OVERVIEW OF THE RECOMBINANT PROTEINS PURIFICATION BY AFFINITY TAGS AND TAGS EXPLOIT SYSTEMS

R. Norouzi¹,*, Z. Hojati² and Z. Badr³

¹Master of Molecular Genetic, Department of Biology, Division of Genetics, Faculty of Science, Isfahan University, Isfahan, Iran
²Isfahan University Associate professor, Biology Group, Genetics unit, Faculty of Science, Isfahan University, Isfahan, Iran
³Master of Molecular Genetics, Department of Biology, Division of Genetics, Faculty of Science, University of Guilan, Rasht, Iran

Published online: 15 May 2016

ABSTRACT

The advancement in protein expression systems causes yield of each peptide intracellular at least as host. The yield proteins should be purified eventually to provide use possibility or study on them. Therefore rapid and thereby economical purification of an active biological recombinant protein from other cell contents is considered as one of greatest challenges in Biotechnology field.

Purification of target protein or enzyme can be facilitated through the integration of genetic sequences with an affinity tag. In this case the tag to the target protein and the protein are expressed as a single unit and the protein can be isolated through one of purification methods by tag from other cellular contents. The affinity tags are commonly used in research laboratories because they can be used to purify wide range of high

Author Correspondence, e-mail: r.norouzi@sci.ui.ac.ir
doi: http://dx.doi.org/10.4314/jfas.v8i3s.168
Purification. One of the important limits of affinity tags is unintended effect of tag on the natural function of the protein, its physicochemical properties and also the next applications. So separating tag after the purification process of target protein is essential. While the process is done through specific endopeptidase, the incoming cost prevents the technique of large-scale application. In addition, applied peptidase should be separated from protein within isolation process which the same matter increases labor costs further and prevents application of these tags in industrial scale. Therefore proper replacement is emphasized for enzymatic removal of purification tags.

**Keywords:** protein purification; recombinant proteins; self-cleavable tags; Intein tags; non-chromatographic methods.

1. **INTRODUCTION**

Purification of a recombinant protein is done with purposes such as crystallography studies (NMR), immunological properties survey, enzymology and medical applications such as pharmaceutical proteins as well as general studies of protein characteristics such as function, activity and structure (Banki M, 2005. Young C, 2012). Biochemists use different methods for separation of proteins based on solubility characteristic, size and also electrical charge using precipitation with different ammonium sulfate concentrations and chromatographic techniques which perform separation based on size, ionic changes and binding affinity. These methods typically involve several chromatography stages that each of the stages should be optimized for each protein separately. Each of these steps can cause cost consumption, time taking and inevitable reduction in final product amount. In addition, this complexity can delay research on new stage proteins in laboratory scale (Waugh DS, 2005). Purification of a recombinant protein is very expensive on a large, industrial scale, treatment purposes and allocates more than 80% of total yield costs. However these costs may be acceptable for valuable therapy and medication cases but, advance in biology generation requires the use of strategies with lower costs (Banki M, 2005). Recent advance in the field of genomics, proteomics and bioinformatics has been facilitated using recombinant DNA technology to assess the target protein without prior knowledge of the cellular protein position or its function. Simultaneous use of affinity tags with recombinant DNA technique provides possibility of protein simple modification for identification, generation and separation from the host systems. A method which can be easily
used for the purification of any protein is adding affinity tag or Peptide handle into the target protein gene (Fong BA, 2010). Affinity tags can be considered as exogenous sequences of amino acids which have a high binding affinity for a chemical or specific biological ligand. 

The gene encoding the sequences N or C-terminal binds to the target protein and is expressed ultimately with it, in result can be easily isolated and purified through fusion tag into its target (Hearn M. T. W. and D. Acosta, 2001). In the first step the merged protein with the tag is purified and then the tag is isolated through a series of protease enzymes from target protein enzymatically. This simple and reliable technique is for target protein purification with the aim of preserving the natural character and it can be considered as a turning point in biological isolation in research scale (Young C, 2012, Waugh DS, 2005). A progress in this field is obtained through using tag with self-cleavable nature. These tags based on self-cleavable protein elements have been named Intein, do not need into proteolysis stage in merged protein purification. Therefore Intein-based tags provide possibility of autolytic reaction in affinity matrix level that is considered as a simple method to purify recombinant protein. We can control simply purification conditions and tag function through pH and temperature status that is of economic benefits and easiness of the technique. However, there are problems contain precipitation of protein, conformational changes, the structural stability and flexibility such as lower in product availability and potential toxicity effect into the host cell should be resolved while using these systems in microbial hosts for expression of target protein expression (Li Y, 2011).

2. PROTEIN BINDING TAGS APPLICATION

Protein binding tags have many applications, including the following items:


2- These tags recognize areas in protein to detect target protein through techniques such as ELIZA1 (Enzyme –linked immune sorbent assay) and quantitative study (Murby M, 1991). These tags are areas in proteins which the target protein can be detected through techniques such as ELISA and quantitative studies that some tags can react such Pseudo-chaperone and causes protein refold
3-  Some the other tags can increase pharmacokinetics and biology half-life of target binding protein significantly (Nygren, P.A. and M. Uhlen, 1994).
4-  Fusion tags specifically influence on target protein purification by low concentrate of cell contents (Brewer SJ, 1991).
5-  Fusion tags have application potential in larger scale such as help to refresh and purified valuable pharmaceutical proteins and also enzymes (Ford CF, 1991).
6-  Using these tags with recombinant proteins is notable in functional-structural studies in low amount scale (Crowe J, 1995).

**Purified tags effects on target protein**

Using an affinity tag can have different effects biochemical protein profile. The mentioned positive effects are access level improvement into protein product, inhibition of proteolysis reactions, facilitation target protein refold process, protection of antigenicity integrated protein, enhance of solubility in protein as well as increase binding techniques sensitivity for SCFV tags. It can be noted the negative effects are modification the target protein conformation (solubility and activity of protein), a decrease in the availability of a protein product (due to defect in finished cleavage process), inhibit enzymatic activity, changes in biological activity of the protein, structural studies disorder by creating unwanted flexibility in protein, and toxicity and how to separate tags in the final phase of the process is expensive (Young C, 2012. Waugh DS, 2005. Li Y, 2011. Arnau J, 2006).

**Classification of purification tags:**

In general fusion-protein tags can be classified into two main groups based on performance:

-1  Solubility tags: these tags increase solubility, target protein expression and stability. Among these, the mentioned tags are GST, MBP, NusA, Trx and SUMO.

-2  Affinity tags: These tags trap target protein using a binding affinity matrix. Among these, the mentioned tags are Poly His, Flag, HA, Strep, CBP1, CBD2.

Using binding affinity tags is enough alone to express complete soluble but for proteins with low solubility or low expression levels, soluble and binding tags should be used together to achieve protein proper expression and purification. For example, MBP and GST fusion proteins can be purified using binding affinity through Glutathione and amylose Resins effectively.
Tags facilitates purification process through affinity methods

Affinity tags are high efficient tools to separation protein, its differentiation and purification. An epitope is known a short sequence of amino acids that is normally characterized as an antigen or area, antibody attached into. So epitope tagging is technique which in a short sequence such as epitope binds to target protein through DNA alternative methods. This technique has wide applications such as western blot analysis, Immune precipitation (IP), Co-immune precipitation (co-IP), immune fluorescence (IF) and purification through affinity. Epitopes provide possibility to determine the protein size, its concentration, modification type after translation, the interaction between proteins and Intracellular trafficking. In some cases, also are used as instrument to achieve purified product. In general purification tags are divided into two groups Chromatographic tags and Non chromatographic tags (Moon AF, 2010).

Purification chromatographic tags

Purification tags are powerful instruments used to purify wide range of target proteins so that has converted them to appropriate options to determine the main origin policy to design commercial kit of purification protein. In this way, affinity tag binds as a fusion part into our target sequence and is expressed with it. In the next step the tag binds fixed ligand into a solid supportive level tight and specifically and other cell impurity and debris are rinsed and separated.

Affinity chromatography provides purification more than 90% within only one column stage normally. In these ways, the main problem is application Resin with heavy costs, limited binding capacity and low recovery power (Banki M, 2005. Young C, 2012).

The different tags are mentioned as following (Chaga et al, 1999; Lee J. And S-H Kim, 2009; Dong et al, 2009):

- FIAG, HA, c-myc, S-tags, SOFtag3, Halo tags, 1D4, Streptavidine binding tags, Calmodulin binding peptide, Chitin binding domain, Cellulose binding domain

Non chromatography purification tags

For example these tags are highly regarded since need into presence one affinity Resin has been removed. As an example of these tags can be noted Elastin-like polypeptides (ELP). ELP precipitates in response to temperature shift and salty buffer specifically and reversibly. This leads to remove of soluble and insoluble impurity by mechanical instrument such as centrifuge apparatus or filtration. This tag is applied to purify expressed proteins in Ecoli and also transgenic plants yet (Fong BA, 2010). Another system without need to chromatography column
is using intracellular granules Hydroxy Alkanoat (PHA), also among them poly hydroxyl butyrate (PHB) is mentioned as Resine. PHA is yield in wide level from bacterial types naturally. Its encoding gene is expressed in Ecoli heterologous and induces generation of polymeric granules which can replace binding affinity classic Resins and are applied to protein purification. Yield of PHB granules and tagged protein simultaneously using Phasin can isolate produced protein.

From cell content exploit simply after cell lysis and centrifuge. Phasin tag mediates binding of target protein into PHB granules ultimately isolate PHB granules with tagged protein by Phasin within centrifuge (Banki M, 2005. Zhou X, 2011). Annex in B protein tag also is used as purification tag in ELP process similar. These tags participate in concentration 20 Mm Calcium and allow isolating specifically of target protein from other cell lysis components. After isolation, tag is isolated of fusion protein (Fong BA, 2010).

Comparison of generated products amount by different purification methods

Regularly purified product amount is considered using column chromatography affinity tags to 90% and non-chromatographic tags such as ELP and PHB with purification lower about 75% but in some cases also is reported about 95% for these tags.

In general, proper tag type of applied purification is determined by the value and importance of target protein (biological value, pharmaceutical and laboratory studies) and requires certain level of purity. For instance, an affinity tag may be suitable to purify an expensive therapeutic protein because of value and need to drug purification level depend on applied Resin type. However, non-chromatographic tags may be suitable to lower costs in related to enzymatic proteins (Fong BA, 2010. Arnau J, 2006).

Tag exploit systems

All tags, small or large have interfering potential with the biological function of the protein, change in physicochemical properties, immunization and prevention the protein crystallization therefore tag exploit have to be performed in final stage of protein purification. This process of integration has always been considered as the Achilles heel. The importance of the issue is in large-scale production of a protein; the greater the costs are found in removing the tags combined with protein. Isolation of fusion protein tags is done by protein according to 3 methods chemically, enzymatically and the use of self-cleavage tags (Fong BA, 2010. Young C, 2012. Arnau J, 2006).
Chemical methods

The used Chemical substances for this purpose include cyanogen bromide (CNBr), formic acid and amid hydroxyl. Chemical isolation undergoes lower costs but these methods have less specificity than enzymatic cleavage and requires more difficult situation which can induce modifications in protein side chain. In this method a specific residues of Methionine has to be between proteins and tags. The frequency of recognition site within the target protein of these chemicals, limits their application. Thus often such chemical cleavages are used to release small peptides or fusion proteins resistant to enzymatic cleavage. Whereas using splitter factors, there is need to chromatographic additional phase to separate target protein from yield composition of cleavage mixture (Arnau J, 2006. Li Y, 2011).

Enzymatic method

Enzymatic methods are performed to tag removal using two groups of proteases contain exoprotease and endoprotease. Exopeptidases do removal reaction and separation of amino acids of target protein terminal and thus can use them to remove small tags such His6. To provide isolation possibility of tag in final stage of target protein a linker area may place normally between tag and protein natural sequence. The linker area may contain specific sequence to cleavage using endoproteases. In this method tag isolation possibility is provided in specific site accurately. The first mentioned advantage is at least misplaced and unwanted cleavage in target protein and in many cases, quick cleavage and separation of protein under mild condition. The one of mentioned problems for this method is its expensive cost and also limit in availability of the specific proteases are inefficient to industrial production in large scale. In addition, many of the proteases are biologically high active such as thrombin and factor Xa and hae to be isolated from target protein through a separate stage. Notable notes are buffer features and situation reinforces protease activity and its function specificity which might are not suitable for fusion protein and its solubility. It is essential to note output of protease cleavage may be unpredictable in any merged protein. Hynderase stearic or inappropriate of residues in around of cleavage site can disturb cleavage process or low output (Young C, 2012. Fong BA, 2010. Li Y, 2011. Arnau J, 2006). The mentioned proteases are as following:

Enterokinase, Factor Xa Sumo protease, Tobacco etch virus (TEV), Thrombin, 3C, preScission.
Use of self-cleavable tags

Self-cleavable tags are specific group of fusion tags based on manipulation and modification of merged protein condition in order to induction of proteolytic cleavage. The tags fusion protein can impress reaction condition of self-cleaving tag by molecular weight and modifications in protein final conformation. In comparison of other proposed options, the mentioned tag provides purification possibility of target protein in one stage. The mentioned tags are as following (Achmuller et al., 2007): FrpC, Cysteine protease domain (CPD), N-Terminal protease (NPRO), Intein SrtAc

Intein

In comparison with methods based on using proteases to extract tag, the Intein based on approaches need lower stages that are important economically and time aspect. Inteins are self-cleavable protein elements have unique ability in extract self of advanced protein outside and facilitate binding the protein fragments together (Banki M, 2005). The target protein can fused in both terminal of an Intein. However sometimes two approaches can have different results of obtained product. For example, as Intein RecA is fusion of N-terminal the yield protein is ATPase inactive dependent to DNA. For another example, a product is made in result of fusion anti-microbial SMAP-29 peptide into Sce Intein N-terminal that has no neutralization activity of toxic peptides and not living clone is not differentiable. In contrast, binding of the peptide into Intein C-Terminal covers effectively dangerous effects of toxic peptides and the protein of interest also is expressed in high level. Studies shows cleavage in both C and N terminal can be affected significantly through residues of Intein around areas. For instance, as the target protein is fused of C-terminal into Sce Intein, if the protein in vivo terminates into Aspartic acid, glutamic acid, arginine, histidine and or threonine, undergo uncontrolled cleavage. In contrast, proteins in C-terminal contain Aspargine, cysteine and or Prolin, encounter inefficient cleavage in non-living medium. In addition different Inteins may have different effects. Therefore select of best condition depend on series of factors. Also as alternative, residues can enter to beside of cleavage area to reform unexpected and uncontrolled cleavages in vivo and low efficiency in non-living environment. As a key note, to aim gain an appropriate cleavage dependent functional Inteins, the fusion protein must be expressed in a soluble environment and also be folded correctly. In general, Inteins causes lower of solubility of target protein and the detagged protein with Intein shows higher solubility level.
The culture in medium with low temperature is considered to overcome the problem as robust strategy widely. If there is not possibility expression in soluble solution, the target protein should gain itself correct folding prior cleavage reaction performance. Regardless of need to additional stage of refold, the expression in insoluble environment is considered some times as advantage to protect fusion protein of intracellular proteases and inhibits of immature cleavage in vivo. The cleavage by Inteins occurs in among Intein and target protein. The full length of gene can put Intein beside without any residue or either additional sequence, using specific incision site and or Overlapping PCR to fusion Intein in both N and or C terminal of protein, in addition, the proteins fusion C- terminally has an additional Methionine in itself N terminal. In addition as mentioned, the cleavage efficiency is impressed amino acids beside Intein terminal widely. In several reports has been seen N- terminal of target protein can inhibit cleavage in C terminal. Therefore adding additional amino acids and or mutation in the main amino acids pare proposed as a resolve in order to increase of cleavage efficiency. So however Intein system has potential of yield protein with itself natural sequence leads to diversity in product. In another report, inhibition of Intein cleavage process has been shown whereas around amino acids was suitable to cleavage area; this process can be inhibited through stearic hyndrase. In addition to cleavage remove or defect in cleavage occurrence, incision is reported in unwanted sites. Another defect of Intein system is uncontrolled cleavage in vivo which can degrade fusion protein within expression process and decrease availability amount to product (Banki M, 2005. Waugh DS, 2005. Li Y, 2011. Arnau J, 2006. Smyth DR, 2003).

In general, Inteins are divided into two groups

1- Inteins mediate cleavage in C- terminal region affected by pH and temperature
2- Inteins affected by Thiol mediate cleavage reaction in both itself C and N terminal region (Young C, 2012).

Inteins which mediate itself cleavage process impressed of pH are more economic because of to mediate cleavage reaction, which just needs to a shift in application buffer pH. Ssp DnaB Intein which is available as part of produced PTWIN1 and PTWIN2 vector by new England biological laboratory is mthe most applied and most well-known Intein impressed by pH shift. The tag is applied associated with many of affinity tags and also non-chromatographic tags. One either tag impressed self-cleavable by pH is ∆I-cm mini Intein which is derived of a ricaminase Intein of Mycobacterium tuberculosis (mtu RecA). This Intein is applied successfully to purification of
fusion proteins using non- chromatographic tags such as ELP, Phasin and also chromatographic tags such as CBD, MBP (Li Y, 2011). One of the problems in related to this group of Inteins is early and immature cleavage of target tagged protein within expression stages. Thiol mediated Inteins in vivo environment show lower immature cleavage (Banki M, 2005. Young C, 2012. Zhou X, 2011). The mentioned these Inteins is sceVMA which is obtained from Saccharomyces Cerviceae. Cleavage in N-terminal region of these Inteins category is induced by compounds such as 2- Mercapto ethane sulfonate sodium (MESNA), Hydroxyl Amin, Thio phenol, betamercapto ethanol, 1,4 di- thio thritol and or free Cysteine. SceVMA Intein is available commercially as part of IMPACT Kit and is used normally with binding affinity tag to kitin. This kit provides possibility of fusion of target protein to C terminal region and also associates with short peptide sequence to N- terminal Intein. Adding Thiol induces cleavage in both terminal of Intein and finally this peptide short sequence is removed through dialyzes. Also IMPACT kit contains mxe GyrA modified Intein which its N- terminal cleavage is in response to add Thiol. Despite of mentioned advantage for lower immature cleavage of this Intein category disadvantages are considered also for this category of Inteins. Many of Thiol compounds are toxic and their separation can increase purification costs in final stage. In addition, Thiol compounds can interfere with Di- sulfide bond which this leads to inappropriate purification for many proteins. In general, for target proteins sensitive to reagents and Reduction reagents, cleavage inducing temperature and pH condition is proposed (Young C, 2012. Waugh DS, 2005. Li Y, 2011).

Fig.1. Overview of structure of 5 self- cleavage tags
a) the main structure of fusion Intein, in this structure Intein N-terminal is bond to target protein.
b) srtAc fusion tag structure contain Histidine sequence in N-terminal and is bond to target protein N-terminal by mediated LPXTG motif. C) NPro fusion structure to target protein N-terminal whereas X of any amino acid may be except Proline.
d) FrpC fusion structure which second SPM causes catalyze of cleavage reaction in Asp414-Pro415 position. This tag pairs the second bond with kitin and or his-tag and after degradation releases target protein associated with additional Asp in C-terminal.

Comparison of different approaches of fusion tags isolation

Protein purification method are very expansive and diverse based on cleavage efficiency, control enable of cleavage reaction, reaction performance condition and yield product modification. Every these factors can affect time factor, economic value and value potential of protein purification assumption. Every factors has to be considered in select time of best appropriate method with target protein (Fong BA, 2010). The required time for performance cleavage reaction dependent proteases, Inteins, FrpC and SrtAc is about 4 to 8 hours. However some proteases such as Tobacco Etch Virus protease (TEV) and Di Peptidil peptidase (DPPI) can complete cleavage reaction within only 1 hour. This matter has converted protease dependent tags into notable options in while the time factor is one of the main limitations. subtilisin. Reports have expressed 80% cleavage efficiency for similar Intein FrpC in related to cleavage rate. This rate is expressed about 95% for proteases. Control enable of cleavage reaction is considered one of main cases in tag selection because of unwanted and immature cleavage causes lower in achieving product. In all considered tags among, Intein tags performing cleavage reaction impressing pH shift, has minimum control on cleavage reaction and contain high level of immature cleavage in expression stages (Banki M, 2005). Dependent tags into protease are most controllable considered case because of not adding protease to environment; the cleavage does not occur but as mentioned using these tags has difficulty. If a tag control enable of cleavage reaction is significant as the main characteristics in select of a tag, Thiol Inteins and or FrpC protein can be a suitable selection. FrpC protein cleavage is controlled intensely and also Thiol
affiliated Inteins show low immature cleavage. In both mentioned cases, the cleavage reaction is more inexpensive and more specific intensively in compartment. For target unstable proteins with Di-sulfide bonds, The cleavage reaction condition can be considered as either factor in selection method of interest to tag extraction. For example, in cases which cleavage reaction is impressed by pH shift, it is necessary the target protein is kept several hours in room temperature and in expose of pH: 6/5. FrpC is cleaved in room temperature but needs DTT and EDTA in purification buffer which may lead to incompatibility with target protein with Di- sulfide bond and or needed to metal cofactors. Similarly, cleavage condition for protease tags is rather mild (temperature 4 to 30 degree in pH: 5/5, 8/5 for TEV, Temperature 37 degree in pH: 7/2 to 7/5 for Subtilisin and 37 degree in pH:8 for DPPI). In addition, one of most important is has to be considered is obtained modification on target protein. For instance, any abnormal amino acid should not be added into pharmaceutical protein in purification time because even one additional amino acid can lead to immune response and changes in pharmaceutical properties of pharmaceutical protein. Some proteases remain one or two amino acid in beside target protein in incision site position such as Thrombin, TEV and PreScission. The same matter is also considered similarly for FrpC and SrtAc. In contrast, any Inteins, Thiol Inteins and of pH impressed do not modify on itself target protein at al. However, their cleavage reaction is impressed by beside amino acid of incision site in target protein, thus before selection of an appropriate strategy to purify, all mentioned cases must be considered (Arnau J, 2006).

3. DISCUSSION AND CONCLUSION
In this study, specific advantages and disadvantages of each fusion tags with protein was considered. Application of each considered protein dependent to our final target is of purified protein. Some tags causes just target protein solubility enhancement and some other inducing protein expression accordance participation particles causes’ protection protein of intracellular proteases. The tags additional, another group facilitate purification process of recombinant protein by mediate binding affinity to resin or independent of it. Purification tags are robust instruments are applied to purify wide series of target. The fusion tags should be isolated biological effects and physicochemical changes in proteins in final stage. The fusion tag removal is done chemically, enzymatically and auto catalysis activity dependent such as Inteins. Inteins are considered as the first self- cleavage fusion systems and are used to yield diverse proteins in
laboratory scale. In comparison of common purification methods based on using proteases, Intein dependent process causes decrease in stages and costs. It is predictable the Intein dependent purification systems improve the expression condition of fusion recombinant protein in large-scale. Despite high potential and high costs for reform of these systems defects, there is not any report based on using Intein dependent systems in industrial large-scale.

Properly two great barriers limit Inteins application. At the first, Inteins decrease expression level and target protein solubility which is unfavorable for their application to yield insoluble and low expressed proteins. The second, Intein systems may causes uncontrolled cleavage in vitro medium and or incomplete cleavage in non-living environment which limits availability to product. Therefore the main focus has to be solubility and cleavage reaction control aspect in order to improve into Inteins efficiency.

4. ACKNOWLEDGEMENT
The article is provided using part of the Master of Science thesis course. Hereby is friendly appreciated of Biology group of Isfahan university for provision facilities of this study.

5. REFERENCES


**How to cite this article:**