Published in final edited form as:

Protein Expr Purif. 2013 October ; 91(2): 169–174. doi:10.1016/j.pep.2013.08.006.

Ligation-independent cloning and self-cleaving intein as a tool for high-throughput protein purification

Tiana D. Warren¹,

Johns Hopkins University, Department of Chemical and Biomolecular Engineering, 3400 North Charles Street, Baltimore, MD, 21218 twarre12@jhu.edu

Michael J. Coolbaugh, and

The Ohio State University, William G. Lowrie Department of Chemical and Biomolecular Engineering, 140 W. 19th Avenue, Columbus, OH 43210 coolbaugh.3@osu.edu

David W. Wood

The Ohio State University, William G. Lowrie Department of Chemical and Biomolecular Engineering, 140 W. 19th Avenue, Columbus, OH 43210 wood.750@osu.edu

Abstract

The rapid production of purified recombinant proteins has become increasingly important for countless applications. Many purification methods involve expression of target proteins in fusion to purification tags, which often must be removed from the target proteins after purification. Recently, engineered inteins have been used to create convenient self-cleaving tags for tag removal. Although intein methods can greatly simplify protein purification, commercially available expression vectors still rely on conventional restriction/ligation cloning methods for target gene insertion. We have streamlined this process by introducing Ligation-Independent Cloning (LIC) capability to our intein expression plasmids, which provides a simple method for constructing self-cleaving tag-target gene fusions. In this work, we demonstrate efficient gene insertion via this system, as well as target protein expression and purification consistent with previously reported results. Through this newly developed system, arbitrary protein genes can be rapidly incorporated into self-cleaving tag expression vectors, and their products purified using convenient platform methods.

© 2013 Elsevier Inc. All rights reserved.

Correspondence to: David W. Wood.

¹Present address: Johns Hopkins University, Department of Chemical and Biomolecular Engineering, 3400 North Charles Street, Baltimore, MD, 21218

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest

The authors do not have any conflicts of interests.

Contributors

DWW conceived of the study. TDW and MJC designed the study and completed the cloning experiments. TDW completed the protein expression and purification work, while MJC performed the statistical analysis of transformed colonies and quantification of intein cleaving kinetics. TDW drafted the manuscript. TDW and MJC responded to reviewer comments. All authors read and approved the final manuscript.

Keywords

ligation-independent cloning; molecular cloning; intein; protein purification; self-cleaving affinity tag

Introduction

The need for rapid and reliable methods to produce purified recombinant proteins has become more urgent as the study of gene products for various applications increases. During the last two decades, several methods have been developed for streamlined molecular cloning and protein purification. In particular, Invitrogen's Gateway® [1] and TOPO® [2] cloning technologies have provided a means for parallel cloning of a single target gene into several vectors. Although these methods are highly efficient, they require the purchase of commercial kits, which can increase the cost of finding the optimal vector for expressing a specific target protein. A more inexpensive cloning technique is ligation-independent cloning (LIC). This method allows the simple and precise cloning of an amplified gene into an expression vector by taking advantage of the 3' to 5' exonuclease activity of T4 DNA Polymerase (Figure 1). In comparison to conventional molecular cloning, LIC relies on extended segments of complementary DNA to facilitate DNA strand annealing and highly efficient ligation *in vivo*, thereby eliminating the need for conventional restriction endonuclease digestion and *in vitro* ligation [3].

Several university researchers and commercial companies have incorporated LIC into simplified protein purification systems. The majority of these systems employ affinity purification methods, which require an affinity tag to be expressed in fusion to the target protein. The expressed fusion protein is then passed over a compatible affinity resin, which selectively binds the affinity tag and allows all other cell components to pass through. Once the fusion protein is purified and recovered, the affinity tag is often cleaved to produce a pure native product protein. Many of the existing LIC/protein purification systems encode polyhistidine (His₆) purification tags, and include convenient protease target sequences for tag removal [4–6]. Although these systems allow simple cloning and purification of proteins, the protease enzymes can be expensive and must also be separated from the product protein in an additional downstream purification step.

One alternative to the use of proteases for tag removal is self-cleaving inteins. Inteins are naturally occurring proteins that excise themselves from a native host protein through self-cleavage and ligation of their flanking peptide bonds. Over the last 20 years, inteins have been reengineered to exhibit isolated N-, C- or double terminal cleavage for the purification of proteins [7–11]. In practice, inteins are expressed as part of a fusion protein, where the intein separates the affinity tag from the target protein. This fusion is then purified using the included tag, after which the tag is induced to self-cleave. This method thus provides all the power and generality of affinity tag methods, but produces a native target without the need for proteolytic tag removal.

We have previously reported the engineered I-CM self-cleaving intein, derived from the *Mycobacterium tuberculosis* RecA intein [12, 13]. Its three primary advantages are its small size (18kDa), its pH and temperature sensitivity for cleavage, and the ability to produce a native target protein without additional amino acids at the N-terminus [14]. In previous work, the I-CM intein was reengineered for use with Invitrogen's Gateway® and TOPO® cloning technologies [15, 16]. These systems were successfully used to express and purify proteins using conventional affinity methods via Maltose Binding Protein (MBP) and Chitin

Binding Domain (CBD) tags, as well as non-chromatographic purification methods via the Elastin-Like polypeptide (ELP) tag.

In this work, two new vectors were constructed that facilitate LIC with intein-mediated protein purification. The utility of the modified intein system was verified by the expression and purification of two different target proteins using both conventional affinity and non-chromatographic purification methods. In addition, this newly developed system yielded 70% successful gene-insertion, comparable to other LIC systems, with negligible affect on intein cleavage efficiency. By combining LIC with the I-CM self-cleaving intein, we have developed an inexpensive, streamlined method for simplified cloning and rapid expression and purification of proteins.

Materials and Methods

Plasmid Construction (pET:CBD-I-LIC, pET:ELP-I-LIC)

A pET vector encoding the Chitin Binding Domain (CBD), the I-CM intein and the acidic Fibroblast Growth Factor (aFGF) genes [16] was digested with EcoRI and XbaI to remove the intein and *aFGF* genes and serve as the backbone for the LIC construct (Figure 1A). PCR was performed using Phusion Polymerase (New England Biolabs) to generate a fragment encompassing all but the last 7 nucleotides of the intein gene, followed by the complete LIC site using forward primer 5'-ATCGGAATTTCGCCCTCGCAGAGGGCACT-3' and reverse primer 5'-CGATTCTAGAGAGTTCGTCTTCTTCTCCATGGACAACAACCCCTTCGGCGACG-3'. In this case, the annealing portion of the reverse primer was ACAACAACCCCTTCGGCGACG-3' and the remaining sequence was the downstream LIC site. The PCR reactions were performed in a total volume of 50 μ l consisting of 5 μ l 10 $\times$ Phusion Polymerase buffer, 5 μ l of each 10 μ M forward and reverse primer, 2 μ l 1:50 diluted template plasmid, 1 μ l dNTP mix (10 mM each dNTP), 0.5 μ l Taq Polymerase and 32 μ l Milli-Q water. The PCR program was designed as follows: incubation at 94°C for 5 minutes, followed by 30 cycles of 94°C for 1 minute, 50–55°C for 1 minute and 72°C for 1 minute/kB target protein gene, followed by 72°C for 10 minutes. The PCR product was digested with EcoRI and XbaI and ligated to the similarly digested backbone with T4 DNA Ligase. This plasmid is referred to as pET:CBD-I-LIC.

A second plasmid encoding the ELP purification tag fused to the intein and LIC site was prepared using the pET:CBD-I-LIC construct as the starting vector. pET:CBD-I-LIC was digested with NdeI and EcoRI to remove the CBD tag. A fragment consisting of the Elastin-Like Polypeptide (ELP) tag was obtained from an NdeI and EcoRI digest of the plasmid pET:EI:MBPD containing the ELP tag DNA sequence, the I-CM intein and the Maltose Binding Protein (MBP) [17]. The ELP tag and the NdeI/EcoRI digested pET:CBD-I-LIC vector were ligated using T4 DNA ligase to create the plasmid pET:ELP-I-LIC.

LIC Procedure

NcoI (5 units per μ g DNA) was used to digest each LIC vector for 3 hours at 37°C. The resulting linearized vector was recovered by ethanol precipitation. Compatible overhangs were generated by treating 0.2 pmol of linearized vector with 1.5 units of T4 DNA Polymerase (Invitrogen) and 100 μ M deoxyadenosine triphosphate (dATP) in a total volume of 20 μ l for 30 minutes at 22°C. The polymerase was inactivated by the addition of 1 μ L of 25 mM EDTA and incubation at 75°C for 20 minutes. Additionally, 0.2 pmol of each insert was similarly treated with 1.5 units of T4 DNA Polymerase and 100 μ M deoxythymidine triphosphate (dTTP). The optimal vector to insert molar ratio was found experimentally for each of the 3 inserts and used for all future ligations. The treated vector was mixed with

each of the treated inserts, incubated for 5 minutes at 22°C and transformed into competent DH5 *E. coli* cells. After verification of successful target gene insertion by restriction mapping, correct clones were re-transformed into BLR *E. coli* cells for protein expression.

Statistical Analysis

To analyze the cloning efficiency of the LIC-Intein system, two different measures were used, as shown in Table 1. A crude determination was made by determining the ratio of the number of transformant colonies resulting from a full LIC reaction to those of a control LIC reaction lacking the target protein insert gene. This allowed us to approximate the amount of undigested or reclosed vector that was transformed into the cells, thus providing a rough indication of the potential background of a given cloning attempt. For each clone, 3 separate trials were completed for statistical accuracy.

In addition, Colony PCR was performed on 10 random colonies from each LIC cloning trial using Taq Polymerase (New England BioLabs) with forward primer 5'-ATCGGAATTTCGCCCTCGCAGAGGGCACT-3' and reverse primer 5'-CGTTATTGATCGTATTGGGGAAC-3'. Each 20 µl PCR screen reaction consisted of 2 µl 10× Taq Polymerase Buffer, 1 µl each of 10 µM forward and reverse primer, 0.2 µl dNTP mix, 0.2 µl Taq Polymerase and 16 µl water. The amplification program was as follows: 1 cycle of 94°C for 5 minutes followed by 35 cycles of 94°C for 1 minute, 57°C for 1 minute and 72°C for 1 minute/kB pair, followed by 72°C for 10 minutes. The PCR products were analyzed using DNA gel electrophoresis. Negative clones were identified by the appearance of DNA fragments 500 base pairs in size, while positive clones were identified by product bands at 3600, 1300, and 1600 base pairs, equal to the size of the intein with the -galactosidase (-gal), -lactamase (-lac) and Maltose Binding Protein (MBP) genes, respectively. A total of two positive clones and two negative clones were further characterized by DNA sequencing.

Expression and Purification

Chitin Purification—A single colony from the plated transformation reaction was used to inoculate 2 ml of Luria Bertani (LB) Media supplemented with 50 µg/ml ampicillin for 12–16 hours with shaking. A 500 µL aliquot of the overnight culture was used to inoculate 50 ml of LB liquid media supplemented with 50 µg/ml ampicillin in a 250 ml baffled flask and the culture was grown to OD₆₀₀ = 0.5–0.6 in an air shaker at 37°C and 150 rpm. The culture was then brought to 15°C in a water shaker for 30 minutes. Expression was induced by the addition of 100 mM isopropyl-β-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The precursor protein (affinity tag-intein-target gene fusion) was expressed for 6 hours at 15°C and 150 rpm. After expression, the cells were harvested by centrifugation at 4°C and 5000×g for 15 minutes. Harvested cells were resuspended in ice-cold pH 8.5 column buffer (20 mM AMPD, 20 mM PIPES, 1 mM EDTA, 500 mM NaCl, 1 mM DTT, pH 8.5) and frozen at –80°C for 30 minutes. The pellet was thawed at room temperature, lysed by sonication and clarified by centrifugation at 10,000×g for 30 minutes at 4°C. The supernatant was taken and diluted 1:5 in ice-cold pH 8.5 column buffer.

An Econo-Pac® Chromatography Column (Bio-Rad Laboratories) was packed with 2 ml of Chitin resin (New England Biolabs) and kept at 4°C for the entire purification process. The column was washed with 20 ml of pH 8.5 Column Buffer and 10 ml of the diluted clarified lysate was passed through the column. This was followed by 24 ml of pH 8.5 column buffer and 16 ml of pH 6.5 Column Buffer (20 mM AMPD, 20 mM PIPES, 1 mM EDTA, 500 mM NaCl, 1 mM DTT, pH 6.5) to induce on-column cleavage. After purifying the precursor protein, the column was brought to room temperature to facilitate self-cleavage. Resin samples were taken at time t = 0 hrs and t = 20 hrs for the cleavage reaction. After 20 hours,

the pure product protein was eluted off the column in two 1 ml aliquots using pH 6.5 column buffer. Intein cleavage efficiency and product purity were analyzed via SDS-PAGE.

Quantification of Intein Cleaving Kinetics—One hundred milliliter cultures of pET:CBD-I-MBP derived from both traditional and LIC derived cultures were grown and expressed as above. The clarified lysate from each culture was diluted 1:5 in pH 8.5 column buffer and 1/3 of the diluted lysate was loaded on to a 2-ml chitin column maintained at 37°C with a jacketed column using an AKTA Purifier 10. A step pH gradient to a pH of 6.2 was then passed slowly through the column to induce intein self-cleavage. The intein cleaving rate constant was determined by fitting the release profile of cleaved MBP to a first order decay model, as described previously [14]. This procedure was repeated two times with the remaining lysate, to give a total of three replicates. A student's T-test was used to compare the average rate constant for both LIC and non-LIC derived constructs.

ELP Purification—A single transformed colony was used to inoculate 2 ml of LB liquid media supplemented with 50 µg/ml ampicillin, which was then incubated with shaking at 37°C for 12–16 hours. A 500 µl aliquot of the overnight culture was used to inoculate 50 ml of Terrific Broth (TB) media (12 g/l BactoTryptone, 24 g/l BactoYeast, 2.31 g/l Monopotassium Phosphate, 12.54 g/l Dipotassium Phosphate) supplemented with 50 µg/ml ampicillin in a 250 ml baffled flask, and the culture was grown to OD₆₀₀ = 0.5–0.6 in an air shaker at 37°C and 150 rpm. The culture was then brought to 15°C in a water shaker for 30 minutes. Expression was induced by the addition of 100mM IPTG to a final concentration of 1 mM. The precursor protein was expressed for 22 hours at 15°C and 150 rpm. After expression, the cells were harvested by centrifugation for 15 minutes at 4°C and 5000×g.

Harvested cells were resuspended in 2.5 ml of pH 8.5 lysis buffer (20 mM Tris, 2 mM EDTA) and frozen at –80°C for 30 minutes. The ELP fusion was purified using previously reported methods [18]. Briefly, the pellet was thawed at room temperature, lysed by sonication and clarified by centrifugation at 10,000×g for 30 minutes at 4°C. The soluble proteins, including the precursor protein, were recovered in the supernatant. Ammonium sulfate from a 1.6 M stock solution was added to the clarified lysate to a final concentration of 0.4 M ammonium sulfate and agitated at room temperature for 10 minutes. This caused the ELP-tagged target protein to precipitate. The solution was centrifuged at room temperature for 10 minutes at 10,000×g and the precipitated ELP fusion protein was recovered in the pellet. The pellet was washed in 2.5 ml of pH 8.5 lysis buffer and a final concentration of 0.4 M ammonium sulfate and then resuspended in 0.75 ml of ice-cold pH 6.5 Column Buffer (see above). The solution was centrifuged at room temperature for 2 minutes at maximum speed to remove additional cell debris.

The fusion protein was recovered in the supernatant and allowed to cleave for 20 hours at room temperature. After cleavage, ammonium sulfate was added (as described above) to a final concentration of 0.4 M, and agitated for 10 minutes at room temperature to precipitate the cleaved tag. The solution was centrifuged at maximum speed for 5 minutes and the pure product protein was recovered in the supernatant. Cleavage products were then analyzed for cleavage efficiency and purity via SDS-PAGE.

Results

LIC-Intein Sequence Design

There are four primary DNA sequence requirements for LIC to be effective [3, 19]. First, the LIC overlap sequences in the plasmid backbone (vector) and gene of interest (insert) should be at least 10 nucleotides in length to ensure adequate base pairing between the two fragments to be ligated (Figure 1C). Second, each LIC overlap sequence should have high

GC content to increase the stability of the base-paired fragments. Third, the overlap sequences should be deficient in either adenine or thymine in the final 10–12 bases of the 3' end of each antiparallel DNA strand. In particular, if the linearized vector is deficient in thymine at its 3' ends, the insert should be deficient in adenine, or vice versa, to allow for base pairing between the two strands. Finally, the LIC site should contain a unique restriction site that can be used to linearize the vector for T4 DNA polymerase treatment.

Successful incorporation of LIC into an existing intein expression plasmid requires that the reading frame at the intein-target protein junction be preserved (Figure 1), and that the intein retain its controllable cleaving activity. Furthermore, in order for the target protein to be recovered with no additional amino acids, the end of the intein coding sequence must immediately precede that of the target protein gene. We identified a region near the end of the intein (from Glycine 163 to the thymine in Valine 166, shown in Figure 1A) that fulfills the LIC requirements and allows for minimal alteration of the intein sequence. The LIC site in our vectors was then constructed using PCR to generate an intein-LIC fragment. The last 7 nucleotides of the intein were omitted from the intein-LIC PCR product to expose the appropriate sequence for T4 DNA polymerase treatment (Figure 1B and 1C). Using these primers, we also introduced a unique NcoI restriction site within the LIC sequence. This allows linearization of the plasmid with the NcoI restriction enzyme prior to polymerase treatment of the vectors and inserts (Figure 1B and 1C).

Protein Purification via LIC Cloning and Self-Cleaving Intein Tags

The ability of the intein cleaving reaction to deliver a native target protein is dependent on the intein amino acid sequence being directly appended to the first codon of the target protein gene. Therefore, the forward primer used to PCR amplify the target protein must include the LIC annealing sequence as well as the intein nucleotides that reside between the intein LIC site and the start codon of the target protein. This results in an amplified insert sequence comprised of the LIC upstream sequence, the end of the intein sequence, the target protein gene, and the downstream LIC annealing sequence, in that order (Figure 1C). More specifically, the sequences 5'-GGGGTTGTTGTACACAAC and 5'-CTCGTCTTCTTCTCCA preceded the annealing portion of the forward and reverse PCR primers, respectively, to introduce the sequences that flank the target protein gene. After T4 DNA Polymerase treatment of the inserts, the overlapping regions of the vector backbone and insert were mostly complementary. As shown in Figure 1C, there was one mismatch (lowercase and italicized) and one gap (lowercase and underlined). The final expression plasmid included all of the bases from the insert, and the vector backbone was corrected for any mismatches and gaps. Therefore, the overhangs always occurred on the vector fragment. There were two vector overhangs: (1) GGTA on the 5' upstream end and (2) CAT on the 5' downstream end.

To assess whether the intein maintained its ability to act as a self-cleaving module after LIC, target proteins from two of the six clones (pET:CBD-I-MBP and pET:ELP-I- lac) were expressed, purified and analyzed using SDS PAGE (Figures 2 and 3, respectively). The behaviors of both of these targets with self-cleaving tag methods have been extensively characterized in previous work [15–18, 20], thus providing convenient benchmarks for the cloning and purification methods. One limitation of the I-CM intein is premature cleavage during protein expression [21]; however, expression for 6 hours at 15°C for the CBD fusion constructs and 22 hours at 15°C for the ELP fusion constructs minimized pre-cleavage and provided satisfactory expression of the tagged target protein. In both cases, the target proteins could be produced with yields and purities similar to those of previous work, demonstrating that the LIC cloning has not affected the ability of the intein to act as a self-cleaving module.

In order to further confirm that the introduction of the LIC cloning site did not affect intein cleaving, the cleaving rate constant for an LIC derived construct was compared to the rate constant for a previously constructed non-LIC derived construct. The rate constant for CI-MBP (LIC) at pH 6.2 and 37°C was found to be $1.09 \pm 0.01 \text{ hr}^{-1}$ while that of CI-MBP was found to be $0.99 \pm 0.06 \text{ hr}^{-1}$. These values are in line with previously reported cleaving rates for the I-CM intein [14]. A student's t-test was performed to determine if the two sample means were significantly different. The two means were not significantly different, to a significance level of 0.05.

Statistical Analysis of Cloning Efficiency

The cloning efficiency of the LIC-Intein system was evaluated by analyzing ten transformation colonies at random via colony PCR from each of the six independent LIC cloning reactions. Results indicate that the correct target protein sequence was successfully inserted in approximately 70% of the transformed clones (Table 1), which is consistent with previously reported LIC cloning efficiencies for other systems [4]. Two positive and two negative clones from the colony PCR screen were sequenced to confirm the presence/absence of the desired target protein. All positive clones had the target protein inserted in frame with the tag-intein and the critical intein-target protein junction as intact. All negative (false-positive) clones were due to reclosed LIC vector.

Discussion

In order for LIC to be effective, the insert and the linearized vector are treated with T4 DNA Polymerase in the presence of free dATP or dTTP. In each case, the added nucleotide corresponds to the one that is absent at the 3' end of each respective strand. The 3' to 5' exonuclease activity of the polymerase removes nucleotides one by one until it reaches an adenine or thymine in the sequence. At this point, the T4 DNA polymerase begins to incorporate the free nucleotide supplied in the reaction mixture, effectively stopping the exonuclease reaction. This generates complementary overhangs of at least 10 nucleotides in length on both the vector and insert. The two fragments are then mixed together, allowing them to anneal via base pairing, and then they are transformed into *E. coli*. Once transformed, the DNA repair mechanism restores the phosphodiester bonds between the two fragments, resulting in an intact expression plasmid. The final plasmid will include all of the bases from the insert, and the vector will conform to bases of the insert by correcting any mismatches or gaps and removing the vector overhangs [22].

A statistical analysis of our LIC-intein system revealed an observed cloning efficiency of 70%. Although this efficiency is comparable to previously reported studies, one potential way to improve the efficiency is to gel purify the digested vector, as described in [4, 23, 24]. Furthermore, a previous study has shown that increasing the length of the complementary overhangs can increase the number of transformants, but does not increase the percentage of clones with successfully inserted target genes [25]. In some cases, this increase in the number of transformants may give the impression of increased cloning efficiency due to a comparative decrease in the percentage of transformants arising from transformed undigested vector.

When designing the LIC-intein vectors, it was crucial to incorporate LIC into the I-CM intein system without deleterious effects on intein activity. Analyses of the expression and purification of two target proteins via SDS-PAGE indicate that LIC cloning has negligible effects on cleavage efficiency.

Similar LIC-intein vectors can easily be constructed, with a variety of purification tags, using the NdeI and EcoRI restriction sites at the beginning and end of the tags in the vectors

used in this work. Furthermore, these vectors can be used for 96-well plate LIC cloning, as demonstrated in previous studies [5, 26]. Thus, it is expected that the LIC-intein combination will be useful for high-throughput cloning and purification of a variety of target proteins in additional LIC-intein vectors.

Conclusion

In this work, we designed two new vectors that combine ligation-independent cloning with intein-mediated protein purification. Specifically, our LIC vectors encode the I-CM intein fused to either a conventional affinity tag or a non-chromatographic ELP purification tag. With these vectors, we demonstrated about 70% correct target gene-insertion, without the need for conventional restriction/ligation cloning methods. Furthermore, we were able to verify intein function through successful expression and purification of two target proteins. This LIC-intein system provides an efficient, streamlined method for parallel cloning and rapid expression and purification of different target proteins.

Acknowledgments

The authors would like to thank Richard Lease and the Wood lab group for technical support. Additionally, Tiana Warren was supported by funds from the NIH Postbaccalaureate Research Education Program (R25GM089571).

Abbreviations

LIC	Ligation-Independent Cloning
MBP	Maltose Binding Protein
CBD	Chitin Binding Domain
ELP	Elastin-Like polypeptide
AMPD	2- Amino- 2- methyl- 1,3-propanediol.

References

1. Hartley JL, Temple GF, Brasch MA. DNA cloning using in vitro site-specific recombination. *Genome Res.* 2000; 10:1788–1795. [PubMed: 11076863]
2. Shuman S. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J Biol Chem.* 1994; 269:32678–32684. [PubMed: 7798275]
3. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* 1990; 18:6069–6074. [PubMed: 2235490]
4. Cabrita LD, Dai W, Bottomley SP. A family of E. coli expression vectors for laboratory scale and high throughput soluble protein production. *BMC Biotechnol.* 2006; 6:12. [PubMed: 16509985]
5. Stols L, Gu M, Dieckman L, Raffin R, Collart FR, Donnelly MI. A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein expression and purification.* 2002; 25:8–15. [PubMed: 12071693]
6. Luna-Vargas MP, Christodoulou E, Alfieri A, van Dijk WJ, Stadnik M, Hibbert RG, Sahtoe DD, Clerici M, Marco VD, Littler D, Celie PH, Sixma TK, Perrakis A. Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. *J Struct Biol.* 2011; 175:113–119. [PubMed: 21453775]
7. Gogarten JP, Senejani AG, Zhaxybayeva O, Olendzenski L, Hilario E. Inteins: structure, function, and evolution. *Annual review of microbiology.* 2002; 56:263–287.
8. Chong S, Shao Y, Paulus H, Benner J, Perler FB, Xu MQ. Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J Biol Chem.* 1996; 271:22159–22168. [PubMed: 8703028]

9. Chong S, Williams KS, Wotkowicz C, Xu MQ. Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J Biol Chem.* 1998; 273:10567–10577. [PubMed: 9553117]
10. Mathys S, Evans TC, Chute IC, Wu H, Chong S, Benner J, Liu XQ, Xu MQ. 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.* 1999; 231:1–13. [PubMed: 10231563]
11. Southworth MW, Amaya K, Evans TC, Xu MQ, Perler FB. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques.* 1999; 27:110–114. 116, 118–120. [PubMed: 10407673]
12. Davis EO, Sedgwick SG, Colston MJ. Novel structure of the recA locus of *Mycobacterium tuberculosis* implies processing of the gene product. *J Bacteriol.* 1991; 173:5653–5662. [PubMed: 1909321]
13. Wood DW, Wu W, Belfort G, Derbyshire V, Belfort M. A genetic system yields self-cleaving inteins for bioseparations. *Nat Biotechnol.* 1999; 17:889–892. [PubMed: 10471931]
14. Wood DW, Derbyshire V, Wu W, Chartrain M, Belfort M, Belfort G. Optimized single-step affinity purification with a self-cleaving intein applied to human acidic fibroblast growth factor. *Biotechnol Prog.* 2000; 16:1055–1063. [PubMed: 11101334]
15. Gillies AR, Hsui JF, Oak S, Wood DW. Rapid cloning and purification of proteins: gateway vectors for protein purification by self-cleaving tags. *Biotechnol Bioeng.* 2008; 101:229–240. [PubMed: 18727029]
16. Wu WY, Gillies AR, Hsui JF, Contreras L, Oak S, Perl MB, Wood DW. Self-cleaving purification tags re-engineered for rapid Topo(R) cloning. *Biotechnol Prog.* 2010; 26:1205–1212. [PubMed: 20945478]
17. Banki MR, Feng L, Wood DW. Simple bioseparations using self-cleaving elastin-like polypeptide tags. *Nat Methods.* 2005; 2:659–661. [PubMed: 16074986]
18. Fong BA, Wu WY, Wood DW. Optimization of ELP-intein mediated protein purification by salt substitution. *Protein expression and purification.* 2009; 66:198–202. [PubMed: 19345265]
19. Wang T, Ma X, Zhu H, Li A, Du G, Chen J. Available methods for assembling expression cassettes for synthetic biology. *Applied microbiology and biotechnology.* 2012; 93:1853–1863. [PubMed: 22311648]
20. Wu WY, Fong BA, Gilles AG, Wood DW. Recombinant protein purification by self-cleaving elastin-like polypeptide fusion tag. *Curr Protoc Protein Sci.* 2009 Chapter 26, Unit 26 24 21-18.
21. Fong BA, Wu WY, Wood DW. The potential role of self-cleaving purification tags in commercial-scale processes. *Trends Biotechnol.* 2010; 28:272–279. [PubMed: 20359761]
22. Li C, Evans RM. Ligation independent cloning irrespective of restriction site compatibility. *Nucleic Acids Res.* 1997; 25:4165–4166. [PubMed: 9321675]
23. Bardocz V, Geczi V, Sawasaki T, Endo Y, Meszaros T. A set of ligation-independent in vitro translation vectors for eukaryotic protein production. *BMC biotechnology.* 2008; 8:32. [PubMed: 18371187]
24. Weeks SD, Drinker M, Loll PJ. Ligation independent cloning vectors for expression of SUMO fusions. *Protein expression and purification.* 2007; 53:40–50. [PubMed: 17251035]
25. Aslanidis C, de Jong PJ, Schmitz G. Minimal length requirement of the single-stranded tails for ligation-independent cloning (LIC) of PCR products. *PCR Methods Appl.* 1994; 4:172–177. [PubMed: 7580902]
26. Dieckman L, Gu M, Stols L, Donnelly MI, Collart FR. High throughput methods for gene cloning and expression. *Protein expression and purification.* 2002; 25:1–7. [PubMed: 12071692]

Highlights

- We have engineered a modified intein for ligation-independent cloning (LIC).
- The engineered intein provides a self-cleaving tag for simple protein purification.
- The LIC intein was demonstrated by cloning and purifying two target proteins.
- The LIC method showed approximately 70% positive clones upon ligation.
- The LIC modifications do not appear to affect the intein self-cleaving characteristics.

		1	2	...	161	162	163	164	165	166	167	168											
E	F	A	L	...	A	E	G	V	V	V	H	N	M	F	N	S	D	*	*	S	R	V	
GAA	TTC	gcc	ctc	...	gcc	gaa	ggg	gtt	gtt	gta	cac	aac	atg	ttc	aat	...	tct	gac	taa	tag	TCT	AGA	gtc
CTT	AAG	cgg	gag	...	cgg	ctt	ccc	caa	caa	cat	gtg	ttg	tac	aag	tta	...	aga	ctg	att	atc	AGA	TCT	cag
		← Intein →											← aFGF →										→ Vector Backbone

E	F	A	L	A	E	G	V	V	V	H	G	E	E	D	E	L	S	R	V	
GAA	TTC	gcc	ctc	...	gcc	gaa	ggg	gtt	gtt	gTC	CAT	Gga	gaa	gaa	gac	gaa	ctc	TCT	AGA	gtc
CTT	AAG	cgg	gag	...	cgg	ctt	ccc	caa	caa	caG	GTA	Cct	ctt	ctt	ctg	ctt	gag	AGA	TCT	cag

(C) LIC

NcoI Digest of Vector, PCR for Insert, T4 DNA Polymerase Treatment of Vector and Insert

Diagram illustrating the LIC (Ligation Independent Cloning) process:

The process involves the NcoI Digest of Vector, PCR for Insert, and T4 DNA Polymerase Treatment of Vector and Insert.

The resulting DNA fragments are shown below:

Top Fragment (Left):

Sequence: E F A L A E
GAA TTC gcc ctc ... gcc gaa-3'
CTT AAG cgg gag ... cgg ctt ccc caa caa caG GTA-5'

Regions: Intein, Upstream LIC Site

Top Fragment (Right):

Sequence: H G E E D E L S R V
5'-CAT Gga gaa gaa gac gaa gtc TCT AGA gtc
3'-ag AGA TCT cag

Regions: Downstream LIC Site, Vector Backbone

Bottom Fragment:

Sequence: G V V V H N M K ... S *
5'-ggg gtt gtt gta cac aac atg aaa ... tcg tga t-3'
3'-t gtg ttg tac ttt ... agc act acc tct tct tct gct c-5'

Regions: Intein, MBP, Downstream LIC Site

E	F	A	L	A	E	G	V	V	V	H	N	M	K	...	S	*	W	R	R	R	R	G	T	L	*	S		
GAA	TTC	gcc	ctc	...	gcc	gaa	ggg	gtt	gtt	gta	cac	aac	atg	aaa	...	tcg	tga	tgg	aga	aga	aga	cga	gga	act	c TC	TAG	Agt	
CTT	AAG	cgg	gag	...	cgg	ctt	ccc	caa	caa	cat	gtg	ttg	tac	ttt	...	agc	act	acc	tct	tct	tct	gct	cct	tga	g AG	ATC	Tca	
Intein												MBP					Vector Backbone											

Protein Expr Purif. Author manuscript; available in PMC 2014 October 01.

(D) Final intein-target protein junction after LIC and transformation. In this case, MBP is shown as an example target protein.

Figure 2. SDS-PAGE analysis showing satisfactory expression and purification of Maltose Binding Protein from pET/CBD-I-MBP clone generated by LIC. M = Molecular weight marker, WL = Whole Lysate, CL = Clarified Lysate, LF = Load Flow-through, W1 = Wash 1 with pH 8.5 column buffer, W2 = Wash 2 with pH 6.5 column buffer, R1 = Resin sample taken just after inducing cleavage with pH 6.5 column buffer, R2 = Resin sample taken 20 hours after induction of the cleavage reaction, E1 = First product protein elution, E2 = Second product protein elution.

Figure 3. SDS-PAGE analysis showing satisfactory expression and purification of β -lactamase from pET/ELP-I-Blac generated by LIC. M = Molecular weight marker, WL = Whole Lysate, CL = Clarified Lysate, SP = Soluble contaminant proteins, Pre = Sample taken just after inducing cleavage with pH 6.5 column buffer, Post = Sample taken 20 hours after induction of the cleavage reaction, Pd = Pure product protein.

Table 1

Results from Statistical Analysis of LIC-Intein Efficiency. The overall cloning efficiency of the LIC-Intein system was calculated to be 70%.

Name of Clone	LIC Transformant Colonies: Control Transformant Colonies	Percentage of Positive Clones from Colony PCR (n=10)
CBD-I- gal	2.3 ± 0.3 (n=3)	90
CBD-I- lac	6.1 ± 0.9 (n=3)	80
CBD-I-MBP	3.3 ± 0.1 (n=2)	50
ELP-I- gal	7.0 ± 0.8 (n=3)	70
ELP -I- lac	9.2 ± 1.0 (n=3)	80
ELP -I-MBP	2.8 ± 0.1 (n=2)	60