Appl Microbiol Biotechnol (2008) 77:1175–1180
DOI 10.1007/s00253-007-1233-0

METHODS

Purification of green fluorescent protein using a two-intein system

Zhonglin Zhao · Wei Lu · Baoqing Dun · Dan Jin · Shuzhen Ping · Wei Zhang · Ming Chen · Ming-Qun Xu · Min Lin

Received: 3 June 2007 / Revised: 27 September 2007 / Accepted: 28 September 2007 / Published online: 1 November 2007

Abstract A two-intein purification system was developed for the affinity purification of GFPmut3*, a mutant of green fluorescent protein. The GFPmut3* was sandwiched between two self-cleaving inteins. This approach avoided the loss of the target protein which may result from in vivo cleavage of a single intein tag. The presence of N- and C-terminal chitin-binding domains allowed the affinity purification by a single-affinity chitin column. After the fusion protein was expressed and immobilized on the affinity column, self-cleavage of the inteins was sequentially induced to release the GFPmut3*. The yield was 2.41 mg from 1 l of bacterial culture. Assays revealed that the purity was up to 98% of the total protein. The fluorescence and circular dichroism spectrum of GFPmut3* demonstrated that the purified protein retains the correctly folded structure and function.

Z. Zhao
College of Biological Sciences, China Agricultural University,
Beijing 100094, China
e-mail: gill2015@hotmail.com

W. Lu · S. Ping · W. Zhang · M. Chen · M. Lin (✉)
Biotechnology Research Institute,
Chinese Academy of Agricultural Sciences,
Beijing 100081, China
e-mail: linmin57@vip.163.com

B. Dun
Institute of Crop Science,
Chinese Academy of Agricultural Sciences,
Beijing 100081, China

D. Jin
Biotechnology Research Center, Southwest University,
Chongqing 400716, China

M.-Q. Xu
New England Biolabs,
Beverly, MA 01915, USA

Keywords Intein · Protein splicing · Green fluorescent protein · Protein purification

Introduction

Protein splicing is a post-translational process analogous to RNA splicing. In this process, inteins catalyze their self-splicing from the precursor protein with the concomitant ligation of flanking polypeptide sequences, termed exteins, via a native peptide bond. Intein-mediated protein purification is a convenient and cost-effective approach, and it utilizes the inducible self-cleavage activity of an engineered intein to isolate the target protein (Chong et al. 1998). The method is able to isolate recombinant proteins by affinity chromatography without the use of exogenous proteases (Guo et al. 2004; Wu et al. 2002; Katragadda and Lambris 2006; Xu and Evans 2003; Singleton et al. 2002). In addition, this approach was applied to express protein ligation, to protein semisynthesis, and to specific labeling of proteins and peptides (David et al. 2004; Muir et al. 1998; Evans et al. 1999a). One drawback of the system is that in vivo cleavage activity of the intein can affect the recovery of the target protein, leading to a low yield. Intein-mediated purification with two intein tags (TWIN system) was originally developed for generation of cyclic proteins or peptides and multimers (Evans et al. 1999a, b; Xu and Evans 2001). The system incorporates two intein tags as fusion partners. The target protein is fused between two self-cleaving inteins, *Synechocystis* sp. DnaB intein (intein1) and *Mycobacterium xenopi* GyrA intein (intein2). Each of the inteins has a chitin-binding domain (CBD) as a fusion partner, which allows the purification of the precursor protein on a chitin resin. Intein1 is engineered to undergo temperature- and pH-dependent cleavage at its

C-terminal splice junction (Mathys et al. 1999). Intein2 is modified for thiol-induced cleavage at its N-terminal by dithiothreitol or 2-mercaptoethanesulfonic acid (MESNA) and releases the target protein (Evans et al. 1999a). Following cleavage of the two intein tags, the target protein is released from the chitin resin (as shown in Fig. 1).

In this report, we have successfully expressed and purified GFPmut3*, a mutant of wild-type green fluorescent protein (GFP; Andersen et al. 1998; Cormack et al. 1996), using the two-intein system. GFPmut3* has highly shifted excitation maxima, the folding was more efficient than that of the wild-type GFP, and its fluorescence intensity was significantly enhanced. The presence of the two intein fusion partners appeared to increase the final yield by overcoming the problem associated with in vivo cleavage activity of a single intein tag. The expression conditions for active GFPmut3* were determined by examination of fluorescence of the *Escherichia coli* host cells, and the folding was further characterized by the fluorescence and circular dichroism spectrum of the purified GFPmut3*.

Materials and methods

Bacterial strains, plasmids, and media

T7-expressed *E. coli* ER2566 (New England Biolabs, Beverly, MA, USA) was used as a host for protein expression. *E. coli* strain JM109, the host strain used for all plasmid construction, was cultivated in Luria–Bertani (LB; 1% tryptone, 1% NaCl, 0.5% yeast extract; pH 7.5). pTWIN1 (New England Biolabs, USA) and pGEM T-easy vector (Promega, USA) were used as expression and cloning vectors, respectively. The plasmid pJH2 containing *gfpmut3** gene was obtained from the Institute of Plant Protection, Chinese Academy of Agricultural Sciences.

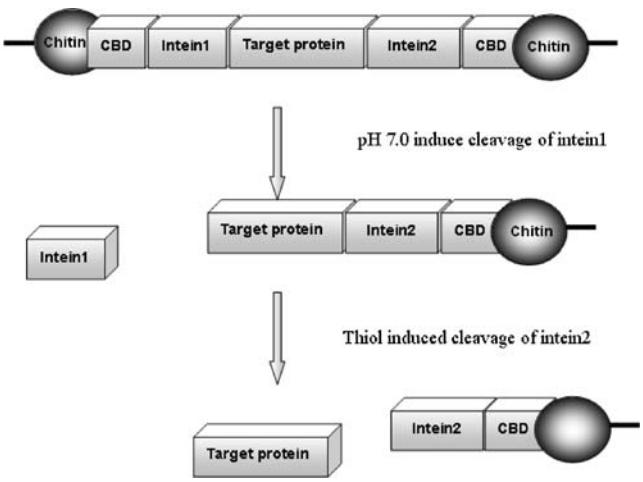

Fig. 1 Protein purification process by the TWIN system. The target protein–intien fusion protein is bound to chitin beads. Intein1 is first induced to cleave at its C-terminal splice site. Then thiol-induced excision of intein2 releases the target protein. The target protein is eluted from the column, whereas the two intein-CBD tags remain bound to chitin beads

Reagents and enzymes

Restriction enzymes, T4 DNA ligase, RNase, and chitin beads were purchased from NEB; Ex Taq DNA polymerase was from Takara, Japan; MESNA was obtained from Sigma, USA. Other chemicals used in this study were of analytical or higher grade.

Construction of expression plasmid pTG

The *gfpmut3** gene bearing the mutations S2R, S65G, and S72A (Andersen et al. 1998), was amplified from the plasmid pJH2 by polymerase chain reaction (PCR) with the following two primers: 5′-GGGCTCTTCCAACTGTT CAGCAGGAACAATGCGTAAAGGAGAAG AA-3′ and 5′-CCGCTCTTCCGCAATAACCCCTTCCCTC GATTTTGTATAGTTCATCC ATG-3′. The PCR-amplified fragments were then cloned into pGEM T-easy vector and verified by DNA sequencing. The clone was digested with *Sap*I and was subcloned into the expression vector pTWIN1 using *Sap*I sites, resulting in the plasmid pTG. The *gfpmut3** gene was sequenced at the Chinese Academy of Agricultural Sciences.

Expression and purification of GFPmut3*

pTG was transformed into *E. coli* ER2566. The cells were grown in 1 l of LB medium containing 100 μg/ml ampicillin at 37°C. When the cell density reached OD_{600} of 0.5–0.7, expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.3 mM, followed by an additional incubation at 15°C for 16 h. Cells were then collected by centrifugation at 4,000×*g* for 10 min and lysed by sonication in buffer A [20 mM Tris–HCl, pH 8.0, with 500 mM NaCl and 1 mM ethylenediaminetetraacetic acid (EDTA)]. The cell extract was centrifuged at 19,000×*g* for 30 min, and the supernatant was applied to a chitin-binding column pre-equilibrated with buffer B (20 mM Tris–HCl, pH 7.0, with 500 mM NaCl and 1 mM EDTA). The column was washed with 200 ml of buffer B and incubated at 23°C for 20 h. Following incubation, the column was washed with ten column bed volumes of buffer B. The column was quickly flushed with three column volumes of buffer C (20 mM Tris–HCl, pH 8.5 containing 50 mM 2-mercaptoethanesulfonic acid, 500 mM NaCl). The column flow was stopped and left at 4°C for 16 h to induce the cleavage of intein2. Then, the target protein was eluted with buffer B.

Approximately 100 μl of samples from each step were saved and analyzed by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Determination of protein concentration and purity

The concentration of the eluted protein was determined by Bradford Assay and analyzed by Coomassie Blue-stained 12% SDS-PAGE gels. Protein purity was estimated by comparing the intensity of samples run on SDS-PAGE. The stained SDS-PAGE gel was quantified by gel document scanning with Bio-Rad Quantity One 4.6.1 1-D Analysis software.

Confocal microscope scanning and spectroscopy of GFPmut3*

Leica TCS SP2 laser scanning spectral confocal microscope (Germany) was used for detecting fluorescence in *E. coli* ER2566 cells that contain pTG. Objective is 100× oil-immersion lens. The excitation wavelength was 488 nm emission ranges, and detection bandwidth was 500 to 600 nm.

To calculate the recovery of GFP during purification, fluorescence spectrums were recorded with Hitachi F-4500 spectrofluorometer at 23°C. A cuvette with 1 cm path length was used. The sample for the emission and excitation scans was prepared in 20 mM Tris–HCl, pH 7.0, with 500 mM NaCl and 1 mM EDTA. Absorbance was measured on a Hitachi U-3010 spectrophotometer at 23°C, 10 mM Tris–HCl, pH 8.0.

Circular dichroism measurements

Circular dichroism (CD) wavelength scans were recorded with a Jasco-715 spectropolarimeter. The protein sample was prepared in 10 mM Tris–HCl (pH 8.0). The concentration of the sample was 0.25 mg/ml. The data was normalized to molar ellipticity with a path length of 0.1 cm.

Results

Construction of two-intein *gfpmut3** fusion construct

The TWIN system allowed the purified protein to possess an N-terminal residue other than methionine and certain amino acid residues which are unfavorable for the intein cleavage reaction (Xu and Evans 2001). To clone the *gfpmut3** gene, *Sap*I sites were incorporated into the forward and reverse primers. The gene was cloned into two *Sap*I sites in pTWIN1. *Sap*I digestion results in noncomplementary overhangs, and there is no self-ligation at the two *Sap*I sites. Another advantage of using *Sap*I sites is that the *Sap*I sites will not be regenerated after the gene was inserted into pTWIN1. This allowed us to isolate the native protein without vector-derived amino acid residues. The *gfpmut3** was fused between two self-cleaving inteins, intein1 and intein2. In this study, a threonine residue was added to the N-terminal residue of *gfpmut3** by PCR. Each intein was also fused with a CBD for affinity isolation of the precursor protein.

Expression and purification of protein and SDS-PAGE analysis

Different growth and induction conditions, such as the induction temperature and time, were tested to optimize the protein expression. We used the conditions of 15°C for 16 h to induce the GFPmut3* expression. The cells showed bright green fluorescent color under confocal microscope with 550 photomultiplier tube (PMT) volts. Additionally, if the culture containing pTG was incubated in a shaker at 4°C for several days without addition of IPTG, the expression of GFP was still detectable. However, when the strains were incubated at 37°C for 2 h or 30°C for 3 h, the expression level of GFPmut3* was very low. The cells showed no visible green color. The fluorescent signal was detected with 700 PMT volts by confocal microscope, and it was found that some *E. coli* cells lacked GFPmut3* expression (Fig. 2). We conclude that the induction temperature was critical, and low temperature was required for expression of correctly folded GFPmut3*. It is assumed that the precursor protein has to fold correctly to retain the intein-mediated cleavage activity. In general, folding of a precursor protein is influenced by different target proteins, and the production process must be optimized.

After GFPmut3* was expressed, the clarified cell extract was loaded onto a chitin column. Each step of the purification processes was monitored by UV lamp. The clarified cell extract containing the GFPmut3* fusion protein displayed green fluorescence. After passage of the cell extract through the chitin column, the fluorescence was not detectable in the flow-through fraction, whereas the column showed bright green fluorescence. The column was washed with buffer B to remove the unbound protein, and then incubated to induce on-column cleavage of intein1 at 23°C pH 7.0 for 16 h. Following incubation, the column was then washed with buffer B to remove any unbound proteins. Then cleavage of intein2 was induced in buffer C overnight, and GFPmut3* was eluted with buffer B. SDS-PAGE (Fig. 3, lane 3) analysis showed that the crude cell extract contained the unspliced full-length precursor protein CBD-Ssp-GFPmut3*-Mxe-CBD (81 kDa) and unexpected splicing products CBD-Ssp-GFPmut3* (53 kDa) and Mxe-CBD (28 kDa), corresponding to their molecular mass. The

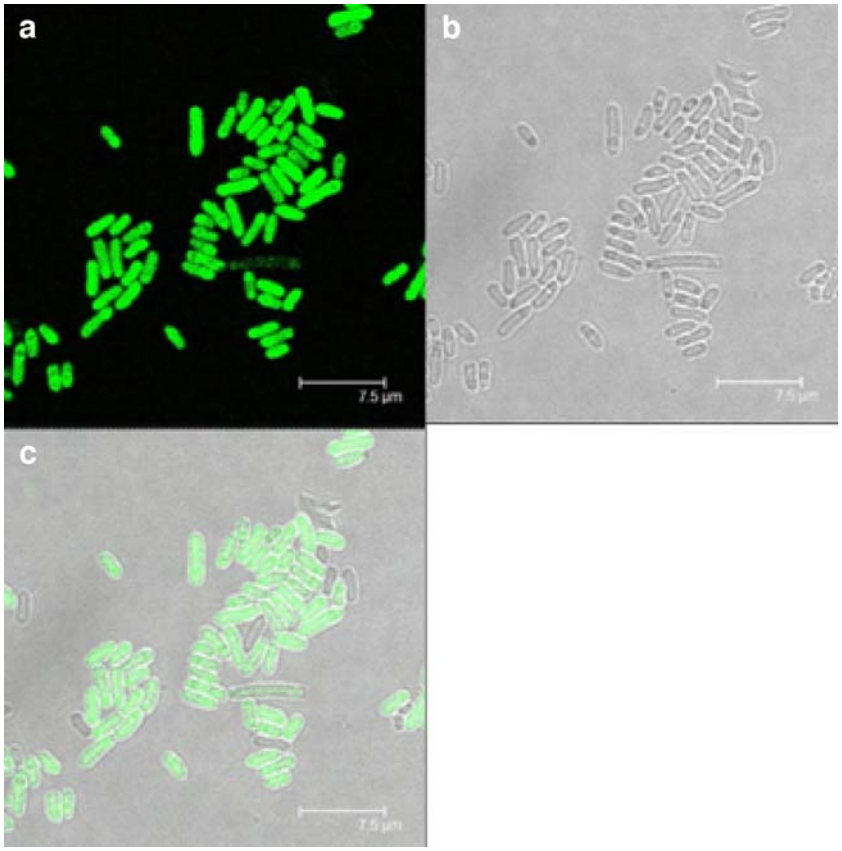

Fig. 2 Expression of GFPmut3* in *E.coli* ER2566 strains, which were incubated at 37°C for 2 h. **a** *E. coli* ER2566 cells with GFPmut3* expressing exhibited bright green fluorescence under UV lamp of 488 nm. **b** General optical image of *E. coli* ER2566 cells. **c** Overlay image of **a** and **b** to identify the GFPmut3* expression

CBD-Ssp-GFPmut3* and Mxe-CBD fragments were generated by cleavage at the N-terminal junction of the Mxe, and both bound to the chitin column (Fig. 3, lane 4). The elution fractions from the chitin column had a bright green fluorescence, indicating that the GFPmut3* protein was correctly folded. There was only one major band of 29 kDa detected in the eluted fraction (Fig. 3, lane 5), corresponding to the molecular mass of GFPmut3*. The purity of the fluorescent proteins was determined by estimating the absorbance ratio at two wavelengths. The absorbance maximum of GFPmut3* was at 502 nm, which has highly shifted form the maximum at 395 nm in parental GFP (Cormack et al. 1996). The absorbance ratio of 502 to 280 nm is about 1.04, which is a good indication of active GFPmut3* (amount). The yield was 2.41 mg from 1 l of bacterial culture. The purity was estimated to be approximately >98% by SDS-PAGE.

Spectroscopic characterization of GFPmut3*

The fluorescent properties of purified GFPmut3* were determined (Fig. 4a). The fluorescence excitation and mission spectra were very close to the class 4 GFP (Tsien 1998). GFPmut3* contained the mutations S2R, S65G, and S72A. The fluorescence of GFPmut3* was more intense than the wild type GFP and has a markedly shifted excitation maxima. The excitation and emission maxima wavelengths of GFPmut3* were 501.8 and 510.4 nm, respectively (Fig. 4a). However, the excitation maxima and emission maxima of the wild-type GFP was 395 and 504 nm. Iwai et al. (2001) obtained cyclic GFP, which showed excitation maxima at 399 nm and emission maxima at 508 nm. GFPmut3 * has a highly shifted excitation maxima from 395–399 to 501.8 nm, comparing these two GFP species. It also fluoresces more intensely when excited at 488 nm. Longer wavelengths of excitation and emission would be useful for multiple labels and reporters, as well as to serve as resonance energy transfer acceptors. The minor excitation peak at 280 nm is presumably due to resonance energy transfer from a single tryptophan residue with an excitation maximum at 280 nm. The spectroscopic characterization of GFPmut3* is a result of both its chromophore formation and folded structure. The GFPmut3* purified by the technique presented has the same spectroscopic and structural characteristics as the reported GFPmut3*. The fluorescence spectrum indicated that the protein has the correctly folded three-dimensional structure.

The CD spectroscopy was used to identify the secondary structure of purified GFPmut3* (Fig. 4b). The known secondary structure of GFP is an 11-stranded

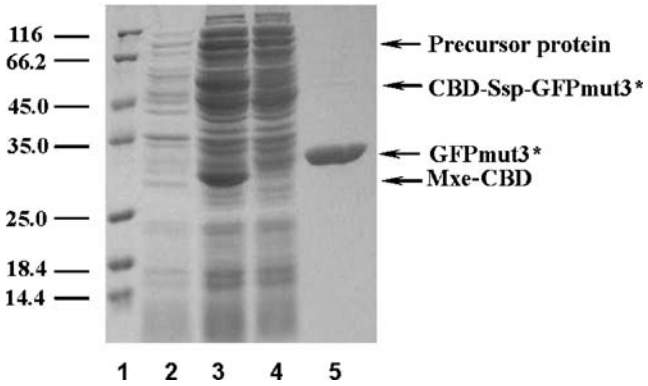

Fig. 3 Examination of purification of GFPmut3* by SDS-PAGE. *Lane 1* Protein marker; *lane 2* uninduced crude cell extract; *lane 3* crude cell extract following induction at 15°C contains the precursor CBD-Ssp-GFPmut3*-Mxe-CBD (81 kDa), Mxe-CBD (28 kDa), and CBD-Ssp-GFPmut3* (53 kDa); *lane 4* clarified cell extract from *lane 3* following passage over chitin column (note that precursor CBD-Ssp-GFPmut3*-Mxe-CBD and CBD-Ssp-GFPmut3* binds to chitin column); *lane 5* proteins eluted from chitin resin. All reactions were performed as described under "Expression and purification of GFPmut3*" and analyzed on a 12% SDS-PAGE gel

β-barrel threaded by an α-helix running up the axis of the cylinder (Tsien 1998). The result showed that there was a peak shoulder between 216 nm, which was the contribution of β-barrel and was similar to the reported GFP structure (Iwai et al. 2001). The secondary structure of the purified GFPmut3* appeared to have no major structural change. The data strongly suggests that the purified protein by the TWIN system retains the correctly folded structure and function.

Discussion

In this study, we have utilized the two-intein system to purify active GFPmut3*. The result of the SDS-PAGE indicated that C-terminal intein (intein2) fusion was unstable. Intein2 underwent substantial in vivo cleavage, which produced CBD-Ssp-GFPmut3* and Mxe-CBD fragments (Fig. 3, lane 3). This property makes it unsuitable as a single C-terminal fusion partner, which is the common strategy employed by its users. In vivo cleavage activity of an intein tag is usually unpredictable, leading to a reduced yield of the target protein. By employing the two-intein strategy, the target protein can still be absorbed onto the affinity resin after a single intein cleavage event, thereby increasing the final yield of the product. Thus, this method presents a practical approach to solve the premature cleavage problem associated with a single-intein system.

The presence of N- and C-terminal CBD tags made it possible to isolate the GFPmut3* product in a single chromatographic column. However, with the CBD tag, impure proteins could be copurified (Chong et al. 1998; Singleton et al. 2002). Use of other affinity tags, such as polyhistidine (his-tag) or *E. coli* maltose-binding protein, complementary to CBD may ensure intactness of the product. Engineering two complementary tags, such as CBD and six-His tag, would require use of both chitin and Ni-nitriloacetic acid (NTA) affinity columns in order to recover the product. Furthermore, production of certain enzymes and pharmaceutical proteins often requires protein processing and activation to remove amino terminal methionine residue. The use of a fusion partner facilitates protein purification, but the tag may exert adverse effects on the structure and function of the proteins. The two-intein system may provide a means to overcome these problems and is capable of isolating proteins with an amino terminal residue other than methionine and free of a fusion tag.

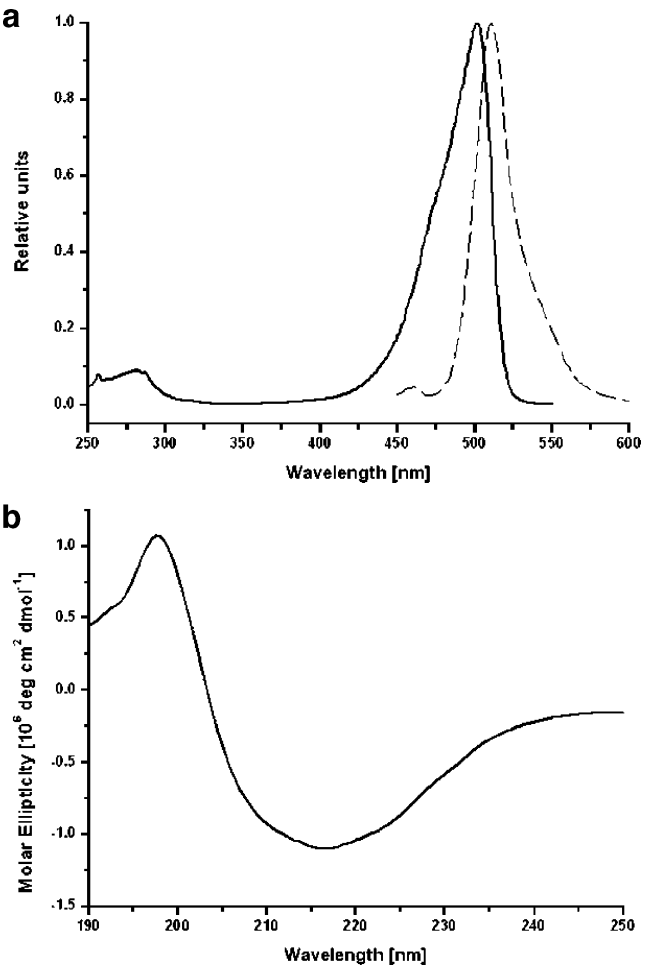

Fig. 4 Spectroscopic characterization of GFPmut3*. **a** Fluorescence excitation and emission spectra of GFPmut3* (*solid and dashed lines*, respectively). Fluorescence excitation spectrum of GFPmut3* was recorded at an emission wavelength of 510 nm at pH 7.0. **b** CD spectra of GFPmut3*. The spectrum was measured at 23°C, 10 mM Tris–HCl, pH 8.0

Acknowledgments This work was supported by Foundation of 863 Project of China.

References

Andersen JB, Sternberg C, Poulsen LK, Bj, rn SP, Givskov M, Molin S (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64:2240–2246

Chong S, Montello GE, Zhang A, Cantor EJ, Liao W, Xu MQ, Benner J (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. Nucleic Acids Res 26:5109–5115

Cormack BP, Valdivia RH, Falkow S (1996) FACS optimized mutants of the green fluorescent protein (GFP). Gene 173:33–38

David R, Richter MP, Beck-Sickinger AG (2004) Expressed protein ligation method and applications. Eur J Biochem 271:663–677

Evans TC, Benner J, Xu MQ (1999a) The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. J Biol Chem 274:3923–3926

Evans TC Jr, Benner J, Xu MQ (1999b) The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. J Biol Chem 274:18359–18363

Guo C, Li Z, Shi Y, Xu M, Wise JG, Trommer WE, Yuan J (2004) Intein-mediated fusion expression, high efficient refolding, and one-step purification of gelonin toxin. Protein Expr Purif 37:361–367

Iwai H, Lingel A, Pl, ckthun A (2001) Cyclic green fluorescent protein produced in vivo using an artificially split PI-*Pfu*I intein from *Pyrococcus furiosus*. J Biol Chem 276:16548–16554

Katragadda M, Lambris JD (2006) Expression of compstatin in *Escherichia coli*: Incorporation of unnatural amino acids enhances its activity. Protein Expr Purif 47:289–295

Mathys S, Evans TC Jr, Chute IC, Wu H, Chong S, Benner J, Liu XQ, Xu MQ (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. Gene 231:1–13

Muir TW, Sondhi D, Cole PA (1998) Expressed protein ligation: A general method for protein engineering. Proc Natl Acad Sci U S A 95:6705–6710

Singleton SF, Simonette RA, Sharma NC, Roca AI (2002) Intein-mediated affinity-fusion purification of the Escherichia coli RecA protein. Protein Expr Purif 26:476–488

Tsien RY (1998) The green fluorescent protein. Annu Rev Biochem 67:509–544

Wu W, Wood DW, Belfort G, Derbyshire V, Belfort M (2002) Intein-mediated purification of cytotoxic endonuclease I-*Tev*I by insertional inactivation and pH-controllable splicing. Nucleic Acids Res 30:4864–4871

Xu MQ, Evans TC Jr (2001) Intein-mediated ligation and cyclization of expressed proteins. Methods 24:257–277

Xu MQ, Evans TC Jr (2003) Purification of recombinant proteins from *E. coli* by engineered inteins. Methods Mol Biol 205: 43–68