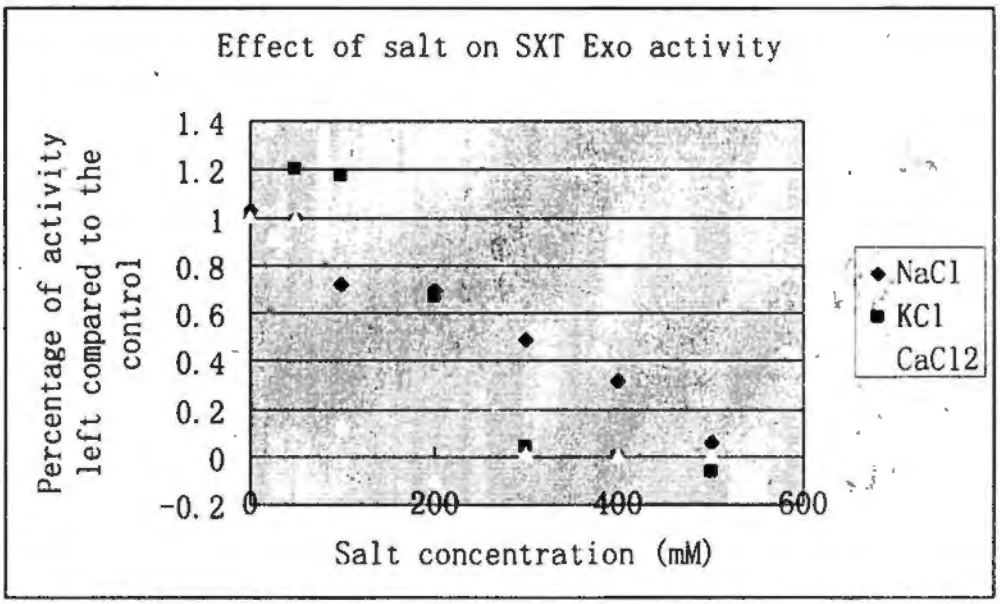


Figure 43. Salt effects on SXT Exo activity

Higher concentration of Tris-HCl buffer inhibited SXT Exo activity, although the pHs were the same (Figure 44). The enzyme almost had no activity at 200mM Tris-HCl, although it still had rather much activity at the buffer of 25mM Tris-HCl, pH7.4, 200 mM NaCl buffer. Sodium phosphate buffer pH7.4 inhibited SXT Exo activity much more than Tris-HCl buffer, the enzyme almost had no activity in 100

mM sodium phosphate buffer. SXT Exo had no activity in 25mM MOPS buffer (pH7.4) and 25 mM potassium phosphate buffer (pH7.4) (data not shown). Sodium sulphate buffer and potassium sulphate buffer inhibited SXT Exo activity (Figure 45).

In the reaction mixture in which there was enough SXT Exo to saturate all the ends of linear dsDNA, increase of enzyme could not further increase slope of digestion curve if the amount of the SXT Exo protein was not less than 6pmol, as shown in Figure 46, which indicates that all the ends of linear dsDNA were bound by the Exo protein and digestion was started. From the slopes, it could be calculated that the digestion velocity is 23.8±4.3 nucleotides per minute.

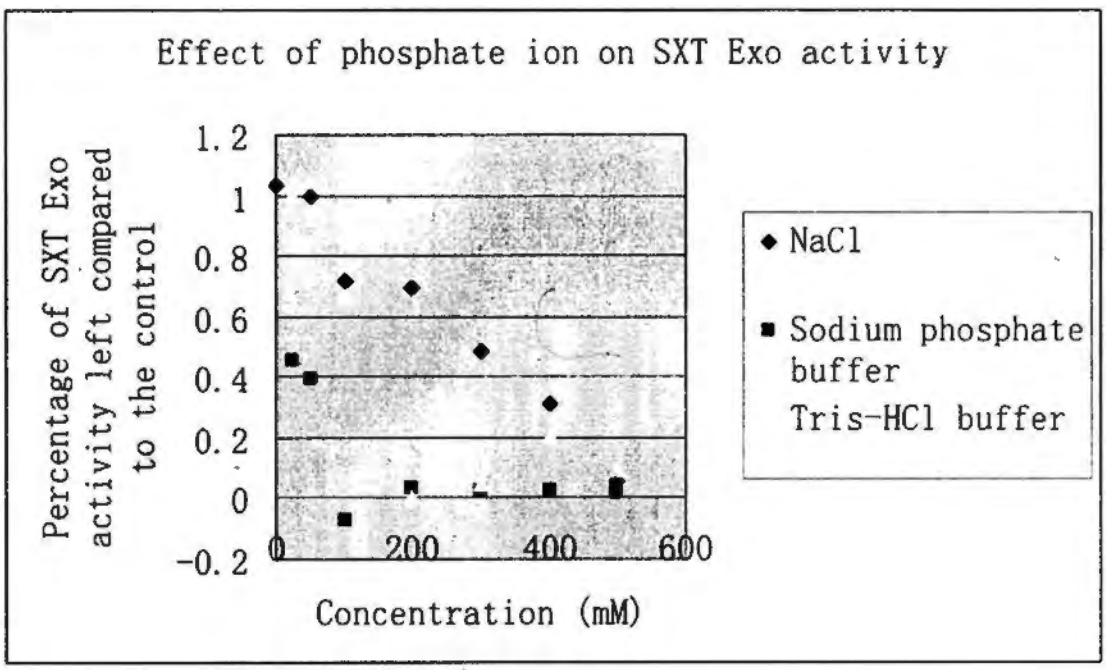


Figure 44. Phosphate (PO_4^{3-}) ion effect on SXT Exo activity

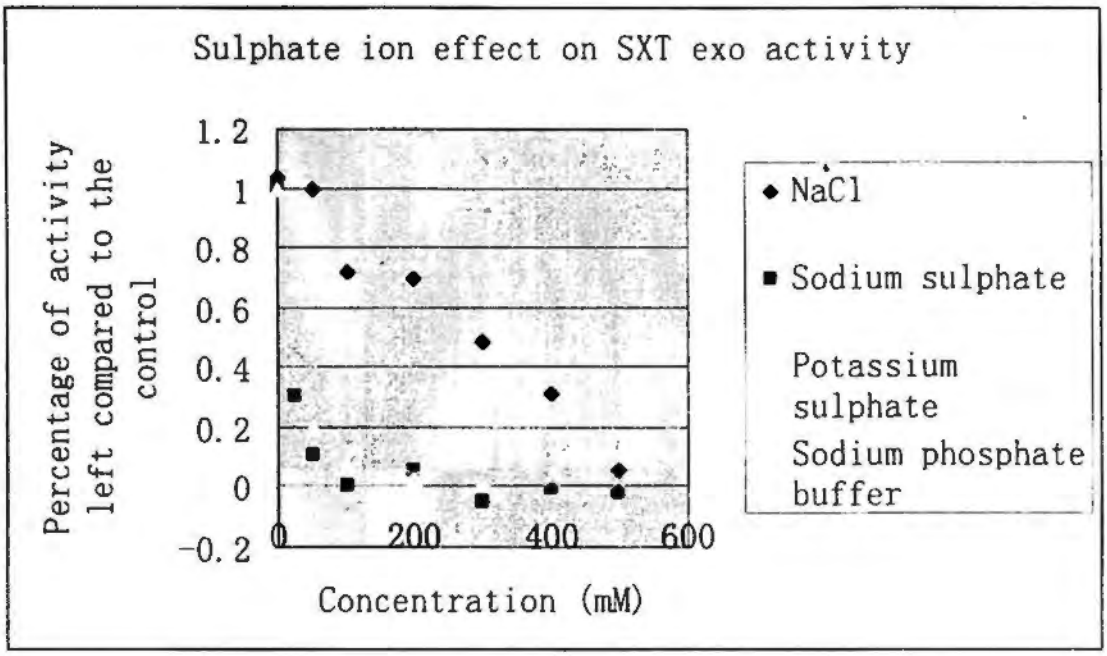


Figure 45. Sulphate (SO_4^{2-}) ion effect on SXT Exo activity

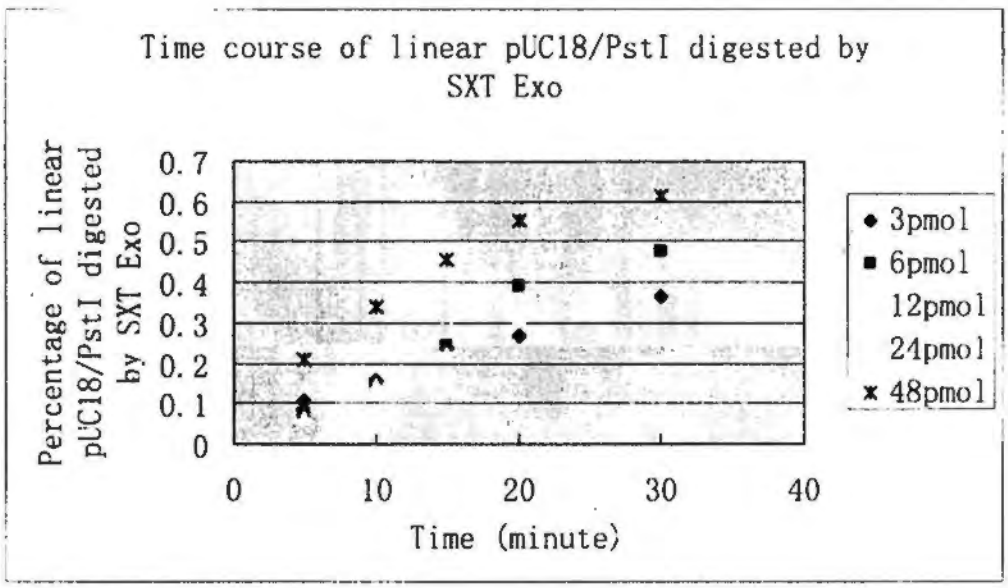


Figure 46. Time course study of defferent linear dsDNA digested by SXT Exo

Part Four, biophysical assay of SXT and LHK proteins

From the cross-linking, it seems that LHK Exo (Figure 47) could be dimer, but crystalogrphy-X-ray study (perdonal communication with Yan Wen in Nankai University, China) and gel filtration study showed that it was a homotrimer, and LHK Bet (Figure 48), SXT Bet (Figure 49), SXT SSB (Figure 50) seems to be dimer.

Gel filtrations show that apparent molecular weight of native SXT Exo is 145 kDal (~3.5mer, Figure 51) and apparent molecular weight of native bacteriophage lambda Exo is 91 kDal (3.35mer, Figure 52), apparent molecular weight of native LHK Exo 55kDal (~2.17mer, Figure 53).

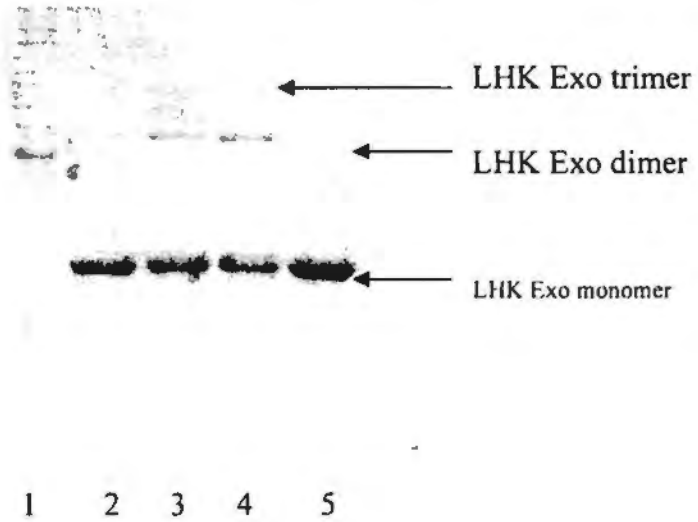


Figure 47. Cross-linking of LHK Exo

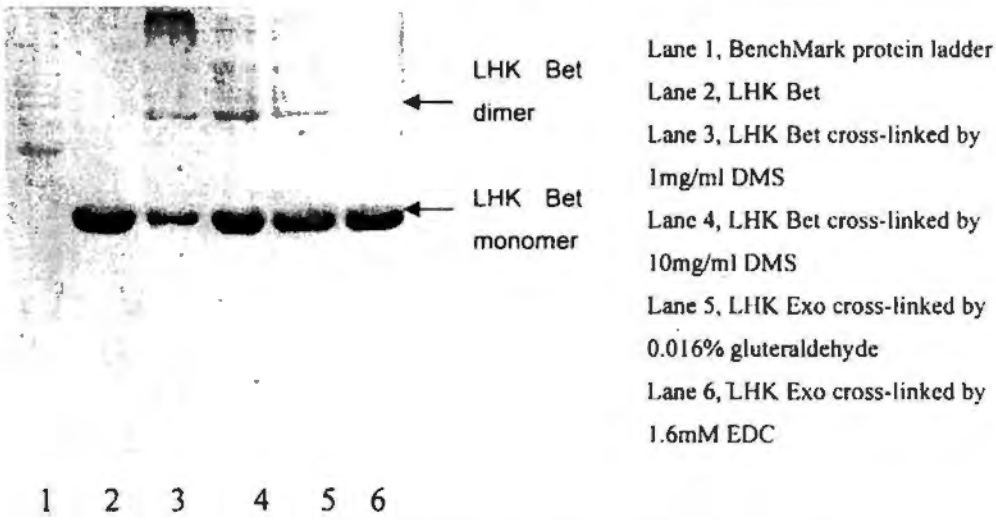


Figure 48. Cross-linking of LHK Bet

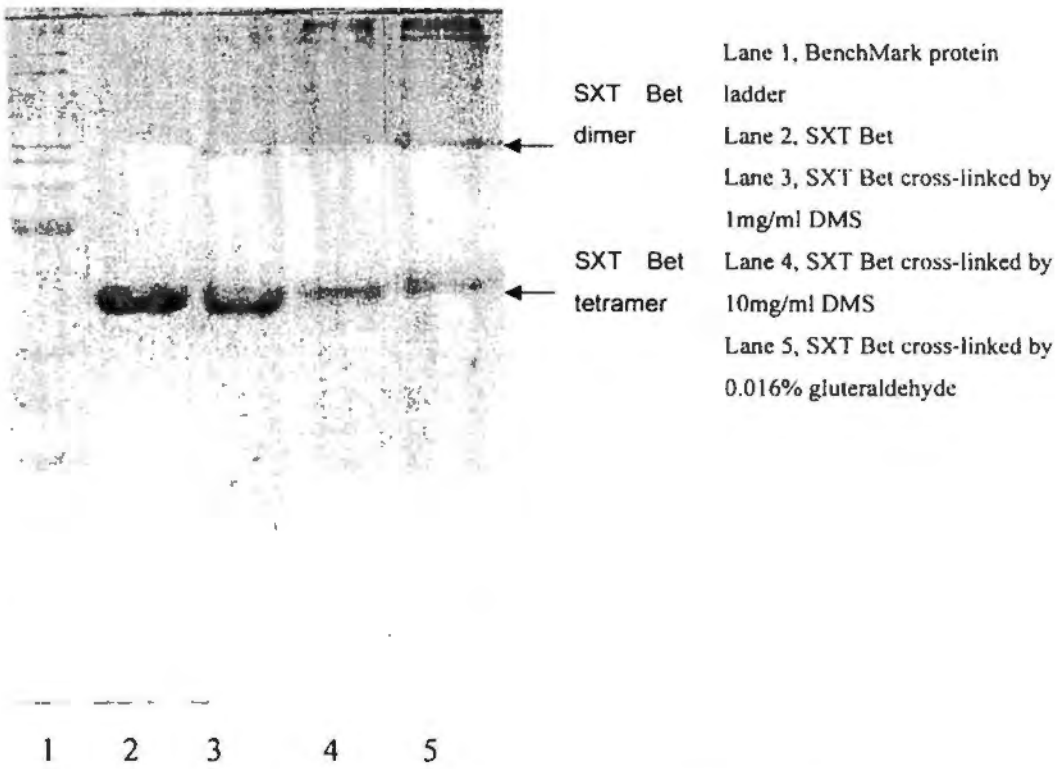


Figure 49. Cross-linking of SXT Bet

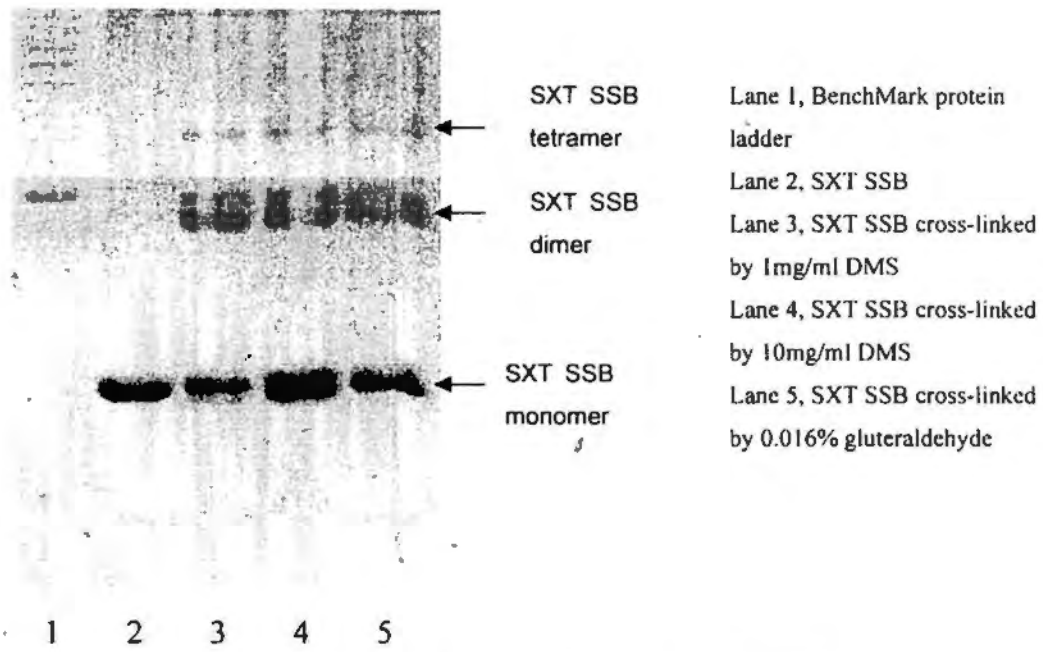


Figure 50. Cross-linking of SXT SSB

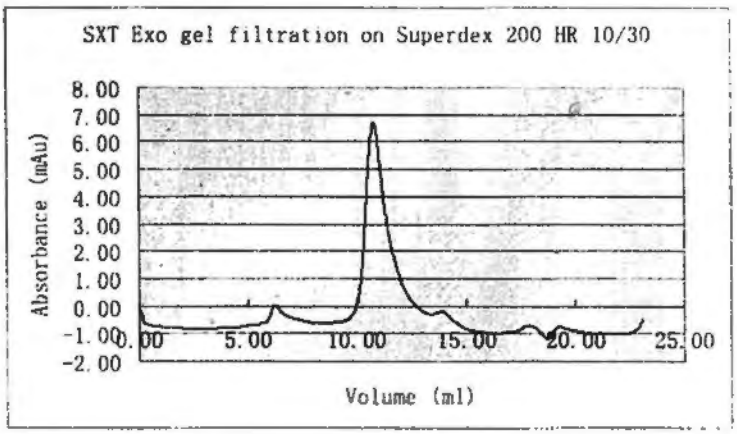


Figure 51. SXT Exo gel filtration on Superdex 200 HR10/30 column

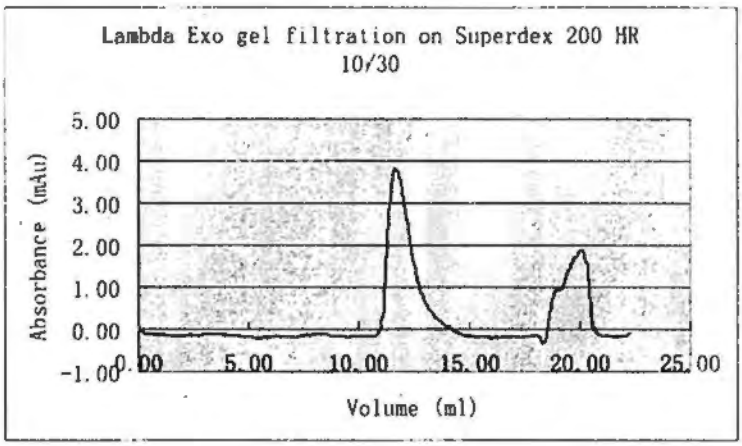


Figure 52. lambda Exo gel filtration on Superdex 200 HR10/30 column

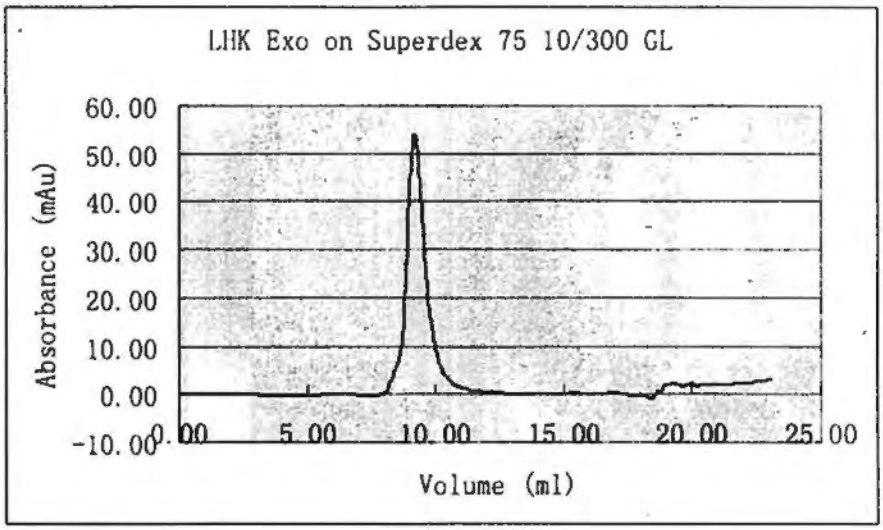


Figure 53. LHK Exo gel filtration on Superdex 75 10/30 GL

The apparent molecular weight of native LHK Bet is about 399kDal in gel filtration in PBS buffer (Figure 54). In high salt concentration buffer (25mM Tris-HCl, pH7.4, 1M NaCl), the native state of SXT Bet is different from the that in low salt buffer (25mM Tris-HCl, pH7.4, 150mM NaCl) buffer (Figure 55, 56). At the pH7.4 and pH 8.0 sodium phosphate buffers, the native state of SXT does not have much difference (Figure 57,58). SXT Bet seems to aggregate in a large molecule. SXT Bet forms larger molecules in the presence of Mg^{2+} ion, and the overnight incubation of Mg^{2+} ion with SXT Bet make SXT Bet forms more large molecule than short time incubation (Figure 59, 60, 61).

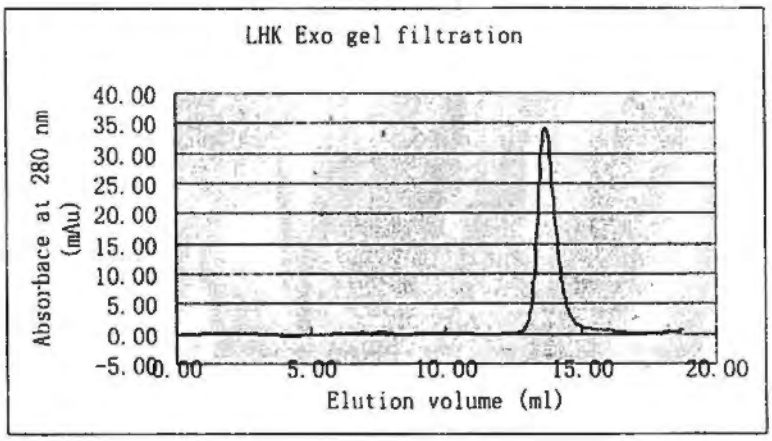


Figure 54. LHK Exo gel filtation on Tricorn Superdex 200HR 10/300GL column on an AKTA-FPLC pre-equilibrated 25mM Tris-HCl pH7.4, 150mM NaCl, 1mM EDTA, 5mM imidazole

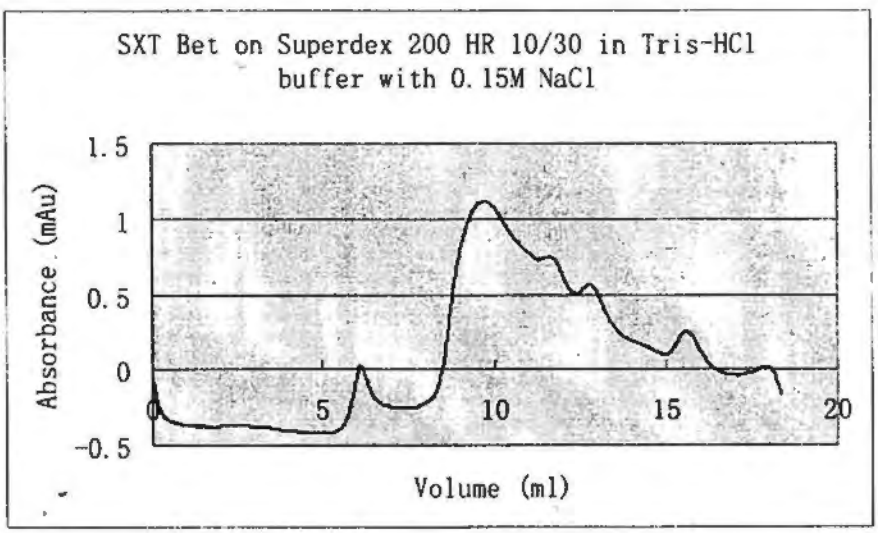


Figure 55. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM Tris-HCl, pH7.4, 150mM NaCl

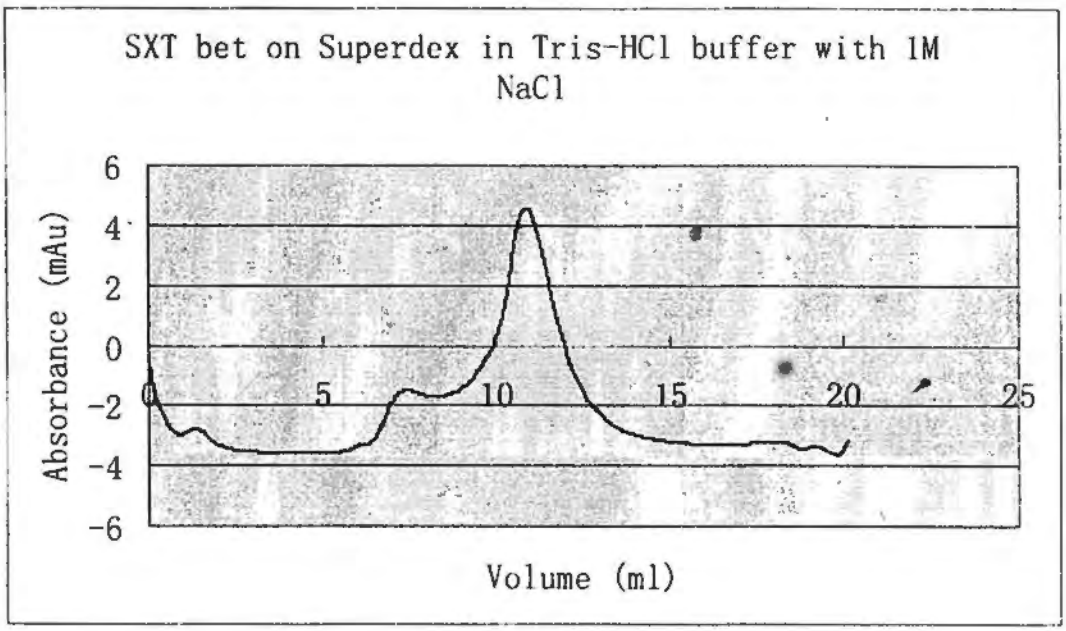


Figure 56. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM Tris-HCl, pH7.4, 1M NaCl

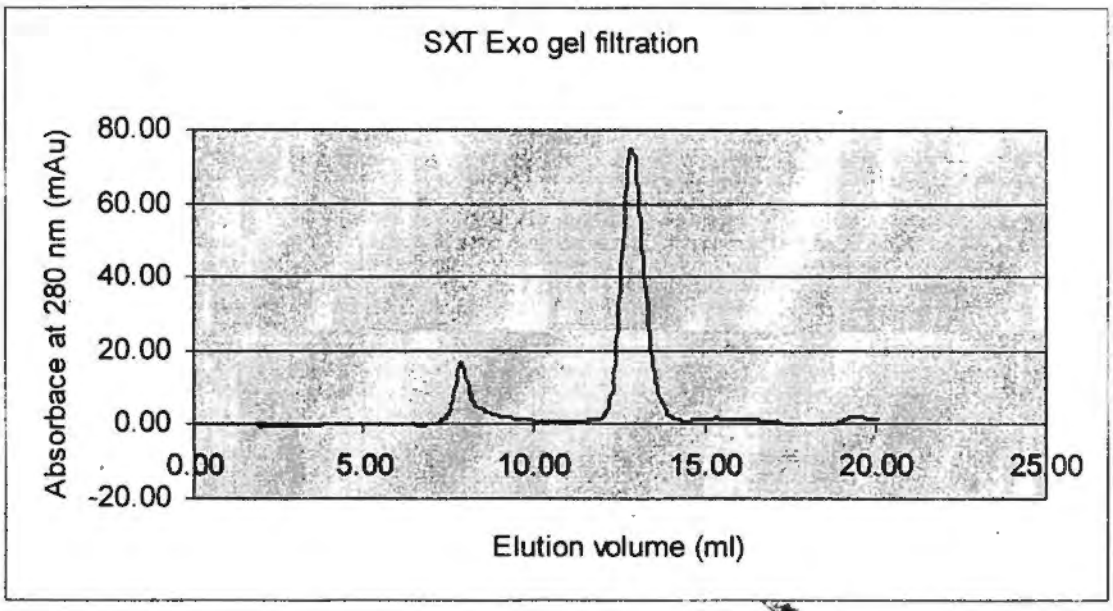


Figure 57. SXT Exo gel filtation on Tricorn Superdex 200HR 10/300GL column on an AKTA-FPLC pre-equilibrated 25mM Tris-HCl pH7.4, 150mM NaCl, 1mM EDTA, 5mM imidazole

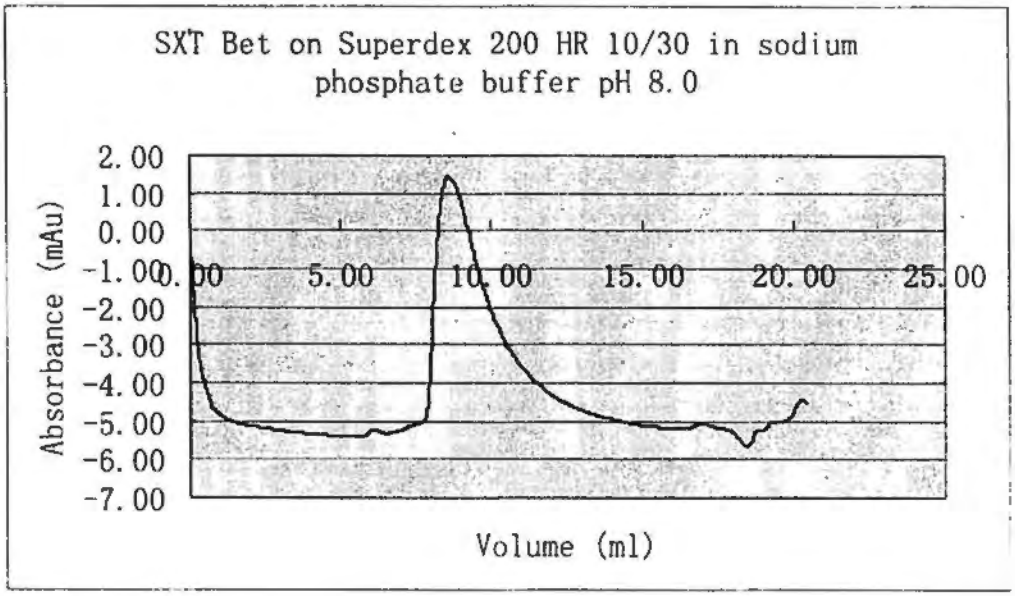


Figure 58. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM sodium phosphate buffer, pH8.0, 150mM NaCl

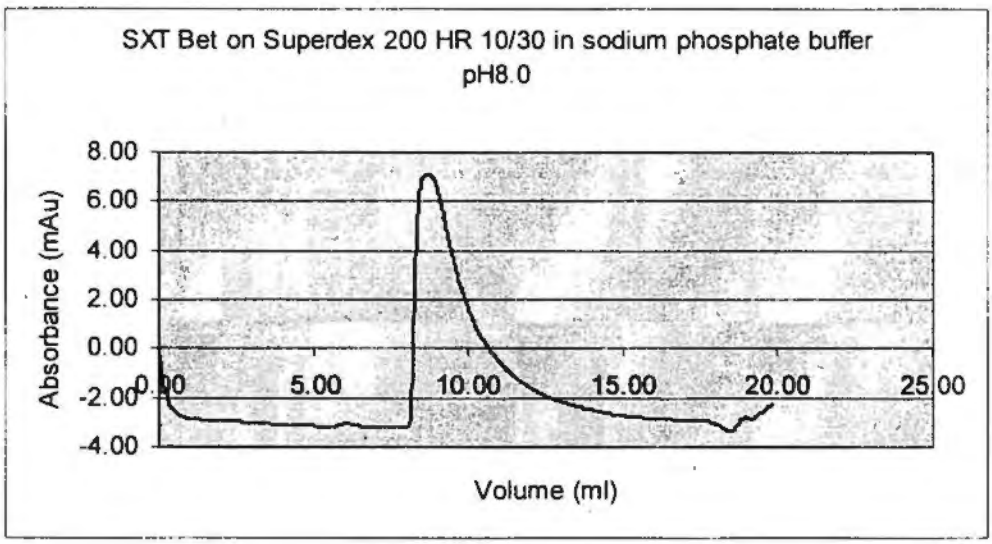


Figure 59. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM sodium phosphate buffer, pH7.4, 150mM NaCl, 1mM EDTA

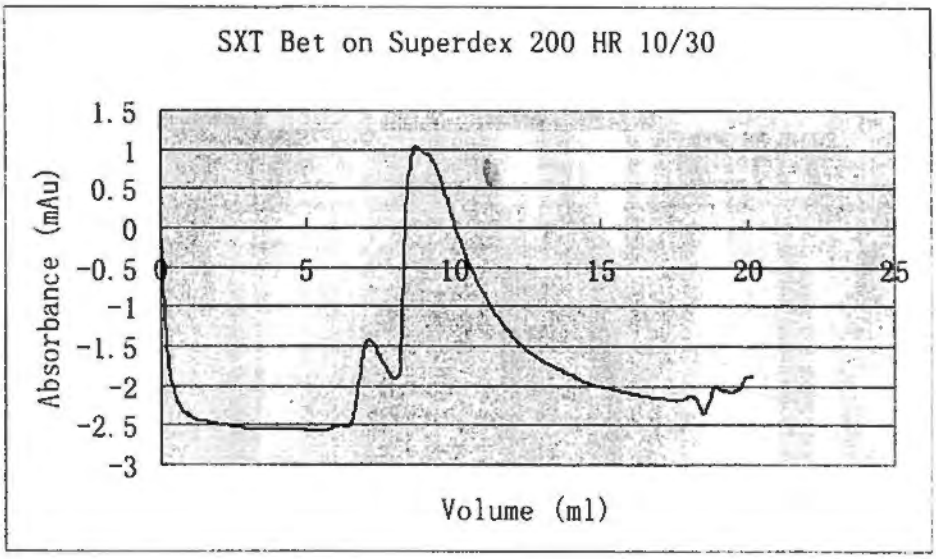


Figure 60. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM sodium phosphate buffer, pH7.4, 150mM NaCl, 1mM MgCl2, (protein has been incubated in baffer for 15 minutes before chromatograph column)

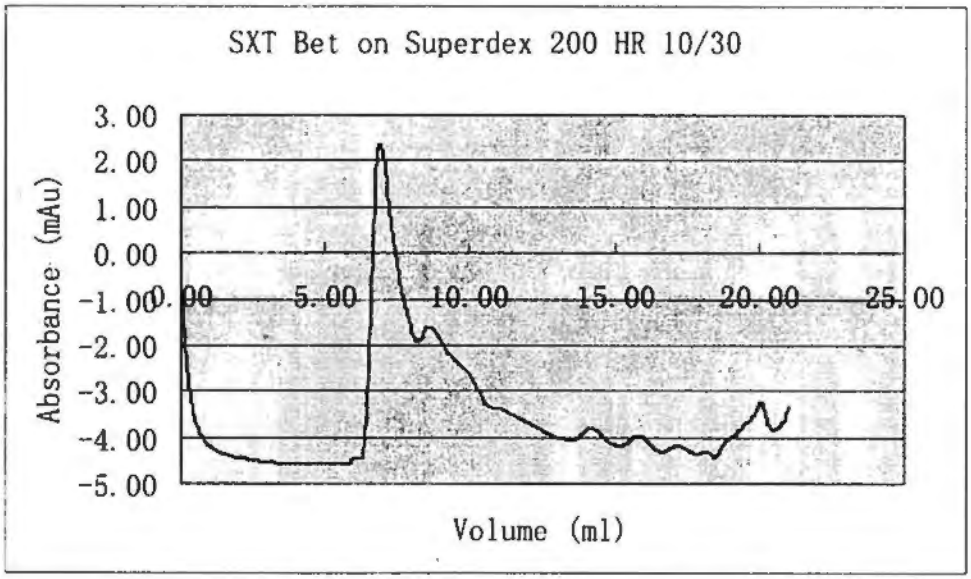


Figure 61. SXT Bet gel filtration on Superdex200 HR 10/30 in 25mM sodium phosphate buffer, pH7.4, 150mM NaCl, 1mM MgCl2, (protein has been incubated in buffer overnight before chromatograph column)

The interaction between LHK Exo and LHK Bet protein was confirmed *in vitro* (Figure 62) and *in vivo* (Figure 63), respectively.

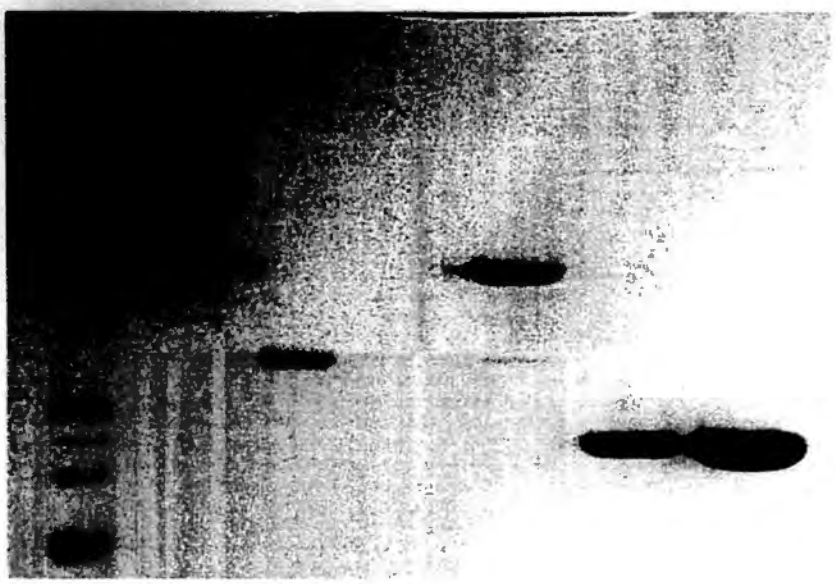


Figure 62. LHK Exo and LHK Bet interaction *in vivo*

Lane 1, BenchMark Protein ladder

Lane 2, LHKexo-pGex-4T1/BL21(DE3) extract loaded to 1ml GSTrap FF column and eluted with 10mM GSH in buffer

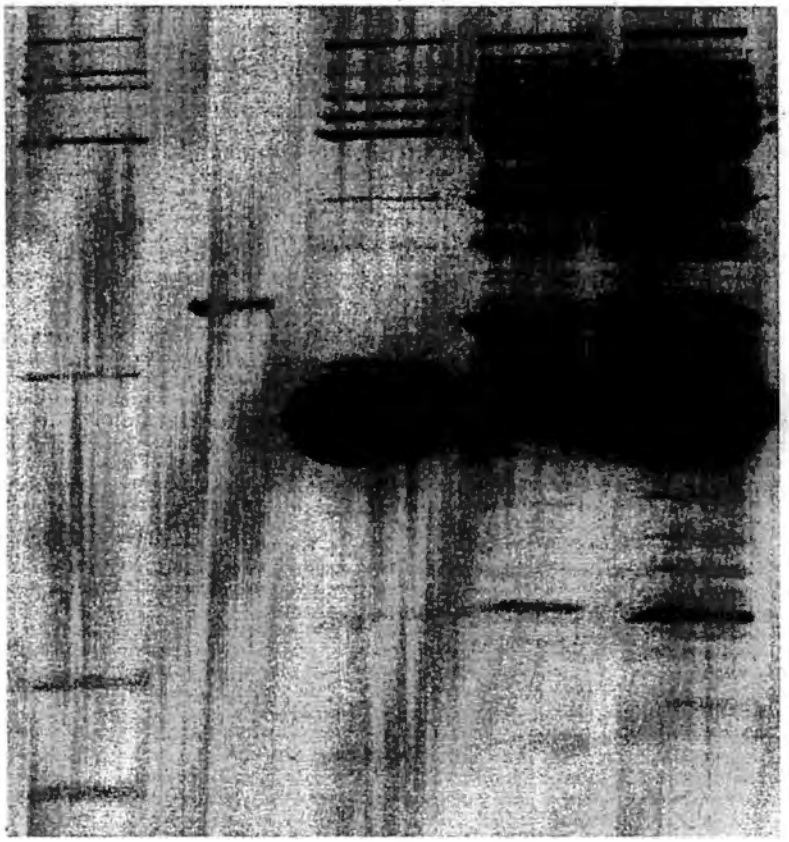
Lane 3, purified His-tagged LHK Bet proteins

Lane 4, LHKbet-pET28a/BL21(DE3) extract loaded to 1ml GSTrap FF column and eluted with 10mM GSH in buffer

Lane 5, LHKexo-pGex-4T1/BL21(DE3) extract loaded to 1ml GSTrap FF column, and then LHKbet-pET28a/BL21(DE3) extract loaded to 1ml GSTrap FF column, finally eluted with 10mM GSH in buffer

Lane 6, pGex-4T1/BL21(DE3) extract loaded to 1ml 1ml GSTrap FF column and then eluted with 10mM GSH in buffer

Lane 7, pGex-4T1/BL21(DE3) extract loaded to 1ml 1ml GSTrap FF column and then eluted with 10mM GSH in buffer, and then LHKbet-pET28a/BL21(DE3) extract loaded to 1ml GSTrap FF column, finally eluted with 10mM GSH in buffer



1 2 3 4 5

Figure 63. LHK Exo and LHK Bet interaction *in vitro*

Lane 1, protein ladder (Biorad broad range proteinladder)
Lane 2, purified His-tagged LHK Bet protein
Lane 3, purified His-tagged LHK Exo protein
Lane 4, LHKbet-exo-pET32a elution 2
Lane 5, LHKbet-exo-pET32a elution 3

Part Five, biological activity of the recombineering systems and their components

LHKBet-Exo-SSb function had highest efficiency, and LHKBet-Exo function had lower activity, pB1E4A, pBADETγ had similar activities, and pBX2B, pGX2B, pEXS1A, pαKX2A, pBex4b1 has lowest activity (Figure 64A &B and Table 2).

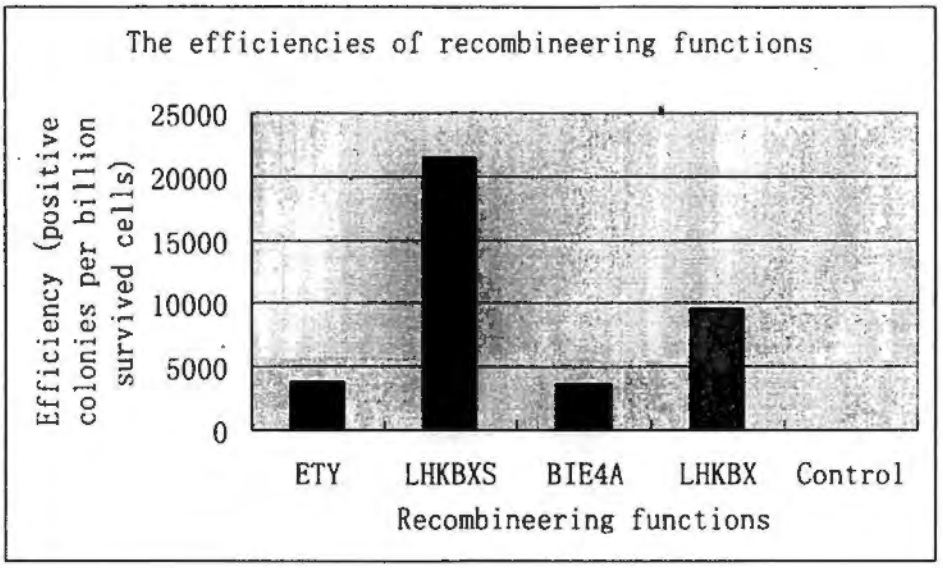


Figure 64. Efficiency of recombineering functions (A)

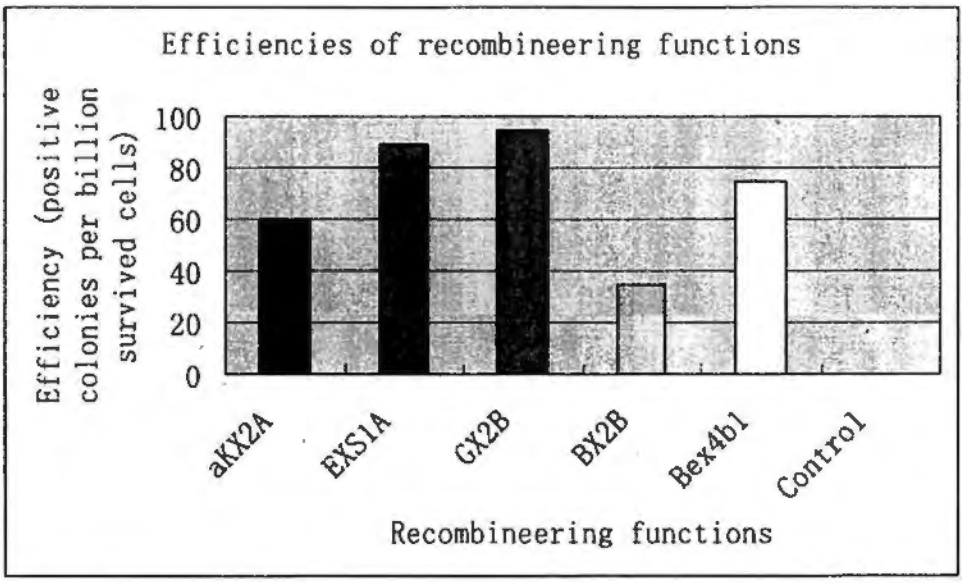


Figure 65. Efficiency of recombineering functions (B)

Table 2. Efficiency of recombineering functions

Recombineering system	Bet/RecT homologous	Exo/Rec Ehomologous	SSB/Gam	Cm^r per 10^9 survive cells
pBADETγ	RecT	RecE	Lambda Gam	3860
B1E4A	Lambda Bet	Lambda Exo	N/A	3700
LHKBX	LHK Bet	LHK Exo	N/A	9560
LHKBXS	LHK Bet	LHK Exo	LHK SSB	21500
Bex4b1	SXT Bet	SXT Exo	N/A	74.8
BX2B	SXT Bet	SXT Exo	SXT SSB	35
GX2B	SXT Bet	SXT Exo	Lambda Gam	94.2
EXS1A	SXT Bet	SXT Exo	N/A	89.7
αKX2A	SXT Bet	Lambda Exo	N/A	59.7

LHK Bet had the highest repair efficiency in the Bet proteins (Figure66). LHK Bet and RecT proteins in pBAD vector had higher efficiencies than in pDH vector, while bacteriophage lambda Bet and SXT Bet had lower efficiencies than in pDH vector. Heat-shock of LHK Bet, *lambda* Bet, *S.aureus* Bet and RecT in pBAD vector decreased their efficiencies.

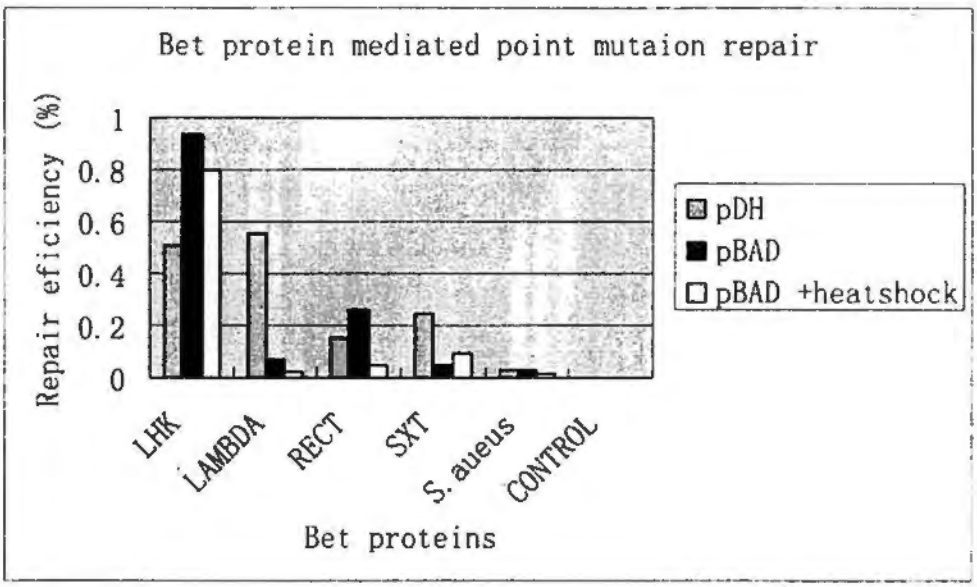


Figure 66. Point mutation repair efficiency mediated by Bet/RecT homologous

In the bacteriophage lambda Bet truncations, only N-terminal truncation residue 56-261 had 40% activity of full length protein left (Figure 67). RecT C-terminal

truncation residue 1-250 had even higher activity than full length RecT, while N-terminal truncation residue 50-269 had no activity (Figure 68). SXT Bet N-terminal truncation residue 54-272 had about 20% activity left while N-terminal truncation residue 35-272 had no activity (figure 69). The repairing efficiency of His-tagged LHK Bet expressed in pBAD was higher than expressed in pDH vector (Figure 70), while His-tagged LHK Bet protein expression in pBAD was lower than in pDH vector (Figure 71).

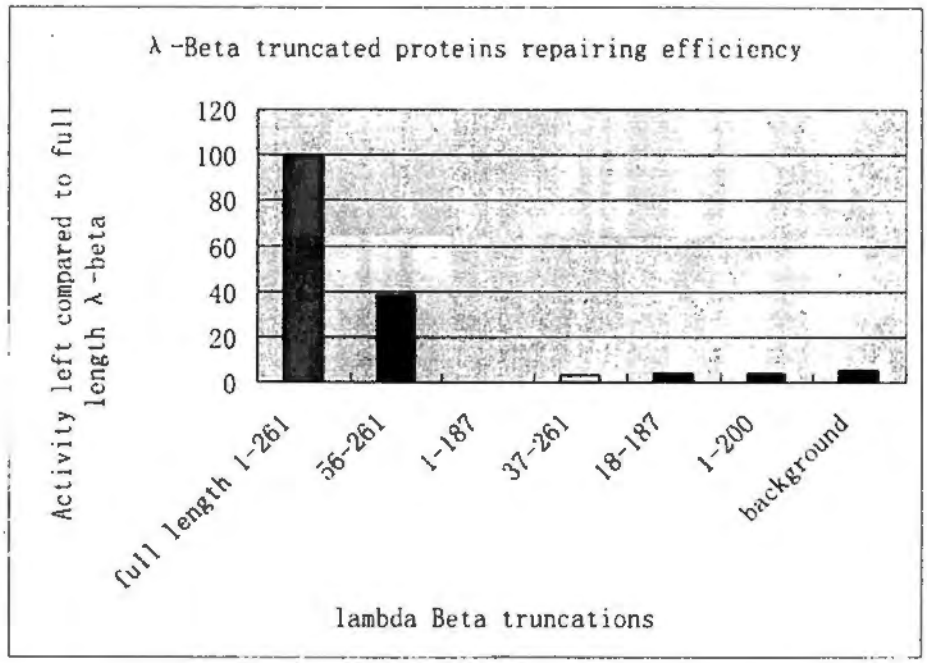


Figure 67. Point mutation repair efficiency mediated by lambda Bet truncations

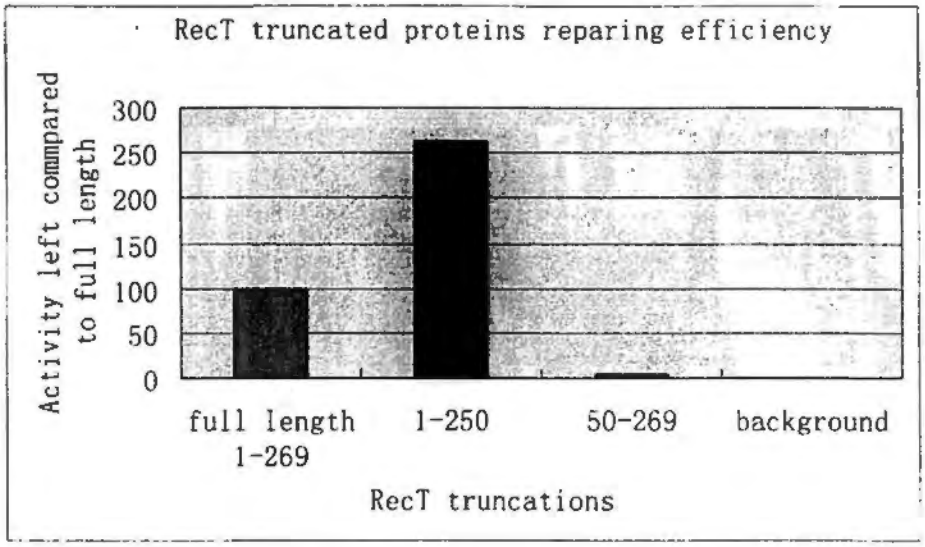


Figure 68. Point mutation repair efficiency mediated by RecT truncations

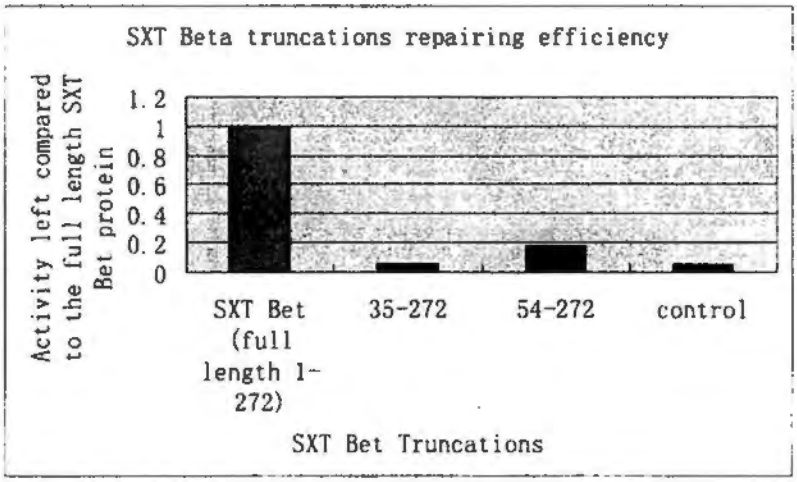


Figure 69. Point mutation repair efficiency mediated by SXT Bet truncations

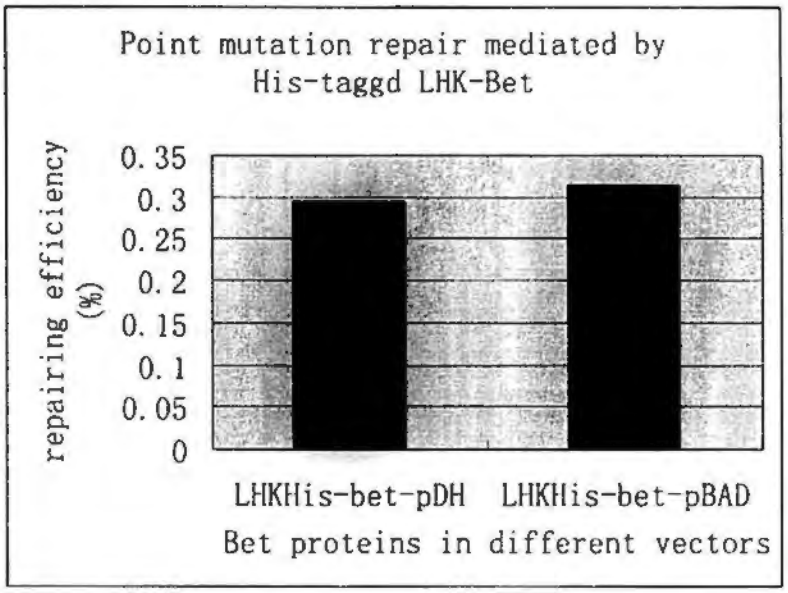


Figure 70. Point mutation repair efficiency mediated by His-tagged Bet

Lane 1, BenchMark protein ladder

Lane 2, His-tagged LHK Bet expressed in pDH vector

Lane 3, His-tagged LHK Bet expressed in pBAD vector

Figure 71. His-tagged LHK Bet expression in pDH and pBAD vectors

In the presence of 1.24ng (~4.12×10^{-16}molar) pLysS plasmid, 1.04×10^{-14} molar of the ssDNA could have 50% of the deletion efficiency mediated by lambda Bet protein (Figure 72). As in point mutation repair, LHK also had the highest deletion efficiency (Figure 73), while RecT and lambda Bet had lower efficiencies, SXT Bet and *S.aureus* Bet had lowest activity. Heat-shock also decreased the efficiencies of LHK Bet and RecT expressed in pBAD. Lambda bet almost had less activity when expressed in pBAD, while it has rather more activity when expressed in pDH. When the single point mutation repair efficiencies and DNA fragment deletion efficiencies

from the same proteins were compared and plotted (Figure 74), there is an exponent of 2 relationship.

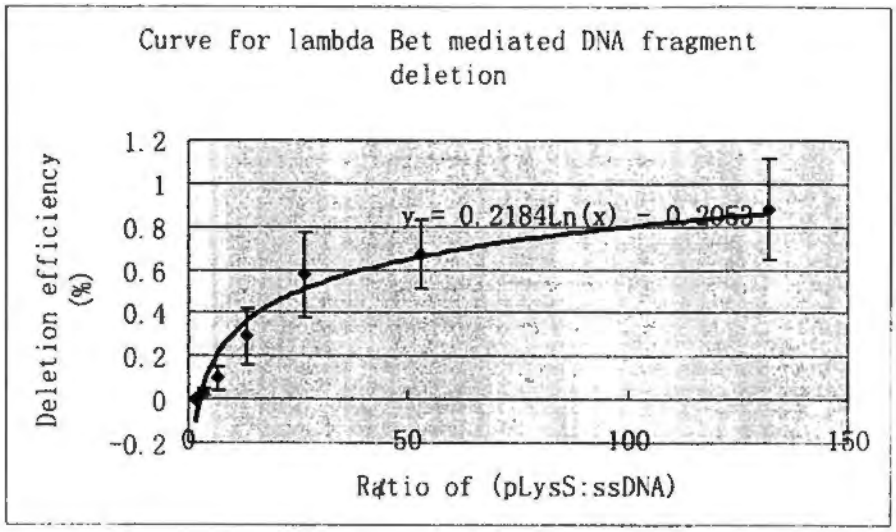


Figure 72. Curve for DNA fragment deletion mediated by lambda Bet

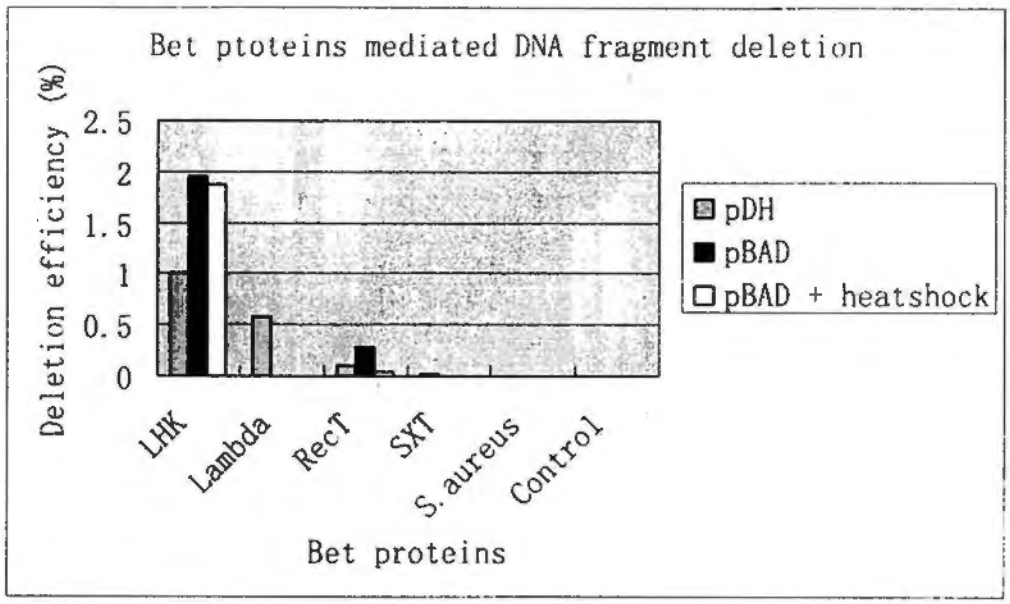


Figure 73. DNA fragment deletion mediated by Bet/RecT homologous

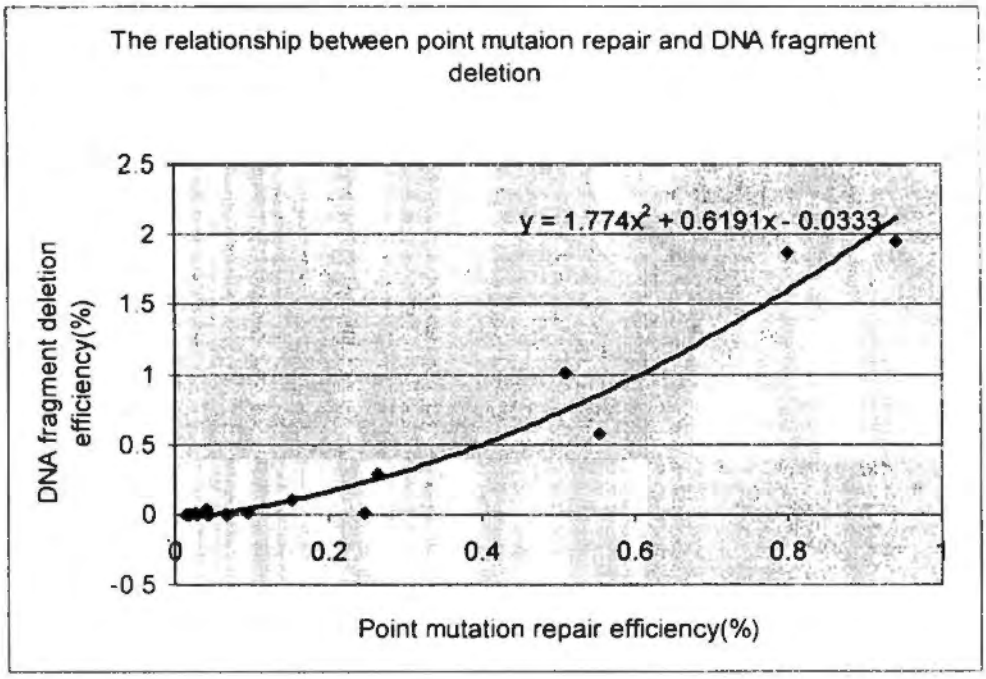


Figure 74. The relationship between efficiency of point mutation repair and efficiency of DNA fragment deletion

Chapter Four Discussion

Part One. The quantification of linear dsDNA

The Picogreen fluorescent assay is a very useful tool for the accurate quantification of dsDNA, which is relatively unaffected by the presence of contaminating molecules (with the notable exception of Manganese). When Picogreen binds to RNA or single stranded DNA, the intensity of the fluorescent emission at 525nm is very low (excitation 485nm). However, when Picogreen binds to double strand DNA, the fluorescence reading is enhanced greatly [108]. The intensity of the fluorescence reading (in relative units) is proportional to the amount of dsDNA present. When linearized double stranded DNA is degraded by the exonuclease, the fluorescence reading will decrease proportionally. If assumed that the ssDNA regions produced do not self-anneal to form a large amount of dsDNA, then it will not has significant effect on the fluorescence reading. It is also possible to use Picogreen assay to perform real time exonuclease assays [47, 108]. However, the binding of Picogreen molecules to the dsDNA may affect the enzyme activity to an unknown extent, as the enzyme will have to actively displace these interchelating molecules in order to move along while unwind the dsDNA helix and digest the dsDNA.

Part Two. Characterization of SXT Exo protein

4.2.1. SXT Exo expression and purification

The SXT-Exo protein was expressed with highest levels from the pET28a vector in *E. coli* BL21(DE3) (Figure 19). However, the vast majority of expressed protein was in the form of insoluble inclusion bodies. There are 6 cystine residues in each SXT Exo monmeric subunit. These may form several disulfide bonds that may make the correct protein folding process more difficult. Another possible reason is that the correct folding of the SXT Exo protein needs divalent metal ions, such as Mn^{2+} or Mg^{2+}, which may be at lower concentration in *E.coli*. Thirdly, the environment in *E. coli* may be very different from the native host, *Vibrio cholerae*, which hinders the protein correct folding. It is also possible that the N-terminal domain plays an important role in the protein folding process. The recombinant protein has N-terminal His-tag which may hinder the formation of the correct protein folding intermediates. A recombinant SXT Exo protein with an N-terminal GST-tag was also prepared, and its expression and activity were investigated. Results showed that the total and soluble protein expression levels were much lower than for the His-tagged protein (data not shown). Lower total protein expression levels may be explained by the presence of a Ptac promoter on the pGEX4T1 plasmid, compared with the stronger T7 promoter on the pET28 vector. However, it may also be (partially) due to the large GST-Tag (ca. 26.5 kDa) at the N-terminus interfering with the folding process.

Noticeably, in the alignment with bacteriophage lambda Exo and LHK Exo, SXT Exo has an extra C-terminal domain. This is interesting, as it is also different from the other uncharacterized exonuclease homologues in the NCBI databases (data not shown). The role of this region is unknown. It is possible that it influences the protein stability and overall expression levels. It may be noted that in the conserved

(N-terminal 200aa) region, SXT-Exo shares much lower levels of amino acid sequence homology with Lambda Exo than LHK Exo does.

4.2.2. Biophysical characterization of SXT Exo

The results from the gel filtration experiments (Figure 51) clearly show that the recombinant SXT Exo protein forms a stable multimer, which has a mass corresponding (on average) to 3.5 monomers (calculated from the migration of protein standards). The recombinant Lambda Exo protein (Figure 52) multimer is calculated and that it seems to have 3.35 monomers. Obviously, a multimer cannot have whole and fractional monomers, and so this is most likely results from a deviation from standard Stokes radius. It should be noted that lambda Exo is not a globular protein; it has a toroidal (ring-like) shape which may make it move faster than a spherical protein through the gel filtration matrix. This would make its apparent molecular mass appear higher than its real (multimeric) molecular mass. The recombinant lambda Exo has an apparent molecular mass of 91 kDa, which corresponds to 3.35 monomers. However, x-ray diffraction data reveals that it crystallizes as a trimer [38]. Since the gel filtration data indicates that SXT Exo contains an average of 3.5 monomers, by comparison with the finding for the lambda Exo protein, this suggests that SXT Exo is also a trimer in solution.

4.2.3. Biochemical characterization of SXT Exo

When the purified recombinant SXT Exo protein is stored in 25mM Tris-HCl pH7.4, 50mM NaCl and 1mM DTT buffer, it maintains high levels of activity for 2-3 days at 0-4°C. SXT Exo loses all its activity after one freeze/thaw cycle. By direct comparison to the lambda and LHK Exo proteins, SXT Exo is not very stable and the soluble expression levels are low. This makes performing reproducible sets of exonuclease assays very difficult.

SXT Exo processiveley degrades linear dsDNA from both ends and moves inward towards the centre. It does not digest linearized dsDNA internally, as shown in Figure 41, which means it has no endonuclease activity. In the presence of 1mM EDTA, the enzyme has no activity. SXT Exo activity is dependent on the presence of Mn^{2+} or Mg^{2+} metal ions as shown in Figure 33 and Figure 34. Although the dsDNA band disappears in the presence of SXT Exo and Fe^{2+} ion, it cannot be concluded that SXT Exo activity is also dependent on Fe^{2+} ions, because the dsDNA band also disappears in the absence of SXT Exo. The most probable reason for this is that the Fe^{2+} ion chelates with the DNA promoting its rapid degradation (Lewis acid catalysis), or that it strongly promotes the formation of hydroxyl radicals, which rapidly and non-specifically degrades the DNA molecules.

Similar to lambda Exo, SXT Exo is an alkaline exonuclease. However, SXT Exo prefers to use Mn^{2+} ions as a cofactor, rather than Mg^{2+} ions. The positive metal ion most likely facilitates the binding of phosphate oxygen atoms, making them more labile to (enzyme-catalyzed) hydrolysis. Its optimal temperature is 41°C, and SXT Exo activity decreases rapidly at temperatures higher than 41°C. One possible reason is the SXT Exo protein is not stable at higher temperatures. Another possible reason is that the dsDNA melts (strands separate) especially the ends. This may impede the enzymatic digestion (processivity), or the melted ends may hinder Exo binding. The optimal concentration of Mn^{2+} and Mg^{2+} ions are much higher than physiological concentrations. High concentrations of Mn^{2+} are toxic to the cell. However, the enzyme still has good activity at physiological concentrations of Mn^{2+} and Mg^{2+} ions, demonstrating that it should have good activity within the cell under native conditions.

Figure 39 shows that the initial activity of SXT Exo varies when linear dsDNA substrates with different types of overhangs are used as substrates. Blunt ended linear dsDNA, or dsDNA with 4nt 5'- or 4nt 3'- overhangs at each termini were tested, which were generated by digesting plasmid DNA with different restriction enzymes

(SspI, BamHI or PstI, respectively). It can be seen that SXT has a slight preference for linear dsDNA with a 5'-overhang, whilst lambda Exo prefers dsDNA with 3'-overhangs. Previous research [36] has shown that 5'-overhangs interfere with the binding of lambda Exo to dsDNA, but SXT Exo does not exhibit such effects. One possible reason is that the overhang is too short, being only 4 nucleotides, while the substrate is too long, 2686 base pairs. Consequently the differences in rate due to the different overhangs are quite small, and may not be significant. Figure 39 shows only a small portion (~10%) of dephosporylated linear dsDNA was degraded by adding limiting amounts of SXT Exo (compared with phosphorylated control). However, figure 40 shows that a similar amount of dephosporylated linear dsDNA was totally degraded by providing additional SXT Exo. This indicates that it is difficult for SXT Exo to bind dephosphorylated linear dsDNA and digest it. Similar to the situation observed for lambda Exo, the 5'-phosphate appears to be important for the binding of SXT Exo and for the initiation of exonuclease activity [36]. The arginine residue at position 28 (R28) is involved in 5'-phosphate recognition in lambda Exo. In the alignment of lambda Exo vs SXT Exo (Figure 7), R17 in SXT Exo corresponds to R28 in lambda Exo. Future research may prove that R17 in SXT Exo is involved in the 5'-phosphate recognition.

Experiment data have revealed that SXT Exo has high levels of activity at low NaCl concentrations, even as low as 1mM. This suggests that the protein can assemble into a stable trimer at low salt concentrations. High concentrations of salt may disrupt hydrogen bond and electrostatic interactions. This may be especially important because the interactions between the surface residues at the subunit interfaces are disrupted, the protein multimer may be destabilize and enzymatic activity may be inhibited. Salt ions can also interact with the residues in the active sites and at metal binding sites. This may interfere with the enzyme-substrate interaction and the metal ion cofactor binding, thus inhibit enzyme activity. Noticeably, KCl salt concentration has more potent inhibitory effects on SXT Exo (Figure 43) than NaCl. SXT Exo has much less activity in buffer containing 100mM KCl than in buffer containing

100mM NaCl, whilst it has slightly higher activity at low KCl concentrations (25-50mM). One possible reason is that the K^+ ion has a stronger interaction with residues in the interfaces between the SXT Exo subunits, or with residues in the active sites and metal binding sites. This will compete for binding with substrate or metal ion, and inhibits the exonuclease activity. $CaCl_2$ has much more obvious inhibitory effects on SXT Exo than NaCl or KCl. Beside disrupting the hydrogen bonds and electrostatic interaction, the Ca^{2+} ion may compete for the metal binding site with Mn^{2+} much strongly than monovalent ions, such as Na^+ and K^+. Ca^{2+} ions can bind to carboxylate groups i.e. Aspartate and Glutamate residues, possibly at the active site, forming stable (inactive) chemical complexes. It may also destabilize the protein, and facilitate its precipitation at high concentrations.

Since the recognition of the 5'-phosphate group on linear dsDNA is important for SXT Exo activity; the high concentrations of free phosphate ions (PO_4^{3-}) in phosphate buffer may compete for the phosphate recognition site in SXT Exo and hence inhibit SXT Exo activity (Figure 44). Consistent with this hypothesis, 100mM sodium phosphate buffer (pH 7.4) totally inhibits SXT activity, although SXT Exo still has high activity in 100mM Tris-HCl buffer (pH 7.4). In the potassium phosphate buffer, SXT Exo has 50% activity at low concentrations of phosphate (e.g. 25mM). Figure 45 shows that high concentrations of the sulfate ion (SO_4^{2-}) also completely inhibits SXT activity, similar to that observed for the phosphate ion. However, SXT Exo retains significant activity at low concentrations of potassium sulfate (e.g. 25 mM), which are similar to those observed in in 25mM potassium phosphate buffer. One possible reason for this is that the chemical structure of sulfate is similar to that of phosphate, so that it can mimic the phosphate group and compete for the 5'-phosphate group binding site in SXT Exo. This would inhibit the binding of dsDNA ends and consequently reduce exonuclease activities. However, it seems likely that the protein does not have such high affinity for either sulfate or phosphate, as since SXT Exo still has (reduced) activity at low concentrations of these ions. The hetero-polysaccharide molecule heparin is DNA mimic and competes with DNA for

Exo binding, thus inhibits the exonuclease activity. In an inhibition assay, 50% dsDNA exonuclease activity of 60nM of SXT Exo trimers was inhibited when they were incubated with 0.12nM ends of dsDNA substrate and ~2777nM of heparin.

It was therefore planned to investigate whether SXT Exo had a distributive or processive mode of exonuclease activity. A processive digestion mechanism means that once the SXT Exo protein binds the DNA substrate and initiates digestion, it will remain bound until the digestion of all the phosphodiester bonds is completed, or until the enzyme dissociates prematurely. A distributive enzyme disassociates from the dsDNA substrate after every catalytic event (i.e. after the cleavage of one phosphodiester bond). If SXT Exo digestion is distributive, then the higher amount of enzyme present, then the higher the DNA digestion rate. However, if the Exo exonuclease activity is processive, then above a certain threshold, the addition of more enzyme will not further increase the digestion rate. Figure 46 shows that 12pmol SXT Exo in the reaction mixture is enough to saturate the ends of linear dsDNA substrate, and the addition of more enzyme could not further increase the digestion activity. The slopes of all the curves were calculated and the digestion rate of SXT Exo was 23.8±4.3 nucleotides per minute. According to Arrhenius law, $k=Ae^{-Ea/RT}$, the rate of digestion will increase logarithmically when the temperature is increased. However, complex enzymatic reactions are not as simple as common chemical reactions, and generally do not follow such relationships. dsDNA digestion is not a single step process, which involves many discrete processes, e.g. the 'melting' of the DNA helix prior to hydrolysis, each with individual kinetic and thermodynamic parameters. It should also be mentioned that Mn^{2+} ions promote the formation of hydroxide ions and peroxide radicals in aqueous solution, both of which can degrade DNA in a non enzyme-dependent manner.

Part Three. Characterization of LHK Exo protein

4.3.1. Biophysical characterization of LHK Exo

The migration of the recombinant LHK Exo protein on the gel filtration column (Figure 53) is consistent with it forming a stable trimmer in solution. Preliminary results of x-ray crystallization studies have shown that C-terminal His-tagged LHK Exo forms a toroidal trimer analogous to lambda Exo (personal communication from Dr. Wen Yang, Nankai University, China). However, chemical cross-linking showed that LHK Exo could be dimmer. Taken together, this suggests that the LHK Exo protein is most likely a trimer, which has anomalous migration through the gel filtration matrix due to its donut shape. It is also possible that in the chemical cross-linking experiments, only two of the 3 subunits are linked together, to be analyzed by SDS-PAGE.

4.3.2. Biochemical characterization of LHK Exo

The protein maintains the activity for 2-3 weeks when stored at 4°C in buffer containing 25mM Tris-HCl, pH7.4, 50mM NaCl. Even after one freeze-thaw cycle, the protein still has 43% (It is supposed that the protein forms trimer) activity. The protein is therefore very stable, and the expression levels are high, which facilitates the investigation of its activities and biophysical properties.

Similar to lambda Exo, LHK Exo is an alkaline exonuclease. It has very low activity at acidic pH, whilst it has optimum activities under basic conditions (pH 8.25). The optimal Mg^{2+} concentration of 7.5mM is within in the physiological concentration range. Figure 27 shows that LHK Exo digestion is processive. The addition of more than 0.272pmol LHK Exo subunits (which saturates the dsDNA ends) does not further increase the digestion activity. Calculated from the slopes of the lines in Figure 28, the rate is 7.29±1.44 nucleotides per second at room temperature (22°C).

Figure 30 shows that LHK Exo prefers linear dsDNA substrates that have 3'-overhangs. However, (as mentioned before) there are only 4 nucleotides of overhang, while the full length of the linear dsDNA molecule is 2686 base pairs, so the effect of this substrate preference may be not obvious. Dephosphorylated dsDNA is not favored substrate for LHK Exo; the Figure 30 shows that only about 10% of the all dsDNA is digested. Further study shows that LHK Exo is able to digest all the dephosphorylated dsDNA if a large excess of enzyme is supplied as SXT Exo (data now shown). From the amino acid sequence alignments, it may be seen that the arginine at position 14 (R14) in LHK Exo corresponds to R27 from SXT Exo, and R28 in Lambda Exo. This may be the residue that plays a key role in recognizing the 5'-phosphate group at the termini of linear dsDNA.

LHK Exo cannot internally digest linear dsDNA (Figure 31), i.e. it is not an endonuclease. The sizes of DNA digested (intermediate products) decrease gradually, which suggests that LHK Exo degrades the DNA from the ends and moves inward.

Part Four. Biophysical characterization of the ssDNA annealing and binding proteins

As shown in Figure 48, the multimerities of LHK Bet protein is very complicated and it be of high molecular weights. The gel filtration in Figure 54 shows that the molecular mass of the recombinant LHK Bet in 25mM phosphate buffer (pH7.5, 150 mM NaCl) is ca. 399 kDa, which corresponds to a multimerity of 12.2 monomers. The multimericity of the recombinant Bacteriophage lambda Bet is 12 and in the presence of Mg^{2+} ion, and 12 subunit of lambda Beta form a ring [54]. So it could be supposed that LHK Bet also form a ring in the buffer, even in the absence of Mg^{2+} ion.

The results from chemical cross-linking experiments shown in Figure 49 indicate that the SXT Bet protein can form a stable dimer in solution. However, there are additional products of high molecular weight observable on the gel, which suggests that the SXT Bet protein can adopt complicated multimeric forms. Figure 55 shows the gel filtration profile of SXT Bet in 25mM Tris-HCl buffer (pH7.4, 150mM NaCl). It also suggests that SXT Bet forms very complicated mixture of various multimeric forms. The highest peak (major component) is at 9.74 ml, corresponding to a molecular weight of 238.5 kDa, which suggests a multimericity of 7.18. Figure 56 shows the gel filtration profile of SXT Bet in 25mM Tris-HCl buffer (pH7.4, 1M NaCl). Compared to the gel filtration chromatogram obtained in low salt conditions, the distribution of SXT Bet multimers in high salt conditions (1M) is much simpler. The highest peak (major component) elutes at 10.86ml in figure 56, corresponding to a molecular weight of 143.2 kDa, which suggests that the multimericity is about 4.3. At high salt concentrations, the electrostatic interactions between protein molecules are much weak, so the protein monomers are less likely to aggregate, and hence lower molecular weight species and a reduced number of multimers are observed in gel filtration.

Figure 57 shows that SXT Bet forms relative simpler multimerity in sodium phosphate buffer, compared to those observed in Tris-HCl buffer (with same pH and NaCl concentration). This suggests that phosphate ions may affect the multimericity of the SXT-Bet. It also shows that SXT Bet has the same distribution of multimericity at pH 7.4 and 8.0 (Figure 58), indicating that (over this pH range) pH does not greatly influence its multimerity. Figure 59, 60 and 61 shows that Mg^{2+} ions also affect the multimerity. The interaction between the Mg^{2+} ions and the protein occurs rather slowly (at room temperature) and requires many hours to go to completion. If Mg^{2+} ions are incubated with SXT Bet for a short time (e.g. 15minutes), only a small portion of SXT Bet interacts with Mg^{2+} ions, forming larger molecules. However, if Mg^{2+} is incubated with SXT Bet for a long time (e.g. overnight), most of the SXT Bet interacts with Mg^{2+} ion and form larger multimers.

The SXT SSB protein was chemically cross-linked to examine its multimeric status. On the SDS-PAGE gel of the cross-linked protein mixture (Figure 50), SXT SSB forms dimers and tetramers. The closely-related SSB protein from *Escherichia coli* is a homotetramer [109]. The likely reason for this observation is that the SXT SSB tetramer has a dimer-dimer structure, similar to that found for the *Escherichia coli* SSB protein [109]. In the cross-linking experiments, the four component monomers of the tetramer are not always coupled to one another, sometimes just dimer pairs, leading to the observation of both multimers on the SDS-PAGE gel.

Figures 62 and Figure 63 show the interaction between LHK Exo and LHK Bet protein *in vitro* and *in vivo*, respectively. In Figure 62, the GST-tagged LHK Exo protein is used as the 'bait' to capture the His-tagged LHK-Bet protein (the 'prey'). Lane 4 indicates that no LHK Bet protein sticking to the glutathione resin in the absence of LHK Exo protein was detected in SDS-PAGE. Lane 7 shows that no LHK Bet protein sticking to the GST protein bound to the glutathione resin was detected in SDS-PAGE, which demonstrats that there is a specific binding interaction between the LHK Bet and Exo proteins. In Figure 63, the LHK Bet and Exo proteins were

cloned into a pET32 expression vectors in a 'one-operon' arrangement, analogous to that found in the bacterium (i.e. mimicking the native transcription and translation arrangement). In this vector construct, only the LHK Exo protein has His Tag, enabling it to bind to the Ni-impregnated column. LHK Bet cannot bind to the resin as it does not have His-tag. The only way it can be retained on the column is that it has a binding interaction with the LHK Exo protein.

The roles of SXT Ssb and LHK Ssb proteins remain to be established. Numerous phage and prophage regions that have been sequenced to date have Bet (RecT), Exo (RecE) and Ssb proteins encoded adjacent to one another, or in very close proximity [56]. This strongly suggests that the SSB protein plays some role in the Bet-Exo recombination process, however, this has not been investigated (there are no reports in the literature to date). The Bet and SSB proteins may potentially affect the activities of the Exo protein via several different mechanisms: they may interact with the Exo protein directly, or they may bind to regions of ssDNA, therefore stabilizing the Exo degradation product, or they may do both. Because the expression level of the SXT Exo protein was very low in *E.coli*, it was very difficult to get a clear result from the pull down assay, to determine whether there was an SXT Exo-SXT Bet interaction.

Part Five Evaluation of the biological activities of the ssDNA annealing Proteins

4.5.1. Point mutation repair by Bet/RecT homologous

Figure 66 shows a summary of the efficiency by which the (plasmid encoded) LHK Bet; Lambda Bet; RecT; SXT Bet and *S. aureus* RecT proteins could correct a 'loss of function' point mutation within a kanamycin resistance gene on the *E. coli* chromosome using a 71nt correction oligonucleotide. This assay measures the efficiency by which Bet/RecT recombinase can promote the oligonucleotide-encoded substitution of a single nucleotide on the chromosome. The proteins were expressed from arabinose-inducible plasmids (pBAD, red color) or from heat shock-inducible cI857/PL plasmids (pDH, blue color). Both plasmids have a similar copy number. A heat shock control was performed for each of the pBAD experiments (yellow color), to detect any differences that may arise from elevating the cell incubation temperature to 42°C for 15 minutes.

The repair efficiency of LHK Bet is just under 1% when expressed from either pDH or pBAD plasmids. The high efficiencies enable recombinant clones to be directly screened without the use of a selectable marker (e.g. by PCR screening). The recombination efficiencies of LHK Bet and RecT in the pDH vectors are lower than those in the pBAD vectors. However, the efficiencies of lambda Bet and SXT Bet in pDH are higher than those in pBAD vectors. The recombination efficiency of *S. aureus* RecT is approximately the same in pDH and pBAD vector. The point mutation repair efficiency of His-tagged LHK Bet and His-tagged LHK Bet does not appear to be directly related to protein expression level (as long as it is above some threshold level) Figure 70, 71). The 15 minute 42°C heat-shock lowers the recombination efficiencies of LHK Bet, *S. aureus* RecT, lambda Beta and RecT in the pBAD vector, whilst it enhances the efficiency of SXT Bet. One possible reason

for this effect is that the heat-shock also induces the expression of the Gam protein, since the *gam* gene is still integrated in the host chromosome under the control of the cI857/P_L in the DY380Δ*bet mKan* reporter strain. Gam not only inhibits nucleases in the cell, but may also interfete with the binding of Bet protein to DNA, because the Gam protein mimics DNA [26]. Also, heat-shock may induce considerable changes in the protein profile of the cell, when compared with the non-heat shocked cells. Consequently, the overall effects of the heatshock are very complicated, and hard to predict or define.

4.5.2. DNA fragment deletion mediated by Bet/RecT homologous

It was planned to investigate the efficiency forBet/RecT proteins mediate oligonucleotide-directed deletion of a large section of DNA on a low-copy plasmid, pLysS. Plasmid pLysS (commercially available from Novagen), was created by inserting T7 lysozyme gene into the BamHI site of the tetracycline resistance gene in plasmid pACYC184, thereby inactivating it (creating an insertional mutant). The correction oligonucleotide straddles the original BamHI insertion point, and thereby directs the 'looping out' of the inserted T7 lysozyme gene, to create the original plasmid pACYC184, which has an active tetracycline resistance gene. A co-electroporation strategy was used, as this meant that both oligo and plasmid would have a high chance of being electroporated into the same cell at the same time. 'Recombination-competent cells' were used i.e. cells that contained the episomally-established recombinase gene on a pBAD or pDH plasmid, which has been induced, immediately prior to making competent cells. The use of this strategy would also mean that the ratio of oligo to plasmid within the same cell would probably be higher than if the pLysS plasmid was first established as a stable episome, and was then electroporated with oligo.

Figure 72 shows the saturation curve for increasing concentrations of correction oligonucleotide with a constant amount (1.24 ng) of pLysS (the correction target).

The plasmid and oligonucleotide were incubated with the 'recombination-competent' *E. coli* DH10B/pDH-lambda-Bet competent cells for 15 minutes prior to electroporation, to promote cellular association. The data shows that the correction efficiency increases with increasing amount of correction oligonucleotide, up to a certain point (ca. 100:1 ratio of oligo: plasmid). The maximum correction efficiency obtained was ca. 1%, which is very good compared with previous literature results. This indicates that above a certain threshold concentration, the amount of oligo does not limit the repair event within the 'recombination-competent' cell.

Figure 73 shows the olionucleotide-directed deletion efficiencies mediated by the various Bet/RecT proteins. The general trend of the results was similar to that obtained from the point mutation repair experiments (described above). However, the differences in the repair efficiencies between the various recombinase proteins are more pronounced than those obtained from the point mutation repair experiments. In order to delete the T7 lysozyme gene from plasmid pLysS, the same correction oligonucleotide must anneal to a region of single stranded DNA at both the 5'- and 3'-flanking regions. However, in the point mutation repair, the oligonucleotide anneals with its complementary ssDNA sequence in one single annealing event, with a centrally-positioned single base-base mismatch. It seems intuitive that the efficiency (or probability) of the double annealing event occurring should be substantially lower than the single annealing event. Consequently, there may be a more pronounced difference, or drop in efficiencies obtained, going from the 'best' recombinase to the least effective one. Figure 74 shows the relationship between the deletion efficiency of lambda-Bet versus the square of the efficiency obtained for the Lambda Bet-mediated correction of the point mutation.

Previous biophysical studies[50, 55] combined with regions of homology shared between the various Bet and RecT recombinases[56] suggests that the N-terminal ca. 200-220 amino acids contain all the amino acids essentially required for activity. Consequently, a set of deletion truncations for several recombinases was created, and

was used to investigate their activities. The positions of the truncations were guided by secondary structure prediction software. The disordered loops were placed between predicted secondary structural unites (i.e. alpha helices or beta sheets). Figure 67, 68 and 69 summarize the point mutation repairing efficiencies of various truncated lambda Bet, RecT and SXT Bet proteins, respectively. The lambda Bet N-terminal truncation comprising residues 56-261 has 20% the activity of the full length protein, while all the C-terminal truncations had no activity. However, Wu et al. proposed that this region of the N-terminal domain is part the stable core for DNA binding, while the C-terminal domain is disordered and does not play an obvious role in ssDNA binding [55]. The work in this thesis shows that the first three helixes in the predicted structure (Figure 75) of bacteriophage lambda Bet are not essential for activity, but the highly variable C-terminal 60aa domain is essential. This is not consistent with the model proposed by Wu et al [55].

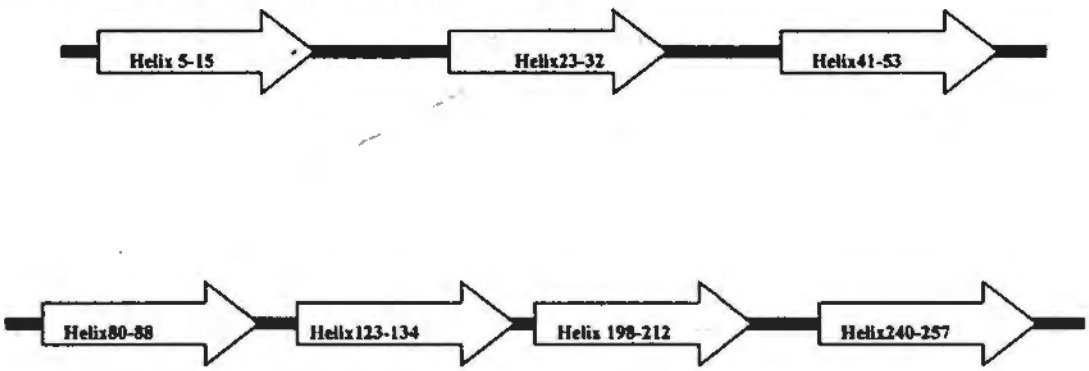


Figure 75. The predicted secondary structure of lambda Bet via www.predictprotein.org

Interestingly, it was found that the lambda Bet protein truncation that lacked the first three N-terminal alpha helixes had a higher activity than the truncation mutant that lacked only the first two helices. One possible reason for this finding is that the third helix interferes with the folding or stability of the residues 37-261 truncation. Another possible reason is that the 'extra' residues contained within helix 3 (residues 38-53) interfere with the ssDNA-protein interaction, or some other key process.

Figure 69 show and Figure 76 show that the first three helixes of SXT Bet protein are not essential for its point mutation correction activity, similar to the situation found for the lambda Bet protein. The alignment of SXT Bet with lambda Bet (figure 77) shows that they share a highly similar N-terminal domain (including the first three helixes in the predicted secondary structure; as shown in Figures 75 and 76). However, the C-terminal domain is essentially required for activity, even though it shares very low levels of amino acid similarity.

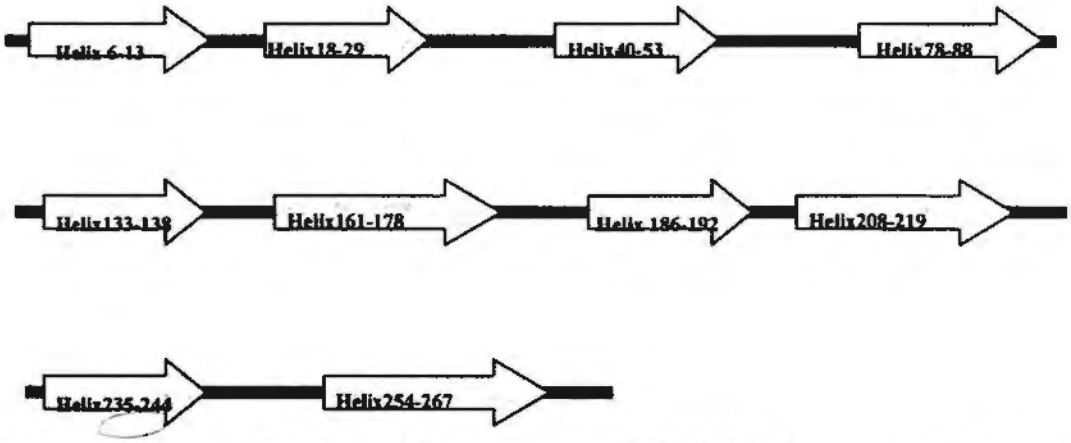


Figure 76. The predicted secondary structure of SXT Bet via www.predictprotein.org

Figure 68 and Figure 78 show that the N-terminal truncation protein of RecT (residues 50-269) has no activity, whilst a protein that contains a small C-terminal truncation (residues 1-250) has 2-fold higher activity than the full length protein. This result was highly unexpected. Taken together, the results suggest that the N-terminal domain (residues 1-49) is essential for RecT activity. The C-terminal-most alpha helix (residues 255-265 in Figure 78) within the variable C-terminal domain is not essential for RecT activity. Furthermore, it appears to interfere with the DNA recombination process. One possible reason for its existence is that this portion of the C-terminal domain interacts with its partnering exonuclease, RecE.

Part Five. Evaluation of the dsDNA recombination activities of the recombinase-exonuclease proteins

The relative dsDNA recombination activities of (plasmid-based) LHK, SXT, RecET and Lambda systems were directly compared in *E. coli* DH10B. LHK Bet-Exo-SSB (LHKBXS) was found to have the highest activity; about 5-fold higher than that of RecETγ (pBADETY) and lambda Bet-Exo (pB1E4A, see Figure 64). LHK Bet-Exo (LHKBX) had dsDNA recombination activity that was about 2.5 fold higher than that of RecETγ fand lambda Bet-Exo. This is extremely promising, as these plasmids are sold commercially by the GeneBridges company, and are the most efficient recombinase-exonuclease pairs characterized to date.

Figure 77. The alighment of SXT Bet and lambdaBet

As shown in Figure 65, it may be seen that all the various permutations and arrangements of the SXT bet, exo and ssb genes exhibited lower activities than either lambda Bet-Exo or RecETγ.

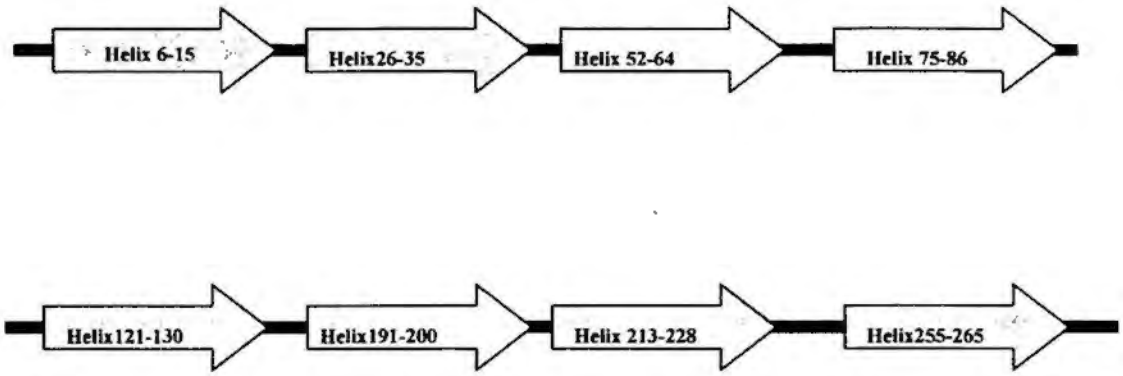


Figure 78. The predicted secondary structure of RecT via www.predictprotein.org

In pBex4b1 the SXT *bet-exo* genes are expressed from pBAD in a one operon arrangement that mimics the native system (minus the ssb gene). In the native arrangement, there is a gap of ca. 250 bp between the terminus of the *bet* gene and the initiation codon of the *exo* gene. The efficiency by which the downstream *exo* gene is transcribed remains to be established. However, when the *exo* gene is placed under the control of the strong (synthetic) constitutive P_{EM7} promoter (as in pEXS1A) there is still low dsDNA recombination efficiency. The efficiency is not enhanced by the inclusion of the gam gene (as in pGX2B), indicating that host nucleases are not greatly affecting the recombination process. Furthermore, the addition of the *ssb* gene in the native *ssb-bet-exo* gene operon arrangement (in plasmid pBX2B) does not lead to a significant increase in recombination levels. The most likely reason for these results is that the SXT Exo protein is not very stable, or is not expressed very well in a soluble form within *E. coli* cells. When the SXT exo gene (in plasmid pEXS1A) is replaced with the lambda *exo* gene (to form plasmid paKX2A) there are still reasonable, albeit low, levels of activity. This suggests that there is relatively poor functional cooperation between the SXT Bet and Lmabda Exo proteins. However, this is still a very significant result, as previously, it has been reported that the Lambda Bet/Exo and RecE/RecT pairs of proteins are not interchangeable [12].

The results indicate that there may be inter-species 'cross-partnering' of Bet and Exo recombination functions. This is an important finding, as many Bet or RecT recombinases annotated in the DNA sequence databases lack an identifiable Exo/RecE exonuclease partner. The results here may help future efforts to identify regions within the Bet/RecT and Exo/RecE proteins that play critical roles in the specific 'partnering' binding interactions

It is extremely difficult to compare actually efficiencies between studies (for a multitude of technical and genetic reasons); it much better to compare the general trends observed. Furthermore, it is extremely interesting to observe that LHK Bet-Exo-SSB has 2-fold higher activity than LHK Bet-Exo. As mentioned above, there are several plausible reasons to explain this observation. Firstly, the presence of the LHK *ssb* gene in the three gene operon may promote the production and correct folding of the (co-transcribed) downstream Bet and Exo proteins. Secondly, LHK Ssb protein may protect the 3'-ssDNA tails produced by the Exo protein. Thirdly, it may physically interact with LHK Exo and/or LHK Bet, to promote the recombination process (e.g. by actively helping the Bet protein anneal to the 3'-ssDNA overhangs produced by the Exo protein).

Perhaps most encouragingly, the results suggest that the LHK recombination proteins have the potential to be used alongside or even supersede the Lambda-Red and RecET systems, which are currently the 'gold standard' recombineering proteins is *E. coli*. As such, they have great potential to be used in the future development of molecular recombineering 'tools' and procedures to be used in this and other bacterial organisms.

Chapter Five Summary

The ability to manipulate or 'engineer' DNA molecules in a precise and efficient manner is of fundamental importance in modern biological research. Recombineering, or DNA engineering using homologous recombination, has recently emerged as an extremely powerful *in vivo* method for the modification of both chromosomal and plasmid DNA.

To date, the most efficient bacterial recombineering systems have been established within *Escherichia coli*, and utilize the Lambda-Exo/Bet or RecE/RecT sets of exonuclease and recombinase proteins. Both sets of proteins can utilize PCR-generated DNA molecules to direct the accurate creation of deletions, insertions and substitutions. The Exo or RecE exonucleases digest linear double stranded DNA (dsDNA) molecules generating 3'-single stranded DNA (ssDNA) tails. Bet and RecT ssDNA annealing proteins (recombinases) bind these ssDNA tails, and promote their recombination with homologous regions on the chromosome or episome. Bet and RecT can utilize oligonucleotides to create analogous genetic alterations with greatly increased frequencies, without requiring the activities of their respective exonuclease partner. However, these two sets of proteins only function in *E. coli* or closely-related bacteria (e.g. *Salmonella*), which limit their widespread application.

As described in this thesis, several novel recombinase and exonuclease proteins were cloned expressed and characterized. These include: 1) *bet* (*s065*) and *exo* (*s066*) homologues from the SXT genetic element that infects *Vibrio cholerae*; 2) *bet* and *exo* homologues from *Laribacter hongkongensis*; 3) a *recT* homologue from a clinical isolate of *Staphlylococcus aureus* USA300. A combination of *in vivo* and *in vitro* techniques was used to compare the biophysical properties and biological activities of these new recombinase and exonuclease proteins with those of Lambda-Exo and Bet, and RecE and RecT. New pairs of recombinase and

exonuclease proteins compose new homologous recombination functions, and the function developed from LHK is much more efficient than Red and RecET functions.

The ssDNA recombination activities of five diverse Bet/RecT homologues in *E. coli* were investigated, to explore their functional diversity. Most notably, it was found that a previously uncharacterized Bet homologue from *L. hongkongensis* (LHK-Bet) has quite exceptional activities, and could create point mutations on the *E. coli* chromosome with efficiencies approaching 1%. This enables recombinant clones to be directly screened without the use of a selectable marker.

Guided by protein sequence homology and secondary structure prediction tools, a set of rationally-truncated recombinase proteins was created, to identify which conserved regions were essentially-required for recombination activity. The results indicate that although some conserved regions of the recombinases are dispensable. Overall there is a complex relationship between protein composition, multimericity and activity.

The results described in this thesis will facilitate the future development of molecular recombineering 'tools' and procedures to be used in bacterial organisms throughout the kingdom. These technologies will be invaluable for a multitude of commercial and clinical applications, in areas such as microbial biotechnology, functional genomics, synthetic biology and medical microbiology.

Chapter Six Bibliography

1. Cohen, S.N., Chang, A.C., Boyer, H.W., and Helling, R.B. (1973) Constructionof biologically functional bacterial plasmid *in vitro*. *Proc. Natl. Acad. Sci. USA* 70(11):3240-3244

2. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1998) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press. NY, Vol 1,2,3

3. Smith, H.O., and Wilcox, K.W. (1970) A restriction enzyme from Hemophilus influenzae.I. Purification and general properties. *J.Mol.Biol.* 51:379-391

4. Kuzminov, A. (1999) Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage and bacteriophage lambda. *Microbiol Mol Biol Rev.* 63(4) 751-813

5. Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., lauder, S.D., and Rehrauder, W.M. (1994) Biochemistry of Homologous recombination in *Escherichia coli .Microbiol. Rev.*58(3) 401-65

6. Nelson, D.L., and Cox, M.(2005) *Lehninger Principles of biochemistry*, fourth edition. Chaper 9, 306-313. Freeman, New York.

7. Shen, P., and Huang, H. V. (1986). Homologous recombination in *Escherichia coli*: Dependence on substrate length and homology. *Genetics* 112, 441–457

8. Parkinson, J.S., and Huskey, R.J. (1971) Deletion mutants of bacteriophage lambda I. Isolation and initial characterization. *J.Mol.Biol. 56*:369-384

9. Taylor, A., and Smith, G.R., (1980). Unwinding and rewinding of DNA by the RecBC enzyme. *Cell* 22:447–57

10. Kolodner, R., Hall, S.D., and Luisi-DeLuca, C.(1994) Homologous pairing proteins encoded by the *Escherichia coli* recE and recT genes. *Mol. Microbial.* 11(1):23-30

11. Chang, H.W., and Julin, D.A. (2001) Structure and function of the *Escherichia coli* RecE protein, a member of the RecB nuclease domain family. *J.Biol.Chem.* 276(49):46004-10

12. Muyrers, J.P., Zhang, Y., Buchholz, F., and Stewart, A.F. (2000) RecE/RecT and Redalpha/Redbeta initiate double stranded break repair by specifically interactive partners. *Genes. Dev.* 14(15): 1971-82

13. Murphy, K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli. J.Bacteriol.* 180(8):2063-71

14. Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 97(12): 6640-5

15. Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., and Copeland, N.G. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli.* (2000) *Proc. Natl. Acad. Sci. USA* 97(11):5978-83

16. Zhang, Y., Buchholz, F., Muyrers, J.P.P., and Stewart, A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli. Nat.Genet.*20:123-128

17. Lee, E. C., Yu, D., Martinez deVelasco, J., Tessarollo, L., Swing, D. A., Court, D. L., Jenkins, N. A., and Copeland, N. G. (2001) A highly efficient *Escherichia coli*–based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73:56–65

18. Court, D.L., Sawitzke, J.A., and Thomason LC (2002) Genetic engineering using homologous recombination. *Annu Rev Genet.*36:361-88.

19. Zhang, Y., Muyrers J.P, Testa, G., and Stewart, A.F. (2000) DNA cloning by homologous recombination in *Escherichia coli. Nat Biotechnol.* 18(12):1314-7

20. Ellis, H. M., Yu, D., DiTizio, T. and Court, D. L. (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. USA* 98:, 6742-6746.

21. Costantino, N., and Court, D.L. (2003) Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc Natl Acad Sci U S A.* 100(26):15748-53

22. Li, X.T., Costantino, N., Lu, L.Y., Liu, D.P., Watt, R.M., Cheah, K.S., Court, D.L., and Huang, J.D. (2003) Identification of factors influencing strand bias in

oligonucleotide-mediated recombination in *Escherichia coli*. *Nucleic Acids Res.* 31(22):6674-87.

23. Kelly, L. E., Davy, B. E., Berbari, N. F., Robinson, M. L., and El-Hodiri, H. M. (2005) Recombineered *Xenopus tropicalis* BAC expresses a GFP reporter under the control of Arx transcriptional regulatory elements in transgenic *Xenopus laevis* embryos. *Genesis* 41:185-191

24. Sarov, M., Schneider, S., Pozniakovski, A., Roguev, A., Ernst, S., Zhang, Y., Hyman, A., and Stewart, F. (2006) A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nature Methods* 3:839 – 844

25. Warren, M., Wang, W., Spiden, S., Chen-Murchie, D., Tannahill, D., Steel, K. P., and Bradley, A. (2007) Allan Bradley A Sall4 mutant mouse model useful for studying the role of Sall4 in early embryonic development and organogenesis. *Genesis* 45:51–58

26. Courta, R., Cooka, N., Saikrishnana, K., and Wigley, D. (2007) The crystal structure of λ-Gam protein suggests a model for recBCD inhibition. *J. Mol. Biol.* 371(1):25-33

27. Unger, R. C., and Clark, A. J., (1972) Interaction of the recombination pathways of bcateriophageλ and its host *Escherichia coli* K12: effects on exonuclease V activity. *J. Mol. Biol.* 70:539–48

28. Murphy, K.C. (1991) Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of *Escherichia coli* RecBCD enzyme. *J. Bacteriol.* 173:5808– 21

29. Karu, A. E., Sakaki, Y., Echols, H., Linn, S. (1975) The gamma protein specified by bacteriophage gamma. Structure and inhibitory activity for the recBC enzyme of *Escherichia coli*. *J Biol Chem.* **250**(18):7377–7387

30. Cassuto, E., and Radding, C.M. (1971). Mechanism for the action of λ exonuclease in genetic recombination. *Nat. New Biol.* 229:13–16

31. Carter, D. M., and Radding, C. M. (1971). The role of exonuclease and protein of phage λ in genetic recombination. II. Substrate specificity and the mode of action of lambda exonuclease. *J. Biol. Chem.* 246, 2502–2512.

32. Cassuto, E., Lash, T., Sriprakash, K. S., and Radding, C. M. (1971). Role of exonuclease and β protein of phage λ in genetic recombination, V. Recombination of λ DNA in vitro. *Proc. Natl. Acad. Sci.USA* 68, 1639–1643.

33. Hill, S. A., Stahl, M. M., and Stahl, F. W. (1997) Single-strand DNA intermediates in phage λ's Red recombination pathway. *Proc. Natl. Acad. Sci. USA* 94:2951–56

34. Little, J.W. (1967) An exonuclease inducedby bacteriophage λ II. Nature of the enzymatic reaction. *J. Biol. Chem.* 242:679–86

35. Sriprakash, K.S., Lundh, N., Huh M-O, and Radding, C.M. (1975) The specificity of λ exonuclease Interactions with single stranded DNA. *J. Biol. Chem.* 250:5438–45

36. Mitisis, P. G., and Kwagh, J. G. (1999) Characterization of the interaction of lambda exonuclease with the ends of DNA. *Nucleic Acids Res* 27(15): 3057-3063

37. van Oostrum, J., White, J. L., and Burnett, R.M. (1985) Isolation and crystallization of λ exonuclease. *Arch. Biochem. Biophys.*243:332–37

38. Kovall, R., and Matthews, B.W. (1997) Toroidal structure of lambda-exonuclease. *Science* 277:1824–27

39. Breyer, W. A., and Matthews, B. W. (2001) A structural basis for processivity. *Protein Science* 10:1699–1711.

40. Joseph, J. W., and R. Kolodner. 1983. Exonuclease VIII of *Escherichia coli*. I. Purification and physical properties. *J. Biol.Chem.* 258:10411-10417.

41. Joseph, J. W., and R. Kolodner. (1983) Exonuclease VIII of *Escherichia coli*. II. Mechanism of action. *J. Biol. Chem.* 258:10418-10424.

42. Kolodner, R., Hall, S. D., and Luisi-DeLuca, C. (1994) Homologous pairing proteins encoded by the *Escherichia coli* recE and recT genes. *Mol. Microbiol.* 11:23-30.

43.Clark, A. J., Sharma, V., Brenowitz, S., Chu, C. C., Satin, L., Templin, A., Berger, I., and Cohen, A. (1993) Genetic and molecular analyses of the C-terminal region of the recE gene from the Rac prophage of *Escherichia coli* K-12 reveal the *recT* gene. *J Bacteriol.* 175(23): 7673–7682.

44. Chang, H. W., and Julin, D.A. (2001) Structure and function of the *Escherichia coli* RecE protein, a member of RecB nuclease domain family. *J. Biol. Chem.* 276 (49): 46004-4601045.

45. Zhang, J., Xing, X., Herr, A.B., and Bell, C.E. (2009) Crystal Structure of E. coli RecE Protein Reveals a Toroidal Tetramer for Processing Double-Stranded DNA Breaks. *Structure*.17(5):690-702.

46. Martínez-Jiménez, M.I., Alonso, J.C., and Ayora S. (2005) Bacillus subtilis bacteriophage SPP1-encoded gene 34.1 product is a recombination-dependent DNA replication protein. *J. Mol Biol.* 351(5):1007-19.

47. Vellani, T.S., and Myers, R.S. (2003) Bacteriophage SPP1 Chu is an alkaline exonuclease in the SynExo family of viral two-component recombinases. *J Bacteriol.* 185(8):2465-74

48. Datta, S., Costantino, N., Zhou, X., and Court, D.L (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl. Acad. Sci. USA*. 105(5): 1626–1631

49. Radding, C. M., Rosenzweig, J., Richards, F., Cassuto, E. (1971) Separation and characterization of exonuclease, β protein, and a complex of both. *J. Biol. Chem.* 246:2510–12

50. Mythili, E., Kumar, K,A., Muniyappa, K. (1996) Characterization of the DNA binding domain of β protein, a component of phage λ Red-pathway, by UV catalyzed cross-linking. *Gene* 182:81–87

51. Muniyappa, K., Radding, C. M. (1986) The homologous recombination system of phage lambda. Pairing activities of β protein. *J. Biol. Chem.* 261:7472–78

52. Karakousis, G., Ye, N., Li, Z., Chiu, S.K., Reddy, G., Radding, C.M. (1998) The β protein of phage λ binds preferentially to an intermediate in DNA renaturation. *J. Mol. Biol.* 276:721–31

53. Li, Z., Karakousis, G., Chiu, S. K., Reddy, G., Radding, C.M. (1998) The Bet protein of phage λ promotes strand exchange. *J. Mol.Biol.* 276:733–44

54. Passy, S. I., Yu, X., Li, Z., Radding, C.M., and Egelman, H. (1996) Rings and filaments of β protein from bacteriophage l suggest a superfamily of recombination proteins *Proc. Natl. Acad. Sci.USA* 96: 4279-4284

55. Wu, Z., Xing, X., Bohl, C.E., Wisler, J. W., Dalton, J.T., and Bell, .C.E. (2006) Domain structure and DNA binding regions of *β* protein from Bacteriophage λ. *J. Biol. Chem.* 281(35): 25205–25214

56. Iyer, L.M., Koonin, E.V., Aravind, L. (2002) Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. *BMC Genomics*. 21;3(1):8. Epub 2002 Mar 21

57. Hall, S. D., Kane, M. F., and Kolodner, R. D. (1993) Identification and characterization of the *Escherichia coli* RecT protein, a protein encoded by the RecE region that promotes renaturation of homologous single-stranded DNA. *J. Bacteriol.* 175:277-287.

58. Hall, S. D., and R. D. Kolodner. (1994) Homologous pairing and strand exchange promoted by the Escherichia coli recT protein. *Proc. Natl. Acad. Sci. USA* 91:3205-3209.

59. Noirot, P., and Kolodner, R. D. (1998) DNA strand invasion promoted by *Escherichia coli* RecT protein. *J. Biol. Chem.* 273(20): 12274–12280

60. Alberts, B. M., and L. Frey. (1970). T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature* 227:1313-1318.

61. Alberts, B., and Herrick, G. (1971) DNA-cellulose chromatography. *Methods Enzymol.* 21:198-217.

63. Raghunathan S, Ricard CS, Lohman TM, and Waksman G. (1997) Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc. Natl. Acad. Sci.USA* 94:6652-6657

64. Williams, K.R., Spicer, E.K., LoPresti, M.B., Guggenheimer, R.A., and Chase, J.W. (1983) Limited proteolysis studies on the *Escherichia coli* single-stranded DNA

binding protein. Evidence for a functionally homologous domain in both the Escherichia coli and T4 DNA binding proteins. *J Biol Chem.* 258(5):3346-55

65. Lohman, T. M., and Overman, L.B. (1985) Two binding modes in *Escherichia coli* single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J Biol Chem.* 260(6):3594-603

66. Bujalowski, W., and Lohman, T. M. (1989). Negative co-operativity in *Escherichia coli* single strand binding protein-oligonucleotide interactions. I. Evidence and a quantitative model. *J.Mol. Biol.* 207:249-68

67. Bujalowski, W., and Lohman, T. M. (1989). Negative co-operativity in Escherichia coli single strand binding protein-oligonucleotide interactions. II. Salt, temperature and oligonucleotide length effects. *J.Mol. Biol.* 207:269-88

68. Bujalowski, W., and Lohman, T.M. (1986). *Escherichia coli* single-strand binding protein forms multiple, distinct complexes with single-stranded DNA.*Biochemistry* 25:7799—802

69. Bujalowski, W., Ovennan, L. B., and Lohman, T. M. (1988). Binding mode transitions of *Escherichia coli* single strand binding protein-single-stranded DNA complexes. Cation, anion, pH, and binding density effects. *J. Bioi. Chem.* 263:4629—40

70. Lohman, T. M., and L. B. Overman. (1985). Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J. Biol. Chem.* 260:3594-3603.

71. Bujalowski, W., and T. M. Lohman. (1986). *Escherichia coli* single strand binding protein forms multiple, distinct complexes with single-stranded DNA. *Biochemistry* 25:7799-7802.

72. Baker, T. A., Funnell, B. E., and Kornberg, A. (1987). Helicase action of dnaB protein during replication from the *Escherichia coli* chromosomal origin *in vitro*. *J. Biol. Chem.* 262:6877- 6885.

73. Sims, J., and E. W. Benz, Jr. (1980). Initiation of DNA replication by the *Escherichia coli* dnaG protein: evidence that tertiary structure is involved. *Proc. Nati. Acad. Sci.USA* 77:900-904.

74. Myers, T. W., and L. J. Romano. 1988. Mechanism of stimulation of T7 DNA polymerase by *Escherichia coli* single stranded DNA binding protein (SSB). *J. Biol. Chem.* 263: 17006-17015.

75. Kunkel, T. A., R. R. Meyer, and L. A. Loeb. (1979) Single strand binding protein enhances fidelity of DNA synthesis in vitro. *Proc. Natl. Acad. Sci.USA* 76: 6331 -6335.

76. Kunkel, T. A., R. M. Schaaper, and L. A. Loeb. (1983) Depurination-induced infidelity of deoxyribonucleic acid synthesis with purified deoxyribonucleic acid replication proteins *in vitro*. *Biochemistry* 22:2378-2384.

77. Kwon-Shin, O., Bodner, J. B., McHenry, C. S., and Bambara, R. A. 1987. Properties of initiation complexes formed between *Escherichia coli* DNA polymerase III holoenzyme and primed DNA in the absence of ATP. *J. Biol. Chem.* 262:2121-2130.

78. Fay, P. J., Johanson, K. O., McHenry, C. S., and Bambara, R. A. (1981). Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli. J. Biol. Chem.* 256:976-983.

79 Anderson, D.G., and Kowalczykowski, S.C. (1997) The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a χ-regulated manner. *Cell* 90: 77–86

80. Kowalczykowski, S.C., and Krupp, .RA. (1987) Effects of the Escherichia coli SSB protein on the binding of *Escherichia coli* RecA protein to single-stranded DNA: Demonstration of competitive binding and the lack of a specific protein–protein interaction. *J. Mol. Biol.* 193: 81–95

81. Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002). Conjugative transposons: the tip of the iceberg. *Mol. Microbiol.* 46: 601–610

82. Taviani, E., Ceccarelli, D., Lazaro, N., Bani, S., Cappuccinelli, P., Colwell, R.R., and Colombo, M. M.(2008) Environmental Vibrio spp., isolated in Mozambique,

contain a polymorphic group of integrative conjugative elements and class 1 integrons *FEMS Microbiol Ecol.*64(1):45-

83. Burrus, V., Waldor, M.K. (2004). Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.* 155:376–386.

84. Ehara, M., Nguyen, B.M., Nguyen, D.T., Toma, C., Higa, N., and Iwanaga, M., (2004). Drug susceptibility and its genetic basis in epidemic *Vibrio cholerae* O1 in Vietnam. *Epidemiol. Infect.* 132:595–600.

85. Iwanaga, M., Toma, C., Miyazato, T., Insisiengmay, S., Nakasone, N., and Ehara, M. (2004). Antibiotic resistance conferred by a class I integron and SXT constin in *Vibrio cholerae* O1 strains isolated in Laos. *Antimicrob. Agents Chemother.* 48, 2364–2369.

86. Waldor, M.K., Tschape, H., Mekalanos, J.J., (1996) A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. *J. Bacteriol.* 178, 4157–4165.

87. Hochhut, B., and Waldor, M.K.(1999). Site-specific integration of the conjugal *Vibrio cholerae* SXT element into *prfC*. *Mol. Microbiol.*32: 99–11

88. Yokota, T., and Kuwahara, S. (1977). Temperature-sensitive R plasmid obtained from naturally isolated drug-resistant *Vibrio cholerae* (biotype El Tor). *Antimicrob. Agents Chemother.*11, 13–20

89. Hochhut, B., Lotfi, Y., Mazel, D., Faruque, S.M., Woodgate, R.,Waldor, M.K., (2001b). Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. *Antimicrob. Agents Chemother.* 45, 2991–3000.

90. Ehara, M., Nguyen, B.M., Nguyen, D.T., Toma, C., Higa, N., and Iwanaga, M., (2004). Drug susceptibility and its genetic basis in epidemic *Vibrio cholerae* O1 in Vietnam. *Epidemiol. Infect.* 132:595–600.

91. Iwanaga, M., Toma, C., Miyazato, T., Insisiengmay, S., Nakasone,N., and Ehara, M. (2004). Antibiotic resistance conferred by a class I integron and SXT constin in *Vibrio cholerae* O1 strains isolated in Laos. *Antimicrob. Agents Chemother.* 48, 2364–2369.

92. Kingston, J.J., Zachariah, K., Tuteja, U., Kumar, S., and Batra, H.V. (2009) Molecular characterization of Vibrio cholerae isolates from cholera outbreaks in North India. *J Microbiol.* 47(1):110-5.

93. Burrus, V., Quezada-Calvillo, R., Marrero, J., and Waldor, M. K. (2006). SXT-related integrating conjugative element in New World *Vibrio cholerae*. *Appl. Environ. Microbiol.* 72: 3054-3057

94. Juiz-Rio, S., Osorio, C.R., de Lorenzo, V., and Lemos, M.L. (2005). Subtractive hybridization reveals a high genetic diversity in the Wsh pathogen *Photobacterium damselae* subsp. *piscicida*:evidence of a SXT-like element. *Microbiology* 151: 2659–2669.

95. Coetzee, J.N., Datta, N., and Hedges, R.W. (1972). R factors from *Proteusrettgeri*. *J. Gen. Microbiol.* 72: 543–552

96. Hochhut, B., Beaber, J.W., Woodgate, R., and Waldor, M.K., (2001). Formation of chromosomal tandem arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share an attachment site. *J. Bacteriol.* 183: 1124–1132

97. Beaber, J.W., Burrus, V., Hochhut, B., and Waldor, M.K., (2002). Comparison of SXT and R391, two conjugative integrating elements: definition of a genetic backbone for the mobilization of resistance determinants. *Cell Mol. Life Sci.* 59: 2065–2070

98. Beaber, J.W., Hochhut, B., and Waldor, M.K., (2002). Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. *J.Bacteriol.* 184: 4259–4269.

99. Boltner, D., MacMahon, C., Pembroke, J.T., Strike, P., and Osborn, A.M., (2002). R391: a conjugative integrating mosaic comprised of phage, plasmid, and transposon elements. *J. Bacteriol.* 184: 5158–5169

100. Beaber, J.W., and Waldor, M.K., (2004). Identification of operators and promoters that control SXT conjugative transfer. *J. Bacteriol.*186, 5945–5949

101. Yuen, K. Y., Woo, P. C. Y., Teng, J. L. L., Leung, K. W., Wong, M. K., and Lau, S. K. P.(2001). *Laribacter hongkongensis* gen. nov., sp. nov., a novel

gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. *J. Clin. Microbiol.* 39:4227–4232.

102. Woo, P.C.Y., Lau, S.K.P., Tse, H., Teng, J.L.L., Curreem, S.O.T., Tsang, A.K.L., Fan, R.Y.Y., Wong, G.K.M., Huang, Y., Loman, N.J., Snyder, L.A.S., Cai, J.J., huang, J.D., Mak, W., Pallen, M.J., Lok, S., and Yuen, K.Y. (2009) The complete genome and proteome of Laribacter hongkongensis reveal potential mechanisms for adaptations to different temperatures and habitats. *PLoS Genetics* 5(3) e1000416

103. Lau, S.K.P., Lee, L.C.K., Fan, R.Y.Y., Teng, J.L.L. Tse, C.W.S., Woo. P.C.Y, and Yuen, K.Y. (2009) Isolation of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, from Chinese tiger frog. *International Journal of Food Microbiology* 129 78–82

104. Lau, S.K.P., Woo, P.C.Y., Fan, R. Y.Y., Lee, R.C.M., Teng, J. L.L., and Yuen, K.Y. (2007) Seasonal and tissue distribution of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, in retail freshwater fish in Hong Kong. *International Journal of Food Microbiology* 113: 62–66

105. Ni, X. P., Ren, S.H., Sun, J.R., Xiang, H.Q., Gao, Y., Kong, Q.X., Cha, J., Pan, J.C., Yu, H., and Li, H.M. (2007) *Laribacter hongkongensis* isolated from a patient with community-acquired gastroenteritis in Hangzhou City. *J Clin Microbiol.* 45(1):255-6

106. Lau, S.K.P, Woo, P.C.Y., Fan, R.Y.Y., Ma, S.S.L., Hui, W.-T., Au, S.-Y., Chan, L.-L., Chan, J.Y.F. , Lau, A.T.K., Leung, K.-Y., Pun, T.C.T., She, H.H.L., Wong, C.-Y., Wong, L.L.L., and Yuen, K.-Y. (2007) Isolation of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, from drinking water reservoirs in Hong Kong *Journal of Applied Microbiology* 103: 507–515

107. Woo, P.C., Kuhnert, P., Burnens, A.P., Teng, J.L., Lau, S.K., Que, T.L., Yau, H.H., and Yuen, K.Y. (2003) Laribacter hongkongensis: a potential cause of infectious diarrhea *Diagn Microbiol Infect Dis.* 47(4):551-6.

108. Tolun, G., and Myers, R.S.(2003) A real-time DNase assay (RedA) based on PicoGreen® fluorescence. *Nucleic Acids Res* 31(18): e111

109. Raghunathan S, Ricard CS, Lohman TM, Waksman G. (1997) Crystal structure of the homo-tetrameric DNA binding domain of Escherichia coli single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc. Natl. Acad. Sci. USA* 94:6652-6657