A publication of
The Williamsburg BioProcessing Foundation

Fall 2006

BioProcessing™ JOURNAL

The Most Trusted Source of BioProcess Technology™

Vol. 5 No. 3

www.bioprocessingjournal.com

A Universal Platform for the Purification of Therapeutic Proteins Using Affinity Tags: *The Use of Engineered Aminopeptidases for His-Tag Removal*

By JOSÉ ARNAU* and JOHN PEDERSEN

Current expression technologies have enabled the production of thousands of recombinant proteins in diverse production hosts. Therapeutic recombinant proteins have been engineered for a variety of purposes including reduced antigenicity, longer half-life, simplified process development, and increased affinity. Protein engineering has relied on various high throughput methods (*e.g.*, directed evolution, phage display) to identify candidate proteins with the desired therapeutic properties. The physiological and biochemical diversity of native and engineered proteins reflects on the abundance of production hosts, expression tools, and different approaches for protein purification.

Notably, a key step in high-throughput protein production is purification, which is a bottleneck where large numbers of samples are involved. Universal purification methods that can be applied to virtually any protein, and that are amenable to automation, can be used to address this problem.[25]

The purification process uses chromatography to separate the protein from the remaining medium and cell components present in the fermentation broth. Typically, chromatographic purification must include several consecutive steps, such as an initial capture step, where the proteins present in the crude extract are bound to the matrix followed by gradual retention/elution of the protein of interest. Several subsequent chromatographic steps are needed to obtain a relatively pure protein. Product loss is inevitably associated with each chromatographic step and typical yields are normally around 30-50%. Accordingly, purification accounts for more than half of the total manufacturing costs for a recombinant protein. A current trend in therapeutic protein production is

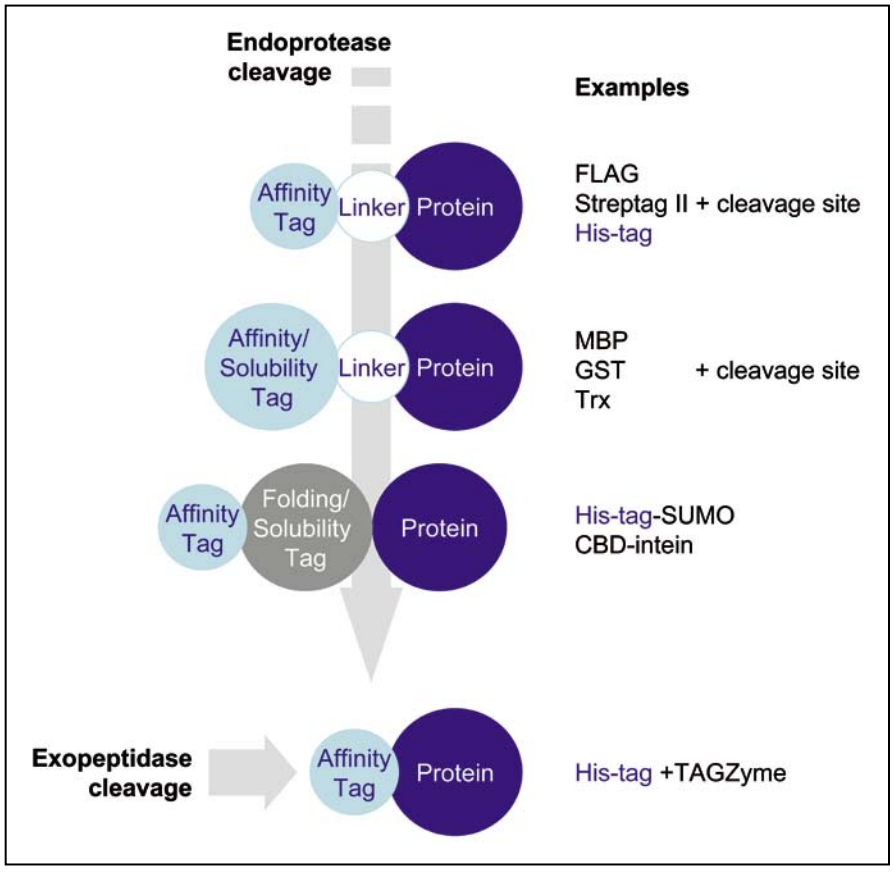

Figure 1. Different Strategies for Affinity Tag and Tag Removal. Various approaches for the design of tagged fusion proteins are shown. Some tags are used to enhance the solubility of the protein and can be also used as affinity tags (*e.g.*, MBP), while other solubility tags require an additional tag for affinity (*e.g.*, SUMO). The use of endoproteases for the cleavage of the tag (the vertical arrow) requires the inclusion of a specific cleavage site in the fusion in a linker region between the tag and the protein. Problems associated with unspecific cleavage are discussed in the main text. Using exopeptidase cleavage (horizontal arrow), just a short affinity tag is required and can be used independently of the protein sequence (*e.g.*, TAGZyme). Examples of different tags are shown to the right.

José Arnau, Ph.D. *(ja@unizyme.dk), is business development manager; and John Pedersen, is research director; Unizyme Laboratories A/S, Hørsholm, Denmark. *Corresponding author.*

to develop ways to reduce costs. One alternative for developing more efficient and economical purification processes is to use affinity tags.

Why Affinity Tags?

Affinity tags are exogenous amino acid (AA) sequences that display a high affinity for a specific biological or chemical ligand. A major group of affinity tags consists of a peptide or protein that binds a small ligand linked on a solid support (*e.g.*, his-tags bind to immobilized metals or maltose binding protein and amylose columns). Another group includes tags that bind to an immobilized protein partner, such as an antibody, or antibody purification using protein A affinity chromatography.[1]

Affinity chromatography is frequently used for projects where numerous proteins are produced, due to the high selectivity of the technique. By using an affinity tag fused to the protein of interest, it is commonly possible to achieve over 90% product recovery; moreover, fewer chromatographic steps are required. For some nonclinical purposes, the recombinant tagged protein can be purified in a single step.[13,16]

Many affinity tags have been developed for protein purification in the last three decades and novel tags are still emerging.[1,7] Some short tags like FLAG and Strep-Tag II (Figure 1) are solely used to enable affinity purification. Larger tags can also be used to enhance solubility (*e.g.*, maltose binding protein, glutathione-S-transferase, or thioredoxin) (Figure 1), to enable solubility and folding of otherwise intractable proteins[4], and/or to allow for the production and purification of toxic peptides as fusions to non-related proteins—as unusual as the insect virus polyhedrin protein.[28] Many refinements of the technology have been developed. One example is the use of intein fusions. Inteins are self-cleavable endoproteases, and vectors are available to enable the production of CBD (chitin-binding domain) intein-protein tripartite fusions. The fusion proteins are first bound to a chitin matrix followed by on-column activation of intein-based cleavage to release the protein of interest.[19] Choosing and/or screening for the most appropriate affinity tag can be a difficult task.

His-tags are the most widely used affinity tags, with over 60% of recombinant proteins produced with a his-tag for structural studies.[11] The use of his-tags for affinity purification has several advantages, including a simple purification scheme, and the reduced likelihood of conformational or biological effects due to the short length of this tag.

His-Tags

Purification of his-tag proteins is based on the use of chelated metal ions as affinity ligands. The metal ion is complexed with an immobilized chelating agent, a technology developed in the early 1970s.[23] Protein separation occurs mainly through the interaction between surface-located histidines in the protein and the metal ion within the immobilized metal chelate. The imidazole side chain of histidine displays a high affinity for chelated metals. The strong binding affinity allows for the purification of proteins under both native and denaturing conditions. Furthermore, the chemical, non-biological nature of the immobilized metal affinity column (IMAC) material allows for more intensive regimes for repeated use.

The use of his-tags, typically located at either the N-terminus or C-terminus of the protein of interest, enables purification of the protein directly from the crude extract of host cells in a single IMAC step. However, caution must be taken about non-specific IMAC binding proteins with histidine or cysteine stretches on their surface, which may copurify with the protein of interest (see below).

Different his-tag amino acid sequences have been used. One of the most widely used is a six-histidine stretch (6xHis), although his-tag sequences with six alternating histidines in different combinations have also been widely used. The binding affinity of the his-tag can be modulated by including, for example, hydrophobic AA in the tag.[2] A 6xHQ tag has been optimized both for high-level expression of proteins in *E. coli* and for effective tag removal by

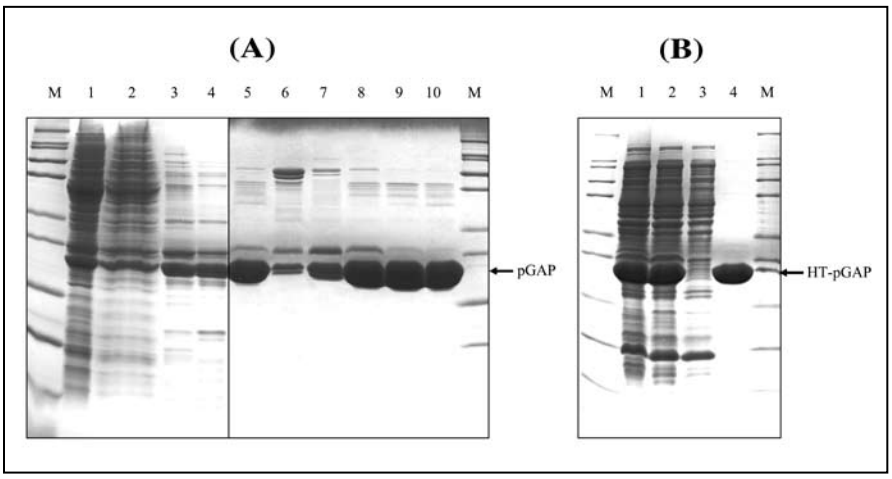

Figure 2. Comparison of Purification Strategies for B. Amyloliquefaciens pGAP Produced in *E. Coli*. The recombinant pGAP was produced with and without an N-terminal his-tag (HT-pGAP, tag sequence: MEP(H6)L). For untagged pGAP, purification included ammonium sulfate precipitation and two consecutive separation steps using phenyl-Sepharose. Subsequently, a desalting step using a Sephadex G-25 F column and a final step using Q Sepharose HP were performed. For HT-pGAP, purification was performed with a single IMAC step. (A) Standard purification of pGAP. Lane M: MWM (Novex); lane 1: cell extract; lane 2: supernatant fraction of cell extract; lane 3: pool from first phenyl-Sepharose step; lane 4: pool from second phenyl-Sepharose step; lane 5: pool after desalting; lane 6-10: fractions from Q Sepharose HP containing pGAP. (B) IMAC purification of HT-pGAP. Lane M: MWM; lane 1: cell extract; lane 2: supernatant fraction of cell extract; lane 3: flow-through fraction from the IMAC; lane 4: eluted HT-pGAP. (Reprinted with permission from Elsevier.[1])

exopeptidase cleavage.[22]

Numerous cloning vectors are available for the expression of proteins with N-terminal or C-terminal his-tag in E. coli and other hosts (yeast, insect cells, mammalian cells). Customized his-tag sequences can also be designed to incorporate a protease cleavage site or engineered to improve expression in the host of choice.[6] Similarly, his-tag sequences can be placed downstream a signal peptide to enable recovery of secreted proteins using IMAC.

Standard vs. Affinity Purification of a Recombinant Protein Produced in *E. Coli*

Pyroglutamyl aminopeptidase (pGAP) from *Bacillus amyloliquefaciens* is an enzyme that catalyzes the removal of N-terminal pyroglutamyl residues from peptides and proteins.[22] A modified version of this enzyme is used in the final step of N-terminal tag removal, by the TAGZyme process, to produce his-tag-free proteins (see below). Purification of recombinant pGAP produced in *E. coli* using the native protein sequence and standard chromatography requires several steps: ammonium sulfate precipitation, two phenyl sepharose columns, a desalting step, and a final separation on Q Sepharose HP. Yet, pGAP purified in this way still contained a number of protein contaminants (Figure 2). Furthermore, the yield for this purification process was 40%; measured by enzyme assay.[1]

In order to improve purity and yield, the same protein was fused to an N-terminal his-tag (HT-pGAP; tag sequence: MEPHHHHHHL) and produced in *E. coli.* HT-pGAP purification was carried out using a single IMAC step, yielding a pure protein (Figure 2). The yield of this process was 96%. Remarkably, no contaminating proteins were evident in the final protein preparation after a single chromatographic step (Figure 2).[1]

The choice of affinity tag depends on the intended use of the protein and also on the biochemistry of the protein of interest. The nature of the affinity tag will have an effect on the design of the purification process. Importantly, especially for proteins intended for clinical use, the presence of exogenous sequences in the final product is undesirable. Also, some tags interfere with protein structural analysis and biological assays.[1] Therefore, a number of enzymatic methods have been developed to remove the affinity tag after the initial affinity purification and to produce proteins with their native sequence and authentic N- or C-termini.

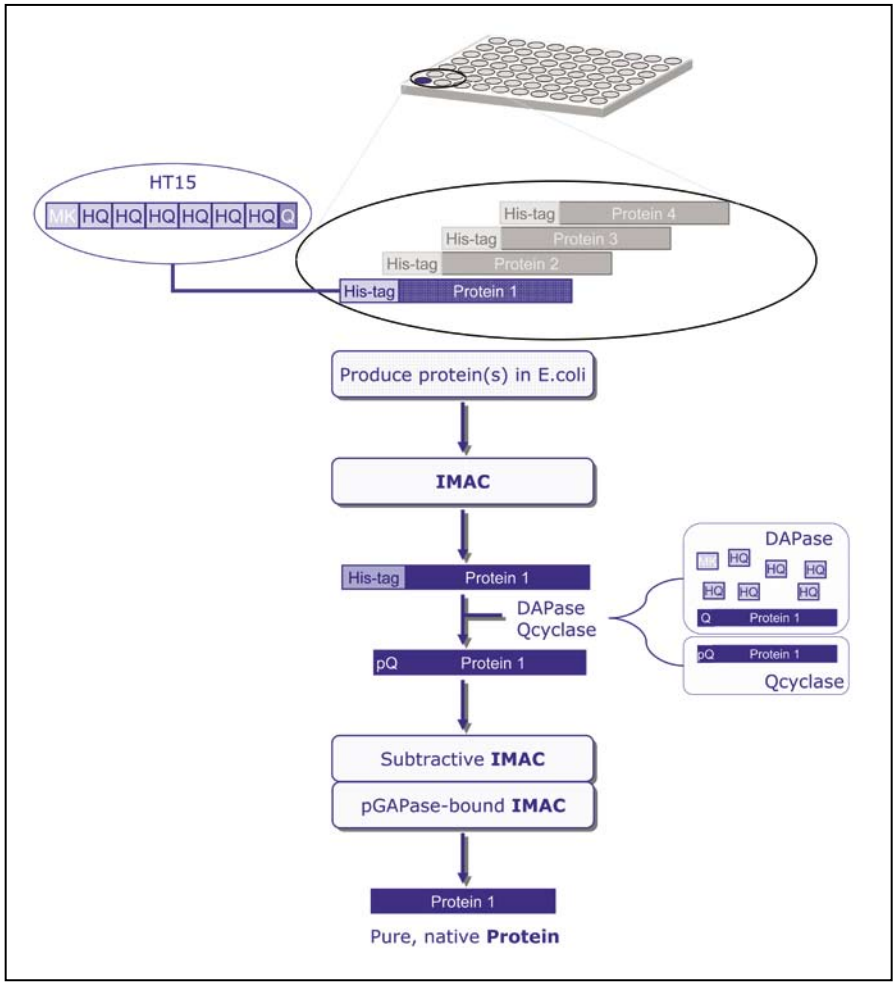

Figure 3. A Universal Strategy for Purification of His-Tag Proteins in *E. Coli.* For high throughput genomics and proteomics projects where numerous different proteins need to be produced and purified, a universal approach can be used for the recombinant, tag-free protein using, for example, the HT15 his-tag (MKHQHQHQHQHQHQ) and IMAC purification followed by his-tag removal using TAGZyme. Briefly, the sequence encoding the HT15 his-tag can be added to the expression vector. After isolating the appropriate collection of E. coli clones, the proteins (*e.g.*, proteins 1 to 4) can be produced at small scale (*e.g.*, in microtiter plates, top) in an automated platform.[25] Cell extracts are directly applied to an IMAC column to obtain the purified his-tagged proteins (*e.g.*, his-tag protein 1). Subsequently, the his-tagged proteins are cleaved with DAPase and Qcyclase, resulting in proteins with a pyroglutamyl residue at the N-terminus. Finally, the mixtures are applied to a second IMAC (subtractive IMAC) column where DAPase, Qcyclase, unspecific IMAC binders and unprocessed molecules are retained. This second IMAC column is set up in connection with a final IMAC containing bound pGAPase (pGAPase-bound IMAC), allowing the removal of the pyroglutamyl and the elution of the tag-free proteins.

Protein Setup Required for the Use of an Affinity Tag and Tag Removal

In the above comparison, the affinity tag was positioned on the N-terminus of the protein of interest. This may represent an advantage for the production of recombinant proteins in bacterial hosts if the tag favors high expression levels due to the coupling of transcrip-

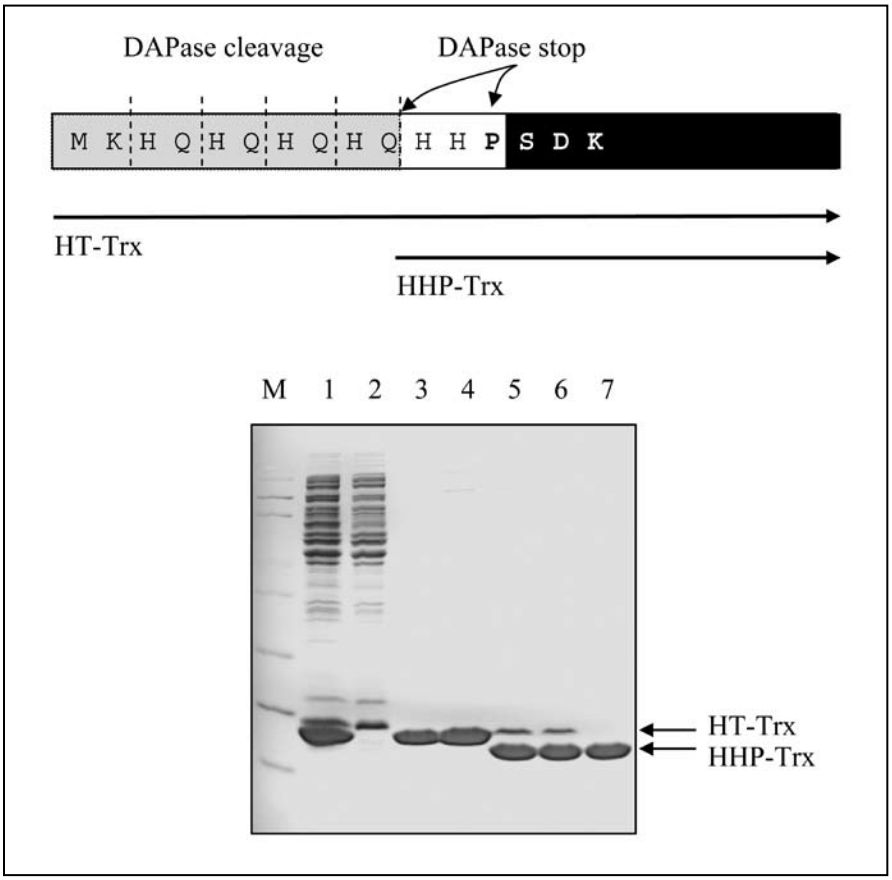

Figure 4. Purification of Recombinant HHP-Trx Using DAPase. HT-Trx was purified from an approximately 2.5 L *E. coli* culture. An overview of the genetic design and the cleavage process is shown at the top. After initial IMAC purification, DAPase cleavage sequentially removes the first five dipeptides (stippled lines depict the position of cleavage, and the grey box, the sequence removed by DAPase), until a P is found at position 3. In this case, DAPase cannot cleave the HH–P bond, resulting in HHP-Trx. The cell extract (in approximately 120 mL of buffer A: 20 mM NaH2PO4, pH 7.5, 300 mM NaCl, and 20mM imidazole) was applied to a Ni2+-chelating Sepharose FF column (5.3 cm x 12 cm) with a flow rate of 5 mL/min. Subsequently, a wash step was performed with 250 mL of buffer A using the same flow rate. A linear gradient from buffer A to buffer B (20 mM NaH2PO4, pH 7.5, 300 mM NaCl, and 1M imidazole) was used for 90 minutes at a flow rate of 2.5 mL/min. Subsequently, an additional 10 minutes with buffer B was used for elution of HT-Trx from the column. Relevant fractions were pooled (50 mL), EDTA (5 mM was added) and the sample was desalted using a Sephadex G-25 F column (5.3 cm x 30 cm) using TAGZyme buffer C (20 mM NaH2PO4, pH 7.0, 150 mM NaCl) at a flow rate of 4 mL/min. The pooled fractions (65 mL) containing 7.3 mg/mL HT-Trx were diluted to 3 mg/mL (474 mg in a final volume of 156 ml). For tag removal, 35 U of DAPase was mixed with 0.3 mL of 200 mM cysteamine in approx. 1.5 mL and pre-incubated. The HT-Trx was pre-incubated with 1.56 mL of 200 mM DTT at 37 °C for 5 minutes. Subsequently, the DAPase mix was added to HT-Trx and incubated at 37 °C for 60 minutes. A second desalting step was performed on a Sephadex G25 F column (19.6 cm x 27 cm) using buffer D (20 mM NaH2PO4, pH 7.5, 300 mM NaCl, and 15 mM imidazole). Imidazole was included, in this case, in the wash buffer to enable elution of HHP-Trx, and fractions containing detagged HHP-Trx were pooled (205 mL containing 443 mg protein). Removal of DAPase was performed by applying the desalted sample using a Ni2+-chelating Sepharose FF column (2 cm x 11 cm) at a flow rate of 2 mL/min and washing at the same rate with buffer D to collect the flow-through fractions. Finally, the sample was desalted on a Sephadex G25 F column (19.6 cm x 54 cm) with buffer E (10 mM NaH2PO4, pH 7.0) at a flow rate of 25 mL/min and fraction pooling yielded 354 mg HHP-Trx (in 295 mL). Lane M: MWM; lane 1: cell extract; lane 2: first IMAC flow-through fraction; lane 3: first IMAC pool; lane 4: desalted first IMAC pool; lane 5: after DAPase treatment; lane 6: desalted pool; lane 7: elution from subtractive IMAC. (Reprinted with permission from Elsevier.[1])

tion and translation. In fact, analysis of the secondary structure of the 5' end of the mRNA permits optimization of the sequences for expression.[6] Alternatively, the affinity tag can be located at the C-terminus, although affinity tags located C-terminal may be embedded in the protein and therefore function as such only under strongly denaturing conditions.[20]

The general design of fusion proteins includes several important choices including the type of tag to be used, its position in the fusion protein, whether to include sequences enabling tag removal by proteolytic cleavage, and which enzyme to use for tag removal (Figure 1).

In the first example shown, the affinity tag is added to the protein with a linker region containing a specific cleavage site for an endoprotease. After an initial affinity chromatography step, the tagged protein is treated with the endoprotease to cleave the tag and linker sequence. Subsequently, the endoprotease is removed by an additional affinity step, as it also contains an affinity tag. A more elaborate example involves the use of a larger affinity tag that provides increased solubility or facilitates protein folding (*e.g.,* maltose binding protein, MBP and glutathione-S-transferase, GST).

Yet another type of fusion includes an affinity tag, a processing protein that either cleaves the fusion protein or facilitates folding of otherwise intractable proteins (intein or small ubiquitin-like modifier [SUMO] are examples of this), and the protein of interest (Figure 1). The SUMO tag can be removed using a protease that recognizes the conformation of the correctly folded SUMO tag rather than a specific sequence.[4] In both cases, an additional affinity tag is needed for purification.

Finally, a simpler type of construction is required for the removal of N-terminal tags using exopeptidase cleavage. Here, a short tag (*e.g.,* his-tag) added to the protein is enough to allow affinity purification, tag cleavage, and production of a tag-free protein, regardless of the protein sequence. TAGZyme is an example of the most widely

used exopeptidase cleavage of tags (Figure 1).

The advantages and considerations involved in the use of endoproteases and exopeptidases for tag removal are briefly described below. Several reviews illustrate in more detail the different affinity tags and tag removal methods available.[1,15,18,27]

Endoprotease Cleavage of the Affinity Tag

A number of recombinant endoproteases have been developed for tag removal purposes. Enterokinase, thrombin, and factor Xa are amongst the most widely used. In addition, several viral endoproteases (*e.g.,* TEV protease, 3C Protease, PreScission, etc.) can be used for tag removal. Other endoproteases include granzyme B, caspase 6, sortase A, and intein. To enable their use and subsequent removal, all these endoproteases include an affinity tag.

A major concern for the use of endoproteases for tag removal is the substrate specificity and the observed unspecific cleavage at secondary sites in the protein that result in protein degradation. Endoprotease-based tag cleavage is often performed using high ratios of protease:protein substrate, and involves long incubation periods. Under these conditions, cleavage at other sites in the protein is favored. Additionally, some endoproteases, like thrombin or viral endoproteases, leave one or two exogenous amino acids (*i.e.*, not present in the native protein sequence) in the protein after cleavage.[1] The presence of added amino acids to a protein intended for use as a therapeutic may require appropriate documentation about the effect of the added amino acids on the protein, including biological activity and toxicity.

With the above-mentioned constraints for the use of endoprotease-based tag removal, one must ensure that the protein of interest is not cleaved by the endoprotease chosen for tag removal. This represents a substantial drawback for high throughput projects in which many different proteins must be produced, and their biological effect investigated.

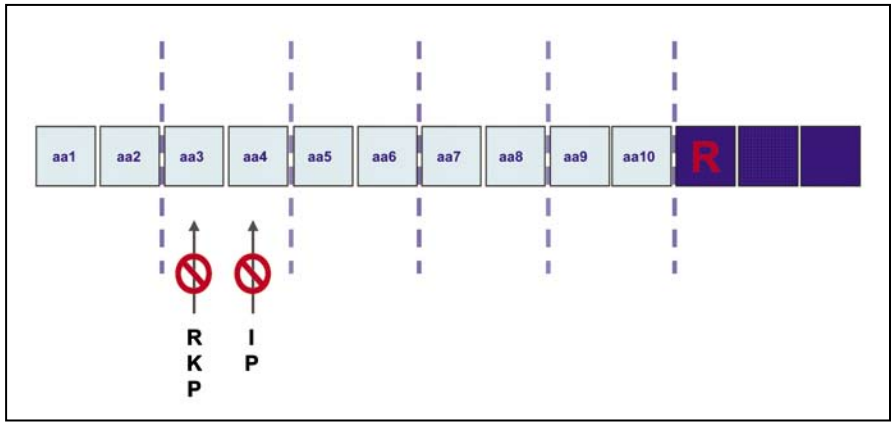

Figure 5. Overview of a His-Tag Sequence and Features Required for DAPase Cleavage. Stop positions for cleavage include R or K at odd positions in the tag, I at even position and P, anywhere in the dipeptide/tag. In the example shown, the native protein sequence (dark blue boxes) includes a suitable stop position (R). Stippled lines depict the position of DAPase cleavage.

Exopeptidase Cleavage of the Affinity Tag Using Engineered Aminopeptidases

An alternative to the use of endoproteases for tag removal is the use of exopeptidases. Although a number of peptidases are available from natural (animal) sources, their use for tag removal from therapeutic proteins is not recommended due to the risk of contamination. Aminopeptidases of bacterial and fungal origin are used for production of current protein drugs.[1,5] DAPase is a recombinant, his-tag version of dipeptidyl aminopeptidase I (DPPI). DAPase is part of TAGZyme, which is a system based on engineered aminopeptidases that enables precise and effective removal of N-terminal short tags, including his-tags.

DPPI cleaves dipeptides sequentially from the N-terminus of proteins and peptides. The past 15 years have produced comprehensive knowledge of DPPI, fueled by the possible use of the human enzyme as a drug target in a number of diseases. The protein structure of this member of the papain family of cysteine proteases has been solved, revealing the presence of the active site on the surface which permits effective cleavage of dipeptides from proteins.[21,26] Interestingly, an elaborate *in vivo* activation mechanism involving processing of the full-length polypeptide into three chains has been elucidated.[8] The presence of an exclusion domain (one of the peptide chains of the mature DPPI) blocks the active site cleft beyond the S2 site and provides the carboxyl group of the Asp1 side chain as the docking residue for the free amino group of the substrate. This biochemical architecture of the active site determines the specificity for dipeptide cleavage and the cleavage efficiency of DPPI for different dipeptides (see below).

DAPase cleavage has been tested on many his-tag proteins of rather different origins. The level of DAPase required for tag removal depends on the tag sequence, but also on protein structure and exposure of the tag. In our experience, his-tagged human proteins (*e.g.*, IL1, TNF or hGH) require a lower amount of DAPase for complete tag removal, whereas bacterial proteins like MBP require nearly a 20-fold higher amount (Table 1). In general, the amount of DAPase required for complete tag removal is 10 mU/mg (0.5 µg/mg) of processed protein for optimized purification processes (Table 1).[1]

Accessory Tagzyme Enzymes: Qcyclase and pGAPase

Qcyclase is a recombinant, his-tag glutamyl cyclotransferase that converts N-terminal glutamine residues present in peptides and proteins to pyroglutamyl

residues. The enzyme has an unusual fivefold propeller structure traversed by a central channel that accounts for the extreme robustness and stability of the protein and its broad substrate specificity.[14] Similarly, pGAPase is a recombinant his-tag version of pGAP.[22]

Overall Features of Tag Removal Using Tagzyme

DAPase is the major component of TAGZyme and can be used alone for his-tag removal in many cases where the protein presents a suitable stop position for DAPase cleavage in the N-terminus. A suitable stop includes N-terminal arginine or lysine, or proline at the second or third position from the N-terminus. Many human proteins include a suitable stop position for DAPase cleavage at the N-terminus, such as a P at position two or three of the mature protein. Examples of these include several interleukins and growth hormone.[1] R or K at position one also represents an effective stop position for DAPase cleavage.[22] Human lactalbumin and lysozyme are examples of this class of proteins. For the above human proteins, a short, even-numbered affinity tag such as a 6xhis-tag may be fused directly to the sequence of the mature protein. Subsequently, tag removal can be effectively performed using DAPase alone, followed by a subtractive IMAC step to elute the detagged protein.

For all other proteins, regardless of their native N-terminal sequence, DAPase is used together with Qcyclase and an odd-numbered tag sequence containing a Q as the last tag residue to cleave the tag and produce a pyroglutamyl-protected protein. As shown in Figure 3, HT15 is an example of a suitable his-tag for TAGZyme cleavage (sequence MKHQHQHQHQHQQ; added Q underlined).[22] The his-tag protein is purified from the crude extract in an initial IMAC step. DAPase cleavage is performed in the presence of an excess Qcyclase. Upon cleavage of the last HQ dipeptide, the N-terminal Q is readily converted to pyroglutamyl, protecting from further DAPase cleavage (Figure 3). Enzyme removal and the final pyroglutamyl cleavage step can be performed using a stacked IMAC and pGAPase-bound IMAC column to directly elute the tag-free protein (Figure 3). The whole purification and tag removal process is fast, simple and amenable for use in high-throughput projects

Custom Tag Design for Imac Purification and Tag Removal of Thioredoxin in *E. Coli*

Thioredoxin (Trx) was chosen in order to study the interaction of metal with histidine in a minimal sequence. A process was designed for the production of his-tag Trx (HTTrx) in *E. coli*. The genetic design included a 10-AA his-tag (MKHQHQHQHQ) for initial IMAC purification, followed by a short HHP sequence adjacent to the first residue (S) of the native Trx (HHP-Trx), as shown in Figure 4.[1] The presence of HHP at the N-terminus of the processed, detagged protein was chosen to enable a detailed study of nickel interaction by paramagnetic NMR (nuclear magnetic resonance) relaxation.[17] The occurrence of a stop position (P) in this sequence allows the removal of the 10-AA tag using DAPase alone, and protects from further cleavage (Figure 4).[1]

The mature protein sequence included two H residues and represented a challenge for IMAC purification where effective elution of HHP-Trx and retention of DAPase (containing a C-terminal his-tag) is needed. In this and similar cases, the subtractive IMAC step is performed using a buffer containing a low concentration imidazole that permits binding of the tag removal enzyme(s) to the IMAC, while precluding binding of the detagged protein (Figure 4). The 10 AA his-tag sequence allows high expression of the recombinant protein in *E. coli*.[22] Additionally, the MK motif present in this tag, as well as in HT15 (Figure 3) serves a double purpose. First, the presence of N-terminal MK results in one of the lowest frequencies of methionine excision in *E. coli*.[10] Second, the presence of K at position two represents a quality control for DAPase processing. Thus, the fraction of HT-Trx molecules where methionine excision occurs is not cleaved by DAPase (an N-terminal K is a natural stop position for DAPase) as shown in Figure 5.[22] Consequently, these molecules retain a functional his-tag and are effectively removed in the subtractive IMAC step. The purified HHP-Trx was successfully used in NMR studies.[17]

In Silico His-Tag and Protein Design

To develop a process for the production of recombinant proteins using his-tag, and tag removal using TAGZyme, it is important to ensure that the sequence of the tag can be correctly and efficiently cleaved. The cleavage of a considerable number of different dipeptide combina-

Table 1. DAPase Requirement for Tag Removal from Selected His-Tagged Proteins of Different Origins.

Protein	DAPase (mU/mg)	Cleavage* (min)	DAPase** required for 1 h cleavage (mU/mg)
IL1β (human)	50	15	12.5
TNFα (human)	50	20	16.6
GH (human)	50	30	25
DNAK (human)	50	30	25
HHP-Thioredoxin (bacterial)[1,17]	75	30	37.5
MBP (Maltose binding protein, bacterial)	200	60	200
GCT (Glutamine cyclotransferase, plant)	200	60	200

*No unspecific cleavage observed using 1000-fold DAPase and overnight treatment.
**Typical requirement for optimized processes is 0.5 µg/mg.
Arnau *et al.*, 2006.[1] Jensen *et al.*, 2004.[17]

tions by the natural DPPI and DAPase has been studied. A number of rules apply not only for stop positions (Figure 5) but also, for example, dipeptides containing hydrophobic amino acids that are not cleaved, or cleaved at a low rate. In order to facilitate the design of the genetic and protein construction and to enable an effective tag removal, a web-based resource is being developed—the TAGDesigner.[29] Using this tool, it is possible to test any protein sequence choosing predetermined or customized his-tags. It also gives recommendations as to which dipeptides need to be modified and whether the native protein contains a suitable stop position, allowing a process using DAPase alone for tag removal. A number of general recommendations for the development of a purification process are also provided.

Removal of Process Enzymes and Validation Tools

One of the long-standing paradigms of traditional protein purification is resistance to the use of affinity chromatography and a tag removal step. A classical argument is that it is undesirable to add 'contaminating' processing enzymes after the initial capture step which yields a relatively pure protein. In purification processes using his-tag and TAGZyme for tag removal, a second IMAC step is used to remove the processing enzymes (a subtractive IMAC step). During the subtractive IMAC, not only TAGZyme components, but also unspecific IMAC binders are removed, and this yields highly purified protein preparations.

The efficiency of removal of the three TAGZyme components during subtractive IMAC has been shown using polyclonal antibodies for each enzyme.[24] No detectable amounts were found in the final protein preparation. Furthermore, given the low DAPase requirement for tag removal, typically a 1:5000 DAPase to tagged protein ratio (corresponding to 0.1-0.2 µg DAPase per mg of protein, or 100-200 ppm) and the high binding efficiency during subtractive IMAC (98%), the estimated residual amount of DAPase in the final product should be 2-4 ppm. This level is in agreement with the lack of detection with polyclonal antibodies.[24]

Additionally, monoclonal antibodies directed against the different TAGZyme components are currently under development and may allow the quantification of residual amounts of process enzymes.

A purification process based on IMAC and TAGZyme is amenable for large-scale production of therapeutic proteins.[12] In addition, the recent expiration of broad patents may encourage manufacturers of protein drugs, especially CMOs, to use his-tag technology and tag removal in purification.

REFERENCES

1. 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 Expr Purif* 2006;47, 48:1-13 (in press).

2. Bernaudat F, Bülow L. Combined hydrophobic-metal binding fusion tags for applications in aqueous two-phase partitioning. *Prot Expr Purif* 2006 (in press).

3. Bhattacharya P, Pandey G, Srivastava P, Mukherjee KJ. Combined effect of protein fusion and signal sequence greatly enhances the production of recombinant human GM-CSF in Escherichia coli. *Mol Biotech* 2005;30:103-116.

4. Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. *Prot Expr Purif* 2005;43:1-9.

5. Cho MS, Lee YP, Chung HS. Selective removal of N-terminal methionine from recombinant human growth hormone by an aminopeptidase isolated by *Aspergillus flavus. J Ind Microbiol Biotech* 1998;20:287-290.

6. Cebe R, Geiser M. Rapid and easy thermodynamic optimization of the 5'-end of mRNA dramatically increases the level of wild type protein expression in Escherichia coli. *Prot Expr Purif* 2006;45:374-380.

7. Chatterjee DK, Esposito D. Enhanced soluble protein expression using two new fusion tags. *Prot Expr Purif* 2006;46:122-129.

8. Dahl SW, Slaughter C, Lauritzen C, Bateman RC, Connerton I, Pedersen J. Carica papaya glutamine cyclotransferase belongs to a novel plant enzyme subfamily: cloning and characterization of the recombinant enzyme. *Prot Expr Purif* 2000;20:27-36.

9. Dahl SW, Halkier T, Lauritzen C, Dolenc I, Pedersen J, Turk V, Turk B. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. *Biochemistry* 2001;40:1671-1678.

10. Dalbøge H, Bayne S, Pedersen J. In vivo processing of N-terminal methionine in *E. coli. FEBS* Lett 1990;266:1-3.

11. Derewenda ZS. The use of recombinant methods and molecular engineering in protein crystallization. *Methods* 2004;34:354-363.

12. Gaberc-Porekar V, Menart V. Potential of using histidine tags in purification of proteins at large scale. *Chem Eng Tech* 2005;28:1306-1314.

13. Glynou K, Ioannou PC, Christopoulos TK. One-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography. *Prot Expr Purif* 2003;27:384-390.

14. Guevara T, Mallorqui-Fernandez N, Garcia-Castellanos R, Petersen GE, Lauritzen C, Pedersen J, Arnau J, Gomis-Rüth FX, Sola M. *Papaya glutamine cyclotransferase shows a singular fivefold β-propeller architecture that suggests a novel reaction mechanism.* Submitted (2006).

15. Hunt I. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Prot Expr Purif* 2005;40:1-22.

16. Jaiswal S, Khanna N, Swaminathan S. High-level expression and one-step purification of recombinant dengue virus type 2 envelope domain III protein in *Escherichia coli. Prot Expr Purif* 2004;33:80-91.

17. Jensen MR, Lauritzen C, Dahl SW, Pedersen J, Led JJ. Binding ability of a HHP-tagged protein towards Ni2+ studied by paramagnetic NMR relaxation: the possibility of obtaining long-range structure information. *J Biomol NMR* 2004;29:175-185.

18. Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S. Comparison of affinity tags for protein purification. *Prot Expr Purif* 2005;41:98-105.

19. Liu XQ. Protein-splicing intein: genetic mobility, origin and evolution. *Ann Rev Genet* 2000;34:61-76.

20. Naested H, Kramhøft B, Lok F, Bojsen K, Yu S, Svensson B. Production of enzymatically active recombinant full-length barley high pI α-glucosidase family 31 by high cell-density fermentation of pichia pastoris and affinity purification. *Prot Expr Purif* 2006;46:56-63.

21. Olsen JG, Kadziola A, Lauritzen C, Pedersen J, Larsen S, Dahl SW. Tetrameric dipeptidyl peptidase I directs substrate specificity by use of the residual propart domain. *FEBS Letters* 2001;506:201-206.

22. Pedersen J, Lauritzen C, Madsen MT, Dahl SW. Removal of N-terminal polyhistidine tags from recombinant proteins using engineered aminopeptidases. *Prot Expr Purif* 1999;15:389-400.

23. Porath J. Immobilized metal ion affinity chromatography. Prot Expr Purif 1972;3:263-281.

24. Schäfer F, Schäfer A, Steinert K. A highly specific system for efficient enzymatic removal of tags from recombinant proteins. *J Biomol Techn* 2002;13:158-171.

25. Steen J, Uhlen M, Hober S, Ottosson J. High-throughput protein purification using an automated set-up for high-yield affinity chromatography. *Prot Expr Purif* 2006 (in press).

26. Turk D, Janjic V, Stern I, Podobnik M, Lamba D, Dahl SW, Lauritzen C, Pedersen J, Turk V, Turk B. Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J.* 2001;20:6570-6582.

27. Waugh DS. Making the most of affinity tags. *Trends Biotech* 2005;23:316-320.

28. Wei Q, Kim YS, Seo JH, Jang WS, Lee IH, Cha HJ. Facilitation of expression and purification of an antimicrobial peptide by fusion with baculoviral polyhedrin in *Escherichia coli. Appl Env Microbiol* 2005;71:5038-5043.

29. TAGDesigner website (http://www1.qiagen.com/products/ protein/tagdesigner/default.aspx).