



Chapter 3

Expression and Purification of Functional XisA Protein

1.1 Introduction

Anabaena PCC 7120 genome is interrupted by three genetic elements; *nifD*, *fdxN* and *hupL*, named according to genes which each one disrupts. These elements are of 11,278 kb, 59,428 kb and 9,419 kb size, respectively and get excised by site-specific recombination mediated by three different excisases (XisA, XisF, XisC, respectively) each encoded within the specific element, during late stages of heterocyst differentiation resulting in the formation of functional genes facilitating nitrogen fixation (Kaneko et al., 2001).

1.1.1 Toxicity hypothesis of XisA protein

Even though the *xisA* gene was cloned and sequenced long ago (Lammers et al., 1986), neither the protein nor transcript has been detected so far in *Anabaena* nor in *E. coli* strains expressing *xisA*. It is interesting to note that in most of the studies on the rearrangement of *nifD* element in *E. coli*, the *xisA* expression was indirectly monitored by *in vivo* excision of the *nifD* element by loss of *lacZ* gene inserted within the *nifD* element. It is believed that XisA protein in overexpressed state is toxic to the *E. coli* expression host hampering its purification and *in vitro* characterization (Lammers et al., 1986; Lammers et al., 1990; Shah et al., 2007; Karunakaran et al., 2008).

1.1.2 Rationale of the present study

The present aimed to identify and address factors hampering *xisA* gene expression and purify functional XisA protein. The present study also deals with providing the first computational insight of secondary and 3D structure of XisA, an important enzyme of *Anabaena* PCC 7120 for nitrogen fixation.

1.2 Materials and Methods

1.2.1 *E. coli* Strains, Plasmids and Oligonucleotide Primers Used in the Study

E. coli strains and plasmid used in the present study are summarized in **Table 3.1** and **Table 3.2**, respectively. Oligonucleotide primers synthesis was out sourced to 1st Base Pvt. Ltd (Singapore) and are summarized in **Table 3.3**.

Table 3.1 Plasmids used in the study.

Name	Characteristics	Reference
pMX25	<i>nifD</i> element with <i>lacZ</i> and Kan ^R cloned in pBR322, Amp ^R	(Lammers et al., 1986)
pAM461	Contains distal and proximal borders of <i>nifD</i> element, Amp ^R	(Lammers et al., 1990)
pJET1.2	pBR322 based positive selection cloning vector with a lethal insert that allows for efficient recovery of blunt-ended PCR products, Amp ^R	Thermo scientific
pJETxisA	pJET1.2 containing full length <i>xisA</i> gene	This study
pET28a (+)	T7 promoter based expression vector, Kan ^R	EMD Bioscience, Darmstadt.
pxisA	pET28a (+) containing full length <i>xisA</i> gene under pT7 promoter, Kan ^R	Chapter 3 & 5

Table 3.2 *E. coli* strains used in the study.

Name	Genotype	Reference
DH5α	<i>fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del) M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	(Hanahan 1983)
BL21 DE3	<i>F⁻ ompT gal dcm lon hsdS_B (r_B⁻ m_B⁻) λ(DE3 [<i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i>])</i>	(Studier and Moffatt 1986)
UN1	BL21 DE3 (pET28a + pAM461)	Chapter 5
UN2	BL21 DE3 (p xisA + pAM461)	Chapter 5

Table 3.3 Oligonucleotide primers used in the study.

Name	Sequence
XisA(F)	5'-GGATCCATGCAAAATCAGGGTCAA-3'
XisA(R)	5'-GAGCTCTCAACTATTCTTATAAGCTATTTCCA-3'
C(F)	5'- GTATCTCTCTACGCTTGCTGGTTGG-3'
C(R)	5'- ACCACCCACTACATCGATAACGCC-3'
X(R)	5'- TGCCGTCGGTAGATGAAAGTGGC-3'

1.2.2 Construction of recombinant *pxisA* (pET28a +*xisA*) expression vector

Strategy for cloning *xisA* gene in pJET1.2 and sub cloning in pET28a is depicted in **Fig. 3.1**.

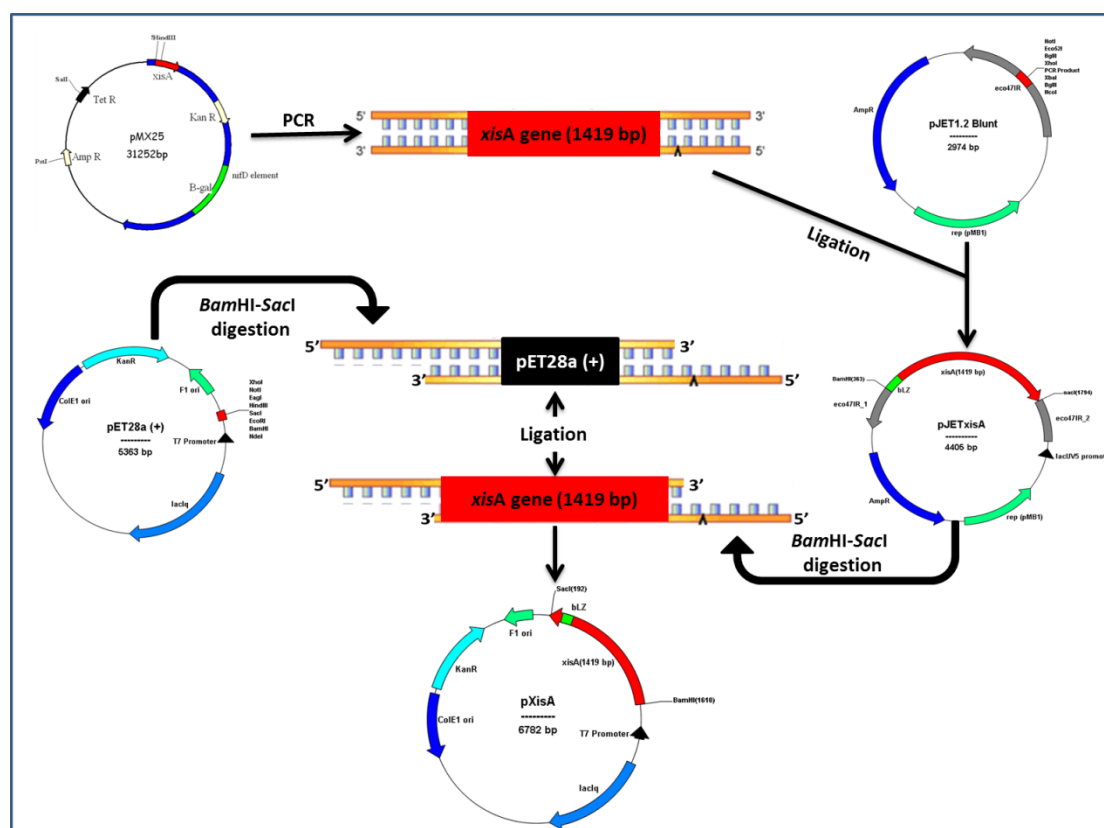


Fig. 3.1 Schematic representation of *xisA* gene cloning strategy.

1.2.2.1 Cloning of *xisA* gene in pJET 1.2 cloning vector

xisA gene was PCR amplified from plasmid pMX25 by Pfu polymerase (Thermo Scientific) using primers XisA(F) forward primer and XisA(R) reverse primer (**Table 2.1**). 1.4 kb *xisA* amplicon was eluted from agarose gel and was ligated with pJET1.2 vector (Thermo Scientific) according to manufacturer's protocol. *E. coli* DH5 α was transformed with this ligation mixture and transformants were screened on Luria Agar plates supplemented with ampicillin to a final concentration of 100 mg L⁻¹. To screen positive clones, next day few transformed colonies were sub cultured to 3 ml Luria Broth supplemented ampicillin to a final concentration of 100 mg L⁻¹ and were allowed to grow as described in section 2.2. On the following day, plasmids were isolated and screened for positive clones by *Bam*HI and *Sac*I double digestion. Restriction analysis was performed on agarose gel. Plasmid isolates displaying 1.4 kb *xisA* insert released were marked as positive and named pJET1*xisA* (**Fig. 3.1**). Fresh

E. coli DH5 α was transformed with recombinant pJET*xisA* and the culture was maintained in -80 °C with a final concentration 50 % glycerol.

1.2.2.2 Sub-cloning of *xisA* gene in pET28a bacterial expression vector

To clone 1.4 kb *xisA* gene in pET28a, recombinant pJET*xisA* and pET28a vectors were double digested with *Bam*HI and *Sac*I. The digested *xisA* gene and pET28a vector backbone were eluted and ligated using T4 DNA Fast Ligase (Thermo Scientific) according to manufacturer's protocol. Recombinant plasmids were screened for positive clones as described in section 3.2.3.2. Positive clones were named as *pxisA* bearing *xisA* gene under strong IPTG inducible T7 promoter (Fig. 3.1).

1.2.3 Toxicity assessment of *xisA* gene overexpression in *E. coli*

The toxicity assessment of IPTG induced *xisA* gene overexpression on the growth of the *E. coli* was monitored by measuring optical density (OD₆₀₀). All measurements were taken from triplicates of each treatment of varying IPTG concentration. *E. coli* BL21 DE3 was transformed with *pxisA*. Transformants were screened on Luria Agar supplemented with Kanamycin 50 mg L⁻¹ final concentration. Transformed colonies were sub cultured in Luria Broth with Kanamycin 50 mg L⁻¹ final concentration and were allowed to grow overnight as described in section 2.2. Following day, fresh LB Kan medium was seeded with 1 % inoculum of *E. coli* suspension and was allowed to grow till mid log phase {O.D₆₀₀ = 0.5 as measured on Shimadzu UV-1800 absorption spectrophotometer (Tokyo, Japan)}. At this point, culture was aseptically supplemented with different concentrations of IPTG (0 mM, 1.5 mM 1.5 mM and 2 mM) and allowed to grow in conditions as described above. These experiments were performed simultaneously with control *E. coli* BL21 DE3 transformed with pET28a. Following the IPTG induction, the growth of *E. coli* was measured at 1-h intervals over a 10-h period.

1.2.4 *In vivo* detection of XisA protein endonuclease activity by restriction digestion

In vivo detection of XisA recombinase activity was carried out according to Brusca et al (1990) using pAM461, a pUC18 based *xisA* deficient plasmid, which contains distal and proximal borders of *nifD* element, used as a substrate plasmid.

Excision of minimal *nifD* element was brought about by providing XisA in trans using *pxisA*. *E. coli* BL21 (DE3) bearing *pxisA* was co transformed with pAM461 and transformants were selected on LB agar plates with 50 µg ml⁻¹ ampicillin and 25 µg ml⁻¹ kanamycin. A single colony of dual transformant was inoculated in 15 ml LB medium with same antibiotics and was allowed to grow overnight in uninduced condition. On the following day, plasmids were isolated and digested with *HindIII* and *EcoRI*. Digestion mixture was analysed on 0.8 % agarose gel. The appearance of 715 bp band indicates rearrangement of pAM461 and XisA activity.

1.2.5 Expression and purification of XisA protein

E. coli BL21(DE3) bearing pNU1 was grown overnight in 100 ml LB medium containing 50 µg ml⁻¹ kanamycin and grown at 37 °C until the OD₆₀₀ reached 0.6–0.7. Expression of *xisA* gene was induced by the addition of 0.8 mM IPTG. XisA protein was detected using SDS-PAGE followed by Coomassie blue R250 staining (Sambrook and Russell, 2006). IPTG induced *E. coli* (pNU1) cells were sonicated in 10 ml of lysis buffer (20 mM Na-phosphate, 0.5 M NaCl, 20 mM Imidazole). Cell debris was removed by centrifugation at 12000 g for 10 min. Histidine-tagged protein was purified from the culture free supernatant by Nickel- nitrilotriacetic acid (Ni-NTA) affinity chromatography using AKTA Prime system (GE Healthcare Life Sciences, Buckinghamshire, UK). Binding of protein to the column was carried out in lysis buffer. Elution was carried using 0.5 M Imidazole in lysis buffer. Eluted fractions were analyzed by SDS PAGE.

1.2.6 Confirmation of purified XisA protein by MALDI-TOF analysis

Purified protein band was excised from SDS PAGE and was analysed by tryptic digestion and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI/TOF MS) at Molecular Biophysics Unit, Indian Institute of Science, Bangalore. Peptide masses generated from the MALDI/TOF MS analysis were used for protein identification by peptide mass fingerprinting (PMF) using the search program MASCOT (Matrix Science). Trypsin was given as the digestion enzyme, one missed cleavage site was allowed. All peptide mass values are monoisotopic and the mass tolerance was set at 100 ppm.

1.2.7 Theoretical modelling of XisA protein

Secondary structure prediction and 3D modelling of XisA protein was performed using online bioinformatics tools. To predict the secondary structure of XisA, the amino acid sequence was submitted to PSIPRED tool (Jones, 1999; Buchan et al., 2013). The domain architecture of XisA was analysed by submitting *xisA* gene nucleotide sequence at Conserved Domain Database tool of National Centre for Biotechnology Information (NCBI) (Marchler-Bauer et al., 2015). The amino acid sequence of XisA was submitted to I-TASSER server to determine 3D structure (Zhang, 2008; Roy et al., 2010; Yang et al., 2015). The output of I-TASSER was analysed using Pymol program (DeLano, 2010). Confidence score (C-score) and Template Modelling score (TM-score) are the parameters to rank the quality of the modelled structure. A C-score value typically ranges from -0.2 to 5 and model with higher C-score describes model with higher confidence. A model with TM score >0.5 is considered to show significant topology with template homolog. XisA protein model with highest C-score and TM-score was selected for structural analysis.

1.3 Results

1.3.1 Cloning and heterologous expression of *xisA* gene

31 kb pMX25 plasmid harboring *nifD* element was the source of *Anabaena* PCC 7120 *xisA* gene. pMX25 was confirmed by restriction digestion analysis with *Hind*III which yielded thirteen fragments upon complete digestion (**Fig. 3.2a**). 1.4 kb *xisA* gene was PCR amplified (**Fig. 3.2b**) and was ligated to pJET1.2 and sub cloned in pET28a to yield pJET*xisA* and *pxisA*, respectively, in a two step cloning procedure schematically represented in **Fig. 3.1**. Recombinant pJET1.2 and *pxisA* were confirmed via restriction digestion analysis (**Fig. 3.4a & b**).

1.3.2 Purification and confirmation of XisA protein

E. coli JM101 (pMX25, pMC71A) did not show high *xisA* expression for the purpose of purification (Karunakaran Thesis, 2000). Hence, *xisA* was cloned in pET28a to yield *pxisA* plasmid with *xisA* is present under strong inducible P₇₇ promoter and the expressed protein would have N-terminal histidine tag. **Fig. 3.4a** is the schematic representation of *xisA* gene cloned in pET 28a and positions of major

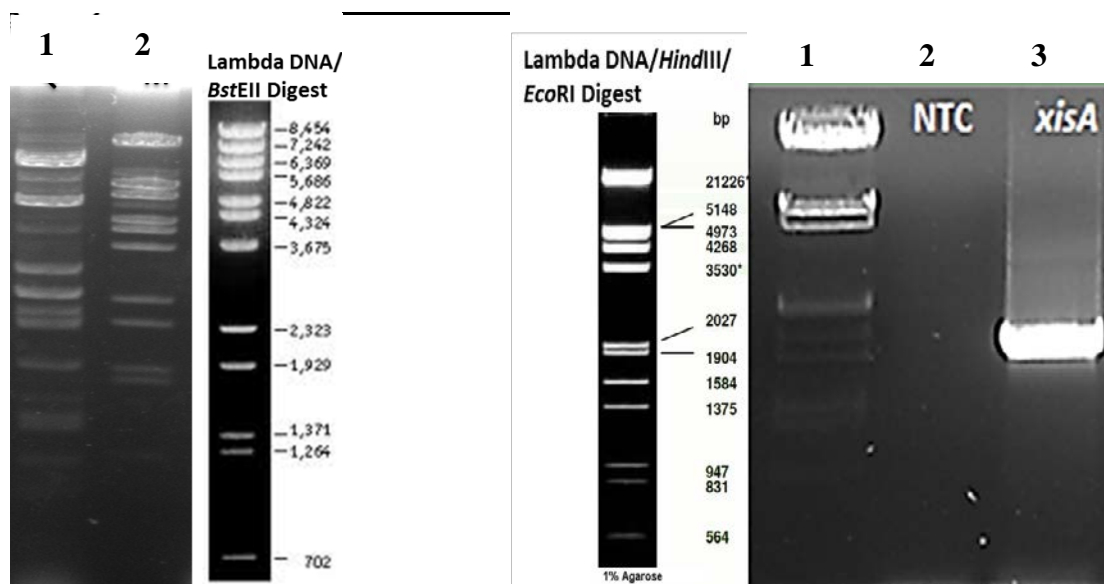


Fig. 3.2 PCR amplification of *xisA* gene. (a) Restriction digestion analysis of pMX25. Lane 1: *Hind*III digested pMX25; Lane 2: Lambda DNA *Bst*EII digest. (b) Agarose gel profile of amplified *xisA* gene. Lane 1: Lambda DNA *Hind*III-*Eco*RI double digest; Lane 2: No template control (NTC); Lane 3: 1.4 kb *xisA* gene amplicon.

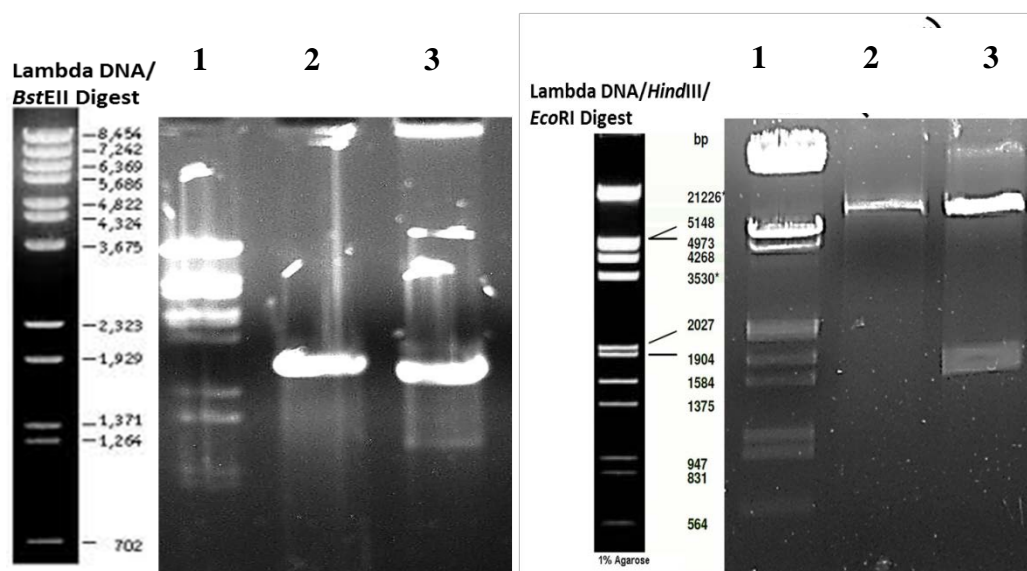


Fig. 3.3 Restriction digestion analysis of *xisA* gene bearing recombinant vectors. (a) Restriction digestion analysis of pJETxisA vector. Lane 1: Lambda DNA *Bst*EII digest; Lane 2: *Bgl*II digested pJET1.2; Lane 3: *Bgl*II digested pJETxisA. (b) Restriction digestion analysis of pxisA vector. Lane 1: Lambda DNA *Hind*III-*Eco*RI digest; Lane 2: *Bam*HI-*Sac*I digested pET28a; Lane 3: *Bam*HI-*Sac*I digested pxisA.

restriction enzyme sites. As seen in **Fig. 3.4b (lane 3)**, strong expression of *xisA* gene was achieved using pxisA plasmid and 52 kDa XisA protein was detected after 3 h of induction with 0.8 mM IPTG in *E. coli* BL21 (DE3) (*pxisA*) grown on LB medium as analysed by SDS PAGE. Ni-NTA affinity chromatography with AKTA prime FPLC system allowed one step purification of the His-tagged 52 kDa XisA protein (**Fig. 3.4c, lane 2**).

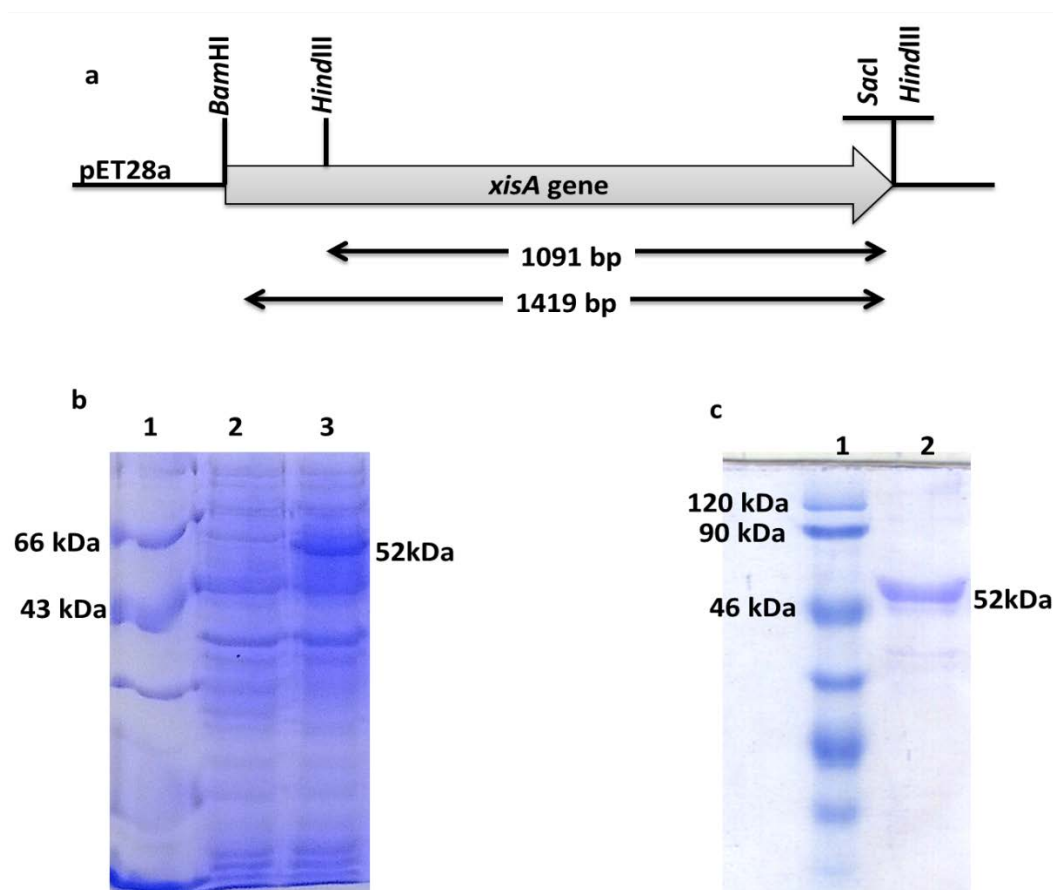


Fig. 3.4 Overexpression and purification of XisA protein expressed using *E. coli* BL21 (DE3) (pxisA) overexpression system. (a) Schematic representation of *xisA* gene expression cassette of pNU1 with recognition sites for major restriction endonucleases. (b) SDS PAGE gel representing overexpression XisA. Lane 1: Protein molecular weight marker; Lane 2 represents crude cell free extract of uninduced *E. coli* BL21 DE3 (pNU1); Lane 3 displays overexpressed 52 kDa XisA protein after 3 h of induction with 0.8 mM IPTG. (C) SDS gel picture displaying purification of XisA protein. Lane 1: protein Molecular weight marker; Lane 2: purified 52 kDa XisA protein facilitated through Ni-NTA affinity chromatography purification.

Purified band was confirmed to be XisA by Protein Mass Finger Printing facilitated by MALDI TOF mass spectrometry. Analysis of m/z values analysis by MASCOT showed a match score of 84 and a query coverage of 36% with XisA protein corresponding to the genome sequence of *Anabaena* PCC 7120 *xisA* gene product (Fig. 3.5).

1.3.3 Effect of *xisA* gene overexpression in *E. coli*

To evaluate the toxicity effect of overexpressed *xisA* gene in *E. coli* BL21 DE3 growth, $O.D_{600}$ of transformants was measured under different IPTG concentrations as described in section 3.2.3. From the growth curve analysis

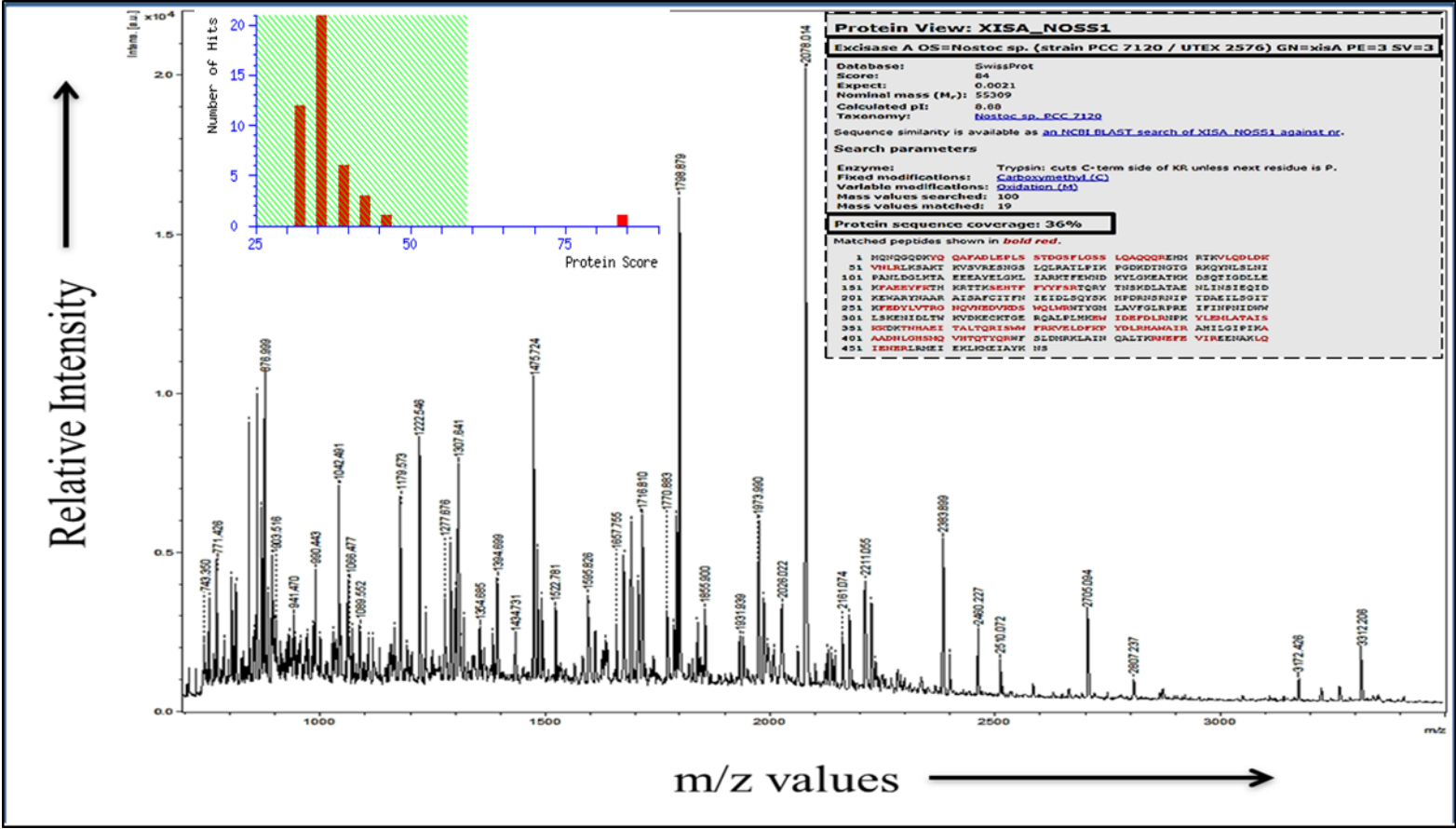


Fig. 3.5 MALDI TOF mass spectrometry analysis of *Anabaena* PCC 7120 XisA protein purified from *E. coli* BL21 (DE3) (*pxisA*). (Insets) Histogram representing the significant score of matched peptides corresponding to XisA protein obtained from MASCOT server.

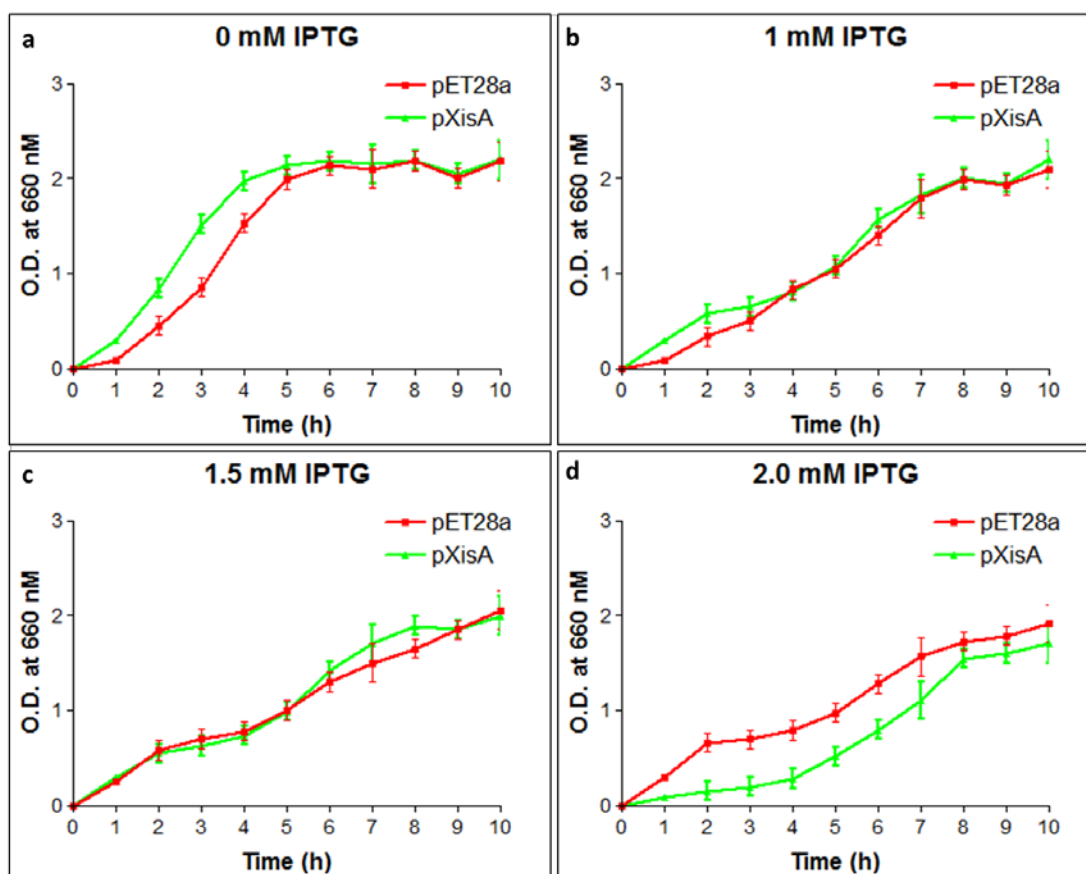


Fig. 3.6 *E. coli* growth curves at different IPTG concentration. (a) 0 mM IPTG. (b) 1 mM IPTG. (c) 1.5 mM IPTG (d) 2 mM IPTG

(Fig. 3.6), it was observed that there was no significant difference in *E. coli* BL21 (DE3) (*pxisA*) growth kinetics up to 1.5 mM IPTG concentration compared to control *E. coli* BL21 (DE3) pET28a which lacks *xisA* gene. However, overall growth was reduced against increasing concentration of IPTG for both control and test groups. The overexpressed *xisA* gene did not hinder the growth of *E. coli*. Only a significant reduction in growth kinetics was observed for *E. coli* BL21 (DE3) (*pxisA*) at 2.0 mM IPTG concentration compared to vector control which may be attributed IPTG toxicity and exhaustion of cellular resources for XisA protein expression under very high IPTG concentration.

1.3.4 *In vivo* detection of XisA functionality

Plasmid pAM461 was confirmed using *Hind*III-*Eco*RI double digestion which yields 1.2 kb insert release as seen in restriction analysis (Fig. 3.7a). In order to determine functionality of XisA protein expressed in pET28a expression system, *E. coli* BL21 DE3 bearing pNU1 was co-transformed with pAM461 substrate plasmid to

obtain *E. coli* strain UN2. Rearrangement of pAM461 was detected by 715 bp fragment released upon *Hind*III and *Eco*RI double digestion of plasmid mixture isolated from *E. coli* UN1 (**Fig. 3.7c, lane 1**) which was not seen in *E. coli* UN1 control which lacked *xisA* (**Fig. 3.7c, lane 2**). Leaky expression of *xisA* gene from *pxisA* was sufficient to carryout complete rearrangement of pAM461.

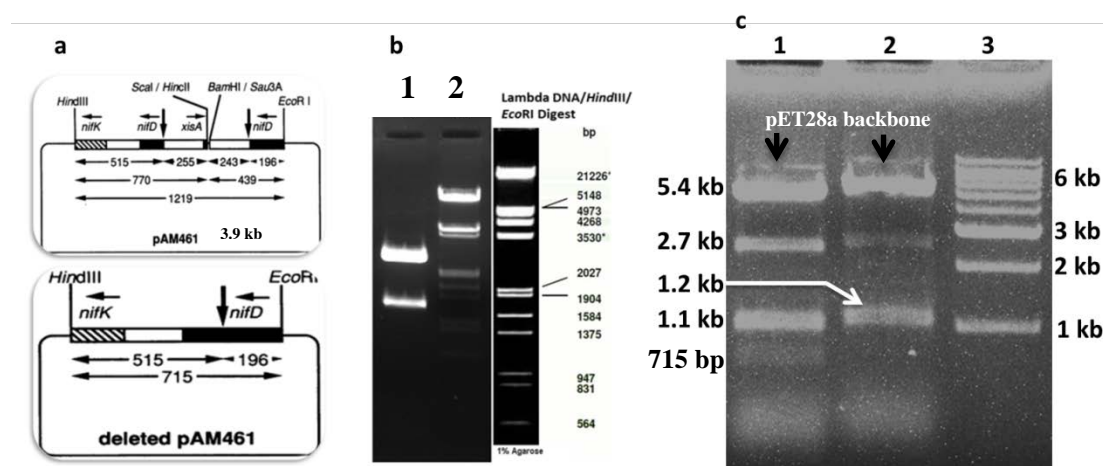


Fig. 3.7 Agarose gel showing *Eco*RI and *Hind*III digestion patterns for confirmation of pAM461 and validation of XisA functionality. (a) Schematic representation of pAM461 rearrangement (Brusca et al., 1990). (b) Restriction digestion confirmation of pAM461. Lane 1: *Eco*RI – *Hind*III double digestion; Lane 2: Lambda DNA *Eco*RI – *Hind*III digest. (c) Agarose gel profile displaying *in vivo* functionality of XisA. Lane 1: plasmid from *E. coli* BL21 (DE3) (pNU1, pAM461); Lane 2: plasmid from *E. coli* BL21 (DE3) (pET28a, pAM461); Lane 3: 1 kb DNA ladder. 715 bp band found exclusively in Lane 1 indicates rearrangement of pAM461.

1.3.5 Predicting secondary and 3D structure of XisA protein

Secondary structure prediction by PSIPRED tool revealed *Anabaena* PCC 7120 XisA is predominantly α -helical in nature. Active site residues (R287, Y381, R384 and Y416) are located in or near close proximity to α -helices (**Fig. 3.8**). Conserved Domain Database analysis of XisA suggested that the region of 268 to 416 residues bears sequence homology to Integrase region of tyrosine recombinase family (**Fig. 3.9a**). Additionally, the region of 435 to 472 showed homology with basic region of leucine zipper family (**Fig. 3.9a and 3.9b**). 3D modeling of XisA protein was performed by I-Tasser tool while model displaying a C-score value of 0.17 and a TM-score of 0.838 was chosen for structural and ligand binding analysis in Pymol (**Fig. 3.9c**). Closest structural homolog of XisA was found to be *Pyrococcus abyssi* Xer recombinase, which is also a member of tyrosine recombinase family. As seen in the XisA-DNA complex model, the active site residues are present in near vicinity of DNA (**Fig. 3B inset**).

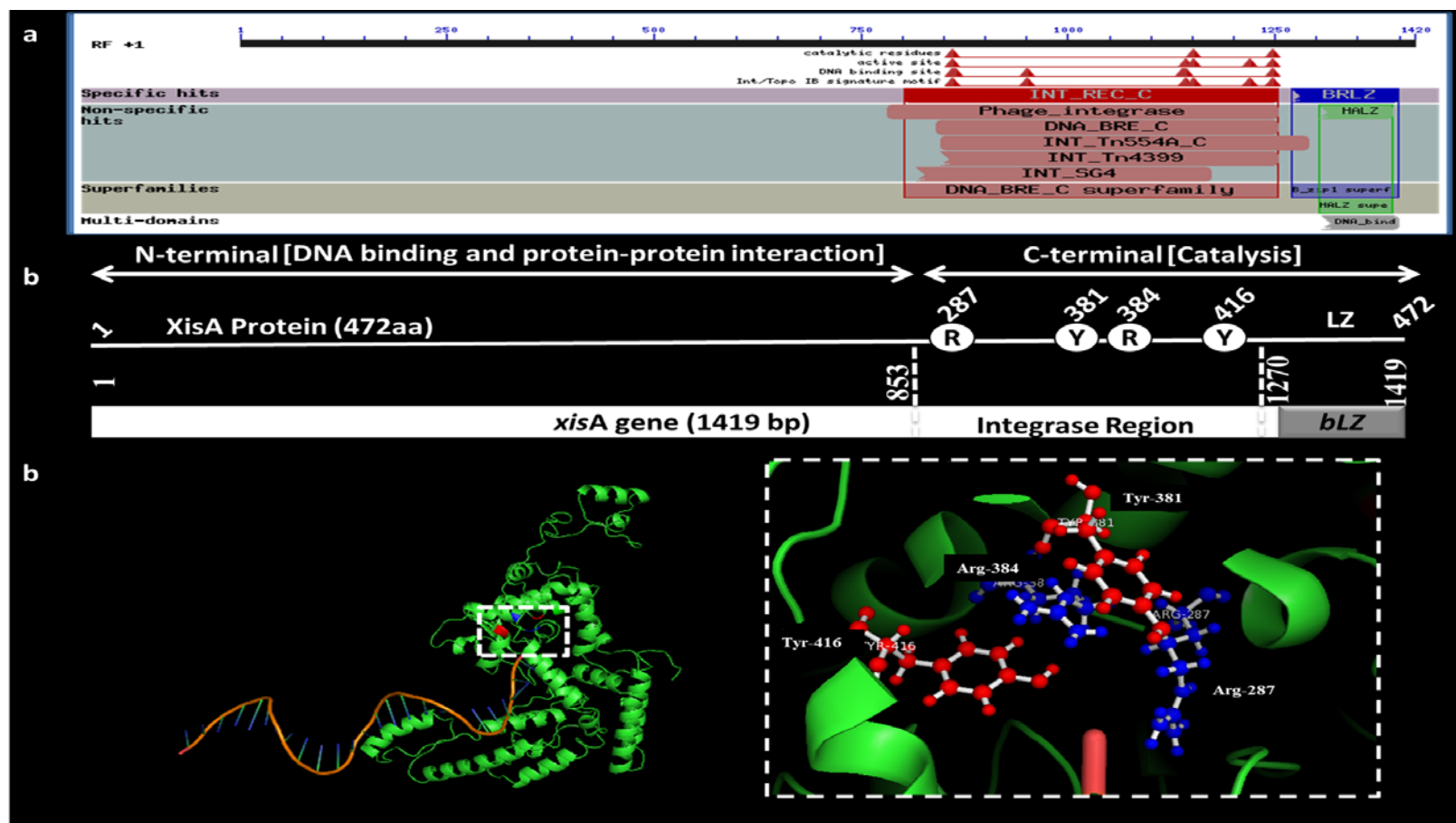


Fig. 3.9 Sequence analysis and theoretical modelling of XisA protein. (a) Sequence analysis of XisA protein. *xisA* gene sequence analyzed using Conserved Domain Database tool of National Centre for Biotechnology Information (NCBI). (b) Schematic representation of the protein and open reading frame of XisA. (c) Theoretical modelling of XisA protein bound to DNA using I-Tasser tool. Dotted box represents XisA active site residues.

purified so far neither it is detected on protein gels and hence it's *in vitro* characterization has been elusive. The present study was aimed at understanding the structural architecture of XisA and to report first purification and identification of XisA.

Expression of *xisA* gene under native promoter seems to be highly restrictive owing to opposite orientation of putative promoters P1 and P2 (Lammers *et al.* 1986). This limitation was addressed by expressing *xisA* gene under strong P_{T7} promoter (pNU1). Overexpression of *xisA* gene was observed at 3 h after induction with IPTG. Expressed XisA protein was His-tagged that facilitated one step purification by Ni-NTA affinity chromatography which was confirmed by MALDI-TOF mass spectrometry. That the overexpressed protein *E. coli* BL21 (DE3) (pNU1) was functionally active was shown by rearrangement of a mini substrate plasmid pAM461(Brusca *et al.*, 1990) by obtaining the restriction enzyme digestion pattern of the rearranged plasmid as reported by Brusca *et al* (1990).

Sequence analysis and modelling studies revealed 3D geometry of XisA is identical to *Pyrococcus abyssi* Xer recombinase, which is also a member of tyrosine recombinase family and bears predominantly α -helical structure. Position of active site residues near DNA ligand is ideal for recombination reaction. Involvement of terminal tyrosine (Tyr416) appears to participate in the formation of phosphor-tyrosine intermediate, a conserved mechanism for the members of Tyrosine recombinase family. Presence of predicted basic region of leucine zipper (bZIP) at extreme C-terminal region of XisA suggests it could play an important role in the XisA dimer formation whereas N-terminal is known to be associated with DNA binding and dimer-dimer interaction for other members of tyrosine recombinase family.

In conclusion, the purified XisA protein can facilitate elucidation of XisA protein structure, which is critical in understanding the excision of *Anabaena* PCC *nifD* element and may also be helpful in unravelling its relation to other recombinases and endonucleases.