



# Intein Based Fusion Proteins: Great Tags for the Soluble Production and Convenient Purification of Recombinant Proteins

Fatemeh Shafiee<sup>1,2\*</sup>, Shima Sharifi<sup>1</sup>, Armin Amini<sup>3</sup>

<sup>1</sup>Department of Pharmaceutical Biotechnology, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, Iran.

<sup>2</sup>Bioinformatics Research Center, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, Iran.

<sup>3</sup>Student Research Committee, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, Iran.

\*Corresponding author: Fatemeh Shafiee, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Hezar Jarib Ave. Isfahan, Iran. Tel: +98-3137927058, Fax: +98-3136680011, E-mail: [f\\_shafiee@pharm.mui.ac.ir](mailto:f_shafiee@pharm.mui.ac.ir)

Received: 2023/07/12 ; Accepted: 2023/03/10

**Background:** The main problem in the recombinant protein expression in *E. coli* strains, especially for high-yield production, is the accumulation in un-folded and inactive inclusion bodies. A suitable solution is the direction into the soluble cytoplasmic products by solubilizing tags. The use of inteins with self-cleaving ability, in addition to increase the chance of soluble protein expression, facilitates their purification process.

**Evidence Acquisition:** In this review article, papers related to the use of intein tags for soluble expression or protein purification were collected regardless the time limit. Available databases including Pubmed, google scholar, ScienceDirect, Web of Science, Scopus, and Embase was searched. The best condition for soluble expression or purification was focused in all articles.

**Results:** There are various intein tags commercially available in expression vectors that results in gaining our goal in facilitating the recombinant protein solubilization as well as its simple purification. It is enough to induce the self-cleavage property of the intein, which varies according to the type of intein used. In this way, the target protein is easily separated from the purification tag without the need to add protease enzymes such as enterokinase or treatment with various chemicals. The most common affinity tag in intein-based systems is Chitin Binding Domain attached to the chitin resin.

**Conclusions:** In this review article, we introduced proteins or peptides which produced in fusion to intein tags and discussed about their expression condition and purification process in order to enhance the chance of soluble expression and intein cleavage in a single stage, respectively.

**Keywords:** Chitin Binding Domain; Intein; Soluble Expression; Purification

## 1. Background

Recombinant DNA technology started in 1970s, is playing a vital role in the promotion of human health by developing therapeutic peptides and proteins. *Escherichia coli* as the simplest and the least expensive expression system is widely used for this purpose.

However, high-yield recombinant protein expression in this system results in the accumulation of un-folded and inactive proteins as inclusion bodies (1). There are two different strategies to overcome this problem including direction into the soluble cytoplasmic products or inclusion bodies refolding (2). However, because there

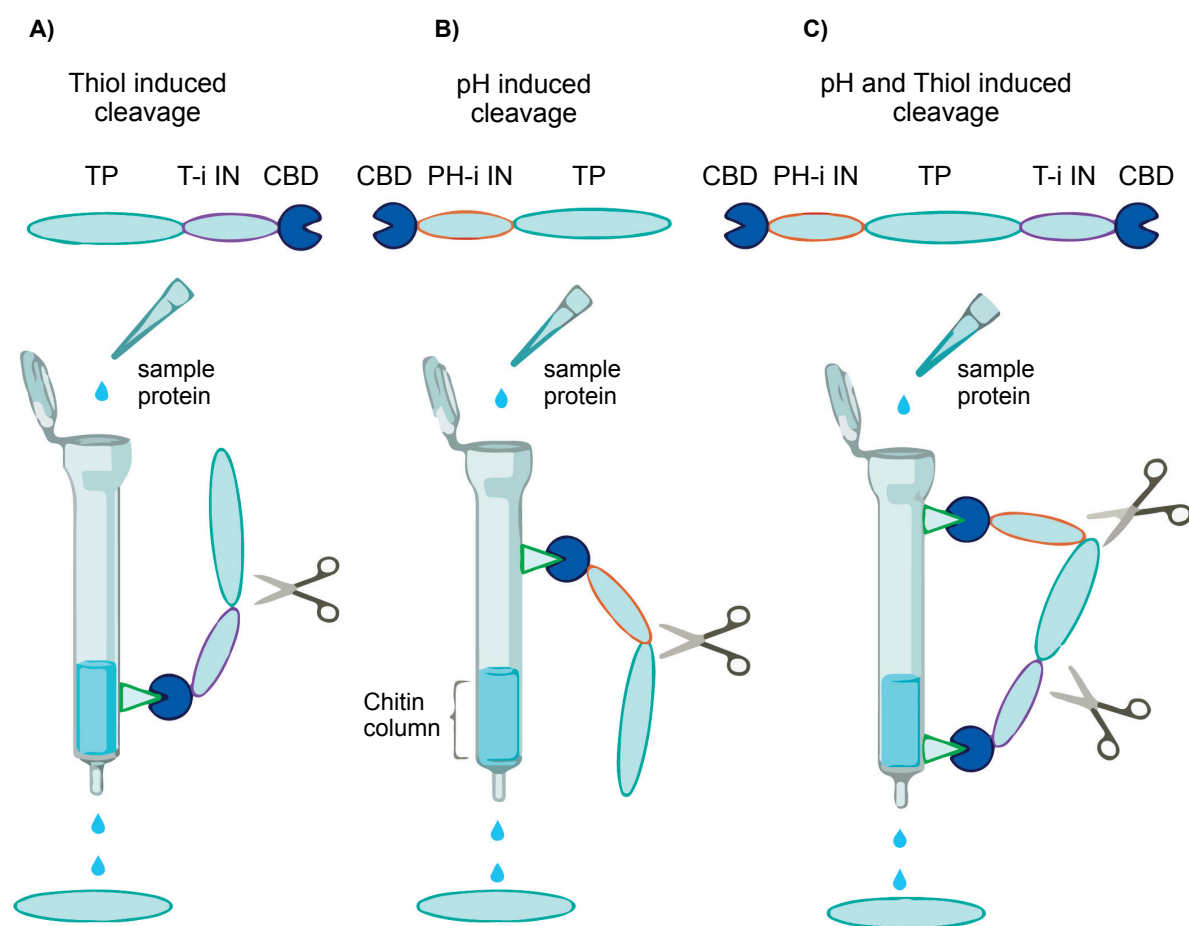
is no established procedure for the efficient folding of the inclusion bodies, the first one is more reliable; attempts to produce the recombinant proteins in soluble form (2). In order to achieve this goal, various strategies has been introduced over the time. Changing the culture conditions for example the incubation time, growth temperature, culture ingredients is one of the simplest solutions for this purpose. On the other hand, co-expression with chaperone molecules, change in host strain, and also the usage of various peptide tags can lead to the increase in the protein solubilization (3). Thioredoxin (Trx) is the most well-known peptide tag that leads to the enhancement in the soluble protein expression in *E. coli* especially for mammalian growth factors and cytokines previously produced as inclusion bodies. Another example in this respect is Maltose-binding protein (MBP) with the major solubilization effect on recombinant proteins in comparison to the GST (Glutathione S-transferase) and Trx. Furthermore, it has been reported that this peptide tag showed chaperone-like behavior (4).

The use of peptide tags, in addition to increase the chance of soluble protein expression, also facilitates their purification process; affinity chromatography can be used based on the substrate with high affinity to the tags added to the N- or C- terminal of the peptide or protein. There are a lot of examples of affinity tags used for the characterization and purification of recombinant proteins in microbial systems. Poly-histidine and poly-arginine are two the most well-known peptides in this regard; however, all tags used for enhancing the solubility of recombinant proteins can lead to facilitate their purification procedure (5). An important approach in the using of any type of peptide tags for the protein soluble expression or purification is this fact that tag deletion is a vital stage for its medicinal use in the biopharmaceutical evolution. For this purpose, various protease including Enterokinase, Factor Xa, Thrombin and so on have been introduced (6). On the other hand, the use of some chemicals can help to reach to this goal. CNBr (Cyanogen bromide) is the most well-known chemical on this matter in order to remove the poly-histidine tag (7). However, the addition of a protease and also a chemical leads to complicate the purification procedure and also increase the purification total price. Another issue, at least in the case of CNBr, is the problem of its high toxicity when used in the pharmaceutical industry (8). So, the usage of tags with

the ability to spontaneously delete themselves, is a suitable approach to overcome these problems. FrpC- and SrtAC-based tags are two example on this point, but a more well-known one in this case is intein based tag (9). Inteins can be used for the soluble expression in addition to help the removal of purification tags from the recombinant proteins (10). In the present review article, we focus on the examples in which various intein tags used for enhancing the solubility of expressed proteins and also with fusing to the recombinant protein, the removal of purification tags is easily done.

### 1.2. Inteins

Inteins contain two main functional domains including splicing and endonuclease domains so excises itself from mature proteins. With this ability, they can separate the affinity tags from the expressed proteins and create a mature molecule (10). On the other hand, mini-inteins with diminished size due to the deletion of the majority of their endonuclease domain and conserved splicing domain successfully function in tag removal (11). The natural splicing domain contains a conserved cysteine and asparagine at its N- and C-terminus, respectively. Mutation of one of these amino acids to alanine creates a mutant that cleaves only at its C or N terminus, allowing the intein to be used in purification tags (10). The mutation position determines the intein cleavage reaction; thiol-induced inteins retain a cysteine at their N-terminus and the cleaving reaction is induced by addition of thiol-containing molecules such as 1,4-dithiotreitol (DTT). Inteins with an asparagine at the C-terminal are known as pH-induced inteins. Thiol-induced inteins are tightly control cleaving because this process does not occur rapidly in the absence of a strong reducing agent. However, the addition of chemicals to induce intein cleavage probably causes toxic effects especially in large scale. Furthermore, this strategy is not suitable for the purification of proteins with disulfide bonds. pH-induced inteins, on the other hand, present a very simple purification process; lower pH promotes intein cleavage, without the need for the addition of any chemicals and can be used for proteins with disulfide bands. However, these pH-induced inteins are not strongly controlled leads to premature cleavage during the protein expression which can be covered by reducing the incubation temperature (10). These two various types of inteins are the bases of intein usage for recombinant protein expression and purification. In



Created in BioRender.com

**Figure 1. Schema of thiol-induced and pH-induced cleavage of two various inteins using the chitin column.**

**A)** C-terminal fusion of the target protein to the inteins that auto-cleaved by the addition of a reducing agent. **B)** N-terminal fusion of the target protein to the inteins that auto-cleaved by a shift in the buffer pH. **C)** C- and N-terminal fusion of the target protein to the inteins that auto-cleaved by the addition of a reducing agent and a shift in the buffer pH, respectively. TP: target protein, T-i: thiol-induced cleavage, pH-i: pH- induced cleavage, IN: intein, and CBD: Chitin Binding Domain.

the following, the commercial vectors contain inteins are introduced. Furthermore, **Figure 1** schematically, illustrates the cleavage induction of these two various inteins.

## 2. Evidence Acquisition

Related keywords were searched in various database including Pubmed, google scholar, ScienceDirect, Web of Science, Scopus, and Embase. All papers were surveyed for the protein type expressed in fusion to the intein tag. Furthermore, the type of vector and

intein was recorded. The best condition for the soluble expression of each protein and the best condition for the most auto-cleavage the intein was noticed.

## 3. Results

### 3.1. Vectors Designed Based on Intein Tags

For the first time, the New England Biolabs (UK) introduced a purification system abbreviated as IMPACT<sup>®</sup> (Intein Mediated Purification with an Affinity Chitin-binding Tag) in which several vectors

were designed base on various intein tags (12). These vectors named as pTXB1 and pTXB3, pTYB1, 2, 3, 4, 22, and 23, pMXB10, pKYB1, and pTWIN1 and 2 (13). The main difference between these vectors are including: the type and number of intein tag coding sequence and the antibiotic resistance gene which acts as the selectable marker.

pTXB1, as the most widely-used vector in this series, contains a mini-intein from the *gyrA* gene of *Mycobacterium xenopi* with 198 amino acid residues (Mxe GyrA intein). The self-cleavage activity of this intein which is inserted at the C-terminus of the gene of interest is induced by the addition of thiol containing chemicals (14).

The pTYB11 vector, on the other hand, contains an intein from the VMA1 gene of *Saccharomyces cerevisiae* (Sce VMA1) with 454 amino acids. The gene of interest is fused at its N-terminus to this intein with a VMA1 CBD tag at its C-terminus (15).

The other vector in this respect is pTWIN1, an *E. coli* expression vector, designed with an N-terminal cysteine and/or a C-terminal thioester and two modified inteins including Ssp DnaB and Mxe GyrA. Hydrolysis of peptide bond between the Ssp dnaB mini-intein C-terminus and target protein is triggered by shifting the pH at room temperature (16). A unique property of this vector is three independent options for cloning procedure including N-terminal fusion, C-terminal fusion and finally, N- and C-terminal fusion of the target gene to the Ssp DnaB, Mxe GyrA, or both inteins, respectively. Finally, pTWIN2 contains two inteins including Ssp DnaB mini-intein and Mth RIR1 intein. The last one is an intein from Ribonucleoside-diphosphate reductase protein originated from *Methanothermobacter Thermotrophicus*. Similar to the Sce VMA intein, this segment also induces its cleavage activity by a thiol rich chemical.

The control vector, pMXB10, derived from pTXB1, with the same intein tag that carries a control target protein (maltose binding protein: MBP), already inserted upstream of the Mxe GyrA intein-CBD with final molecular weight as 71 kDa. The other control vector is pKYB1 with kanamycin resistance gene instead of ampicillin resistance gene as the selection marker. This vector is based on pTYB vectors with mentioned difference (16).

All of these vectors containing one or two chitin binding domain (CBD) as the affinity tag in order to facilitate its

purification using chromatography column containing chitin resin with high affinity to CBD. As described above, CBD can be easily removed after the induction of self-cleavage activity of the related intein (17).

After the introduction of various vectors containing the intein tags and CBD for enhancing the soluble expression and facilitating the purification of target proteins, in this part we have several examples in which various strategies have been used in order to increase the probability of success of these two processes.

### 3.2. Soluble Expression Potential of Inteins-Tagged Proteins

The success of using inteins in increasing the soluble expression of target proteins has been proven (18). However, in various publications, in addition to the use of this tag, measures have been taken to increase the final product in soluble form when they expressed in bacterial expression systems. We have recently had some recombinant proteins tried to express in soluble form with this system. In general, decrease in incubation temperature leads to lower the protein production speed and decrease the inclusion body formation (19). Furthermore, inducer concentration directly increase the expression level of recombinant protein as insoluble form. So, lower concentration of IPTG (isopropylthio- $\beta$ -galactoside) (as inducer in T7 promoter-based vectors) results in increasing in the soluble protein level (19). Based on two proved facts, in trying to the soluble protein expression, lower temperature after the induction and inducer concentration are routine recommendations. In our pervious study, in attempt to produce G-CSF in fusion to the Ssp DnaB intein of the pTWIN1, inducer final concentration was about 0.4 mM of IPTG at 15 °C, while the expression in the optimum condition for *E. coli* culture (37 °C with IPTG in 1 mM final concentration) led to the inclusion body formation and without any level of soluble protein. We used this strategy as a base for the production of other recombinant proteins with IMPACT system (20). Cecropin B2 (21), the human Interleukin-1 receptor antagonist (22) and a mutant form of the soluble receptor of IL-6 (23) are the other examples recombinantly produced in the same condition. The other example in this regard is the expression of DFF40-iRGD as a novel fusion protein with potential targeting efficacy of tumor cells (24). After testing some temperature to detect the soluble expression, we concluded that the soluble

expression only occurred in very low temperature such as 4 °C; actually, incubation at 7 °C led to the soluble expression but in low amounts. However, it seems that the protein size has an important role in solubilization of expressed proteins. DFF40-iRGD is a 40 kDa protein which when fused to the Ssp DnaB intein, become larger and it is obvious that soluble expression is more difficult for larger proteins than smaller proteins (24). In support of this supposition, in the previous example which the protein was smaller in size, the soluble expression was occurred at a higher temperature (20). The other examples on this point is the expression of IL1-RA and a mutated form of IL-6R which were produced in the same system in fusion to intein 1 of Ptwin1 vector and the soluble protein expression was successfully achieved at 15 °C (22, 23). Again, for DFF40-iRGD, the soluble expression just occurred in 0.1 mM of IPTG as the inducer and a few increase in the inducer led to prevent the soluble expression (24). As a general comment, all related publications reported that IPTG in the 1 mM and higher concentrations lead to increase in the total protein expression but as inclusion body. In this case, a recently published paper for the recombinant production of Z domain of protein A of *Staphylococcus aureus* in fusion to the gp41-1-based intein in *E. coli* expression system that was performed at 37 °C and 1 mM IPTG for 5 h (25).

One study in which various parameters were analyzed including the change in the intein tag (Sce VMA1, Ssp DnaB, and Ssp DnaX) resulted in N- or C- terminal fusion, host strain (BL21 (DE3), Origami, and Shuffle) and also the protein accumulation site to recombinantly produce the human growth hormone (hGH). The expression of each fusion protein were induced by 0.4 mM of IPTG or 16% lactose at 22 °C. Because of the presence of di-sulphid band in the hGH 3D structure, periplasmic expression or special bacterial hosts were used. However, the soluble expression in two various strains including BL21 (DE3) or Origami was not statistically different and as a final conclusion the authors said that the expression in the periplasmic space of the BL21 (DE3) is the most suitable strategy for the production of functional hGH (26).

Although the intein usage leads to the soluble expression of recombinant proteins, there are examples in which protein expression was conducted into the inclusion bodies. For example, in the study of Zhang *et al.*, an insulin analogue was produced as inclusion body by

IPTG induction as the 0.1 mM final concentration for 4 h at 37 °C (27).

The other example on this point is the recombinant production of DFFB106 in fusion to the Ptwin1 intein1 in different inducer concentrations as 0.3, 0.5 and 1 mM and various incubation temperatures including 16, 22, and 37 °C. The SDS-page analysis of the whole cell lysate as well as the cell lysate supernatant revealed that the best condition for the most soluble protein expression is with 0.3 mM of IPTG at 37 °C (28).

So, the other variables which were tested for increasing the soluble expression of recombinant proteins in fusion to the intein tags are including host strains, the expression strategy (cytoplasmic, periplasmic or secretory) and so one. *E. coli* BL21 (DE3) pLysS is the other host cell used for the soluble expression of proteins in fusion to intein tags. In the study conducted by Xie *et al.*, for example, pTWIN1 expression vector was used for the expression of OG2 antimicrobial peptide as fusion to the intein 1 or 2. For both expression strategy, the final concentration of inducer was 0.1 and 0.5 mM and in temperature range 20-37 °C. The results of this project showed that the most amounts of soluble expression is in 20 °C and 0.1 mM IPTG (18). It is obvious that with increasing the incubation time the protein amount may be increased. Actually, there was an increase in the production level in 6 h incubation time in comparison to the incubation for 2 h. For this reason, we incubated the induced cells for about 16-24 h to gain the most soluble protein (20, 22-24).

### 3.3. Various Intein Cleavage Strategies

The self-cleavage induction procedure in intein usage is dependent to the type of intein. As explain in the previous parts, there are two different self-cleavage strategies including the change in pH and incubation temperature as well as the addition of reducing agents such as dithiothritol or 2-mercaptoethanol or even cysteine amino acid (29). In general, change in pH is more convenient than reagents usage. This strategy, on the other hand, is not dependent on the first amino acid residue of the target protein linked directly to the intein sequence unlike to the inteins that the self-cleavage procedure induced by the reducing agents addition is completely related to the C-terminal amino acid of the target protein (30). There is several pieces of advice for every type of amino acid residue in the carboxyl terminal of the protein in fusion to these

**Table 1. Self-cleavage induction condition by the addition of DTT according to the C- or N-terminal fusion to Mxe gyrA and Sce VMA inteins, respectively.**

<b>C-terminal residue of the target protein</b>			
<b>Amino acid residues</b>	<b>Best temperature (°C)</b>	<b>Best incubation time (hrs)</b>	<b>The most efficacy (%)</b>
Tyr, Phe, Gln, Asn, Thr, Lys, Ala, His, Leu, Met	23	16	95
Ile, Arg, Glu, Trp, Cys	23	40	95
Val	23	40	90
Gly	23	40	60
Asp	23	40	30
Ser, Pro	4 or 23	Without difference in various time	15
<b>N-terminal residue of the target protein</b>			
<b>Amino acid residues</b>	<b>Best temperature (°C)</b>	<b>Best incubation time (hrs)</b>	<b>The most efficacy (%)</b>
Met, Ala, Gln	23	16	Up to 95
Gly, Leu, Asn, Trp, Phe, Tyr	23	16	Up to 95
Val, Ile, Asp, Lys, Arg, His	23	40	95
Pro	Without difference in various temperature	Without difference in various time	Below to 10
Thr	23	40	Up to 90

inteins. For gyrA, for example, the most suitable concentration of DTT can be used is in 40 mM (10). Furthermore, the temperature and incubation time is related to the final residue of the target protein. **Table 1** summarized the best condition for the most cleavage percent based on this regard in the case of temperature and incubation time. In this review, we are not going to present all the peptides and proteins produced by this system. Rather, our goal is to provide examples in which special conditions are used to increase the cleavage efficiency of the used intein. We previously purified BR2 antimicrobial peptide, a cancer specific cell penetrating peptide which was synthetically produced for the first time. We used recombinant DNA technology and intein mediated purification by gyr A intein under the DTT treatment and 37 for about 40 h because of the presence of an arginine residue at its carboxyl end (31). The other example on this point was the purification of Entrocin P as inclusion body which was finally induced for cleaving by a buffer containing 50 mM DTT at room temperature, overnightly (32). The other strategy in which only change in pH and incubation temperature leads to induce the self-cleaving activity of inteins in for example Ssp dnaB is

simpler and more convenient (10). However, the most important concern in the usage of these types of inteins is the cleavage induction even in the fusion protein expression stage especially in higher temperatures. This mentioned issue can be prevented by reducing the post incubation temperature and also precise adjustment of pH in *E. coli* culture to 7.4. The recommended conditions for the self-cleavage induction of these inteins is only a shift in pH by two units in cleavage buffer and incubation at room temperature for 24 h. This strategy was used in our study in order to purify Anakinra (22). As mentioned, only with this cleavage buffer we could gain the target protein as about 2.3 mg. L<sup>-1</sup> of medium culture. Again, for our other recombinant protein, IL1RA-ABD, the similar condition was used with an extraordinary efficiency as about 4.5 mg of the purified protein per each liter of culture media (33). However, in our previous experience, in an effort to produce G-CSF, we concluded that with increasing in the incubation time of resin column at room temperature, the final amounts of final protein can be increased. Actually, after 48 h of incubation time, we could earn the target protein with final concentration as about 4.7 mg. L<sup>-1</sup> (20). In another project, various cleavage conditions includ-

ing two different temperature and incubation time (4 and 25 °C for 3 or 24 h of incubation) were used and the results showed that when the column was incubated at refrigerator, increasing in the incubation time (from 3 to 24 h), leads to the more cleaved protein. However, for the incubation at room temperature, there was no difference between the final cleaved proteins in two these incubation time. On the other hand, various pH of cleavage buffer were surveyed for removing the intein 1 and it was concluded that there is no difference between incubation of resin column with pH 7 or 5.6 cleavage buffer (34).

On this point, the other study with the aim to recombinantly produce a mutant form of soluble interleukin 6 receptor in fusion to Ssp dnaB intein, the best condition for its self-cleavage was surveyed for three independent variables including cleavage buffer pH, incubation temperature, and incubation time. The best condition for the most purified protein was determined as pH 4 at room temperature and for 24 h. Actually, the usage of more incubation period didn't lead to the increase in the final protein concentration (23).

There is another example in which a beta defensin peptide (DFFB118) was produced by intein mediated purification in to the intein1 of Ptwin1 vector. A range of pH from 3 to 8 was used for inducing the auto-cleavage activity of intein at room temperature. The results of this project confirmed that the optimal pH for this purpose ranges from 5.0 to 5.5 with the final yield as approximately 0.5-1 mg for liter of LB broth (35).

An important note in the use of the inteins which are cleaved by the change in buffer pH is this fact that the cell lysis and extraction of soluble protein in lower temperatures leads to the minimum un-expected intein self-cleavage. In Amaranto et al study, for example, in order to purify the recombinant growth hormone by this system, the mentioned note has been noticed; cell lysis by the sonication was performed at 4 °C (25).

The other strategy which can help to more efficiently cleave of inteins from the target protein is the usage of detergents especially tween-20 and triton-X100 in the cleavage buffer (24). We previously used a cleavage buffer containing two these detergent in 0.2 and 0.3 % final concentration, respectively, and saw their wonderful effects in increasing the final target protein resulted in this system. Actually, without adding two these reagent to the cleavage buffer, we couldn't see the cleaved target proteins including BIF1-iRGD and

DFF40-iRGD from intein 1 in the SDS-page. Only after the addition them in the mentioned amounts, a noticeable amounts of purified proteins were produced (24). According to the IMPACT protocol, the addition of 0.1–0.5% Triton X-100 or 0.1–0.2% Tween-20 to the column and cleavage buffers can improve the binding or cleavage of target protein as well as improve the fusion protein solubility.

Triton X-100, in the final concentration as 0.1 % V/V was added to the column buffer for the purification of beta defensin 6 (DFFB106). However, the role of this detergent in the washing buffer is different from one in the cleavage buffer. Actually, Triton X-100 help the fusion protein to attach more specifically to the chitin column by CBD. In the mentioned study, cleavage buffer with pH ranged from 4 to 7.5 was used to cleave the intein 1 and finally, it was concluded that the optimal pH for auto-cleavage of intein-DEFB106 ranged from 6.0 to 6.5. In this condition about 80% of DEFB106 fusion protein was cleaved (28).

For the other study in this case, on the other hand, in attempt to produce the purified DFFB136, which above 90% of the target recombinant protein was expressed in soluble form under the 0.3 mM of IPTG and at 16 °C. For this protein, the suitable pH for auto-cleavage of fusion protein ranged from 4.0 to 6.0 and the best one was 5.0 at room temperature (36). The other recent study is the recombinant production of TP4-LYCI in fusion to intein 1 and 2 of the pTWIN-1 vector. For this recombinant protein, pH 6 and 40 mM of DTT for 24 h at room temperature was used to induce the auto-cleavage of intein 1 and 2 (37, 38). These conditions was repeated for the auto-cleavage induction of DFF40-LYCI in the other study (39).

As a final conclusion in this part, it seems that the most preferred pH for auto-inducing the cleavage of Ssp intein is ranged from 5.5 to 6.5; as recommended in the IMPACT manual. And for the inteins which their self-cleavage is induced by the addition of a reducing agent, DTT is the most popular reagent in this way.

#### 4. Conclusion

In the present paper, we reviewed the peptides and proteins attached to the various inteins as urgent tags for protein solubilizing as well as an excellent vehicle for the convenient detachment of purification tags act based on affinity chromatography. We mentioned that CBD is the most prevalent affinity tag, although

it is possible to use any other affinity tags. Inteins are efficient tools for the soluble production of recombinant proteins which favored instead of inclusion bodies formation in bacterial expression systems. Although the use of intein tags is usual for any other expression systems, its benefit in bacterial systems is more obvious due to the more challenges pattern of insoluble expression especially in the higher yields production procedures. Also, we presented the auxiliary solutions for increasing the target protein expression in soluble form especially focused on the change in the culture conditions including temperature after the expression induction, inducer final concentration and the incubation time. In this case, the impact of decreasing the post-induction temperature, and also the usage of less amounts of inducer leads to the more chance of soluble protein production. Actually, in all examples of peptides and proteins discussed in this review, the expression system is *E. coli* BL21 (DE3), the simplest *E. coli* based expression system with the greatest tendency to produce recombinant proteins in insoluble form. However, we know that bacterial expression system change for example to the SHuffle or Origami are well-known strains of *E. coli* with more capability to produce soluble recombinant proteins.

In the cleavage procedure, two general strategies were introduced to cleave the CBD tag from the target protein and several examples in each cases were presented. As shown in the main text, pH exchange and also the addition of reducing agents especially DTT are two main strategies in order to induce the self-cleavage activity of inteins used in the related expression vectors. We briefly mentioned the advantages of each of these strategies over the other. As discussed, the most serious problem in the usage of pH-induced inteins such as Ssp dnaB is its un-wanted self-cleavage in the expression stage which leads to a reduction in the final yield of the target product. However, in some cases, several solutions has been recommended and we presented them. For the usage of reducing agents, on the other hand, the complication of the purification procedure by an extra reagent and impossibility to use in the case of peptides and proteins with di-sulfide band is the main concern. However, we can't deny the advantages of this recombinant protein production system. It seems that with optimizing the expression and purification stages in various cases, we can use it as the most efficient system for the easy production of target proteins using

recombinant DNA technology. Finally, this review article provides a suitable view to clarify the effect of different variables on the soluble expression of proteins and peptides fused to various types of intein and their separation from intein. Because by changing these conditions or adding various reagents, it is easy and economical to gain a partial purified protein.

## References

1. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers Microbiol.* 2014;**5**:172-188. doi:10.3389/fmicb.2014.00172.
2. Sorensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Fact.* 2005;**4**(1):1-8. doi: 10.1186/1475-2859-4-1.
3. Gopal GJ, Kumar A. Strategies for the production of recombinant protein in *Escherichia coli*. *Prot J.* 2013;**32**(6):419-425. doi: 10.1007/s10930-013-9502-5.
4. Paraskevopoulou V, Falcone FH. Polyionic tags as enhancers of protein solubility in recombinant protein expression. *Microorganisms.* 2018;**6**(2):47-63. doi: 10.3390/microorganisms6020047.
5. Arnau J, Lauritzen C, Petersen GE, Pedersen J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Prot Expres Purif.* 2006;**48**(1):1-13. doi: 10.1016/j.pep.2005.12.002.
6. Young CL, Britton ZT, Robinson AS. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnol J.* 2012;**7**(5):620-634. doi: 10.1002/biot.201100155.
7. Rais-Beghdadi C, Roggero MA, Fasel N, Reymond CD. Purification of recombinant proteins by chemical removal of the affinity tag. *Appl Biochem Biotechnol.* 1998;**74**(2):95-103. doi: 10.1007/BF02787176.
8. Freitas AI, Domingues L, Aguiar TQ. Tag-mediated single-step purification and immobilization of recombinant proteins toward protein-engineered advanced materials. *J Advanced Res.* 2022;**36**:249-764. doi: 10.1016/j.jare.2021.06.010.
9. Fong BA, Wu W-Y, Wood DW. The potential role of self-cleaving purification tags in commercial-scale processes. *Trends Biotechnol.* 2010;**28**(5):272-279. doi: 10.1016/j.tibtech.2010.02.003.
10. Coolbaugh MJ, Shakalli Tang MJ, Wood DW. High-throughput purification of recombinant proteins using self-cleaving intein tags. *Anal Biochem.* 2017;**516**:65-74. doi: 10.1016/j.ab.2016.10.016.
11. Wood DW, Camarero JA. Inteins applications: from protein purification and labeling to metabolic control methods. *J Biol Chem.* 2014;**289**(21):14512-14519. doi: 10.1074/jbc.R114.552653.
12. Perler FB. InBase, the New England Biolabs Inteins Database. *Nucleic Acids Res.* 1999;**27**(1):346-347. doi: 10.1093/nar/27.1.346.
13. Chong S, Perler FB. Inteins-mediated protein purification. in: Blot, M. (eds) *Prokaryotic Genomics.* Springer; 2003. p. 172-193.
14. Southworth MW, Amaya K, Evans TC, Xu M-Q, Perler FB. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein.



- Biotechniques*. 1999;**27**(1):110-120. doi: 10.2144/99271st04.
15. Zhao Q, Zhou B, Gao X, Xing L, Wang X, Lin Z. A cleavable self-assembling tag strategy for preparing proteins and peptides with an authentic N-terminus. *Biotechnol J*. 2017;**12**(6):1600656-1600678.
  16. Gwak WS, Choi JB, Han BK, Bae SM, Woo SD. Enhanced Production of Recombinant Protein by Fusion Expression with Ssp DnaB Mini-Intein in the Baculovirus Expression System. *Viruses*. 2018;**10**(10):523-532. doi: 10.3390/v10100523.
  17. Mitchell SF, Lorsch JR. Protein Affinity Purification using Intein/Chitin Binding Protein Tags. *Methods Enzymol*. 2015;**559**:111-125. doi: 10.1016/bs.mie.2014.11.002.
  18. Xie Y, Han F, Luan C, Zhang H, Feng J, Choi Y, *et al.* High-yield soluble expression and simple purification of the antimicrobial peptide OG2 using the intein system in *Escherichia coli*. *BioMed Res Inter*. 2013;**2013**:754319-754324. doi: 10.1155/2013/754319.
  19. Shafiee F, Rabbani M, Jahanian-Najafabadi A. Production and evaluation of cytotoxic effects of DT386-BR2 fusion protein as a novel anti-cancer agent. *J Microbiol Meth*. 2016;**130**:100-105. doi: 10.1016/j.mimet.2016.09.004.
  20. Sima S, Shafiee F, Jahanian-Najafabadi A. Expression and one step intein-mediated purification of biologically active human G-CSF in *Escherichia coli*. *Mol Biol Rep*. 2020;**47**(4):2861-2869. doi: 10.1007/s11033-020-05404-8.
  21. Fang Y, Li S, Hu N, Yang J, Liu J, Liu Y. Study on cecropin B2 production via construct bearing intein oligopeptide cleavage variants. *Molecules*. 2020;**25**(4):1005-1019. doi: 10.3390/molecules25041005.
  22. Adelnia R, Shafiee F. Recombinant Production and One-Step Purification of IL-1Ra in *Escherichia coli* and Evaluation of its IL-1 Antagonizing Efficacy. *Iran J Immunol*. 2021;**18**(2):141-149. doi: 10.22034/iji.2021.89103.1929.
  23. Feghhi-najafabadi S, Shafiee F. Recombinant Production of a Mutant Form of Soluble IL-6 Receptor with Inhibitory Effects against Interleukin-6. *Iran J Biotechnol*. 2022;**20**(1):98-105. doi: 10.30498/ijb.2021.278685.3021
  24. Amrollahi-Nia R, Akbari V, Shafiee F. DFF40-iRGD, a novel chimeric protein with efficient cytotoxic and apoptotic effects against triple-negative breast cancer cells. *Biotechnol Lett*. 2021;**43**(10):1967-1976. doi: 10.1007/s10529-021-03178-y.
  25. Nandy S, Maranholkar VM, Crum M, Wasden K, Patil U, Goyal A, *et al.* Expression and Characterization of Intein-Cyclized Trimer of *Staphylococcus aureus* Protein A Domain Z. *Int J Mol Sci*. 2023;**24**:1281-1297. doi:10.3390/ijms24021281
  26. Amaranto M, Vaccarello P, Correa EM, Barra JL, Godino A. Novel intein-based self-cleaving affinity tag for recombinant protein production in *Escherichia coli*. *J Biotechnol*. 2021;**332**:126-134. doi: 10.1016/j.jbiotec.2021.04.003.
  27. Zhang M, Zhang Y, Wu B, Peng Y, Simair AA, Siegel GW, *et al.* Intein-mediated recombinant expression of monomeric B22Asp desB30 insulin. *BMC Biotechnol*. 2020;**20**(1):1-9. doi: 10.7302/5256.
  28. Xin A, Zhao Y, Yu H, Shi H, Liu H, Diao H, *et al.* Soluble fusion expression, characterization and localization of human  $\beta$ -defensin 6. *Mol Med Rep*. 2014;**9**(1):149-155. doi: 10.3892/mmr.2013.1768.
  29. Li Y. Self-cleaving fusion tags for recombinant protein production. *Biotechnol Lett*. 2011;**33**(5):869-881. doi: 10.1007/s10529-011-0533-8.
  30. Mujika JI, Lopez X, Mulholland AJ. Mechanism of C-terminal intein cleavage in protein splicing from QM/MM molecular dynamics simulations. *Organic Biomol Chem*. 2012;**10**(6):1207-1218. doi: 10.1039/c1ob06444d.
  31. Shafiee F, Minaiyan G, Moazen F, Jahanian-Najafabadi A. Recombinant production and intein-mediated purification of an antimicrobial peptide, BR2. *Inter J Pep Res Ther*. 2017;**23**(4):501-507. doi:10.1007/s10989-017-9583-7
  32. Le TN, Do TH, Nguyen TN, Tran NT, Enfors SO, Truong H. Expression and simple purification strategy for the generation of anti-microbial active Enterocin P from *Enterococcus faecium* expressed in *Escherichia coli* ER2566. *Iran J Biotechnol*. 2014;**12**(4):17-22. doi:10.15171/ijb.1154
  33. Shafiee F, Yazdani A. Recombinant Production of IL-1Ra in Fusion to Albumin Binding Domain for Its Extended Half-Life. *Res J Pharm*. 2024;**19**(3).
  34. Xing L, Wu W, Zhou B, Lin Z. Streamlined protein expression and purification using cleavable self-aggregating tags. *Microbial Cell Fact*. 2011;**10**(1):1-7. doi:10.1186/1475-2859-10-42
  35. Hou J, Liu H-y, Diao H, Yu H. The truncated human beta-defensin 118 can modulate lipopolysaccharide mediated inflammatory response in RAW264. 7 macrophages. *Peptides*. 2021;**136**:170438-170443. doi: 10.1016/j.peptides.2020.170438.
  36. Liu H, Diao H, Hou J, Yu H, Wen H. Soluble expression and purification of human  $\beta$ -defensin DEFB136 in *Escherichia coli* and identification of its bioactivity. *Prot Express Purif*. 2021;**188**:105968. doi: 10.1016/j.pep.2021.105968-73.
  37. Mohammad Pour H, Jahanian-Najafabadi A, Shafiee F. Recombinant Production of TP4-LYC1, A New Chimeric Peptide with Targeted Cytotoxicity to HeLa Cells. *Avic J Med Biotechnol*. 2023;**16**(1):9-15. doi: 10.18502/ajmb.v16i1.14166.
  38. Soltani R, Sadeghi-Koopae F, Shafiee F. The Antimicrobial Effects of TP4-LYC1 Fusion Protein against Multidrug-Resistant *A. Baumannii*. *J Isf Med School*. 2023;**41**(729):615-23. doi: 10.48305/jims.v41.i729.0615.
  39. Shafiee-Ardestani Z, Shafiee F. Production of recombinant DNA fragmentation factor 40 in fusion to an antimicrobial peptide from spider venom and evaluation of its cytotoxic effects. *Res Pharmaceut Sci*. 2024;**19**(1):93-104. doi: 10.4103/1735-5362.394824.