Chapter 2: Role of accessory sequences in the site-specific recombination mediated by XisA

2.1 Introduction:

Genetic information can be reshuffled by inversion, insertion, duplication, deletion and translocation of DNA segments. Rearrangements fulfill specific physiological functions, programmed processes, random events contributing to the genetic diversity of a population. Site-specific recombination describes a variety of specialized recombination processes that involve reciprocal exchange between defined DNA sites. Site-specific recombination is currently described as to involve (*a*) two DNA partners, (*b*) a specialized recombinase protein that is responsible for recognizing the sites and for breaking and rejoining the DNA, and (*c*) a mechanism that involves DNA breakage and reunion with conservation of the phosphodiester bond energy (i.e., lacking a requirement for either DNA synthesis or a high-energy nucleotide cofactor).

Prototypes of site-specific recombination are the integration of bacteriophage λ into the *Escherichia coli* chromosome (Azaro, 2002), the resolution of cointegrates derived from transposition of Tn3-related transposons (Grindley, 2002), the DNA inversions responsible for flagellar phase variation (Johnson, 2002), developmentally regulated gene expression (Carrasco *et al.*, 1995; Craig *et al.*, 2002).

Site specific recombinases fall into two major groups: tyrosine recombinase and serine recombinase on the basis of the amino-acid involved in catalysis (Nunes-Duby *et al.*, 1997; Smith & Thorpe, 2002). In tyrosine recombinase, the active site tyrosine forms a phospho-tyrosine bond with the DNA while in case of serine recombinase serine is the corresponding residue. The tyrosine family is structurally diverse and extremely versatile and includes integrases, resolvases, invertases and transposases. The integrase family of

tyrosine recombinase is well studied and its prototype member is λ integrase (Nunes-Duby *et al.*, 1997).

The recombination site is generally 5-30 bp long and contains an inverted repeat of recognition sequence to which the enzyme binds. Cre and FLP are the examples of such enzymes which require only the recognition sequences. However, in case of others besides the cross-over site, additional sequences are also involved in the recombination such sequences span 100 or more base pairs from the cross over site. This region contains additional sites of protein recognition and may bind more copies of the recombinase or other protein factors encoded by the host cell or by the genetic element (e.g., phage or transposon) associated with the recombination system. Function of these additional DNA bound proteins may be regulatory, structural, or both. They may initiate or stabilize the pairing of recombination sites or inhibit inappropriate pairings; they may deliver recombinase catalytic domains to the crossover site; they may activate the recombinase; and they may determine the directionality of recombination (Grindley *et al.*, 2006)

Anabaena genome in the late stages of heterocyst maturation undergoes three site-specific recombination events (Haselkorn, 2002). All three recombinases reside in the excising element. Out of three, XisA and XisC belong to the integrase family particularly tyrosine recombinase while the XisF belongs to resolvase/invertase family (Carrasco *et al.*, 1995). XisA recognizes 11bp sequence CGGAGTAATCC. Besides the crossover site, the promoter region of the *xisA* shows the presence of additional sequences such as 8 bp repeats next to the 11bp sequences and three NtcA binding sites.

In order to check the role of additional sequences in XisA mediated recombination, different substrate plasmids with different combination of sequences flanking the target sites were used *in vivo E. coli* model system for the XisA mediated recombination.

2.2 Materials and methods

For routine growth and cloning purpose, *E. coli* was grown at 37°C in LB media at 200 RPM. The medium was supplemented with 100μg/ml ampicillin or 40μg/ml of gentamycin for maintenance of plasmids. For endonuclease assay, *E. coli* cells were grown on M9 minimal media whose composition as given below, where antibiotic concentration was used half as compared to Luria Broth.

2.2.1 Following antibiotic concentrations were used in the study:

Table 2.1 Antibiotic stock solutions

Antibiotic	Stock Concentration(1000X)
Ampicilin	100 mg/ml in st. d/w.
Kanamycin	50 mg/ml in st. d/w
Gentamycin	20 mg/ml in st. d/w
Chloramphenicol	40 mg/ml in Methanol
Tetracycline	30 mg/ml in 70% ethanol

M9 medium:

5X M9 Salt/liter (Sambrook *et al.*, 1989)

$Na_2HPO_4.2H_20$: 64.0 g
KH_2PO_4 : 15.0 g
NaCl : 2.5 g
NH_4Cl : 5.0 g
Distilled water to 1000ml;
After autoclaving add
Glucose : 100 mM
1M $CaCl_2.2H_2O$: 0.1 ml
1M $MgSO_4..7H_2O$: 2.0 ml
Thiamine -HCl : 10.0 mg /Liter

2.2.2 Bacterial strains and plasmids used:

Table 2.2

E. coli / Plasmids Strains	Genotype / Features	Source/Reference
JM101	*FΔtraD36lacIqΔ (lacZ)M15 proA$^+$B$^+$/supE thi Δ (lac-proAB)*	Yanisch-Perron *et al.*, 1985
DH5α	*FΔ`endA1hsdR17(r_k^- m_k^+)supE44 thi-1 recA1gyrA96(nal^r) relA1(lacZYA-argF)U169deoR (Δ80dlacΔ (lacZ)M15*	Hanahan, 1985
pMX32	Km^R Amp^R LacZ$^+$ *xisA*$^-$ (defective *nifD* element vector)	Lammers *et al.*, 1986
pxis-gm	Contains *xis* A, Gen^R	This study
pxis71A-gm	Contains *xis A* and *nifA* of *klebsiella pnuemoniae*, Gen^R	This study
pNG1	Contains two 11 bp sites flanking *lacZ*, Amp^R	This study
pNG2	Contains two 11bp + 8 bps flanking *lacZ*,Amp^R	This study

pMX32 plasmid :

pMX32 is *xisA* defective plasmid. It contains entire 11 kb *nifD* element flanked by two 11 bp target sites in the EcoR1 site of pBR322 plasmid. It also contains mini-*mu-lac* transposon inserted in such a way that it makes *xisA* gene non-functional. Size of the plasmid is 31 kb. It has ampicillin, kanamycin, tetracycline resistance and it is *lacZ* positive.

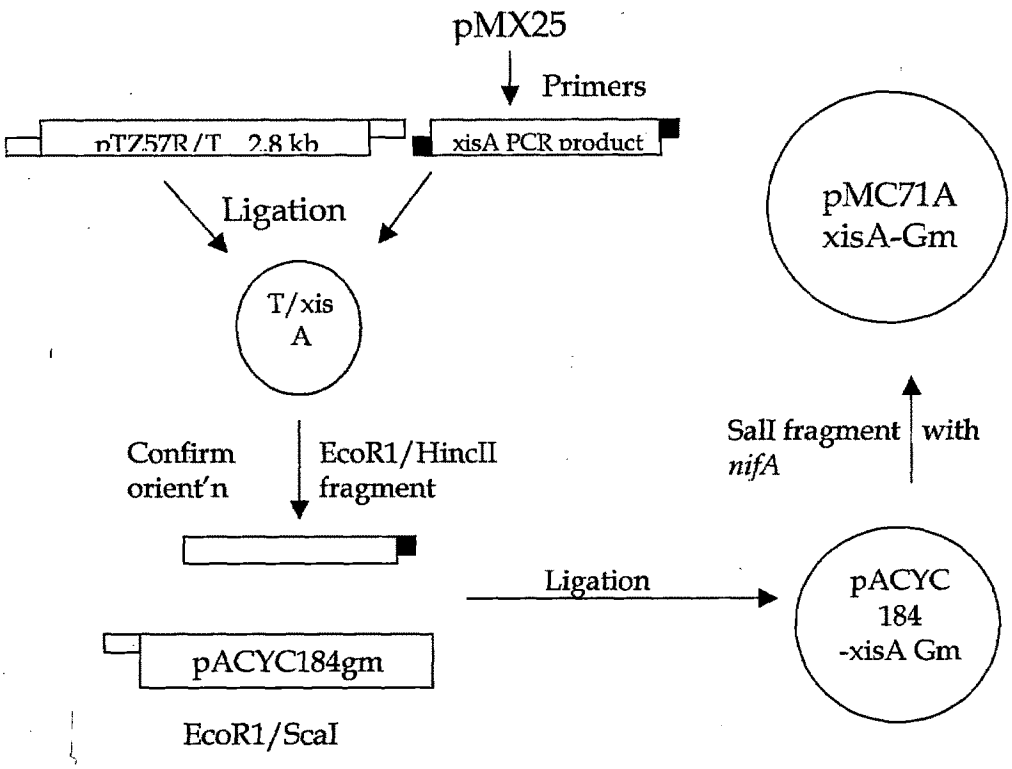
p*xis*-gm plasmid:

This plasmid contains *xisA* gene in EcoR1/ScaI site in the chloramphenicol resistance gene. This plasmid has gentamycin resistance gene and it does not contain 11 bp target site. The size of the plasmid is 7 kb.

p*xis*71A-gm:

This plasmid contains *nifA* gene from *Klebsiella pneumoniae* in the SalI site in the tetracycline resistance gene, under *tet* promoter. Size of the plasmid is 9.9 kb.

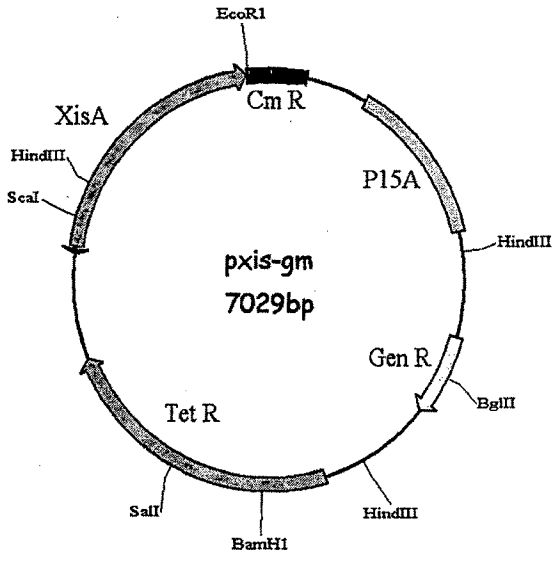
Strategies used for cloning:

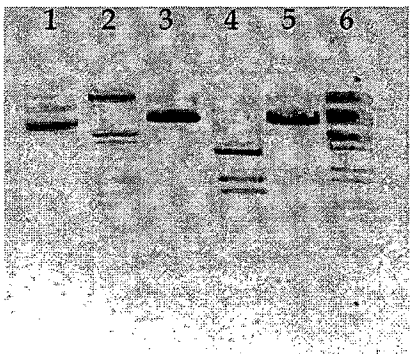


2.2.3 Construction of *xisA* containing plasmids:

xisA gene was amplified using forward primer (5′ TGAGTCGACTGATC AAAGCATTGAGCAGAT 3′ and reverse primer (5′ GGTGATCATAATGTTT GCACTGAGCAGTGT 3′) with pMX25 as template. Reactions were carried out in 50µl system with *pfu* polymerase (B. Genei, India) with primer concentration at 0.25µM and dNTPs concentration of 200nM. Initial denaturation was done for 3 minutes. For initial 5 cycles, annealing temperature was kept at 42°C while in the rest 20 cylces it was kept at 62°C, while the synthesis was carried out at 72°C

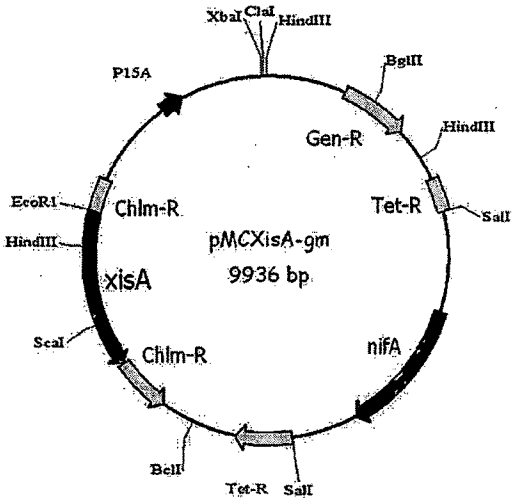
for 1 min throughout. 1.4 kb PCR product was purified using DNA extraction kit by Bangalore genei, India as per the given protocol. It was first cloned in pTZ57R/T vector according to the manual. 100 ng of PCR product was ligated with vector in presence of 40% PEG in 30µl reaction system at 22°C overnight. Transformation was done by chemical transformation method as per the given protocol in the manual. Recombinant plasmid (pTZ-xisA) was selected on the blue-white selection. PTZ-xisA was digested with EcoR1 and HincII enzyme to release a 1.5 kb insert containing *xisA* gene. Plasmids pACYC184-gm and pMC71A-gm were digested with EcoR1 and ScaI enzyme and this linearized vector was ligated with the 1.5 kb insert in 10µl ligation mixture which was carried out overnight at 22°C using 1-3 U of T4DNA ligase from MBI fermentas. Selection of clones was done on gentamycin plates with the loss of chloramphenicol resistance used as indicator for cloning.

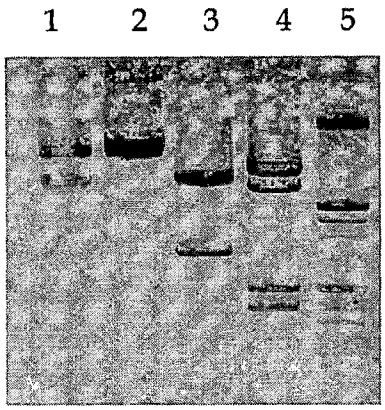




Lane 2 & 6 Mol. Wt. Marker
Lane 1 control plasmid
pCYC184 xis digestted with ?
Lane 3 Eco R1
Lane 4 HindIII
Lane 5 Bam H1

Figure 2.1 Plasmid map and restriction digestion pattern of p*xis*-gm





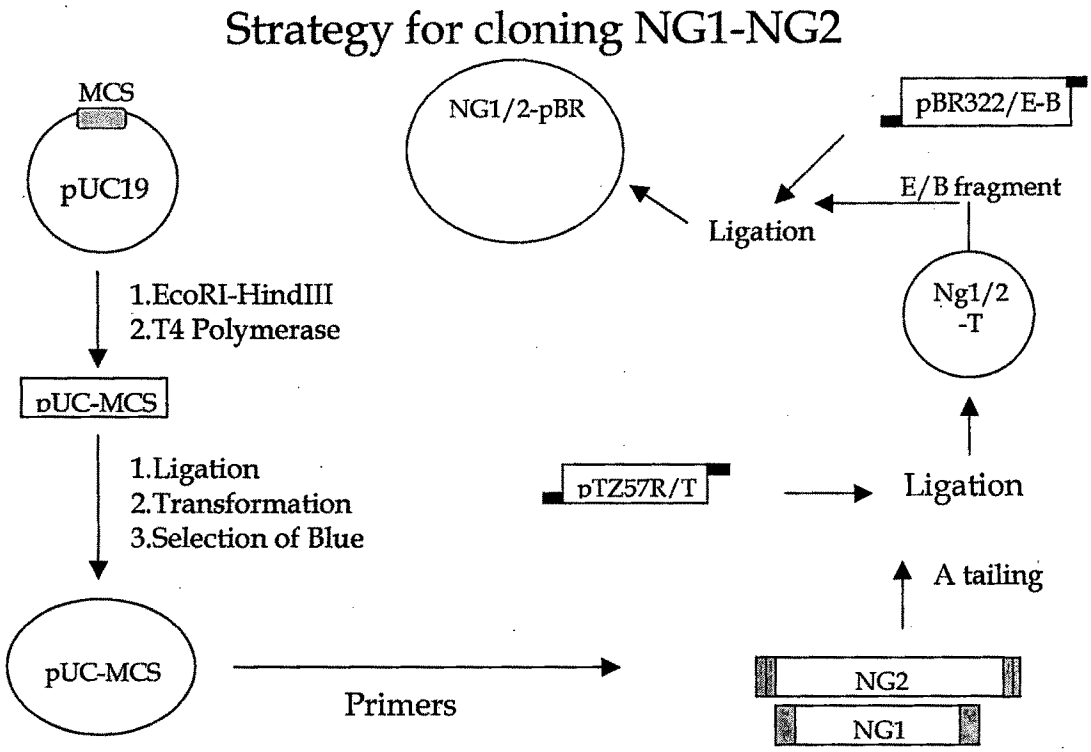
Lane 1 Control plasmid
pMC71AXisgm digested with
Lane 2 Eco R1
Lane 3 Sal I
Lane 4 HindIII
Lane 5 Mol. Wt. Marker

Figure 2.2 Plasmid map and restriction digestion pattern of pMC*xis*-gm

2.2.4 Construction of substrate plasmids NG1 and NG2 with two target sites

Strategy for cloning:

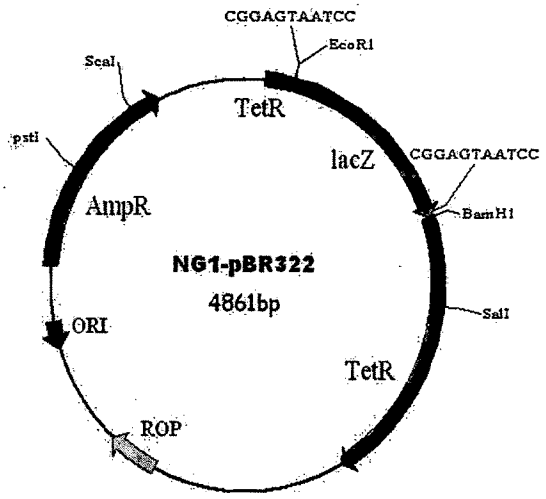
To remove MCS, pUC18 was first digested with EcoR1 and HindIII in 15µl reaction system at 37°C for 3 h. The digested vector was purified using gel extraction kit. To remove the overhangs, linearized DNA was treated with T4 DNA polymerase in 10µl reaction system at 37°C for 15 minutes and reaction was stopped by heating at 75°C for 20 minutes. This blunt DNA was allowed to self ligate in 10µl reaction system to generate pUC18 without MCS. Primers were designed to contain target site within its sequence.

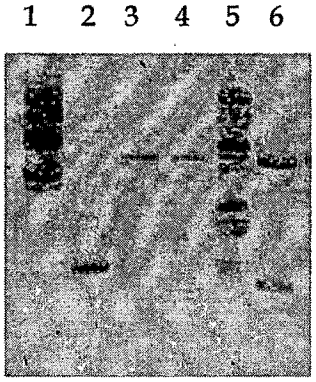


NG1 PCR product contains 11bp target site flanking the *lacZ* while NG2 contains 11bp+8bp flanking lacZ gene. Primers were designed to contain 11bp and 11+8 bp as followed (5′ TACGGAGTAATCCGTCAGCGAGTCAGT GA 3′) as forward and (5′ TAGGATTACTCCGTCCCACCTGACGTCTAA 3′) as reverse primer for NG1 (5′ TAGCGGAGTAATCCTTTTTGTGCAGCGAGTCAGTGA 3′) as forward and (5′ GGATTACTCCGTTTTTGTGCCACCT GACGTC TAAG 3′) as reverse primer for NG2. pUC19 without MCS was used as a template - Temperature profile 94°C for 3 mins, (94°C/30s, 42°C/30s, 72°C 90s) for 5 cycles and (94°C/30s, 62°C/30s, 72°C 90s) for 25 cycles. Cloning was done as described earlier.

2.2.4.1 pNG1 plasmid:

This plasmid contains two copies of the 11bp target site flanking the *lacZ* gene. This *lacZ* gene has been cloned in the EcoR1/BamH1 site of the pBR 322 plasmid. The size of the plasmid is 4.8 kb. It has ampicillin resistance.



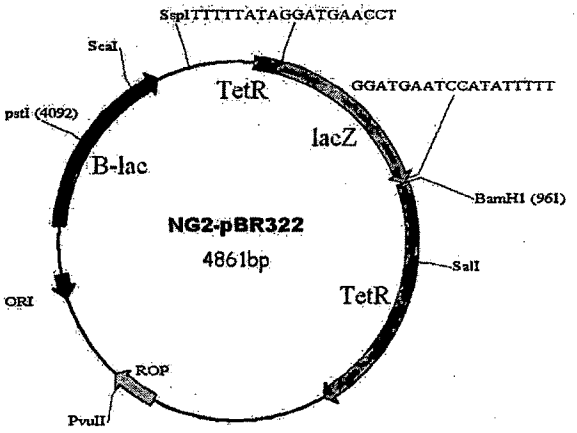


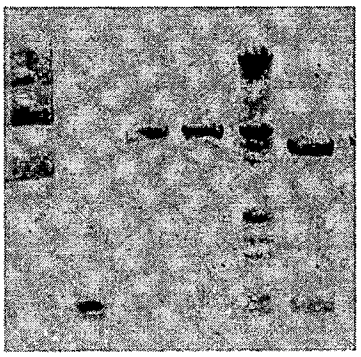
Lane 5 Mol. Wt. Marker
Lane 1 Control plasmid
Lane 2 NG1 PCR product
Lane 3 pBR 322/EcoR1
Lane 4 pBR 322/BamH1
Lane 6 NG1- pBR322/ECoR1/BamH1

Figure 2.3 Plasmid map and restriction digestion pattern of pNG1

2.2.4.2 pNG2 plasmid :

This plasmid contains two copies of 11bp+8 bp sites flanking the *lacZ* gene in pBR 322 plasmid.





Lane 1 Control plasmid
Lane 2 NG2 PCR product
Lane 3 NG2-pBR/EcoR1
Lane 4 NG2-pBR/BamH1
Lane 5 Mol. Wt. marker
Lane 6 NG2-pBR/EcoR1/BamH1

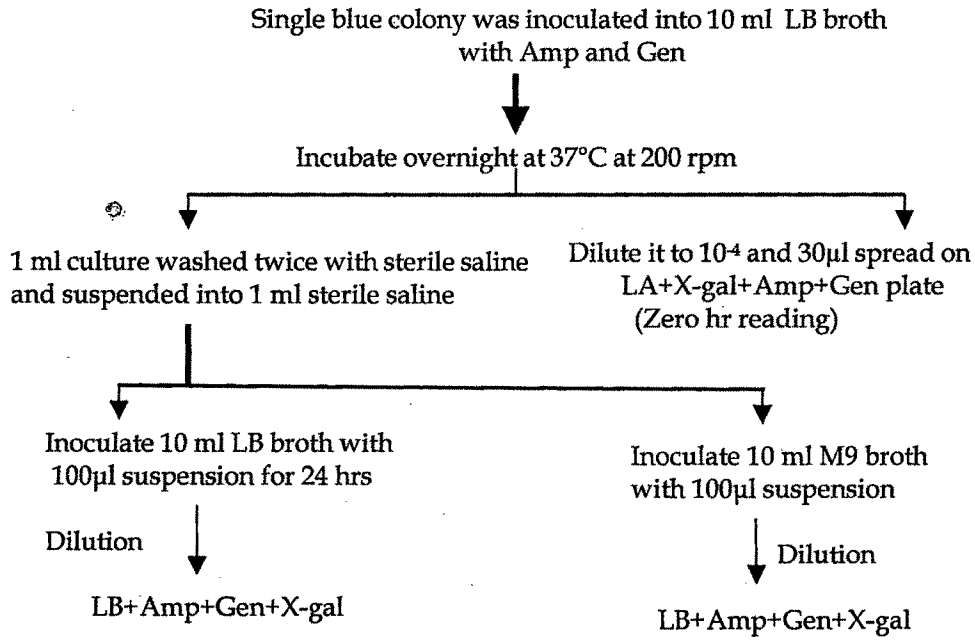
Figue 2. 4 Plasmid map and restriction digestion pattern of pNG2

2.2.5 Rearrangement assay:

E. coli containing pNG1/pNG2 along with *xisA* gene containing plasmids pxisgm/pxis71Agm was grown overnight by inoculating a single colony in LB medium with ampicillin (100μg/ml) and gentamycin (40μg/ml). The overnight grown culture was washed thrice with Saline. From this, proper aliquote was used to prepare dilutions and for the 0 h plating. 100 μl of the washed cells were further inoculated in 10 ml of LB and 10ml of M9 with containing appropriate antibiotics. After 24 h, 1ml of sample was withdrawn from LB and M9 media and washed thrice with saline. It was diluted properly and plated on LA plates containing Ampicilin-gentamycin-IPTG-X-gal. Next day blue and white colonies were calculated to find out the rearrangement frequency using following formula:

$$\text{Excision (\%)} = \frac{\text{Total number of white colonies}}{\text{Total number of colonies}} \times 100$$

Protocol :-rearrangement

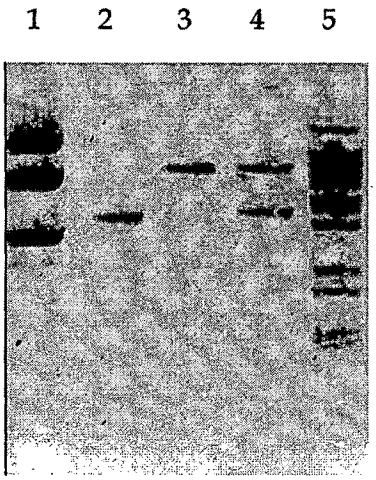


For control experiments, target sites containing plasmids were transformed with pACYC 184 and pMC 71A separately and used for the rearrangements.

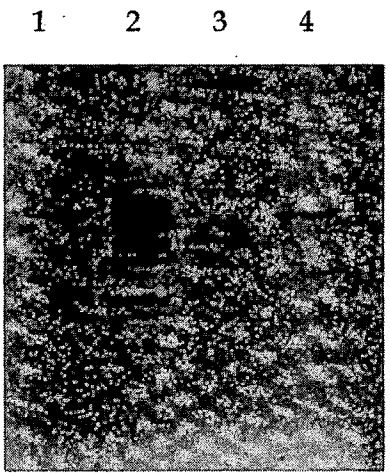
2.3 Results:

In order to check the role of additional sequences different substrate plasmids were transformed with *xisA* gene containing plasmids. *E. coli* containing pNG1, pNG2 and pMX32 plasmids was transformed with pxis-gm and pxis71A-gm plasmids respectively. Transformants were selected on ampicillin and gentamycin plates. Conformation of the double transformants was determined by analysis of restriction digestion pattern (fig 2.5).

Figure 2.5 Restriction digestion patterns of substrate plasmid and *xis* containing plasmid



Lane 1 Control plasmid
Lane 2 NG1-PBR/EcoR1
Lane 3 184xisgm/EcoR1
Lane 4 NG1-pBR+184xisgm/EcoR1
Lane 5 Mol. Wt. Marker



Lane 1 NG2-pBR+184xisgm/EcoR1
Lane 2 Mol. Wt. Marker
Lane 3 NG2-pBR/EcoR1
Lane 4 184xisgm/EcoR1

E. coli containing both substrate and *xisA* plasmids were grown overnight and 100µl was inoculated in LB and M9 media and kept for 24 h after which it was washed, diluted and selected on plates with X-gal for blue-white selection to monitor the rearrangement frequency (Table 2-4).

Media	No. of Blue colonies	No of white colonies	Excision (%)
Presence of *xisA*			
LB	391± 36	1	0
M9	312± 7	2	0
Presence of *xisA* and *nifA*			
LB	652± 39	1	0
M9	313± 15	2	0

Table 2.3 : Rearrangement of NG1-pBR in LB and M9 media

Media	No. of Blue colonies	No of white colonies	% Excision
Presence of *xisA*			
LB	377±22	2	0
M9	347±10	2	0
Presence of *xisA* and *nifA*			
LB	437±12	2	0
M9	413±7	3	0

N=5

Table 2.4: Rearrangement of NG 2 in presence of *xisA* gene in LB and M9 media

Table 2.5 : Rearrangement of pMX32 in LB and M9 minimal media

Media	No. of Blue colonies	No of white colonies	% Excision
Presence of *xisA*			
LB	1203± 30	50± 8	3.99± 1.6
M9	1286± 100	186± 21	12± 3
Presence of *xisA* and *nifA*			
LB	1680 ± 120	720 ± 110	30 ± 4
M9	1286 ± 100	1245 ± 150	49 ± 7

N=5

2.4 Discussion:

XisA is necessary to carry out the *nifD* element excision in the *Anabaena* PCC 7120. Present study has been carried out to find out the role of accessory sequences other than the recognition sequence. NG1 plasmid with only 11bp target sequence flanking the *lacZ* gene in pBR322 plasmid in presence of *xisA* gene did not show any rearrangement. This implies that just 11 bp is not sufficient to carry out the rearrangement by XisA protein.

The promoter region of the *xisA* gene contains GTGTTTTT 8 bp inverted repeat sequence near to the 11bp target sites. pNG2 plasmid containing the 11bp target sequence along with the 8 bp sequence flanking the *lacZ* gene also did not show any rearrangement even in the presence of *xisA* and *nifA* genes. In majority of the site specific recombinases require the presence of accessory proteins binding sites in surrounding region of the recognition sequence (Grindley *et al.*, 2006).

Besides the 8 bp target site, the promoter region of the *xisA* gene contains 3 NtcA/NifA binding sites and 6 more such sites are present at the distal part of the *nifD* element. pMX32 plasmid contains entire *nifD* element with a defective *xisA* gene in presence of *xisA* gene provided in *trans* showed rearrangement upto

~4% in LB medium and ~12% in M9 medium. In presence of *nifA* gene, in LB the rearrangement increased upto 30 % while in M9 medium it was around 50%. These results demonstrate that the role of NifA/NtcA binding sites in the excision of *nifD* element. *Anabaena* PCC 7120 NtcA, which is a transcriptional activator of *xisA* gene, could be playing a similar role by either stabilizing the DNA-XisA complex or bring the two 11bp target site in close proximity by bending the DNA through protein-protein interactions between the bound NtcA.

The rearrangement is higher in case of M9 medium than LB. This implies that the physiological conditions could also play an important role in the rearrangement. Accessory protein/factors interacting with the sequences adjacent to the target site may be produced in higher amounts in M9 medium growth condition. It is supported by the fact that rearrangement in the presence of *nifA* gene was 49% in case of M9 medium whereas 30% in LB.

XisA seems to function in a manner similar to λ integrase and many other members of integrase family of tyrosine recombinases where accessory sequence/sequences play very important role. In *Bacillus subtilis*, along with the cross over site, the recombination site contains a 21 bp inverted repeat sequences which has TTTTTGTA sequence. In case of XerD mediated recombination in *E. coli*, accessory sequence of ~200 bp is necessary for the segregation of plasmids by site-specific recombination. This sequence could be the binding sites of proteins which bring about looping of DNA. In the case of λ integrase, HU protein and IHF plays such role. In case of XisA protein, so far no protein/factor has been determined.