Journal of Microbiology (2015) Vol. 53, No. 12, pp. 837–846
DOI 10.1007/s12275-015-5281-3

In silico analysis and experimental validation of lipoprotein and novel Tat signal peptides processing in *Anabaena* sp. PCC7120[§]

Sonika Kumari and Akhilesh Kumar Chaurasia[#*]

Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai-400085, India
[#]Present address: Department of Molecular Cell Biology, School of Medicine, Sunkyunkwan University, Suwon- 440746, Republic of Korea

(Received Jun 2, 2015 / Revised Nov 2, 2015 / Accepted Nov 4, 2015)

Signal peptide (SP) plays a pivotal role in protein translocation. Lipoprotein- and twin arginine translocase (Tat) dependent signal peptides were studied in All3087, a homolog of competence protein of *Synechocystis* PCC6803 and in two putative alkaline phosphatases (ALPs, Alr2234 and Alr4976), respectively. *In silico* analysis of All3087 is shown to possess the characteristics feature of competence proteins such as helix-hairpin-helix, N and C-terminal HKD endonuclease domain, calcium binding domain and N-terminal lipoprotein signal peptide. The SP recognition-cleavage site in All3087 was predicted (AIA-AC) using SignalP while further in-depth analysis using Pred-Lipo and WebLogo analysis for consensus sequence showed it as IAA-C. Activities of putative ALPs were confirmed by heterologous overexpression, activity assessment and zymogram analysis. ALP activity in *Anabaena* remains cell bound in log-phase, but during late log/stationary phase, an enhanced ALP activity was detected in extracellular milieu. The enhancement of ALP activity during stationary phase was not only due to inorganic phosphate limitation but also contributed by the presence of novel bipartite Tat-SP. The Tat signal transported the folded active ALPs to the membrane, followed by anchoring into the membrane and successive cleavage enabling transportation of the ALPs to the extracellular milieu, because of bipartite architecture and processing of transit Tat-SP.

Keywords: *Anabaena* sp. PCC7120, lipoprotein signal peptide, bipartite signal peptide, protein transport, twin arginines translocase, Sec pathway

Introduction

Majority of secreted or integral membrane proteins are synthesized as pre-proteins that possess N-terminal signal peptides for their translocation (Blobel and Dobberstein, 1975). A significant number of proteins in bacteria are exported from cytoplasm to cell membranes and extracellular milieu. This necessitates the presence of signal peptides which play a crucial role in protein export in all living organisms from archaea, bacteria to eukaryotes and eukaryotic cellular organelles such as mitochondria and chloroplast (Blobel and Dobberstein, 1975; von Heijne, 1990; Dalbey and Kuhn, 2000; Kushnareva *et al.*, 2001; Bolhuis, 2002; Zhang and Glaser, 2002). Signal peptides are characterized by positively charged region in the N-terminus (n-region) responsible for efficient translocation, followed by a hydrophobic region (h-region) crucial for translocation and c-terminus (c-region) with a specific signal peptidase (SPase) recognition and cleavage site, which together give it a very specific architecture (Ling Lin *et al.*, 2007). Various kinds of signal peptides, such as secretory (Sec), twin arginine translocase (Tat), lipoprotein, prepillin, and bacteriocin/pheromone, have been predicted and characterized in both Gram-negative and Gram-positive bacteria (Pugsley *et al.*, 1990; Tjalsma *et al.*, 1999, 2000; Widdick *et al.*, 2006; Rezwan *et al.*, 2007). Signal peptide performs three major functions, namely (a) protection of nascent amino acid chains from misfolding, (b) recognition by signal recognition particle (SRP) or transport system and (c) as a substrate for the secretion machinery (Hartl *et al.*, 1990; Dalbey *et al.*, 2000; Wickner and Schekman, 2005).

The major protein secretion pathway entails the general secretory pathway (Sec) or a more specialized twin arginine translocase (Tat) pathway. The sec-SRP pathway takes place in 3 functional steps: (a) targeting, (b) translocation, and (c) folding and release. The Sec pathway has been subcategorized based on involvement of transport of (a) extracellular proteins possessing typical sec-signal peptides having glycine (G) and proline (P) helix-breaking conserved residues in the h-region (Sec-type) or (b) lipoproteins, wherein the signal has shorter h-region devoid of these conserved P and G residues, and the proteins are integrated into the membrane upon interaction with protein, and/or lipid bilayer exposed towards the extracytoplasmic side (Simonen and Palva, 1993; Tjalsma *et al.*, 2000; Zanen *et al.*, 2005). The Tat SP recognizes longer and less hydrophobic signal peptides carrying a distinctive pattern of two consecutive arginine residues (RR) in the n-region (Tjalsma *et al.*, 2004). The major differences in the two pathways is the translocation of unfolded proteins by Sec pathway through a protein-conducting channel, followed by refolding after translocation, while the Tat pathway exports completely folded proteins (Wickner and Schekman, 2005; Joshi *et al.*, 2010; Palmer and Berks, 2012).

Tat and Sec dependent pathways have been reported in unicellular cyanobacteria, *Synechocystis* sp. PCC6803 (Spence *et al.*, 2003), *Synechococcus* sp. (Nakai *et al.*, 1994), and *Phormidium* (Barbrook *et al.*, 1993). However, the information

*For correspondence. E-mail: chaurasia@skku.edu, chaurasia.ak@gmail.com; Tel.: +82-31-299-6150
[§]Supplemental material for this article may be found at http://www.springerlink.com/content/120956.

available on cyanobacterial signal peptide is scarce and restricted by using non-native cytoplasmic protein; *i.e.* Green Fluorescent Protein (GFP) in *Synechocystis* sp. PCC6803 (Spence *et al.*, 2003) and *Anabaena* sp. strain PCC7120 (hereafter referred to as *Anabaena* 7120) (Mariscal *et al.*, 2007). Tagging GFP provides visualization advantage. However, it cannot provide the information about the state (folded or unfolded) of transported protein because of inherent folding properties of GFP in cytoplasm. The cleavage of a membrane-targeted manganese superoxide dismutase (MnSOD) by signal peptidase resulting in its distribution to both the periplasm and cytosol has been demonstrated in *Anabaena* 7120 (Raghavan *et al.*, 2013).

In this study, as a first step towards understanding translocation of proteins with different signal sequences in *Anabaena*, work was initiated by analyzing the translocation of two specific proteins, namely (i) putative membrane protein All3087 and (ii) putative alkaline phosphatases Alr2234 and Alr4976. The study on these three proteins of *Anabaena* reveals (a) specific signal peptidase cleavage sites in All3087 protein, (b) transport of fully folded active alkaline phosphatase by Tat pathway, (c) bipartite signal peptide mediated successive transport to the membrane and then during the late logarithmic or stationary phase, ALP was secreted to the extracellular medium, and (d) physiological relevance and rationale of ALPs successive transport. In some cases use of foreign signal peptides for the expression of secretory protein can work sub-optimally and need optimization for expression (Lam *et al.*, 2003; Ravn *et al.*, 2003). The photoautotrophic growth of cyanobacteria provides a unique feature to be used as a cellular factory for the production of useful biochemical. The present results on native signal peptides of *Anabaena* 7120 can be utilized without further optimization to express protein extracellularly to reduce the downstream processing of industrially useful enzymes. The twin arginine translocase signal has potential to be employed for pH dependent release of protein/enzyme. These applications are conceivable as cyanobacteria are being extensively used in industrial productions of protein based probiotics and pigments.

Materials and Methods

Strains and culture conditions

Anabaena 7120 was grown photoautotrophically in BG11 medium (pH 7.2) (Castenholz, 1988) with or without combined nitrogen (17 mM $NaNO_3$) at 27°C ± 2°C under continuous illumination 30 μE/m^2/sec under static or shaking culture conditions. The BG11 media was supplemented with 175 μM K_2HPO_4 inorganic phosphate (phosphate replete) or without inorganic phosphate for phosphate deplete conditions. Growth was estimated as chlorophyll *a* content per ml in methanolic extracts (Mackinney, 1941). *Escherichia coli*

Table 1. Primers, plasmids, and strains used in this study

Primers	Sequences	Source/reference
*alr2234*F	ACCACACCATGGATGGTAAATTACCATAATGTTCAG	This study
*alr2234*R	ATACCACTCGAGAACTACTTGTCGTACTCCAGGAATAC	This study
*alr4976*F	AGCATACCATGGGTGAATATGCAATTAATAAATCGCAAC	This study
*alr4976*R	ACCAGACTCGAGCGCTGTTTTTGCGCCTGGTTGCCCGTTTTC	This study
*all3087*F	AGCATACATATGGCGATCGCTGCCTGTCAAAAAGTCCAATC	This study
all3087–N-His- R	ATCATACTCGAGCTAAAATTGAATACGCCCTTCC	This study
*all3087*ΔSCPS- F	ACCATACATATGCAAAAAGTCCAATCTCACAATAATCG	This study
all3087-C-His-R	ATCATACTCGAGAAATTGAATACGCCCTTCC	This study
Plasmids		
pET16b	Expression vector with N-terminal His-tag protein	Novagen
pET29a	Expression vector with C-terminal His-tag protein	Novagen
pET16b-*all3087*	*all3087* with N-terminal His-tag with SPCS	This study
pET16b-*all3087*ΔSCPS	*all3087* with N-terminal His-tag and without SPCS	This study
pET29a-*all3087*	*all3087* with C-terminal His-tag and with SPCS	This study
pET29a-*all3087*ΔSCPS	*all3087* with C-terminal His-tag and without SPCS	This study
pET29a-*alr2234*	*alr2234*with C-terminal His-tag	This study
pET29a-*alr4976*	*alr4976* with C-terminal His-tag	This study
Strains		
E. coli DH5α	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoRnupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ–*	Laboratory strain
E. coli BL21(DE3)	*F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3) pLysS(cmR)*	Laboratory strain
Anabaena sp. 7120	Fresh water photoautotrophic filamentous heterocystous diazotroph	R.Haselkorn
Ec(*alr2234*)	*E. coli* BL21(DE3) expressing Alr2234	This study
Ec(*alr4976*)	*E. coli* BL21(DE3) expressing Alr4976	This study
All308-N-His	*E. coli* BL21(DE3) expressing All3087with N-terminal His-tag and SP	This study
All3087ΔSCPS-N-His	*E. coli* BL21(DE3) expressing All3087 with N-terminal His-tag and without SCPS	This study
All3087-C-His	*E. coli* BL21(DE3) expressing All3087with C-terminal His-tag and SCPS	This study
All3087ΔSCPS-C-His	*E. coli*BL21(DE3) expressing All3087 with C-terminal His-tag and without SCPS	This study

strains were grown in LB broth medium or LB-agar plate at 37°C, supplemented with 50 µg/ml kanamycin or 100 µg/ml carbenicillin and 33 µg/ml chroramphenicol in case of recombinant *E. coli* cells.

Bioinformatics analysis

Sequences of the hypothetical gene (*all3087*) and the annotated putative alkaline phosphatase encoding genes (*alr2234* and *alr4976*) of *Anabaena* 7120 were obtained from cyanobase (http://microbedb.genome.jp/*Anabaena*) (Kaneko *et al.*, 2001). The presence of Sec and twin arginine signal sequences and possible translocation of these proteins were analyzed using online servers, SignalP version 3.0 and http://www.cbs.dtu.dk/services/TatP/, respectively (Emanuelsson *et al.*, 2007). The presence of signals and its type were further verified using other online analysis tools such as Phobius, PSORT, Pred-Lipo etc. The conserved consensus sequences of lipoprotein signal peptide (LppSP) of All3087 was analyzed using the reported sequences of LppSP (Supplementary data Table S1) using WebLogo. The hydropathy analysis of the deduced amino acid for these putative membrane proteins was performed using the BioEdit program (Hall, 1999). The structures of the All3087 and Slr0197 were predicted by Phyre server (Kelley and Sternberg, 2009). Homology modeling, structural alignment, and helix hairpin helix analysis were performed using UCSF Chimera 1.8 program (Pettersen *et al.*, 2004).

Construction of expression plasmids

The DNA fragments coding for All3087 with or without putative signal peptidase cleavage site (SPCS) were cloned into N-terminal (pET16b) and C-terminal His-tag (pET29a) expression vectors (Novagen). The N-terminal constructs were used to assess the N-terminal signal processing while the C-terminal constructs served as the control for All3087 proteins expression with or without signal peptidase cleavage site (SPCS). The *all3087* (1.629 kb) with SPCS having either N-terminus His-tag or C-terminal His-tag was amplified using *all3087*F (Table 1) as the forward primer and either *all3087*-N-His-R or *all3087*-C-His-R (Table 1) as the reverse primer, respectively. For PCR amplification of *all3087* without SPCS, the forward primer used was *all3087*ΔSPCS-F (Table 1) in combination with either *all3087*-N-His-R or *all3087*-C-His-R (Table 1) as the reverse primer to generate *all3087*ΔSPCS with either N-terminal or C-terminal His-tag constructs, respectively. The PCR products were cloned at *Nde*I-*Xho*I restriction sites in pET16b and pET29a, respectively. The resultant plasmids were designated as pET16b-*all3087*, pET29a-*all3087*, pET16b-*all3087*ΔSPCS, and pET29a-*all3087*ΔSPCS (Table 1).

The annotated putative phosphatase *alr2234* (*phoD1*), *alr4976* (*phoD2*) were PCR amplified from *Anabaena* 7120 genome using the primer pairs shown in Table 1. The PCR amplified *alr2234* and *alr4976* (1.593 and 1.575 kb) inserts and pET29a vector were restriction digested with *Nco*I and *Xho*I restriction endonucleases (New England Biolabs), purified using Qiagen gel extraction kit (Qiagen) and ligated to obtain the plasmids pET29a-*alr2234* and pET29a-*alr4976*. These ALP plasmids were transformed into *E. coli* BL21 (DE3)/p*lysS* strain to achieve heterologously overexpressing ALPs strains of *E. coli* (Table 1). The accuracy of the cloned sequences was verified by DNA sequencing.

Protein expression, Western blotting, and immuno-detection

The recombinant proteins were expressed *E. coli* BL21 (pLysS) (DE3) cells upon induction with 1 mM IPTG at 37°C for 3 h unless stated otherwise. Alr4976 protein was induced at 18°C for 16 h for the optimal induction to achieve the protein in soluble fraction. Protein extraction, Western blotting and immunodetection of the C- or N-terminal His-tag proteins using anti-His-tag monoclonal antibody was performed as described earlier (Chaurasia and Apte, 2011).

Alkaline phosphatase assay

Whole cell protein extracts from *Anabaena* 7120 or *E. coli* strains expressing recombinant phosphatase was extracted by ultrasonication on ice bath (Branson Sonifier) for 2 min (pulse on 1 sec, off 2 sec). The sonicated samples were centrifuged at 13,400 rpm for 10 min at 4°C (Eppendorf) and the protein extract were used for zymogram and alkaline phosphatase activity analysis. Alkaline phosphatase activity was measured by using *para*-nitrophenol phosphate (*p*NPP) as substrate, which is dephosphorylated to *para*-nitrophenol (*p*NP) in Tris-HCl buffer (pH 9) as described earlier (Nilgiriwala *et al.*, 2008). The color was developed at the end of 30 min by the addition of 0.2 N NaOH and absorbance was measured spectrophotometrically at 405 nm. Liberation of inorganic phosphate (Pi) from organic substrate (phenolphthalein diphosphate or β-glycerophosphate) by extracellular and cell associated phosphatase activity was measured spectrophotometrically at 820 nm (Fiske and Subbarao, 1925). Briefly, 30% of test solution was mixed with 70% of reaction mix (prepared by mixing one part of 10% ascorbic acid with 6 parts of 0.42% ammonium molybdate) and incubated at 45°C for 20 min followed by measurement of absorbance at 820 nm.

Zymographic (in-gel activity) analysis

The ultrasonicated cells were centrifuged at 13,400 rpm for 10 min at 4°C, and the supernatant was used for whole cell alkaline phosphatase activity (Tris-HCl buffer pH 9.0) by zymographic analysis. Proteins were separated by 10% SDS-PAGE, gel washed with the Tris-HCl buffer pH 9.0 containing 1% Triton-X100, three times for 30 min each. The repeated washing of SDS-PAGE gel using Tris-HCl buffer with 1% Triton-X100 (pH 9.0) helps in removal of SDS, equilibration of protein gel towards alkaline pH range and Triton-X100 refolds the protein for optimal alkaline phosphatase activity. The in-gel activity was detected using NBT-BCIP (nitrobluetetrazolium chloride 5-bromo-3chloro-3-indolyl phosphate) (Roche Applied Sciences) as described earlier (Nilgiriwala *et al.*, 2008).

Membrane and cytosolic separation and activity distribution analysis

The sonicated cell suspension was centrifuged at 18,000 rpm for 10 min at 4°C to remove the unbroken cells and cell

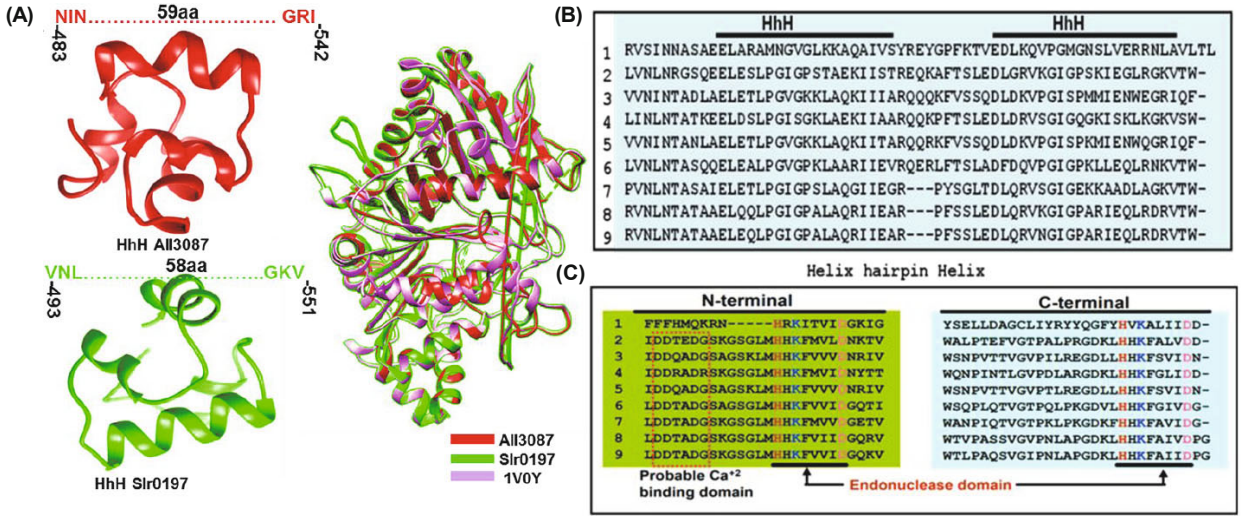


Fig. 1. Comparative bioinformatics analysis of hypothetical protein All3087 of *Anabaena* 7120 with a membrane bound competence protein Slr0197 of *Synechocystis* sp. PCC6803. (A) Structural alignment of All3087 and Slr0197 with phospholipaseD/Nuclease of *Streptomyces* sp. (PDB no. 1V0Y) and the helix-hairpin-helix of All3087 (red) and Slr0197 (green). (B) Helix -hairpin-helix (HhH) of All3087 and its bioinformatics comparison with various membrane proteins (1-9) and (C) Analysis of N- and C-terminal endonuclease HKD domains in competence proteins with probable Ca^{2+} binding (DDxxDG) motif in various bacteria (1-9). Different numbers (1-9) represent *Bacillus subtilis*, ComE (1); competence protein, Slr0197 (2); hypothetical protein of *Anabaena*, All3087(3); *Microcystis aeruginosa*, MAE_23940 (4); phospholipase D/Transphosphatidylase of *Anabaena variabilis*, Ava_4857 (5); phospholipase D/transphosphatidylase of *Acaryochloris marina*, AM1_0550 (6); conserved hypothetical protein of *Synechococcus* sp. PCC7002, SYNPCC7002_A1898 (7); phospholipase D/competence protein ComEA of *Synechococcus* sp. JA-3-3Ab CYA_2151(8); and phospholipase D/competence protein ComEA of *Synechococcus* sp. JA-2-3B'a, CYB_2756 (9).

debris. The supernatant subjected to ultracentrifugation using ultra swinging bucket rotor (TLS55) at maximum speed 55,000 rpm (259,000 × *g*) at 4°C for 1 h (Beckman Coulter's Optima MAX High-Capacity Ultracentrifuge). To avoid cytosolic contamination to the membrane fraction, the membrane pellet was washed thrice with cold saline, followed by ultracentrifugation. The translocation of ALP in membrane was assessed in terms of percent enzyme activities and in-gel activity analysis using cytosolic and membrane fractions.

Results and Discussion

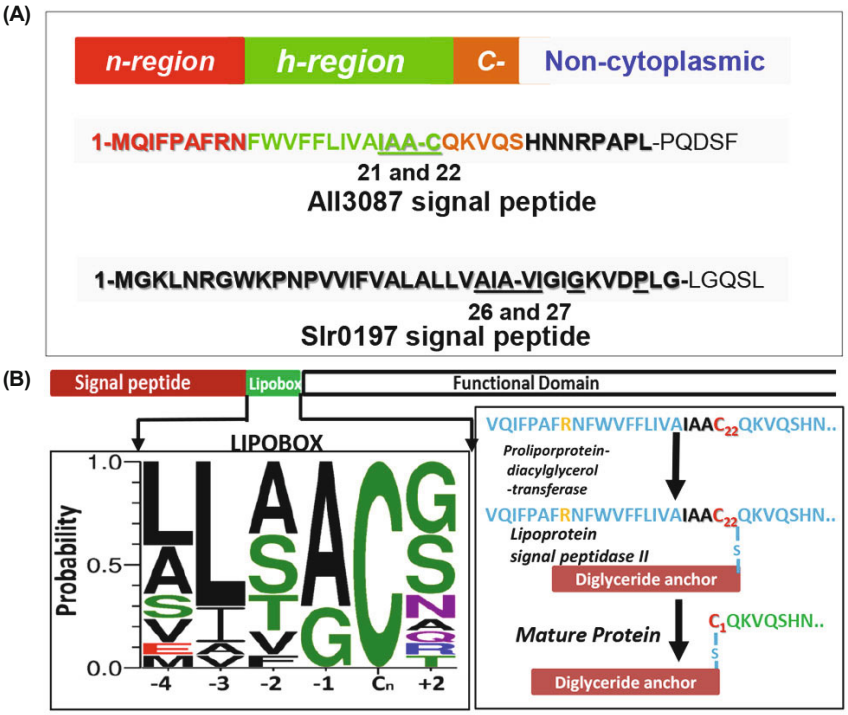
Bioinformatics analysis of All3087 with competence protein Slr0197

Slr0197 encodes for the competence protein, ComEA in *Synechocystis* sp. PCC 6803. The mutational analysis of Slr0197 was shown to be essential for the uptake of exogenously added DNA (Yoshihara *et al.*, 2001). All3087 exhibits 43% sequence identity corresponding to 96% query coverage (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Supplementary data Fig. S1) and nearly 100% structural homology with Slr0197 modelled by using Phyre server (Fig. 1A). The proteins also exhibited 100% and 95% expected precision and structural similarity with phospholipaseD/Nuclease of *Streptomyces* sp. (PDB no. 1V0Y) (Fig. 1A) and *Thermus thermophillus* competence ComEA proteins (PDB no. 2DUY), respectively. The homologs of All3087 and Slr0197 were found to be present across all cyanobacterial species, though in some cases they are also annotated as phospholipase D (http://genome.microbedb.jp/cyanobase).

In addition to the overall structural homology with competence proteins of other bacteria, hypothetical protein All3087 also exhibits several other features typical of competence proteins. These include (i) helix-hairpin-helix (HhH) motif at the C-terminus (483 to 542 amino acid) for DNA binding, similar to that in other competence proteins (Fig. 1A and B), (ii) EF hand (R^{134} to Q^{169} amino acids) typically in Ca^{2+}-binding proteins, (iii) a well conserved HKD endonuclease domain, both at the N-terminus (H^{190} to D^{197}) and the C-terminus (H^{395} to D^{401}) (Fig. 1C), and (iv) calcium binding site DDxxDG ($D^{177}$to G^{182}) motif at the N-terminus (Fig. 1C). Competence proteins are known to bind Ca^{2+} to develop competence enabling uptake of DNA by altering the membrane topography and by creating transient pores in the membrane to facilitate entry of DNA (Mandel and Higa, 1970; Cohen *et al.*, 1972). These characteristics features suggest that in all probability All3087, though designated as a hypothetical protein, presumably codes for the competence protein of *Anabaena* 7120.

Analysis of signal peptide of All3087

Analysis of the amino acid sequence of the putative competence protein All3087 indicated the presence of a distinct lipoprotein signal peptide (Fig. 2A) having a lipobox, with the conserved cysteine residue (C_{22}) (Fig. 2B, Supplementary data Table S1), known to involve in lipidation, with a high probability of cleavage as predicted using LipoPred algorithm (Supplementary data Fig. S2), indicating the existence of All3087 as a membrane protein. The N-terminal signal peptide of All3087 possesses a signal peptidase recognition/cleavage site 'IAA-C' (Fig. 2). The requisite criteria for lipoprotein signal peptide (LppSP) (Sutcliffe and Harrington, 2002) is consistent in case of All3087, which includes (a.)



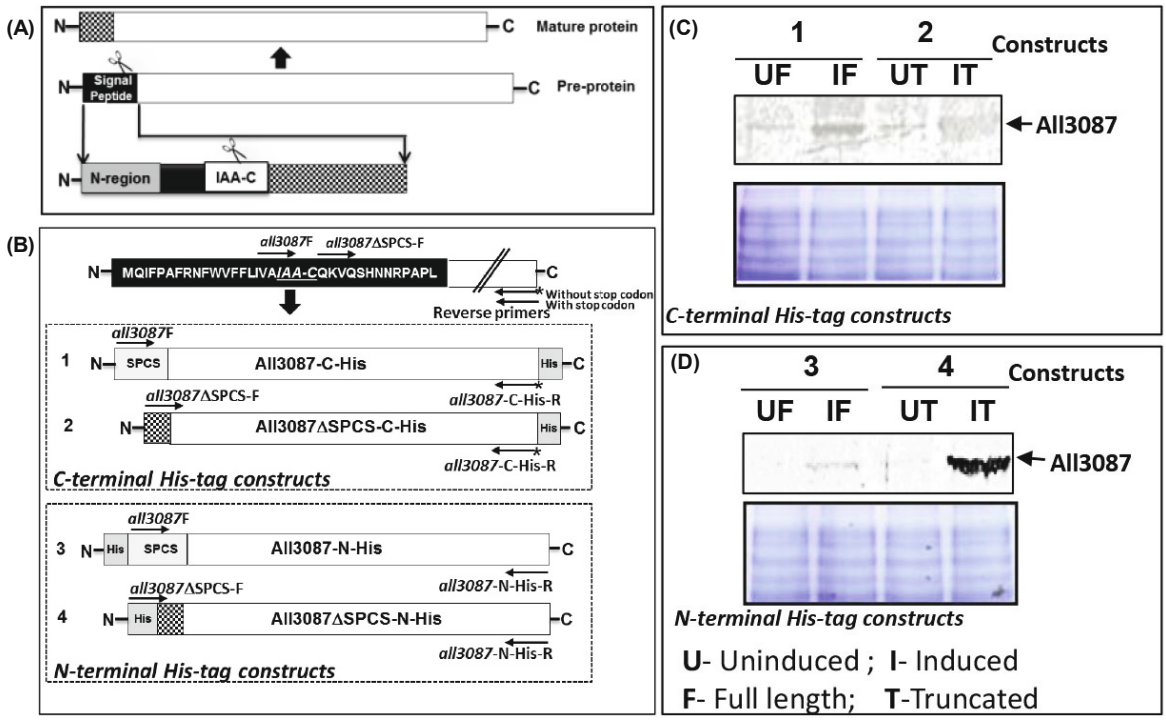


Fig. 2. Bioinformatics analysis of signal peptide of All3087. (A) *In silico* analysis of lipoprotein signal peptides of All3087 protein with probable signal peptide cleavage site (IAA-C). (B) Consensus sequence and probability of amino acids residues in lipobox based on reported lipoprotein signal peptides shown in supplementary data Table S1 and the site of lipoprotein signal peptide cleavage at conserved cysteine (C) residue to integrate into the membrane by lipidation.

Fig. 3. Mapping of signal peptidase II cleavage site in lipoprotein signal peptide of All3087. (A) Schematic representation of the N-terminal lipoprotein signal peptide as preprotein and mature All3087 protein. (B) With or without signal peptide cleavage site (SPCS) constructs design to map the SPCS using C-terminal (constructs, 1 and 2) and N-terminal His-tagging (constructs, 3 and 4). (C-D) Assessment of N-terminal signal peptide processing at signal peptidase cleavage site wherein (C) Western blot detected significantly high All3087 expression in construct (1) and SPCS truncated construct (2) C-terminal His-tag constructs due to intact C-terminal His-tag. The All3087 expression with or without SPCS in C-terminal His-tag constructs (C) served as a control for All3087 expression in N-terminal His-tag constructs (D). (D) Western blot showing extremely feeble Western band in N-terminal His-tag with SPCS (construct 3) due to signal processing (IF, induced full length) while the truncated SPCS (construct 4) showing abundant All3087 protein due to absence of signal peptidase cleavage site (IT, induced truncated) due to no processing of His-tag. Lower panel in C and D show the coomassie brilliant blue stained SDS-PAGE gel for equal protein loading.

conserved cysteine residue located between 15-35 amino acid [All3087: C_{22}], (b.) presence of at least one lysine or arginine in n-region around the 7th position of the LppSP [All3087:R_8], (c.) absence of charged amino acids in h-region of LppSP [No D, E,R, K in LppSP of All3087] (Fig. 2A and B).

Signal peptidase II recognition and cleavage site of signal peptide of All3087

For All3087 to function as a competence protein, it would need to be targeted to the membrane. The signal peptide with high N-terminal hydrophobicity allows its proper localization (Supplementary data Fig. S3). The signal peptidase II (*alr4577*) responsible for cleavage of LppSP in *Anabaena* 7120 is significantly conserved and showed similarity with other bacterial SPaseII including *Escherichia coli* (Supplementary data Fig. S4). The processing of the signal peptide of All3087 protein at the signal peptide cleavage site (SPCS) was confirmed in *E. coli* using His-tag constructs at either the N-terminal or at the C-terminal of the protein with or without the region up to SPCS. The conversion of the preprotein to the mature protein as predicted after the cleavage of the signal peptide at IAA^{21}-C^{22} is shown schematically in Fig. 3A.

The positions of primers used for overexpression of All3087 or (All3087ΔSPCS) with His-tag at C/N-terminal and the different proteins generated are shown in Fig. 3B. All the four proteins, i.e. (1) All3087-C-His, (2) All3087ΔSPCS-C-His, (3) All3087-N-His, and (4) All3087ΔSPCS-N-His (Fig. 3B) were overexpressed in *E. coli* BL21(DE3)(p*lysS*) upon induction with 1 mM IPTG and immunodetected using anti-His antibody (Fig. 3C and D). The protein expression for C-terminal His-tagged constructs with or without SPCS i.e. All3087-C-His and All3087ΔSPCS-C-His served as the control of protein expression with or without SPCS for N-terminal constructs (Fig. 3C). The protein band detected in constructs (All3087-C-His and All3087ΔSPCS-C-His) were found to be similar in content due to no cleavage of C-terminal His-tag in both the constructs and truncation of SPCS do not affects the protein expression significantly. N-terminal His-tagged protein band detected by Western blotting was higher for All3087ΔSPCS-N-His, compared to extremely low in case of All3087-N-His (Fig. 3D). The extremely lower band intensity in N-terminal His-tag construct with SPCS (All3087-N-His) (Fig. 3D; lane IF) compared to SPCS truncated (All3087ΔSPCS-N-His) construct (Fig. 3D; lane IT) proved the processing of N-terminal His-tagged signal peptide at signal peptide cleavage site (Fig. 3D).

Functional characterization of Alr2234 and Alr4976

The putative alkaline phosphatases, *alr2234* (*phoD1*) and *alr4976* (*phoD2*) were heterologously overexpressed in *E. coli* BL21(DE3) to ensure their alkaline phosphatase activities. Alr2234 was detected as a ~59.5 kDa protein after 3 h of induction at 37°C (Fig. 4A). The expression of Alr4976 is detected (~59 kDa) by inducing at 18°C for 16 h (Fig. 4B). The IPTG induced *E. coli* strains expressing Alr2234 and Alr4976 showed enhanced alkaline phosphatase activity compared to their corresponding uninduced controls (Fig. 4C). In-gel phosphatase activity of *Anabaena* whole cell extracts on 10% SDS-PAGE zymogram showed triplet feeble activity bands below conspicuous ~59 kDa ALPs bands, which presumably corresponds to Alr2234 and Alr4976 (Fig. 4D) because the *Anabaena 7120* genome does not possess any other inducible ALP (upon inorganic phosphate limitation) of this molecular weight analyzed by genome wide search for alkaline phosphatase gene. However, the possibilities of unknown protein possessing ALP activity with the aforesaid molecular weights cannot be ruled out without mutational analysis. The detection of the active ALP bands as close triplets suggested possible processing of these proteins which was further confirmed by the presence of twin-arginine signal peptide in both ALPs indicates the possibility of posttranslational processing of the signal peptide.

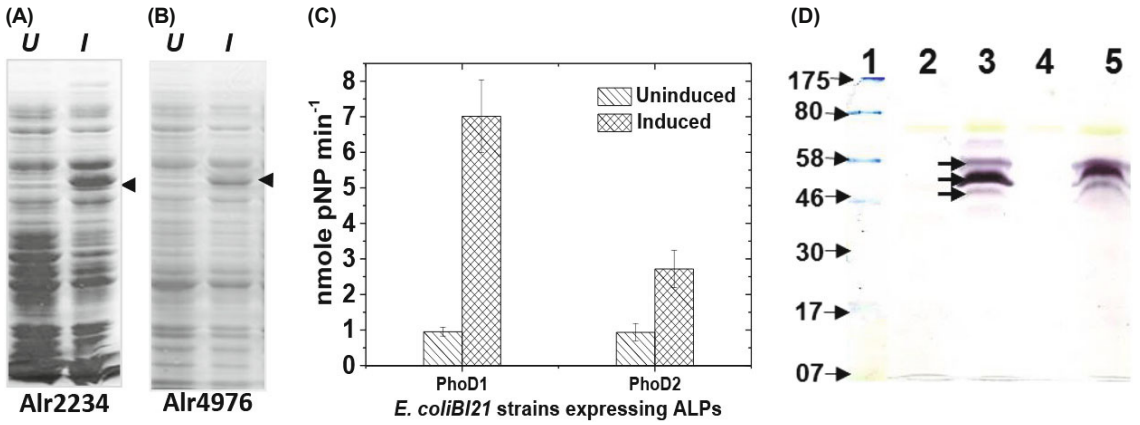


Fig. 4. Whole cell alkaline phosphatase activity in *Anabaena* 7120 and confirmation of alkaline phosphatase activity encoded by putative ALP genes *alr2234* and *alr4976*. (A-B) Coomassie brilliant blue stained SDS-PAGE protein gels showing heterologous overexpression of Alr2234 (A) and Alr4976 (B) where uninduced and induced lanes are marked as U and I. (C) Alkaline phosphatase activity of *E. coli* cells expressing ALPs compared corresponding uninduced control cells. (D) whole cell extract of *Anabaena* 7120 showing total alkaline phosphatase activity under inorganic phosphate replete or deplete conditions. Various lanes show pre-stained protein marker (lane 1), ALP zymographic bands under presence (+) and absence (-) combined nitrogen and inorganic phosphate, N^+P^+ (lane 2), N^+P^- (lane 3), N^-P^+ (lane 4), and N^-P^- (lane 5).

Transport of alkaline phosphatase to membrane and extracellular milieu using bipartite Tat signal peptide

Whole genome analysis of *Anabaena* 7120 showed that the twin arginine translocase dependent protein transport is less frequent than Sec-dependent secretion system (data not shown). Both Alr2234 and Alr4976 alkaline phosphatases contain twin arginines translocase (Tat) signals with high N-terminal hydrophobicity (Supplementary data Fig. S5) with distinct n, h, and c regions (Fig. 5A) (Palmer and Berks, 2012). Twin arginines signal peptides are critical for export of the corresponding protein to the membrane and thylakoid membrane (Stanley *et al.*, 2000). Two signal peptidase I (SPase I) sites have been predicted for both Alr2234 and Alr4976 ALPs i.e. **AIA**-SQ and **AIA**-KS for Alr2234 and at VV**A**-SQ and VM**A**-QT for Alr4976 (Fig. 5A). The substitution of 'A' by 'V' at Tat SPase site has been shown to enhance processing by thylakoid processing peptidase (TPP) in chloroplastic delta pH dependent transport (Shackleton and Robinson, 1991). Such possibility in case of cyanobacterial Alr4976 may also prevail due to presence of distinct thylakoid membrane in *Anabaena* 7120.

The limitation inorganic phosphate induced both Alr2234 and Alr4976 ALPs to dephosphorylate the extracellular organic phosphate to secure the continuous minimal supply of inorganic phosphate for cell survival (Fig. 5B). Both ALPs showed the pNPP dephosphorylation and zymographic phosphatase activity bands even in the cytosolic fraction; suggesting the folding of these ALPs in the cytoplasm before being transported to the membrane. The transport of folded active protein is the characteristic feature of Tat-dependent protein transport pathway (Fig. 5B and C). The cytosolic ALP activity bands prove that the nascent protein is completely conformed and active before translocation (Fig. 5B) most probably using the GroES-EL molecular chaperon, which is a general molecular chaperon in *Anabaena* 7120 (Chaurasia *et al.*, 2013; Rajaram *et al.*, 2014) or any specific molecular chaperon as reported earlier (Hitchcock *et al.*, 2010).

To study the protein transport through bipartite Tat-signal peptide in *Anabaena* 7120, the ALP activity bands were visualized in exclusive membrane fractions to confirm their targeting to the membrane. Six different bands observed in cytosolic fractions were corresponds to pre-protein (p), intermediates (i), and intermediate mature (ic1 and ic2) for both Alr2234 and Alr4976, while the membrane fraction showed only 3 bands intermediate (p), mature (im1), and mature (im2) for Alr2234 and Alr4976, respectively. The membrane transported intermediates (im1 and im2) would further cleaved by membrane bound SPaseI (Fig. 5C). The reduction in molecular size due to SPaseI processing and their match of size with membrane fraction activity bands (im) with intermediate mature (ic) cytosolic activity bands suggest the first cleavage occur at cytosolic side of membrane and then the protein (im) was flipped into the periplasmic side where possibly the second cleavage occur when the delta pH generates upon cyanobacterial growth. The current findings and analysis opens an extension for future studies on mutational analysis of All3087 and ALPs and their complementation with wild type or point mutated alleles at the key residues to investigate membrane topographic change by All3087 due to lipoprotein signal peptide mediated targe-

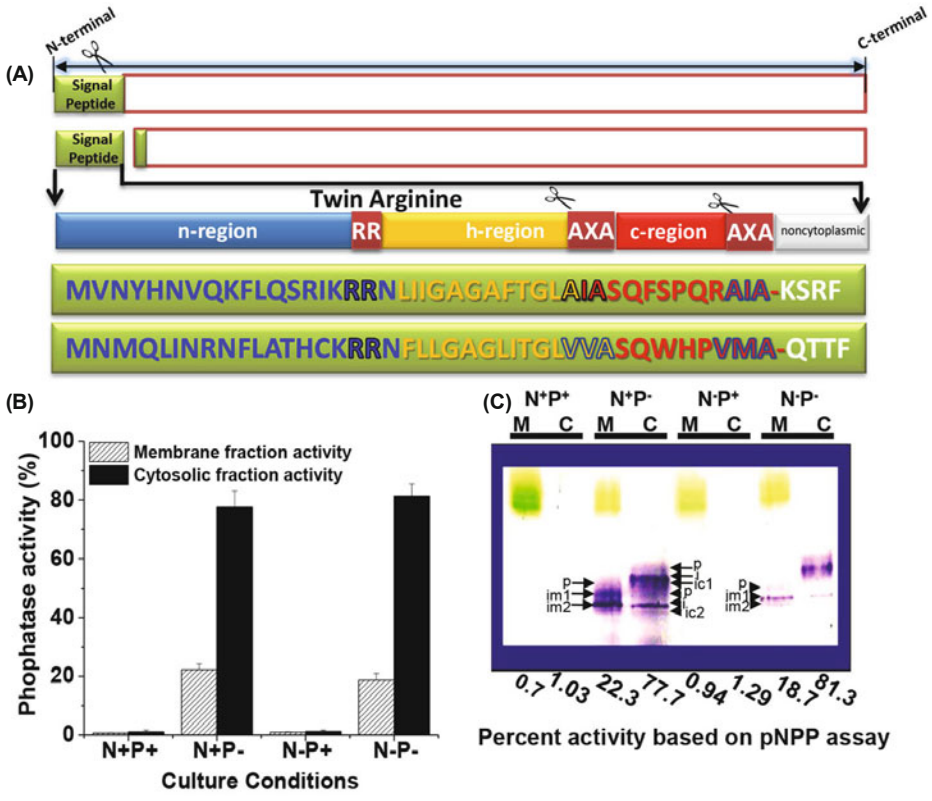


Fig. 5. Novel sequence architecture of Tat-dependent signal peptides and ALPs translocation under replete or deplete conditions of fixed nitrogen and inorganic phosphate. (A) Signal peptides of PhoD1 (Alr2234) and PhoD2 (Alr4976) showing twin arginine residue (RR) and bipartite architecture with two signal peptidase cleavage sites for PhoD1 (AIA-SQ, AIA-KS) and for PhoD2 (VVA-SQ, VMA-QT). (B) Percent alkaline phosphatase activity distribution in cytosolic and membrane fraction. (C) Zymographic analysis of membrane and cytosolic fractions where the cytosolic fraction showed 6 bands [(p, i, ic1 for PhoD1) and (p, i, ic2 for PhoD2)] of different sizes contributed by both the ALPs and only 3 bands of which two distinct intermediate (im1) and mature (im2) proteins in the membrane fraction indicating their transport to the membrane and then to extracellular milieu upon successive processing.

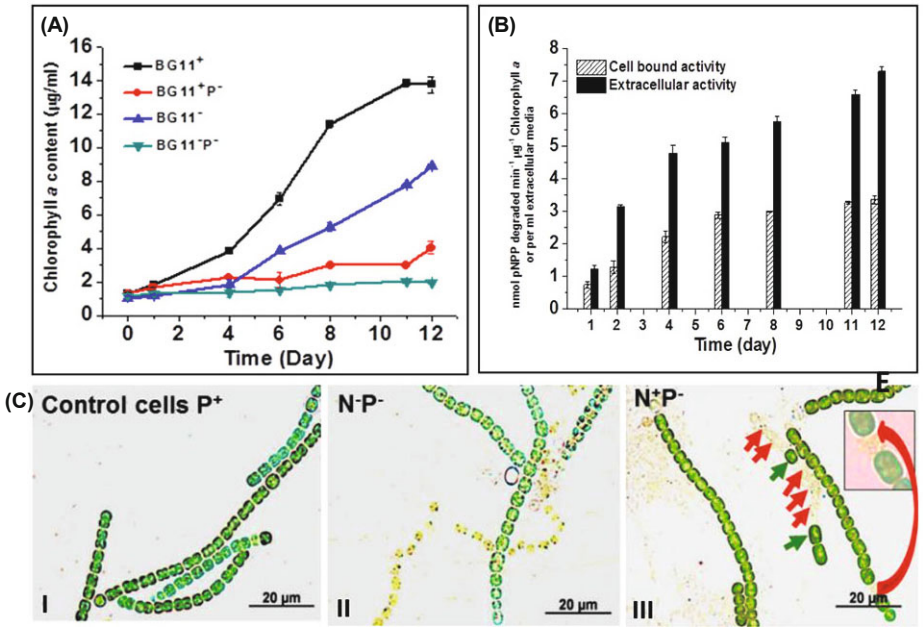


Fig. 6. Physiological relevance of bipartite signal peptide in ALP. (A) *Anabaena* 7120 growth under fixed nitrogen and inorganic phosphate replete or deplete conditions and pH change from 7.2 to 9.0 ± 0.5 at stationary phase. (B) ALP activity with the growth phase. (C) Cellular phenotype of cells grown under presence (+) or absence (-) of fixed nitrogen (N) or inorganic phosphate (P). (I) control cells grown N^-P^+; (II) cells grown in N^-P^- showing significant disintegration of cellular boundaries (cell membrane) and (III) cells grown in media N^+P^- conditions.

ting to the membrane, calcium and DNA binding to All3087 protein, successive transport of ALPs using bipartite signal peptides in wild type *versus* mutants phenotypic analysis.

Physiological relevance of bipartite signal peptide in ALP

The bipartite signal cleavage of precursor (p) generating intermediates (i) and mature (m) processed products have been reported in case of ΔpH dependent transport in thylakoid membrane and HSP70 in mitochondria (Ungermann *et al.*, 1994; Gruhler *et al.*, 1997, Wexler *et al.*, 1998; Chen *et al.*, 1999). The phosphatase activity of *Anabaena* 7120 remained cell-bound during active growth but during late log-phase to stationary phase the cell-bound activity remained static and the extracellular phosphatase activity enhanced not only due to inorganic phosphate limitation but also due to second cleavage of signal peptide and successive transport of ALP to extracellular media (Fig. 6A and B) when the ΔpH generates due to growth of *Anabaena* 7120 (Fig. 6B). Interestingly, the disintegration of *Anabaena* cell boundary creating "empty cells with no visible chlorophyll" was observed under phosphate depleted media irrespective of nitrogen status (Fig. 6C I-III). This observation may be implicated as 'sacrificing strategy of metabolically unfit cells to save community' by releasing their cellular organic phosphate molecules for dephosphorylation using ALPs transported extracellularly to generate inorganic phosphate for the metabolically active alive cells.

Conclusion

In silico analysis of the competence protein All3087 revealed that it possesses N-terminal lipoprotein membrane-targeting signal peptide with conserved cysteine residue. The signal processing at signal peptidase cleavage site was proven experimentally at IAA-C. The heterologous overexpression of *alr2234* and *alr4976* confirmed the alkaline phosphatase activities of these putative ALPs. The inorganic phosphate (Pi) repressible alkaline phosphatases, Alr2234 and Alr4976 are transported to the membrane by Tat dependent bipartite signal peptide. The higher ALP activity in the cytosolic fraction compared to the membrane fraction suggests complete folding of the ALPs prior to targeting to the membrane. It is noteworthy, that by tagging the non-native reporter proteins using Tat signal sequences provides information about targeting only, while by choosing the native ALPs with their signal sequence provides both the membrane transport and the state of transported protein (folded/unfolded), reported for the first time in cyanobacteria.

Acknowledgements

The work reported here was supported by Department of Biotechnology (DBT), New Delhi and Department of Atomic Energy (DAE), India. A.K.C. acknowledges receipt of a DBT-Postdoctoral Research Fellowship. A.K.C. acknowledges Dr. Hema Rajaram for discussion, reading and editing during manuscript preparation.

References

Barbrook, A.C., Packer, J.C., and Howe, C.J. 1993. Components of the protein translocation machinery in the thermophilic cyanobacterium *Phormidium laminosum*. *Biochem. Biophys. Res. Commun.* **197**, 874–877.

Blobel, G. and Dobberstein, B. 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J. Cell Biol.* **67**, 835–851.

Bolhuis, A. 2002. Protein transport in the halophilic archaeon *Halobacterium* sp. Nrc-1: A major role for the twin-arginine translocation pathway? *Microbiology* **148**, 3335–3346.

Castenholz, R. 1988. Culturing methods for cyanobacteria. *Method Enzymol.* **167**, 68–93.

Chaurasia, A.K., Adhya, T.K., and Apte, S.K. 2013. Engineering bacteria for bioremediation of persistent organochlorine pesticide lindane (gamma-hexachlorocyclohexane). *Bioresour. Technol.* **149**, 439–445.

Chaurasia, A.K. and Apte, S.K. 2011. Improved eco-friendly recombinant *Anabaena* sp. strain PCC7120 with enhanced nitrogen biofertilizer potential. *Appl. Environ. Microbiol.* **77**, 395–399.

Chen, X., Van Valkenburgh, C., Fang, H., and Green, N. 1999. Signal peptides having standard and nonstandard cleavage sites can be processed by imp1p of the mitochondrial inner membrane protease. *J. Biol. Chem.* **274**, 37750–37754.

Cohen, S.N., Chang, A.C., and Hsu, L. 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA* **69**, 2110–2114.

Dalbey, R.E., Chen, M., Jiang, F., and Samuelson, J.C. 2000. Understanding the insertion of transporters and other membrane proteins. *Curr. Opin. Cell Biol.* **12**, 435–442.

Dalbey, R.E. and Kuhn, A. 2000. Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu. Rev. Cell Dev. Biol.* **16**, 51–87.

Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. 2007. Locating proteins in the cell using Targetp, SignalP and related tools. *Nat. Protoc.* **2**, 953–971.

Fiske, C.H. and Subbarao, Y. 1925. The colorimetric determination of phosphorus. *J. Biol. Chem.* **66**, 375–400.

Gruhler, A., Arnold, I., Seytter, T., Guiard, B., Schwarz, E., Neupert, W., and Stuart, R.A. 1997. N-terminal hydrophobic sorting signals of preproteins confer mitochondrial hsp70 independence for import into mitochondria. *J. Biol. Chem.* **272**, 17410–17415.

Hall, T.A. 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for window 95/98/nt. *Nucleic Acids Symp. Ser.* **41**, 95–98.

Hartl, F.U., Lecker, S., Schiebel, E., Hendrick, J.P., and Wickner, W. 1990. The binding cascade of *secB* to seca to *secY/E* mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* **63**, 269–279.

Hitchcock, A., Hall, S.J., Myers, J.D., Mulholland, F., Jones, M.A., and Kelly, D.J. 2010. Roles of the twin-arginine translocase and associated chaperones in the biogenesis of the electron transport chains of the human pathogen *Campylobacter jejuni*. *Microbiology* **156**, 2994–3010.

Joshi, M.V., Mann, S.G., Antelmann, H., Widdick, D.A., Fyans, J.K., Chandra, G., Hutchings, M.I., Toth, I., Hecker, M., Loria, R., *et al.* 2010. The twin arginine protein transport pathway exports multiple virulence proteins in the plant pathogen *Streptomyces scabies*. *Mol. Microbiol.* **77**, 252–271.

Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., *et al.* 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res.* **8**, 205-213; 227–253.

Kelley, L.A. and Sternberg, M.J. 2009. Protein structure prediction on the web: A case study using the phyre server. *Nat. Protoc.* **4**, 363–371.

Kushnareva, Y.E., Polster, B.M., Sokolove, P.M., Kinnally, K.W., and Fiskum, G. 2001. Mitochondrial precursor signal peptide induces a unique permeability transition and release of cytochrome *c* from liver and brain mitochondria. *Arch. Biochem. Biophys.* **386**, 251–260.

Lam, S.L., Kirby, S., and Schryvers, A.B. 2003. Foreign signal peptides can constitute a barrier to functional expression of periplasmic proteins in *Haemophilus influenzae*. *Microbiology* **149**, 3155–3164.

Ling Lin, F., Zi Rong, X., Wei Fen, L., Jiang Bing, S., Ping, L., and Chun Xia, H. 2007. Protein secretion pathways in *Bacillus subtilis*: implication for optimization of heterologous protein secretion. *Biotechnol. Adv.* **25**, 1–12.

Mackinney, G. 1941. Absorption of light by chlorophyll solutions. *J. Biol. Chem.* **140**, 315–322.

Mandel, M. and Higa, A. 1970. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**, 159–162.

Mariscal, V., Herrero, A., and Flores, E. 2007. Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium. *Mol. Microbiol.* **65**, 1139–1145.

Nakai, M., Nohara, T., Sugita, D., and Endo, T. 1994. Identification and characterization of the sec-a protein homologue in the cyanobacterium *Synechococcus* PCC7942. *Biochem. Biophys. Res. Commun.* **200**, 844–851.

Nilgiriwala, K.S., Alahari, A., Rao, A.S., and Apte, S.K. 2008. Cloning and overexpression of alkaline phosphatase *phok* from *Sphingomonas* sp. strain BSAR-1 for bioprecipitation of uranium from alkaline solutions. *Appl. Environ. Microbiol.* **74**, 5516–5523.

Palmer, T. and Berks, B.C. 2012. The twin-arginine translocation (Tat) protein export pathway. *Nat. Rev. Microbiol.* **10**, 483–496.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. 2004. UCSF chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

Pugsley, A.P., d'Enfert, C., Reyss, I., and Kornacker, M.G. 1990. Genetics of extracellular protein secretion by Gram-negative bacteria. *Annu. Rev. Genet.* **24**, 67–90.

Raghavan, P.S., Rajaram, H., and Apte, S.K. 2013. N-terminal processing of membrane-targeted mnsod and formation of multiple active superoxide dismutase dimers in the nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC7120. *FEBS J.* **280**, 4827–4838.

Rajaram, H., Chaurasia, A.K., and Apte, S.K. 2014. Cyanobacterial heat-shock response: Role and regulation of molecular chaperones. *Microbiology* **160**, 647–658.

Ravn, P., Arnau, J., Madsen, S.M., Vrang, A., and Israelsen, H. 2003. Optimization of signal peptide sp310 for heterologous protein production in *Lactococcus lactis*. *Microbiology* **149**, 2193–2201.

Rezwan, M., Grau, T., Tschumi, A., and Sander, P. 2007. Lipoprotein synthesis in mycobacteria. *Microbiology* **153**, 652–658.

Shackleton, J.B. and Robinson, C. 1991. Transport of proteins into chloroplasts. The thylakoidal processing peptidase is a signal-type peptidase with stringent substrate requirements at the -3 and -1 positions. *J. Biol. Chem.* **266**, 12152–12156.

Simonen, M. and Palva, I. 1993. Protein secretion in *Bacillus* species. *Microbiol. Rev.* **57**, 109–137.

Spence, E., Sarcina, M., Ray, N., Moller, S.G., Mullineaux, C.W., and Robinson, C. 2003. Membrane-specific targeting of green fluorescent protein by the tat pathway in the cyanobacterium *Synechocystis* PCC6803. *Mol. Microbiol.* **48**, 1481–1489.

Stanley, N.R., Palmer, T., and Berks, B.C. 2000. The twin arginine consensus motif of tat signal peptides is involved in sec-independent protein targeting in *Escherichia coli*. *J. Biol. Chem.* **275**, 11591–11596.

Sutcliffe, I.C. and Harrington, D.J. 2002. Pattern searches for identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology* **148**, 2065–2077.

Tjalsma, H., Antelmann, H., Jongbloed, J.D., Braun, P.G., Darmon, E., Dorenbos, R., Dubois, J.Y., Westers, H., Zanen, G., Quax, W.J., *et al.* 2004. Proteomics of protein secretion by *Bacillus subtilis*: Separating the "secrets" of the secretome. *Microbiol. Mol. Biol. Rev.* **68**, 207–233.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S., and van Dijl, J.M. 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* **64**, 515–547.

Tjalsma, H., Zanen, G., Venema, G., Bron, S., and van Dijl, J.M. 1999.

The potential active site of the lipoprotein-specific (type ii) signal peptidase of *Bacillus subtilis*. *J. Biol. Chem.* **274**, 28191–28197.

Ungermann, C., Neupert, W., and Cyr, D.M. 1994. The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria. *Science* **266**, 1250–1253.

von Heijne, G. 1990. The signal peptide. *J. Membr. Biol.* **115**, 195–201.

Wexler, M., Bogsch, E.G., Klosgen, R.B., Palmer, T., Robinson, C., and Berks, B.C. 1998. Targeting signals for a bacterial sec-independent export system direct plant thylakoid import by the delta pH pathway. *FEBS Lett.* **431**, 339–342.

Wickner, W. and Schekman, R. 2005. Protein translocation across biological membranes. *Science* **310**, 1452–1456.

Widdick, D.A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschröder, M., and Palmer, T. 2006. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci. USA* **103**, 17927–17932.

Yoshihara, S., Geng, X., Okamoto, S., Yura, K., Murata, T., Go, M., Ohmori, M., and Ikeuchi, M. 2001. Mutational analysis of genes involved in pilus structure, motility and transformation competency in the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* **42**, 63–73.

Zanen, G., Houben, E.N., Meima, R., Tjalsma, H., Jongbloed, J.D., Westers, H., Oudega, B., Luirink, J., van Dijl, J.M., and Quax, W.J. 2005. Signal peptide hydrophobicity is critical for early stages in protein export by *Bacillus subtilis*. *FEBS J.* **272**, 4617–4630.

Zhang, X.P. and Glaser, E. 2002. Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci.* **7**, 14–21.