Staphylococcus aureus sortase A cyclization and evaluation of enzymatic activity *in vitro*

Dmitry Zhulenkovs*, Ainars Leonchiks

Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga LV-1067, Latvia

*Corresponding author, E-mail: dmitry@biomed.lu.lv

Abstract

Protein cyclization has been shown to increase the stability of molecule structure by reducing the number of possible conformation variations. Cyclization of an enzyme can significantly enlarge the spectrum of its application in specific conditions, including high temperature, presence of a chaotropic agent, and denaturing environment. In this study the cyclic form of *Staphylococcus aureus* sortase A enzyme (srtA) was obtained by introducing a peptide bond between amino acids at N- and C-termini. The catalytic activity of purified cyclic srtA was tested *in vitro* by protein-protein ligation reaction. In contrast to the linear srtA, the cyclic form of the enzyme was active in presence of chaotropic agent. This finding suggests the potential use of srtA enzyme in protein engineering techniques.

Key words: cyclization, sortase A, *Staphylococcus aureus,* transpeptidation, trans-splicing.
Abbreviations: eGFP. Green Fluorescence protein; GST, glutahione-S-transferase; NMR, nuclear magnetic resonance; srtA, sortase A.

Introduction

Nuclear magnetic resonance (NMR) studies of protein higher structures have shown that protein cyclization can increase the stability of the molecules. Stabilized proteins can remain folded for a longer period of time when in presence of chaotropic agents, such as urea or guanidine hydrochloride at high concentrations. Also, refolding of the denatured cyclic protein molecule can be facilitated, since the number of possible protein folding variations is reduced, in comparison to the linear forms (Iwai et al. 2001).

Staphylococcus aureus sortase A (srtA) is an enzyme naturally located in the cell wall of grampositive bacteria, where it catalyses transpeptidation reactions without the need of ATP. The enzyme ensures specific covalent linkage of surface proteins to pentaglycine bridges of the cell walls, and thus is crucial for bacterial pathogenesis (Schneewind et al. 1992). The ability to infect host organisms is thus greatly reduced for the bacteria with mutations in the *srtA* gene (Mazmanian et al. 1999). The length of srtA is 206 amino acid residues. It consists of only one catalytical domain, which is composed of 59-206 amino acid residues. The recombinant catalytical domain itself exhibits activity similar to the full length protein (Ilangovan et al. 2001). In the new era of postgenomic research, srtA enzyme is a perspective tool in protein engineering applications due to its ability to form peptide bonds in the absence of ATP (Mao et al. 2004). The generation of the cyclic form of srtA can significantly stabilize the enzyme molecule providing the broad spectrum of its application in protein engineering.

Synechocystis sp. PCC6803 intein domains have been used to ensure srtA cyclization (Evans et al. 2000). In the living cells inteins naturally take part in protein post-translational splicing. These domains have been successfully applied to perform *in vivo* cyclization of target recombinant protein objects (Evans et al. 2000). The intein domains are being expressed as a fusion with the target protein (extein), flanking its amino acid sequence in a precursor molecule. During the process of intramolecular trans-splicing, inteins are split out of the precursor after the molecule synthesis is completed in the host cell, leaving the extein termini joint beside a flexible amino acid linker, consisting of nine amino acid residues with an 25 Å length (Williams et al. 2002). Thus, the cyclization of recombinant proteins can be performed *in vivo* and the modified protein can be purified from the host cells.

The aim of this study was to assay the enzymatic activity of two cyclic srtA forms containing linkers with 9 and 13 aa respectively. Two protein substrates for sortase-catalysed protein ligation reactions were used in order to evaluate the enzymatic activity of cyclic srtA in both denaturing and native conditions. The substrates were recombinant *Schistosoma japonicum* glutathione-S-transferase (GST) (Beckett, Hayes 1993) carrying the specific LPETG sequence at the C-terminus, and recombinant *Aequorea victoria* Green Fluorescence protein (eGFP) (Cubitt et al. 1995), possessing a glycine residue at the N-terminus. Both the LPETG motif and N-terminal glycine residue are crucial for transpeptidation reaction, since the catalyst specifically recognizes the amino acid LPETG site and acts as a hydrolase, cleaving the first substrate between T and

G residues forming acyl-enzyme, which in turn is attacked by the second substrate with a N-terminal glycine residue, thus forming a new bond between T of the first substrate and G of the second (Ilangovan et al. 2001).

Materials and methods

Synthesis of pET23a-lsrtA vector for linear srtA expression

A sortase A gene fragment corresponding to 60-204 amino acid residues, was amplified using MSSA476 *S. aureus* cell genomic DNA. A colony of bacterial cells, grown on agar plate, was suspended in water and heated at 95 °C for 10 min to lyse the bacteria. Then lsrtA_dir (5'-GATA TACATATGCAAGCTAAACCTCAAATTCCG-3'; *Nde*I) and lsrtA_rev (5'-GTGGTGCTCGAGTTTGACTTCTGTA GCTACAAAGAT-3'; *Xho*I) primers were used for srtA amplification. The amplified fragment was inserted between *Nde*I and *Xho*I sites of the pET23a vector (Novagen, USA). For expression the plasmid was transformed into *E. coli* BL21(DE3) host cells and the culture was grown in LB medium at 37 °C. Protein expression was induced by adding 1 mM IPTG, carried out at 30 °C for 3 h. The cells were suspended in 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.1% Tween 20 and 2 mM β-mercaptoethanol and lysed by sonication. The supernatant of the centrifuged lysate was applied to a 2-mL Ni-NTA (Qiagen) column. Then, impurities were removed by washing the column with 20 mL of 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.1% Tween 20, 2 mM β-mercaptoethanol and 20 mM imidazole. Linear sortase A was eluted with 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.1% Tween 20, 2 mM β-mercaptoethanol and 300 mM imidazole. Pooled fractions were dialyzed against 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM DTT and the protein was stored at 4 °C.

Synthesis of pNW1120-csrtA+9aa and pNW1120-csrtA+13aa vectors and cyclic srtA enzymes expression

In order to obtain genetic constructs for cyclic srtA isoform expression with 9 (pNW1120-csrtA+9aa) and 13 (pNW1120-csrtA+13aa) amino acid residues long linker between N- and C-termini, the srtA gene was amplified using pET23a-lsrtA vector as a template with the following primer pairs: csrtA_9_dir (5'-TAATGAATTCC AAGCTAAACCTCAA-3'; *Eco*RI) and csrtA_9_rev (5'-AATTACGCGTTTTGACTTCTGTAGC-3'; *Mlu*I); csrtA_13_dir (5'-TAATGAATTCGGATCTGGACAAGCT AAACCTCAA-3'; EcoRI) and csrtA_13_rev (5'-AATTACG CGTTCCTTTGACTTCTGTAGC-3', *Mlu*I), respectively. PCR fragments were inserted between the *Eco*RI and *Mlu*I sites of pNW1120 vector (Williams et al. 2002). Both constructs were designed for direct expression of Int_C-srtA-Int_N fusion using the thermoinducible bacteriophage λ promoter. The plasmids were transformed into *E. coli* BL21 host cells and the culture was grown in LB medium at 37 °C. Protein expression was induced by heating the cultures at 42 °C for 15 min in a water bath. The cells were suspended in 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.1% Tween 20 and 2 mM β-mercaptoethanol and lysed by sonication. The supernatant of the centrifuged lysate was applied to a Sephadex G-25 (GE Healthcare, Sweden) column, with buffer exchange against 20 mM ammonium acetate at pH 5.0. Then the eluted protein solution was applied to a Sepharose Fast Flow SP (GE Healthcare) column and the protein was finally eluted by an increasing NaCl gradient in 20 mM ammonium acetate, pH 5.0. Pooled fractions were applied to a Sephadex G-25 (GE Healthcare) column and buffer exchange against 20 mM Tris-HCl, pH 8.0. The solution was applied to a Source Q15 (GE Healthcare) column and protein was eluted by an increasing NaCl gradient in 20 mM Tris-HCl, pH 8.0. The purity of cyclic sortase in the pooled fraction was enriched by gel filtration on a Superdex 75 (GE Healthcare) column, simultaneously exchanging the storage buffer to 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM DTT.

Expression of inteins by pNW1118 vector

In order to identify the protein band corresponding to covalently-joined mini-intein domains, and to obtain a positive control for the splicing reaction, the pNW1118 (Williams et al. 2002) vector was used for intein expression. The plasmid directed simultaneous expression of both *Synechocystis* sp. PCC6803 DnaB mini-intein domains from one bicistronic mRNA. Downstream of the T7 promoter in pNW1118 there was a strong *E. coli* ribosome-binding site and ATG start codon followed by the Int_C-encoding sequence, terminated by a TAA stop codon. This was followed by a further RBS and start codon preceding the int_N gene, terminated with a stop codon.

To observe the possibility of mini-inteins forming a dimer as a result of posttranslational transpeptidation, the pNW1118 vector was transformed into *E. coli* BL21(DE3) host cells and the culture was grown in LB medium at 37 °C. Protein expression was induced by adding 1 mM IPTG, and was carried out at 37 °C for 2 h. After expression, an aliquot of the cell suspension was taken and analysed by SDS-PAGE.

Synthesis of pGEX-5X-1-LPETG vector and GST-LPETG expression

The vector for expression of GST protein with LPETG motif near C-terminus was synthesized by insertion of a DNA adapter, composed of GST_LPETG_dir (5'-GATCCTGCCGGAAACCGGTC-3') and GST_LPETG_rev (5'-TCGAGACCGGTTTCCGGCAG-3') primers. The primers were mixed at 5 μM final concentration in 1× Y^+/Tango™ buffer (Fermentas), heated at 95 °C for 5 min. Then, hybridization was carried out at 50 °C for 15 min and the mixture was stored at room temperature. The prepared adapter was inserted between *Bam*HI and *Xho*I sites of the pGEX-5X-1 vector (GE Healthcare). The

plasmid was transformed into *E. coli* BL21 host cells and the cell culture was grown in LB medium at 37 °C. Protein expression was induced by adding 1 mM IPTG, and was carried out at 30 °C for 3 h. The cells were suspended in 1× phosphate buffer saline and lysed by sonication. The supernatant then was applied to a 2-mL Glutathione Sepharose™ 4 Fast Flow (GE Healthcare) column. The column was washed with 20 mL of 50 mM Tris-HCl, pH 8.0, 150 mM NaCl and the protein was eluted with 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 15 mM reduced glutathione. Pooled fractions were dialyzed against 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM DTT.

pNW1120-3C-eGFP vector synthesis and 3C-eGFP protein expression

The vector for eGFP protein expression, carrying the N-terminal 3C protease LEVLFQGP site, was constructed as follows: a DNA adapter, composed of pNW1120_3C_dir (5'-TATGCTGGAGGTGCTGTT CCAGGGTCCAAC-3') and pNW1120_3C_rev (5'-CATGGTTGGACCCTGGAA CAGCACCTCCAGCA-3') oligonucleotides was prepared as described above. The adapter was inserted between *Nde*I and *Nco*I sites of the pNW1120 vector (GE Healthcare), resulting in the pNW1120-3C vector. Then, the eGFP gene was amplified using eGFP_dir (5'-ATTACCATGGTGAGCAA GGGCGAGGAGCTG-3'; NcoI) and eGFP_rev (5'-ATATC TCGAGTTACTTGTACAGCTCGTCCATGCCGA-3') primers. Oligonucleotides were added at a 0.2 μM concentration and PCR was performed in the presence of 0.2 mM dNTPs, 3.6 mM $MgCl_2$ and 2.5 U of Taq polymerase (Fermentas). The amplified PCR fragment was inserted between the *Nco*I and *Xho*I sites of pNW1120-3C. The plasmid was transformed into *E. coli* BL21 host cells and the culture was grown in LB medium at 37 °C. Protein expression was induced by heating the cultures at 42 °C for 15 min in a water bath. The cells were then suspended in 50 mM NaH_2PO4, pH 7.5, 300 mM NaCl, 0.1% Tween 20 and 2 mM β-mercaptoethanol and lysed by sonication. The supernatant of the centrifuged lysate was applied to a 2-mL Ni-NTA (Qiagen) column. The column was washed with 20 mL of 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.1% Tween 20, 2 mM β-mercaptoethanol and 20 mM imidazole. Then, the column was equilibrated with buffer, containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 2 mM DTT. In addition, 50 U of PreScission (GE Healthcare) protease was added for immobilized protein cleavage. Proteolysis was carried out at 8 °C for 20 h. Cleaved eGFP protein, possessing a glycine residue at the N-terminus, was eluted from the column and PreScission protease was removed by treating the solution with Glutathione Sepharose™ 4 Fast Flow (GE Healthcare).

In vitro *protein-protein sortase-mediated ligation*

To evaluate transpeptidation activity in native conditions, approximately 15 μM GST-LPETG and 15 μM G-eGFP substrates were mixed with cyclic and linear sortase A enzymes at a 5 μM concentration. Reaction incubation was conducted in the presence of 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM CaCl2 and 2 mM DTT at 25 °C for 24 h. Urea was added at a 2 M final concentration for evaluation of activity under denaturing conditions. The resulting products were analyzed in 12% Tris-Glycine SDS-PAGE.

Results and discussion

Expression of circularly permuted Intein-srtA(60-204) fusion protein

The plasmid pNW1120 was designed for cyclization of *E. coli* dnaB protein, located between Int_C and Int_N domains. In this study we substituted the sequence encoding dnaB by a sequence encoding the *S. aureus* srtA catalytic domain (Fig. 1). Two primer sets were used for *srtA* gene fragment amplification, therefore allowing construction of two plasmids for cyclic sortase A (csrtA) isoforms with different lengths of amino acid linkers between N- and C-termini. The pNW1120-csrtA+9aa plasmid encoded the precursor, which was expected to yield a srtA catalytic domain with N- and C-termini fused via nine-amino-acid-residue-long linker TRESGSIEF, and pNW1120-csrtA+13aa plasmid was expected to yield cyclic srtA with 13-amino-acid-long GSGTRESGSIEFG linker, respectively. Both constructs demonstrated the same expression profile (here and further

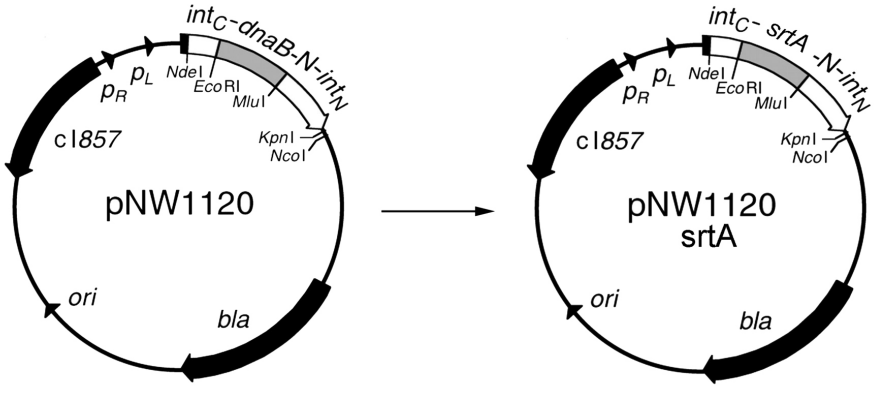

Fig. 1. Schematic diagram of the pNW1120 vector (Williams et al. 2002) used for the srtA gene insertion (described in methods).

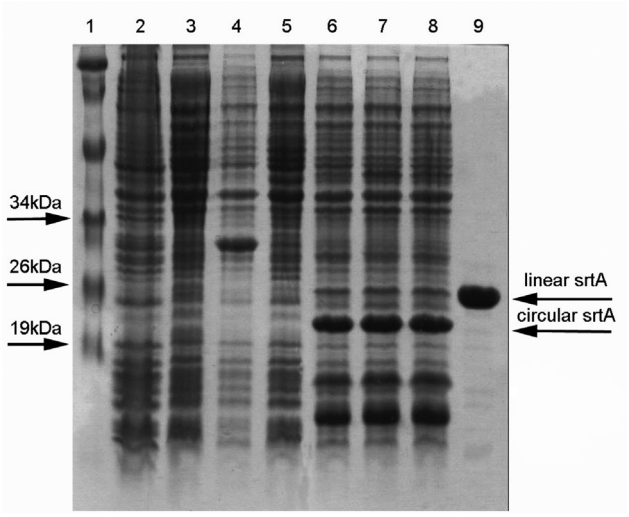

Fig. 2. Expression of circular (csrtA+9aa) and linear srtA (indicated by arrows) in *E. coli* BL21 cells. Sonicated cell lyzates were analysed in SDS/PAGE. Lane 1, prestained Protein MW Marker (Fermentas, Lithuania). Lane 2, control cells (without thermal shock). Lane 3, control cells, grown at 42 °C for 3 h (exposed to heat shock). Lane 4, control cells transformed with pNW1118 vector and induced with 1 mM IPTG. Lane 5, control cells transformed with pNW1120-csrtA9aa vector (without thermal induction). Lanes 6 to 8, cells transformed with pNW1120-csrtA9aa vector and induced at 42 °C for 2, 3, 4 h respectively. Line 9, purified linear srtA.

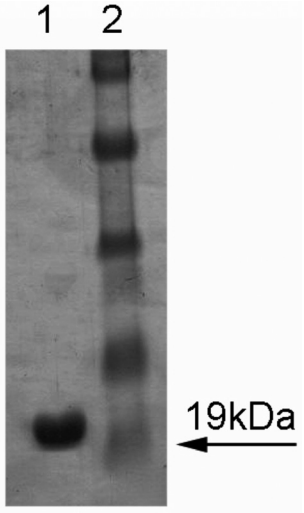

Fig. 3. Purified circular srtA in SDS PAAG (lane 1). The protein was purified from sonicated cell lyzate by ion exchange chromatography and gel filtration as described in the methods. Lane 2, prestained Protein MW Marker (Fermentas, Lithuania). The smallest marker band is indicated by an arrow.

not shown for the pNW1120-csrtA+13aa). The plasmids directed heat-induced expression of several protein species, resolved by SDS-PAGE, which were absent in the controls (Fig. 2, lanes 2 to 5). One of the bands, located approximately at the level of protein standard corresponding to 19kDa, was identified as being a circular form of srtA (Fig. 2, lanes 6 to 8). Electrophoresis showed that cyclic forms of the proteins had significantly greater mobility than linear His-tagged srtA purified by affinity chromatography, although the molecular weights of the proteins were nearly the same (Fig. 2, lane 8). The identified csrtA isoforms were purified from the lysates by consequent cation exchange, anion exchange and gel filtration steps, yielding highly homogenic isolated protein samples (Fig. 3, lane 1) that could be used for enzymatic activity evaluation.

We showed that the mini-intein domains of *Synechocystis* sp. dnaB are well suitable for sortase A cyclization, since active and soluble protein is being expressed in *E. coli* host cells with high yields, comparable to those with linear form. Interestingly, the mobility of cyclic srtA forms was faster in SDS-PAGE than the mobility of the linear form. This can be explained by a smaller radius of gyration of circular molecules in the denatured state. Faster movement in denaturing gel has been reported also for other cyclized proteins (Iwai et al. 2001).

Expression of GST-LPETG and G-eGFP protein substrates

GST substrate carrying a C-terminal LPETG srtA recognition site was expressed under bacteriophage lambda promoter (Hedgpeth et al. 1978) in chemically (IPTG) induced *E. coli* host cells and purified by affinity chromatography as described above. One-step purification yielded highly soluble protein that was stable in solution at 4 °C for a long period without stabilizing agents (Smith et al. 1988). The eGFP protein, carrying the 3C protease site at the N-terminus, was synthesized in thermally induced host cells, transformed with pNW1120-3C-eGFP plasmid. The protein was immobilized on Ni-NTA sepharose and treated by PreScission protease liberating G-eGFP after proteolytic cleavage of N-terminal His-tag (not shown). Both substrates were used for srtA enzymatic activity evaluation.

Evaluation of srtA activity

Both circular and linear srtA isoforms were shown to exhibit specific activity in native conditions. GST-LPETG and eGFP were successfully ligated in the presence of csrtA+9aa, csrtA+13aa and lsrtA enzymes (Fig. 4, lanes 2 to 4). A protein band at the level of the 60 kDa standard was identified as GST-eGFP fusion in these reactions, which was absent in controls (Fig. 4, lanes 8 to 11). The same protein band was detected in denaturing conditions in the presence of 2 M urea (Fig. 4., lanes 5 to 6), although the band intensity was lower than in the case of native conditions. Linear srtA did not exhibit specific activity in the presence of urea, and the band corresponding to GST-eGFP fusion was not found (Fig. 4, lane 7).

The exhibited activity of csrtA isoforms suggests correct folding of the active centre of the enzyme. Increased stability of the modified proteins is confirmed by the ability to catalyze reactions in the presence of 2 M urea, while the linear form is not active in the same conditions. The yielded

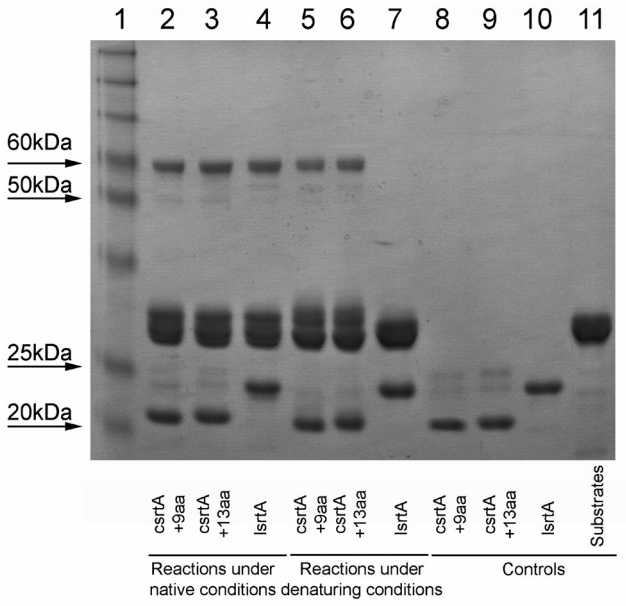

Fig. 4. Protein-protein ligation reactions. Purified proteins GST-LPETG (modified glutathione S-transferase, see in methods) and G-eGFP (modified enhanced green fluorescent protein, see in methods) were used as substrates for ligation reactions catalysed by two variants of circular (csrtA+9aa and csrtA+13aa) and by linear (lsrtA) srtA enzymes. Lane 1, prestained Protein MW Marker (Fermentas, Lithuania), indicated by arrows. Lanes 2 to 4, protein ligation reactions under native conditions. Lanes 5 to 7, protein ligation reactions under denaturing conditions (2 M urea). Lanes 8 to 10, control, containing csrtA+9aa, csrtA+13aa and lsrtA enzymes without substrates, respectively. Lane 11, control, containing both substrates without enzyme.

product quantities were similar for the reactions driven by csrtA isoforms with different linker lengths in both native and denaturing conditions. Thus, the linker length does not influence the conformation of the proteins.

In this study we observed csrtA ability to catalyze reactions in the presence of urea, which can potentially be applied for synthesis of mosaically labelled proteins for NMR needs. Segmental isotope labelling is crucial for liquid-state NMR analysis of large proteins and has been performed using intein-mediated transpeptidation reactions (Otomo et al. 1999). Since the intein application is limited and the reaction can be performed in native conditions, csrtA-driven protein ligation reactions could widen the possibilities of partial protein labeling. Activity exhibited in denaturing conditions can allow to locate a LPETG recognition sequence in any region of the target protein, regardless of its folding, and of the performed ligation of partly or fully denatured molecules with subsequent refolding of the mosaic polypeptide.

Acknowledgements

We thank Gottfried Otting and Nicholas E. Dixon for providing pNW1118 and pNW1120 vectors.

References

Beckett G.J., Hayes J.D. 1993. Glutathione S-transferases: biomedical applications. *Adv. Clin. Chem*. 30: 281–380.

Cubitt A.B., Heim R., Adams S.R., Boyd A.E., Gross L.A., Tsien R.Y. 1995. Understanding, improving and using green fluorescent protein. *Trends Biochem. Sci.* 20: 448–455.

Evans T.C.Jr., Martin D., Kolly R., Panne D., Sun L., Ghosh I., Chen L., Benner J., Liu X. Q., Xu M. Q. 2000. Protein trans-splicing and cyclization by a naturally split intein from the *dnaE* gene of *Synechocystis* species PCC6803. *J. Biol. Chem.* 275: 9091–9094.

Hedgpeth J., Ballivet M., Eisen H. 1978. Lambda phage promoter used to enhance expression of a plasmid-cloned gene. *Mol. Gen. Genet.* 163: 197-203.

Ilangovan U., Ton-That H., Iwahara J., Schneeewind O., Clubb R.T. 2001. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Nat. Acad. Sci. USA* 98: 6056–6061.

Iwai H., Lingel A., Plückthun A. 2001 Cyclic green fluorescent protein produced *in vivo* using an atificially split PI-PfuI intein from *Pyrococcus furiosus*. *J. Biol. Chem.* 276: 16548–16554.

Mazmanian S. K., Liu G., Ton-That H., Schneewind O. 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285: 760-763.

Mao H., Hart S.A, Schink A., Pollok B.A. 2004. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* 126: 2670-2671.

Otomo T., Ito N., Kyogoku Y., Yamazaki T. 1999. NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* 38: 16040–16044.

Schneewind O., Model P., Fischetti V.A. 1992. Sorting of protein A to the staphylococcal cell wall. *Cell* 70: 267–281.

Smith D.B, Johnson K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutatione S-transferase. *Gene* 67: 31–40.

Williams N. K., Prosselkov P., Liepinsh E., Line I., Sharipo A., Littler D. R., Curmi P. M. G., Otting G., Dixon N. E. 2002. *In vivo* protein cyclization promoted by a circularly-permuted *Synechocystis* sp. PCC6803 DnaB mini-intein. *J. Biol. Chem.* 277: 7790–7798.

Received 1 September 2010; received in revised form 20 September 2010; accepted 20 September 2010