

Review

# Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins

José Arnau\*, Conni Lauritzen, Gitte E. Petersen, John Pedersen

*Unizyme Laboratories AIS, Dr. Neergaards vej 17, DK-2970 Horsholm, Denmark*

Available online 28 December 2005

## Abstract

Affinity tags are highly efficient tools for protein purification. They allow the purification of virtually any protein without any prior knowledge of its biochemical properties. The use of affinity tags has therefore become widespread in several areas of research e.g., high throughput expression studies aimed at finding a biological function to large numbers of yet uncharacterized proteins. In some cases, the presence of the affinity tag in the recombinant protein is unwanted or may represent a disadvantage for the projected application of the protein, like for clinical use. Therefore, an increasing number of approaches are available at present that are designed for the removal of the affinity tag from the recombinant protein. Most of these methods employ recombinant endoproteases that recognize a specific sequence. These process enzymes can subsequently be removed from the process by affinity purification, since they also include a tag. Here, a survey of the most common affinity tags and the current methods for tag removal is presented, with special emphasis on the removal of N-terminal histidine tags using TAGZyme, a system based on exopeptidase cleavage. In the quest to reduce the significant costs associated with protein purification at large scale, relevant aspects involved in the development of downstream processes for pharmaceutical protein production that incorporate a tag removal step are also discussed. A comparison of the yield of standard vs. affinity purification together with an example of tag removal using TAGZyme is also included.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** Affinity purification; Affinity tag; His-tag; IMAC; Recombinant protein; Protein purification; Tag removal; Endoproteases; Exopeptidases; Aminopeptidases; Therapeutic protein purification

With the modern advances in genomics, proteomics and bioinformatics, the number of proteins being produced using recombinant techniques is exponentially increasing. High throughput screening approaches are being performed to rapidly identify proteins with a potential application as therapeutic, diagnostic or industrial enzymes [1]. For this purpose, different expression hosts (e.g., *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, insect and mammalian cell lines) have been developed to express heterologous proteins [2–7]. Additionally, genomic approaches are being pursued to solve the structure of numerous proteins [8,9].

The above-mentioned screening approaches would not be feasible if specific purification procedures were to be developed for each individual protein. Here, the use of affinity tags enables different proteins to be purified using a common method as opposed to highly customized procedures used in conventional chromatographic purification.

When designing a downstream processing strategy for a protein, the inclusion of an affinity tag might be attractive for a number of additional reasons. In many cases, the protein candidate may exist as a version that includes an affinity tag from its early research stages where no biochemical characterization or functional assay is yet available. For structural studies, more than 60% of the proteins produced include a polyhistidine tag (his-tag, [10]). Additionally, the fact that affinity purification normally results in high yields—often over 90%—makes this

\* Corresponding author.

E-mail address: [ja@unizyme.dk](mailto:ja@unizyme.dk) (J. Arnau).

Table 1  
Affinity and solubility tags for recombinant proteins

Tag	Size (aa)	Comments	References <sup>a</sup>
His-tag	5–15	Purification under native or denaturing conditions	[10,38,85,109]
FLAG	8	Calcium-dependent, mAb-based purification	[42,43,45]
Streptag II	8	Modified streptavidin, elution with biotin analog	[49,51,52,110]
HA-tag	9	Influenza virus hemagglutinin tag, Ab-based purification	[36]
Softag1, Softag 3	13, 8	Recognized by polyol-responsive mAb	[46–48]
c-myc	10	mAb-based purification	[31]
T7-tag	11–16	mAb-based purification	[56]
S-tag	15	S-protein resin affinity purification	[45]
Elastin-like peptides	18–320	Protein aggregation by temperature shift, intein used to remove tag	[64–66]
Chitin-binding domain	52	Binds only insoluble chitin (see intein, Table 3)	[68,111]
Thioredoxin	109	Affinity purification with modified resin	[17,26,89,112]
Xylanase 10A	163	Cellulose based capture, elution with glucose	[113,114]
Glutathione <i>S</i> -transferase	201	Glutathione or GST-Ab affinity	[9,26,87]
Maltose binding protein	396	Amylose affinity purification	[17,26,53,54,57,60]
NusA	495	Increased solubility in <i>E. coli</i> . Affinity tag needed for purification	[19]

<sup>a</sup> Only a few relevant references are included.

alternative economically favorable. Other economical and practical issues like the number of unit operations needed and the timesaving resulting from a reduction in chromatographic steps are also relevant aspects for consideration.

Introducing an affinity tag may have a positive effect in the biochemical properties of the target protein. A literature survey reveals that affinity tags have been observed to: (i) improve protein yield [11,12], (ii) prevent proteolysis [13], (iii) facilitate protein refolding [14], (iv) protect the antigenicity of the fusion protein [15], and (v) increase solubility [16–19]. Affinity tags have also been used to increase the sensitivity of binding assays for tagged ScFv [20].

On the other hand, adding a tag has also been reported to negatively affect the target protein resulting in e.g., (i) a change in protein conformation [21], (ii) lower protein yields [22], (iii) inhibition of enzyme activity [23,24], (iv) alteration in biological activity [25], (v) undesired flexibility in structural studies [26] and (vi) toxicity [27].

Due to the somehow unpredictable changes that adding a tag may introduce in a protein and its behavior, it is usually desirable to remove the tag. This reflects on the design of the protein fusion. Importantly, removal of the tag needs to be considered when designing a process for the production of a recombinant protein that is intended for human use to enable production of a ‘native’ (i.e., tagless) protein. And consequently, both the enzyme(s) used to cleave the tag and the cleaved fusion partner need to be removed from the purified protein.

Here, a review of affinity tags commonly used for recombinant protein production and the methods available for tag removal are discussed. A comparison of purification processes for a recombinant enzyme with and without affinity tag is also presented together with an overview of a downstream process that incorporates affinity purification and tag removal. Finally, an exam-

ple of process for tag removal is presented for a his-tag thioredoxin (Trx).<sup>1</sup>

### An overview of affinity tags and the design of the protein fusion

Recent reports have included several overviews of the currently available affinity tags for protein production and purification [28–32]. Nevertheless, since the choice of affinity tag and the method for tag removal are mutually dependent, an introduction to affinity tags is given herein for clarity.

Affinity tags can be defined as exogenous amino acid (aa) sequences with 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, discussed below). Another group includes tags that bind to an immobilized protein partner such as an antibody or antibody purification using protein A affinity chromatography ([29], Table 1). The protein A-based methodology, used for e.g., purification of monoclonal antibodies (mAb), has been extensively reviewed elsewhere [33,34] and will not be discussed here.

His-tags are the most widely used affinity tags. Purification of his-tagged proteins is based on the use of chelated

<sup>1</sup> *Abbreviations used:* Trx, thioredoxin; 6×his, a stretch containing six consecutive histidine residues; aa, amino acid(s); AAP, *Aeromonas proteolytica* aminopeptidase; APM, aminopeptidase M; *B.*, *Bacillus*; CPA, carboxypeptidase A; CPB, carboxypeptidase B; DAPase, recombinant dipeptidyl peptidase, part of TAGZyme; *E.*, *Escherichia*; ELP, elastin-like polypeptides; FMN, flavin mononucleotide; FP, green fluorescent protein; GST, glutathione *S*-transferase; His-tag, a polyhistidine tag; IMAC, immobilized metal–ion affinity chromatography; mAb, monoclonal antibodies; MBP, maltose-binding protein; NMR, nuclear magnetic resonance; pGAP, recombinant pyroglutamyl aminopeptidase; pGAPase, an engineered version of recombinant pGAP used in TAGZyme; PHB, polyhydroxybutyrate; Qcyclase, recombinant glutamine cyclotransferase, part of TAGZyme; ScFv, single chain antibodies.

metal ions as affinity ligands. The metal ion is complexed with an immobilized chelating agent (immobilized metal-ion affinity chromatography, IMAC). Protein separation using IMAC occurs on the basis of interactions between certain aa residues, especially histidine and the metal ions within an immobilized metal chelate [35]. The imidazole side chain of histidine shows high affinity for chelated metals [36].

The use of short histidine stretches or his-tags, typically placed as affinity tags at either the N-terminus or C-terminus, enables the purification of the desired protein from the crude extract of the host cells in a single step. Different chromatographic supports and strategies are available for IMAC [35,37]. The most widespread IMAC supports use either nitrilotriacetic acid (NTA) as a ligand for immobilizing metals like nickel in affinity chromatography (Ni-NTA) or different chelating Sepharose matrices. Additionally, a number of alternatives incorporating different metals have been developed [38]. Importantly, the binding specificity enables the purification of proteins under both native and denaturing conditions [35].

Although universally applicable, the use of his-tags and IMAC purification is not recommended for proteins containing metal ions. Similarly, other aa like cysteine and naturally occurring histidine rich regions in host proteins may result in unwanted protein binding during IMAC purification [39].

Since its development, numerous proteins and peptides have been purified using his-tags, and several therapeutic candidates are in clinical studies [40]. His-tags have also been used for purification of proteins using expanded bed adsorption [37,41].

Amongst other widely used tags, a short affinity tag that uses antibody-based purification is FLAG [27,32,42–45]. FLAG is a hydrophilic octapeptide (DYKDDDDK) recognized by the M1 mAb [43]. Recent studies suggested that a shorter FLAG-related peptide (DYKD) could be recognized with similar affinity by M1 [45]. The binding

reaction on an M1-chromatographic support is dependent on calcium, so proteins may be eluted from an affinity matrix by EDTA-containing buffer [42]. Similarly, Softag1 and Softag3 are small peptides (sequences SLAELLNAGLGGG and TKDPSRVG, respectively; Table 1) recognized with high affinity by polyol-responsive mAb. The use of Softag allows for the elution of pNSroteins from antibodies under mild conditions. SNSoftags are especially useful for the purification of labile, multisubunit enzyme complexes and to facilitate the study of protein interactions [46–48].

Another example is Streptag II. It consists of a streptavidin-recognizing octapeptide (WSHPQFEK) that was selected for its improved affinity. Proteins containing a Streptag II can be purified by affinity using a matrix with a modified streptavidin and eluted with a biotin analog [29,49–52]. An inherent advantage of Streptag II is its independence from metal ions in the purification, an aspect often of interest when metalloproteins are to be studied or when paramagnetic impurities must be avoided for NMR [51]. Other commonly used affinity tags are illustrated in Table 1.

The choice of a suitable affinity tag depends both on the type of application for the protein of interest and the stage of development of the protein for e.g., a therapeutic drug candidate. Additionally, the costs of the different chromatographic supports and the scalability of the process may be influential. In general, the cost of affinity tags based on mAb binding might be prohibitive for large-scale processes [29].

Furthermore, the selection of an affinity tag—and particularly if the tag is to be removed in the final product—determines the genetic design of the fusion protein used (Fig. 1).

For affinity purification, the tag may be placed at either end of the protein or in a region with appropriate surface exposure to allow binding or recognition [44]. To enable removal of the tag, a linker region is typically included

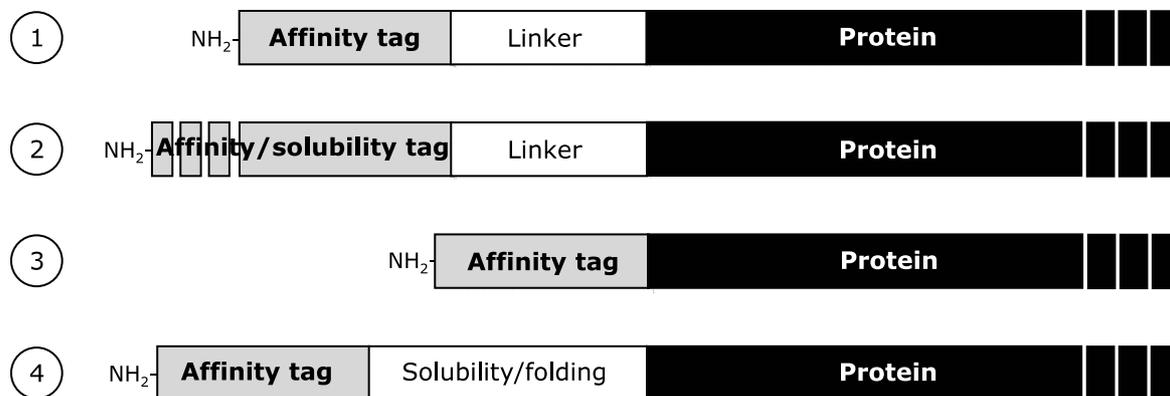


Fig. 1. Different strategies to incorporate and remove affinity tags. Different fusion proteins including the different elements required for affinity purification and tag removal are depicted. For clarity, only N-terminal fusions are shown. (1) A fusion protein consisting of an affinity tag, a linker region including a specific sequence for endoprotease cleavage and the native protein; (2) similar to above but the affinity tag (e.g., GST, MBP) increases the solubility of the protein; (3) a fusion protein designed for exopeptidase removal of the tag (only for N-terminal tags, using TAGZyme); (4) a fusion protein where a solubility and folding partner is fused N-terminal to the target protein (e.g., SUMO, sortase A). An affinity tag at the N-terminus is required for purification.

between the tag and the native protein sequence. This linker region may contain a sequence specific for endoprotease cleavage (construction 1 and 2, Fig. 1). The linker contributes to increased accessibility of the affinity tag and is often required for effective endoprotease cleavage. Although the number of different tags available may reflect the different biochemical properties of proteins and the host cells, it is possible to discriminate tags needed for purification (e.g., his-tag) from those developed to aid, e.g., solubility and folding like maltose-binding protein (MBP) [53], glutathione *S*-transferase (GST) [54] or small ubiquitin modifying protein (SUMO), among others [55,56]. Some tags such as MBP or GST are used for both affinity purification and solubility (construction 2, Fig. 1). Although GST fusions have been widely used in research, the slow binding kinetics of GST to glutathione–Sephacryl resin represents a severe drawback since loading of cell extracts to the resin is extremely time consuming for large-scale production [28].

As mentioned above MBP, GST and thioredoxin (Trx, Table 1) have long been used to increase the solubility of complex proteins in *E. coli* (construction 2, Fig. 1) [16,18,32,53,57–59]. Remarkably, not all fusion proteins containing MBP can be purified on amylose [58]. MBP fusions have also been reported to result in protein degradation [60]. His-tags may be used to aid solubility and folding since both native and denaturing conditions can be used during purification [16]. For example, his-tags may be used as the sole tag for purification and folding when located at the N-terminus of the protein [61], e.g., if exopeptidase cleavage is used for tag removal [62]. Thus, the simplest genetic design includes the affinity tag and the native protein sequence without a linker region (construction 3, Fig. 1).

### Other emerging methodologies for affinity purification and tag removal

Novel fusion tags to increase solubility such as a small chaperone, Skp, have successfully been tested for production of otherwise insoluble proteins in *E. coli* [56]. Other tags used to increase solubility in *E. coli* include NusA and T7 protein kinase (Table 1) [56,59].

Ubiquitin-based tags have been used to increase expression levels but also to aid in protein folding. One of these technologies (SUMO) has emerged as an alternative for the production, solubility and correct folding of otherwise intractable proteins (construction 4, Fig. 1). When fused to SUMO, insoluble proteins have been observed to fold properly and become soluble [55]. The SUMO tag can be removed using a specific protease (e.g., the yeast SUMO protease-1 Ulp1) that recognizes the conformation of the ubiquitin partner rather than a specific sequence [55]. However, an additional affinity tag is required for initial purification of the fusion protein when using SUMO. The use of SUMO is mostly constrained to *E. coli*, since highly conserved SUMO proteases are present in eukaryotes that may cleave the fusion protein during fermentation. A split

SUMO system has been proposed to overcome this problem [55]. Additionally, the high ratio of SUMO protease to fusion protein required to process fusion proteins remains a challenge for the upscaling of production processes based on this technology.

Elastin-like polypeptides (ELP) consist of several-to-numerous repeats of a peptide motif that undergo a reversible transition from soluble to insoluble upon temperature upshift (Table 1) [63–66]. ELP tags have been developed that allow purification of the fusion protein by temperature-induced aggregation, separation by centrifugation and resolubilization. Importantly, the length of the ELP tag sequence has been shown influential on the yield [64]. Thus, no chromatographic support is needed for initial purification of the fusion product. Further refinement of this technology includes a tripartite system that incorporates intein to enable tag removal (see below) in *E. coli* [66]. The robustness of ELP tags remains to be shown for large-scale processes of commercially relevant proteins. However, comparable protein yields have recently been shown for selected proteins using ELP and his-tag at small scale [65].

Inteins are self-cleavable proteases [67]. The intein-based IMPACT system uses a protein fusion consisting e.g., of an N-terminal chitin-binding domain (affinity tag), a central intein and a C-terminal target protein [68]. Binding to a chitin matrix is followed by on-column cleavage using either a thiol reagent or pH and temperature shift to yield intein cleavage and elution of the target protein [67,69,70].

Yet another system for tag cleavage that allows for one-step, on-column purification employs a protein fusion including the catalytic core of the *Staphylococcus aureus* sortase A, which recognizes the sequence LPETG and cleaves the TG peptide bond (Table 1). The suggested protein design includes an N-terminal his-tag for IMAC purification, followed by the sortase A domain, the cleavage sequence and the target protein at the C-terminus. This allows for sortase-mediated, on-column tag cleavage in the presence of e.g., calcium [71].

For purposes such as following the fate of the target protein through expression and purification, a number of tags have been developed in recent years [72,73]. Rainbow tags constitute such a tool that can be used to visualize correctly folded proteins and to track the protein during purification [72]. For this purpose, the flavin mononucleotide (FMN)-binding domain of cytochrome P450 reductase (displaying a blue-green or yellow color depending on the oxidation state of the FMN cofactor) and the red colored, heme-binding cytochrome *b5* have been used [72,74]. Opposite to other visual tagging system like GFP that require an external energy source, rainbow tags are visible with the naked eye. The use of rainbow tags requires an additional affinity tag for purification.

One remarkable example of more untraditional affinity tags is the use of polyhydroxybutyrate (PHB). The system includes production of PHB granules in *E. coli* together with a fusion containing (i) a PHB-binding domain, (ii) an intein for tag removal and (iii) the protein of interest [75].

## Combining different affinity tags

In recent years, a number of approaches have been published that utilize two different affinity tags fused to the target protein (dual affinity tags, also referred to as tandem affinity purification or TAP) for purification. In this approach, a protein of interest is fused in-frame with an N- or C-terminal tag comprised of two affinity tags surrounding an endoprotease cleavage site [76]. Applying two different affinity purification regimes sequentially, a highly purified protein is obtained [7,32,49,77,78]. These approaches are therefore directed to the production of a highly pure protein for e.g., crystallization [7,49] but are also widely used for the isolation of protein complexes [76,78,79] to study protein interactions. Originally developed in yeast, the method utilizes protein A and calmodulin-binding peptide separated by a specific TEV protease cleavage site and fused in-frame to the target protein. The TAP-tagged protein is expressed at physiological conditions to form a complex with its natural proteins partners. The TAP-tagged protein is first immobilized on IgG-Sepharose via the protein A moiety and gently washed. Immobilized protein complexes are incubated with TEV protease to release the protein complex. In the second step, the calmodulin-binding peptide of the TAP-tagged protein binds to calmodulin-Sepharose in the presence of calcium. After washing, calcium chelation leads to the release of the complex [76]. Once the protein complex has been purified through two affinity chromatography steps as above, it is resolved by SDS-PAGE into its components and protein bands are digested in-gel and identified by mass spectrometry [79]. A number of additional tag combinations have been used in other organisms for TAP and for the study of protein networks [76].

Additionally, dual tagging has been developed to monitor expression (via green fluorescent protein, GFP) and purify proteins via his-tags located within GFP [30].

## Comparison of a standard purification process with affinity purification

Numerous reports have been published where different purification procedures are compared [1,7,12,17,29,30,52,67,78,80]. Generally, the yield and efficiency of any specific purification procedure depends on the level of optimization developed for the individual protein and method. It is therefore recommended to use the data presented in different comparisons as indicative rather than definitive since it is unlikely that e.g., identical elution conditions are optimal for different proteins. As an example, suboptimal elution of his-tag proteins has been reported that is likely due to the low imidazole concentration used. Similarly, impurities—that may be accounted for by a certain level of metal leakage from the IMAC support—are observed for the purified his-tag proteins. This is a likely consequence of the use of  $\beta$ -mercaptoethanol during extraction [29]. As mentioned above, unspecific protein

binding may occur during IMAC purification and several approaches have been used to reduce this problem, including the use of higher imidazole concentration in the wash buffer [39]. Moreover, the concentration of imidazole used for wash and elution is a powerful tool to fine-tune the purification process for increased purity and specificity. Typically, elution of a ‘detagged’ protein after the his-tag removal step is performed on an (subtractive) IMAC column in the absence of imidazole to facilitate flow through of the protein while the his-tag process enzymes and other unspecific binders are retained in the column.

Standard chromatographic methods often include 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 relative pure protein. This results in time-consuming procedures and a relatively low yield of recovery, typically below 50% of the starting material for optimized processes. Remarkably, the yields obtained in purification of proteins using affinity chromatography are typically over 90% and include a reduced number of steps, adding to the reduction in costs of manufacturing. Importantly, IMAC are chemically stable to the prolonged cleaning-in-place procedures widely used in pharmaceutical production.

To illustrate this, we used a *Bacillus amyloliquefaciens* pyroglutamyl aminopeptidase (EC 3.4.19.3, pGAP) as a model protein. This protein catalyzes the removal of N-terminal pyroglutamyl residues from peptides and proteins. For comparison purposes, pGAP was produced in *E. coli* with and without an N-terminal his-tag (HT-pGAP, tag sequence: MEP(H)<sub>6</sub>L). For untagged pGAP, a purification process was performed that included ammonium sulfate precipitation, two consecutive separation steps using phenyl-Sepharose and a final step using Q Sepharose HP. As shown, many proteins co-purified with pGAP during the phenyl-Sepharose separation steps (Fig. 2A). In spite of an effort for the optimization of these chromatographic steps, especially with regard to the pH used in the final Q Sepharose HP step, a pGAP preparation was obtained where several contaminating bands were apparent, as observed by SDS-PAGE (Fig. 2A). Additionally, activity measurement showed a 40% recovery for this process with significant protein loss at step one and three, respectively (Table 2). For HT-pGAP, a higher amount of protein was observed in the cell extract. The reasons for this difference with the untagged pGAP remain unclear but may result from a higher expression level for this protein [80,81]. However, yield discrepancies due to intrinsic biological variation or culture conditions cannot be excluded. Remarkably, a very pure protein was obtained in a single IMAC step (Fig. 2B) and the yield of the process reached 96% without optimization (Table 2). Moreover, especially for high throughput approaches where no knowledge of the biochemical properties of the proteins is available, affinity tags—and his-tag in particular—may represent a universal platform for purification [1,43,62,82].

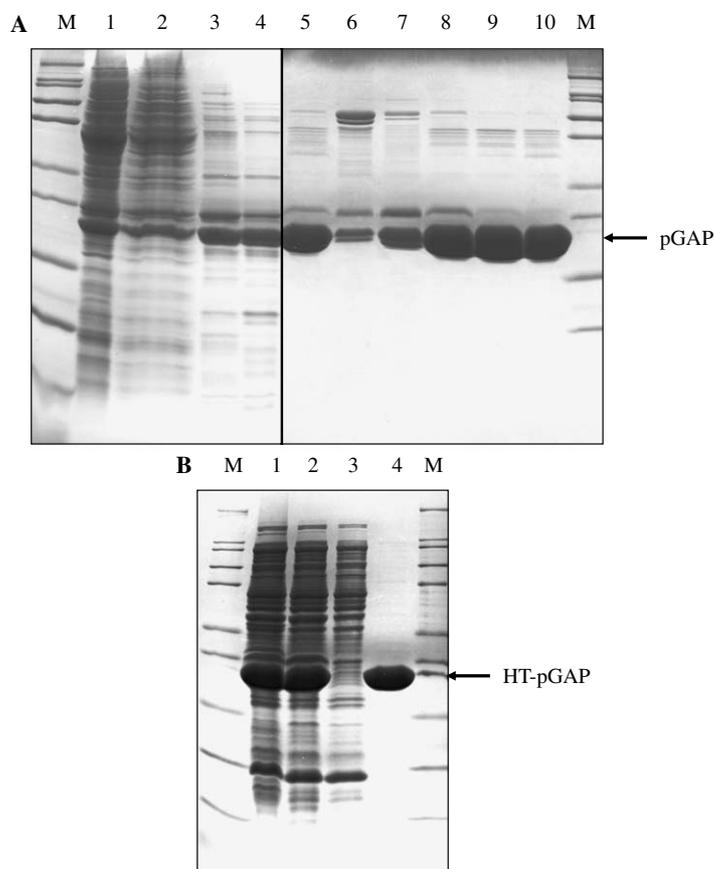


Fig. 2. Comparison of purification strategies for recombinant pGAP. *A. B. amyloliquefaciens* pyroglutamyl aminopeptidase (pGAP) was produced in *E. coli* with and without an N-terminal his-tag (HT-pGAP, tag sequence: MEP(H)<sub>6</sub>L). For untagged pGAP, purification included ammonium sulfate precipitation 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 the second phenyl-Sepharose step; lane 5: pool after desalting; lanes 6–10: several 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 the cell extract; lane 3: flow through fraction from the IMAC; lane 4: eluted HT-pGAP. See Table 2 for process yields.

Table 2  
Comparison of purification strategies for recombinant pGAP

Purification step	Total volume (ml)	Activity (U/ml)	Total activity (U)	Yield (%)
<i>Standard purification (untagged pGAP)</i>				
Cell extract	750	2.0	1463	100
Phenyl-Sepharose HS	400	2.6	1040	71
Phenyl-Sepharose LS	160	5.6	888	61
Q Sepharose HP	57	10.3	587	40
<i>IMAC purification (his-tag pGAP)</i>				
Cell extract	120	19.0	2280	100
Ni-NTA Sepharose	84	26.0	2184	96

See also Fig. 2.

### Methods for affinity tag removal

In a few specific cases, the affinity tag can be removed by harsh chemical treatment with e.g., cyanogen bromide or hydroxylamine. For this purpose, a unique methionine residue must be placed at the junction between the fusion partner and the protein of interest [83,84]. Chemical methods are rather unspecific and may lead to protein denaturation

and side chain modification of amino acids in the target protein [30]. Enzymatic methods are more specific and cleavage is achieved usually under mild conditions.

Most of the available methods for affinity tag removal include enzymatic cleavage of the tag followed by specific removal of the process enzyme(s) by an affinity chromatographic step to yield the detagged protein. Several affinity-tagged process enzymes are therefore used for tag removal

(Table 3) [30,37,62,67,85–87]. The rationale behind this approach is the need of a step for the efficient removal of the processing enzyme after the tag is removed to ensure that no contaminating protease activity is present in the protein preparation (illustrated for his-tag purification in Fig. 3).

Several endoproteases have been developed for tag removal (Table 3). Enterokinase, thrombin and factor Xa are among the most widely used enzymes for this purpose [37,42,88–93]. Enterokinase and factor Xa are examples of proteases that are able to cleave without requiring a specific aa or sequence at the C-terminus of the cleavage site (Table 3), allowing for the complete removal of the tag and yielding the native protein. Production of a tag-free protein is especially important for the manufacture of therapeutic proteins. Cleavage of the tag using other endoproteases like Thrombin, and virus-derived TEV protease, 3C protease or similar enzymes results in one or two aa residues from the cleavage site left at the N-terminus of the processed protein (Table 3). In these cases, documentation for the effect of the residual exogenous aa left in the therapeutic protein may be required.

Other classes of endoproteases have also been developed for tag removal. Granzyme B is a lysosomal serine protease involved in a number of protein activation processes in higher eukaryotes. An engineered, autoactivable granzyme B has recently been developed for tag removal [94], although its broad substrate specificity (Table 3) may represent a disadvantage if unspecific degradation of the target protein occurs. Caspase-6, a protease involved in protein processing during apoptosis, has also been used for tag processing of GST-fusions [94].

A general drawback of endoprotease-mediated tag cleavage is the need for high ratios of enzyme to protein and long incubation time required for complete tag removal. A further complication is the fact that under the

above severe conditions, endoprotease cleavage often leads to undesired, non-specific cleavage of the protein at cryptic sites (e.g., secondary GR sites for factor Xa or other basic aa for enterokinase, Table 3), resulting in protein degradation and lower yields [59,88,89,95–97]. In some cases, cleavage is observed at secondary sites but not at the introduced cleavage site [99].

Generally, processes for large-scale production of therapeutic proteins would encourage the use of on-column affinity purification and tag removal since a clear benefit applies to the reuse of columns and process enzymes for the production of large amounts of protein. Consequently, a number of on-column processes have been developed for tag removal by combining the most commonly used endoproteases with the most efficient affinity tags like his-tag, GST, FLAG, etc. [12,19,30,42,71,85,86,90,91,97,98]. Using relatively high amounts of endoprotease for tag removal must be accompanied by a process where efficient capture/immobilization of the enzyme is achieved. Several issues related to process validation have to be considered when adopting such a strategy to ensure e.g., no leakage of the endoprotease from the column during purification.

Furthermore, ensuring that no cryptic sites are present in the native sequence of the protein is a pre-requisite when using endoproteases. Contaminating proteases derived from either the host organism or the endoprotease preparation may also contribute to product degradation during extended incubation or incubation with high concentration of protease.

### The use of exopeptidases for tag removal: TAGZyme

An alternative strategy to endoprotease cleavage of affinity tags is the use of exopeptidases for tag removal. Several aminopeptidases and carboxypeptidases are available from natural sources. Examples of these are aminopep-

Table 3  
Enzymatic methods for tag removal

Enzyme	Cleavage site	Comments	References <sup>a</sup>
Enterokinase	DDDDK*	Secondary sites at other basic aa	[43,88,89,91,96]
Factor Xa	IDGR*	Secondary sites at GR	[37,59,88,96]
Thrombin	LVPR*GS	Secondary sites. Biotin labeled for removal of the protease	[88,92,96]
PreScission	LEVLFG*GP	GST tag for removal of the protease	[29]
TEV protease	EQLYFQ*G	His-tag for removal of the protease	[19,86,95,115]
3C protease	ETLFG*GP	GST tag for removal of the protease	[57,78,96]
Sortase A	LPET*G	Ca <sup>2+</sup> -induction of cleavage, requires an additional affinity tag (e.g., his-tag) for on column tag removal	[71]
Granzyme B	D*X, N*X, M*N, S*X	Serine protease. Risk for unspecific cleavage	[94]
Intein	Self-cleavable	Artificial aa left after cleavage in some applications. On column (chitin-beads) cleavage	[67,68,70,75]
SUMO	Conformation	No affinity purification per se (requires His-tag)	[55]
DAPase (TAGZyme)	Exo(di)peptidase	Cleaves N-terminal. His-tag (C-terminal) for purification and removal	[37,62]
<i>Aeromonas</i> aminopeptidase	Exopeptidase	Cleaves N-terminal, effective on M, L. Requires Zn	[100]
Aminopeptidase M	Exopeptidase	Cleaves N-terminal, does not cleave X-P	[99]
Carboxypeptidase A	Exopeptidase	Cleaves C-terminal. No cleavage at X-R, P	[99]
Carboxypeptidase B	Exopeptidase	Cleaves C-terminal R, K	[99]

The position of endoprotease cleavage is indicated with an asterisk (\*). Residues in bold remain in the protein after endoprotease cleavage.

<sup>a</sup> Only a few relevant references are included.

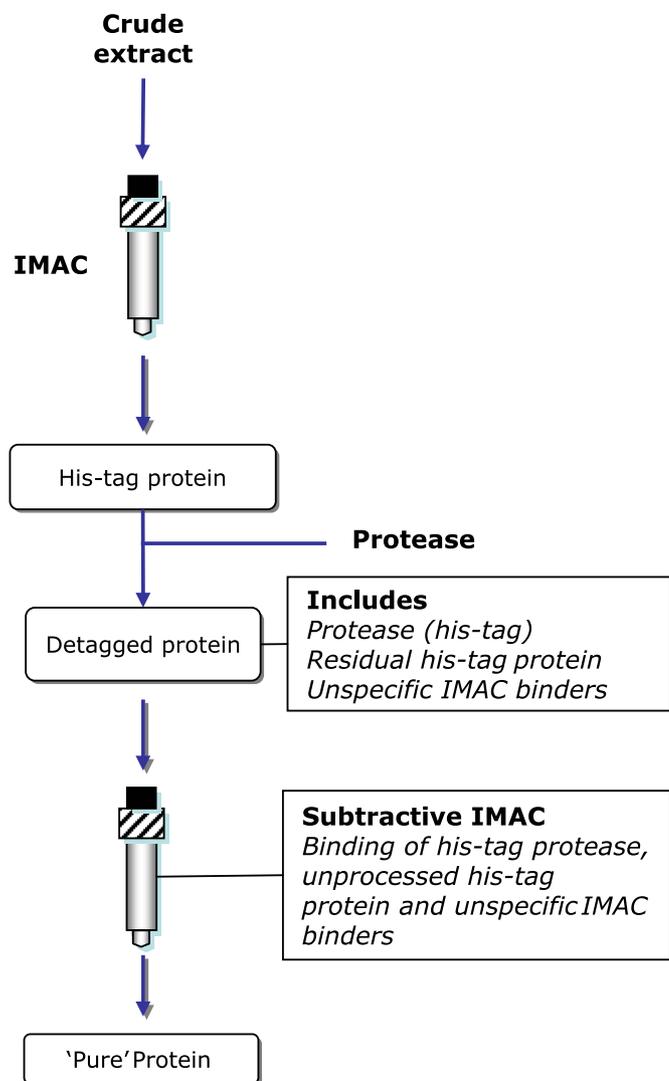


Fig. 3. Overview of a process for the purification of his-tag proteins and subsequently tag removal. The process is designed for the initial purification of the his-tag protein using IMAC; a tag removal step where a protease (or peptidase)—that also contains a his-tag—is added to the his-tag protein and a final subtractive IMAC purification. In this last step, process impurities such as unprocessed his-tag protein, the protease (or peptidase) and any unspecific binder are retained in the column while the detagged ‘pure’ protein is eluted.

tidase M (APM) obtained from porcine kidney and carboxypeptidase A and B (CPA and CPB) from bovine pancreas (Table 3) [99]. APM releases a single aa sequentially from the N-terminus of a protein with P being released slowly. CPA and CPB display different substrate preference, e.g., R is not released by CPA and rapidly processed by CPB. In general, with the growing concern on contamination from animal sources, the potential use of these enzymes for tag removal in pharmaceutical production is somewhat limited. Additionally, since these natural proteins lack an affinity tag, their removal represents an additional challenge for the purification of the target protein after tag cleavage.

Caution must also be used when choosing an amino- or carboxypeptidase that is a member of a metalloprotease

family for the removal of his-tags. The metal ion required for peptidase activity may be sequestered by the his-tag in the cleavage step, yielding the peptidase inactive. *Aeromonas proteolytica* aminopeptidase AAP is a small zinc metalloenzyme that has been used for the removal of a short natural tag in the production of human growth hormone [100,101]. However, no detectable cleavage is observed during the treatment of different his-tag proteins with AAP (unpublished data).

A method to circumvent the above limitations and to exploit the specificity of exopeptidase cleavage is provided by DAPase, a recombinant rat dipeptidyl aminopeptidase I (DPPI) that is part of the TAGZyme system [62]. DPPI is a cysteine peptidase that sequentially removes dipeptides from the N-terminus until a stop position is encountered in the sequence [62,99,102]. Using DAPase alone—for proteins that include a suitable stop position at the N-terminus—or in combination with two accessory enzymes acting on the N-terminal aa, for all other proteins [62,103,104], efficient and precise cleavage of short ( $\leq 25$  aa) N-terminal affinity tags is obtained. This approach enables the purification of any protein and its production with the native N-terminus [62,99,104]. Characteristic for DPPI and DAPase and different from typical proteases and other peptidases, the active site is located on the surface of the protein enabling rapid processing of peptides and proteins [102]. This feature reflects on the low DAPase requirement for tag removal, typically a 1:5000 DAPase to tagged protein ratio (corresponding to 0.2  $\mu\text{g}$  DAPase per mg of protein or 200 ppm). Moreover, complete tag removal is achieved using short incubation times, typically under 1 h [62].

Many human proteins do 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 (IL-2, IL-3, IL-5, IL-10 and IL-13) and growth hormone. R or K at position 1 represents also an effective stop position for DAPase cleavage [62]. 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 6 $\times$ his-tag maybe fused directly to the sequence of the mature protein. Subsequently, tag removal can be effectively performed using DAPase alone [62].

For all other proteins without a stop position at the N-terminus, an additional Q is added as the last residue of the affinity tag (placed at an uneven position) adjacent to the desired N-terminal aa of the mature protein. Upon DAPase cleavage in the presence of Qcyclase, the occurrence of an N-terminal Q leads to the effective formation of a N-terminal pyroglutamyl that prevents further DAPase cleavage. Subsequent removal of DAPase and Qcyclase by IMAC and treatment with pGAPase (e.g., on-column) removes the pyroglutamyl residue yielding the protein with the native N-terminus [62]. Typical yields for tag removal using TAGZyme are 90–98% for optimized processes.

## Process design for affinity purification and N-terminal his-tag removal

As mentioned above, most pharmaceutical applications necessitate that recombinant proteins do not contain any exogenous aa residue. A combination of the high yields obtained using for e.g., IMAC purification, the need for efficient tag removal and the fact that the latest generation of process proteases are typically tagged to enable their removal in the purification process has to be considered when designing a production and purification process.

Evidence for better yield using N-terminal tag compared to a C-terminal has been reported [22] although this is protein-dependent to some extent. Placing the tag at the N-terminus may also aid to increase expression levels at least in a bacterial host. Moreover, tag sequence optimization can be performed to match the host preferences. A simple process for purification and tag removal of a his-tag protein (N-terminal) includes a first IMAC step, a tag removal step followed by a second subtractive IMAC where the native protein is eluted at high purity. One example of this class of processes is illustrated below.

### A process integrating production of his-tag Trx in *E. coli* and tag removal

Trx was chosen as a well-characterized model protein to attempt the study of the molecular interaction of metal with histidine in a minimal sequence [105]. A process was therefore designed for the production of his-tag Trx (HT-Trx) 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, Fig. 4). The occurrence of a P in this sequence allows the removal of the 10-aa tag using DAPase alone and protects from further degradation (Fig. 4). The presence of HH at the N-terminus of the processed, detagged protein was decided to enable a detailed study of nickel interaction by paramagnetic NMR relaxation [105].

The fact that the desired mature protein sequence includes two H residues represented a challenge for IMAC purification where effective elution of HHP-Trx and retention of DAPase (containing a C-terminal his-tag) [62] is needed. In this and similar cases, the subtractive IMAC step is performed using a buffer containing a low concentration imidazole that allows for binding of the tag removal enzyme(s) to the IMAC while precluding binding of the detagged protein (Fig. 4).

The 10 aa his-tag sequence allows high expression of the recombinant protein in *E. coli* [62]. Additionally, the MK motif serves a double purpose. First, the presence of N-terminal MK results in a low frequency of methionine excision in *E. coli* [106,107]. Second, the presence of K at position 2 represents a quality control for DAPase processing. Thus, the fraction of HT-Trx molecules where M excision occurs is not cleaved by DAPase (an N-terminal K is a natural

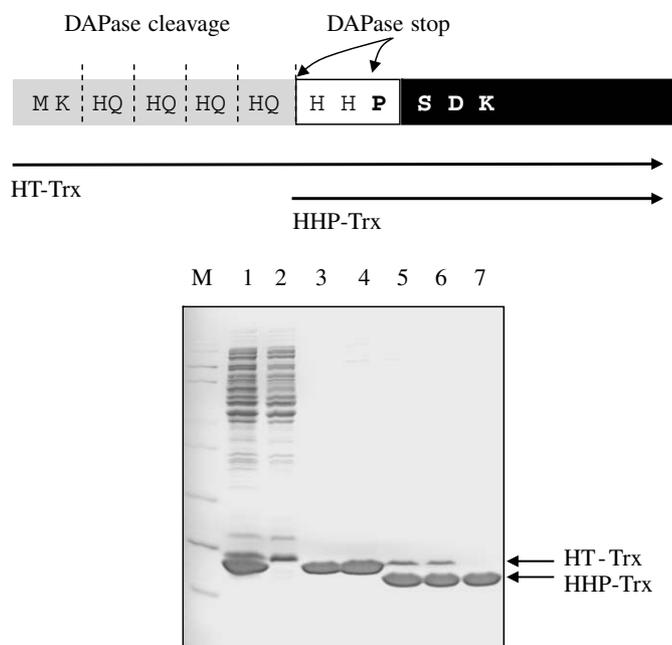


Fig. 4. Purification of recombinant HHP-Trx using DAPase. HT-Trx was purified from approx. 2.5 L *E. coli* culture. An overview of the genetic design and the cleavage process is shown at the top. After initial IMAC purification, DAPase cleavage sequentially removes the first five dipeptides (stippled 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 approx. 120 mL buffer A: 20 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.5, 300 mM NaCl, and 20 mM imidazole) was applied to a  $\text{Ni}^{2+}$ -chelating Sepharose FF column ( $5.3 \text{ cm}^2 \times 12 \text{ cm}$ ) with a flow rate of 5 mL/min. Subsequently, a wash step was performed with 250 mL buffer A using the same flow rate. A linear gradient from buffer A to buffer B (20 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.5, 300 mM NaCl, and 1 M imidazole) was used for 90 min at a flow rate of 2.5 mL/min. Subsequently, an additional 10 min 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 ( $5.3 \text{ cm}^2 \times 30 \text{ cm}$ ) using TAG-Zyme buffer C (20 mM  $\text{NaH}_2\text{PO}_4$ , 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 DAPase were 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 min. Subsequently, the DAPase mix was added to HT-Trx and incubated at 37 °C for 60 min. A new desalting step was performed on a Sephadex G25 F column ( $19.6 \text{ cm}^2 \times 27 \text{ cm}$ ) using buffer D [20 mM  $\text{NaH}_2\text{PO}_4$ , 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 (see main text for further information))] 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  $\text{Ni}^{2+}$ -chelating Sepharose FF column ( $2 \text{ cm}^2 \times 11 \text{ cm}$ ) 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 ( $19.6 \text{ cm}^2 \times 54 \text{ cm}$ ) with buffer E (10 mM  $\text{NaH}_2\text{PO}_4$ , 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.

stop position for DAPase) [62]. Consequently, these molecules retain a functional his-tag and are effectively removed in the subtractive IMAC step.

HT-Trx was initially purified on an IMAC (Fig. 4). After subsequent removal of the tag using DAPase and desalting, approx. 443 mg protein was obtained from approx. 2.5 L *E. coli* culture. Subsequent DAPase removal was achieved by a subtractive IMAC step. The final HHP-Trx preparation (354 mg) was obtained after a last desalting step (Fig. 4). No unspecific cleavage of HHP-Trx was observed, emphasizing the lack of endoprotease activity and the efficient protection from further exopeptidase cleavage of the tag provided by P (Fig. 4).

The use of the highly pure HHP-Trx in NMR studies demonstrated that HH forms a stable dimer complex with nickel, where each metal ion binds two HHP-Trx molecules with different orientation [105].

### Emphasis on the design of the tag is essential

In the example above, a tag was designed to enable a process for the purification and removal of a his-tag to yield a pure protein. In this case, the tag design enabled removal by cleavage with DAPase alone. As described above, for proteins which do not include a suitably placed DAPase stop position, a his-tag may be used that includes Q at the C-terminus of the tag and at an odd position. Virtually, any protein can therefore be produced using this approach. However, a critical concern is the design of the tag and the sequence used to ensure that tag removal is effective.

One important issue is the compatibility between affinity tag sequence and the specificity and sequence constraints of the enzyme chosen for tag removal. An example is the FLAG sequence that includes a cleavage site for enterokinase (DYKDDDDK, cleavage site underlined) [42,43]. Recent studies suggest that a shorter peptide, DYKD, can also be recognized with almost the same affinity by the M1 mAb used for the octapeptide FLAG version [43]. A tag sequence that includes DYKD can be used for protein purification and tag removal using TAGZyme only if the K residue is located at an even position in the fusion protein, e.g., by adding an uneven number of aa at the N-terminus and ensuring a stop position is placed at the appropriate location. Thus, a recommended FLAG tag designed for production in *E. coli* and effective removal using TAGZyme could be e.g., MK AD YK DQ Q (added residues underlined). MK is added to avoid processing of protein molecules where methionine excision has occurred and to facilitate high expression in *E. coli*. A is added for effective DAPase cleavage and to avoid K at an uneven position. Finally, QQ is added for effective DAPase cleavage (of the DQ dipeptide) and to introduce a stop position that can be removed using pGAPase after Qcyclase conversion of the second Q to pyrroglutamyl.

To design an effective process for tag removal using endoprotease cleavage, the linker region and the cleavage sequence need to be carefully considered together with the sequence of the protein to be produced. This is required to

avoid unspecific cleavage at cryptic/secondary sites that is often encountered [88,89].

### Conclusion

The market for protein drugs is expected to increase significantly from 2004 (\$34 bn) to 2010 (\$52 bn). New protein drugs currently under development will be introduced and yet novel protein drugs will be developed [108]. Generic protein drugs are also a competitive market where the use of more effective purification processes will add a significant advantage. In a competitive pharmaceutical market where production costs are high, the development of more economically sound production and purification processes may be a trend to follow.

Genome-based projects directed to the elucidation of protein structure or function and also strategies that make use of directed evolution of existing proteins will also benefit from a universal platform for production and purification. In this scenario, the use of affinity tags may represent a valid alternative to costly standard chromatographic methods for purification.

A variety of affinity tags and methods for tag removal have been developed. Most of these methods are suitable for laboratory scale applications and their use in pharmaceutical production is therefore not straightforward. In this scenario, an integrated view of the production and purification process that includes choice and design of the affinity tag together with the choice of tag removal strategy is required.

Although several tags have been extensively used over the years, his-tags are the unquestionably preferred affinity tag for protein purification. Other tags are also generally used for enhanced solubility and folding. The major advantages of using a his-tag compared to standard purification were illustrated with a model protein, *B. amyloliquefaciens* pGAP. A significant increase in yield and a reduction of chromatographic steps are achieved using a his-tag. Moreover, the impurities observed in the pGAP preparation obtained using standard chromatography were effectively avoided during IMAC purification of HT-pGAP.

For therapeutic applications, the tag has to be removed from the fusion protein. Here, the choice of affinity tag has to be considered as an integral process that includes a method for tag removal. Methods for tag removal are also abundant for research applications. Most of these use endoproteases and a cleavage site in the fusion protein. Major concerns associated with these approaches for pharmaceutical applications are the relatively high amounts used and the need for subsequent effective removal and unspecific cleavage of the target protein. Also, residual aa after cleavage may be unacceptable for therapeutic use in humans. An alternative to circumvent these limitations is the use of exopeptidases for removal of N-terminal tags. The use of a purpose-designed his-tag in *E. coli* was illustrated for the production of HHP-Trx. The N-terminal tag was effectively removed using DAPase, and enzyme removal using subtractive IMAC resulted in a pure HHP-Trx.

For affinity purification and tag removal, a careful consideration for the design of the tag is necessary to ensure an effective process. Finally, the potential application of TAG-Zyme technology is also illustrated for FLAG.

Using his-tag purification and tag removal is bound to become the method of choice for pharmaceutical production in the years ahead as well as a universal platform for protein structural studies and genomic projects. The different methods available discussed here should assist in designing a more effective production process.

## References

- [1] U. Korf, T. Kohl, H. van der Zandt, R. Zahn, S. Schleegeer, B. Ueberle, S. Wandschneider, S. Bechtel, M. Schnolzer, H. Ottleben, S. Wiemann, A. Poustka, Large-scale protein expression for proteome research, *Proteomics* 5 (2005) 3571–3580.
- [2] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol.* 22 (2004) 1399–1408.
- [3] L. Borsig, E.G. Berger, M. Malissard, Expression and purification of His-tagged beta-1,4-galactosyltransferase in yeast and in COS cells, *Biochem. Biophys. Res. Commun.* 240 (1997) 586–589.
- [4] L.J. Byrne, K.J. O'Callaghan, M.F. Tuite, Heterologous gene expression in yeast, *Methods Mol. Biol.* 308 (2005) 51–64.
- [5] R. Bhikhabhai, A. Sjöberg, L. Hedkvist, M. Galin, P. Liljedahl, T. Frigård, N. Pettersson, M. Nilsson, J.A. Sigrell-Simon, C. Markeland-Johansson, Production of milligram quantities of affinity tagged-proteins using automated multistep chromatographic purification, *J. Chromatogr.* 1080 (2005) 83–92.
- [6] D. Daly, M.T. Hearn, Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production, *J. Mol. Recognit.* 18 (2005) 119–138.
- [7] B. Prinz, J. Schultchen, R. Ryzewski, C. Holz, M. Boettner, U. Stahl, C. Lang, Establishing a versatile fermentation and purification procedure for human proteins expressed in the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* for structural genomics, *J. Struct. Funct. Genomics* 5 (2004) 29–44.
- [8] Z.J. Liu, W. Tempel, J.D. Ng, D. Lin, A.K. Shah, L. Chen, P.S. Horanyi, J.E. Habel, I.A. Kataeva, H. Xu, H. Yang, J.C. Chang, L. Huang, S.H. Chang, W. Zhou, D. Lee, J.L. Praissman, H. Zhang, M.G. Newton, J.P. Rose, J.S. Richardson, D.C. Richardson, B.C. Wang, The high-throughput protein-to-structure pipeline at SECSG, *Acta Crystallogr. D Biol. Crystallogr.* 61 (2005) 679–684.
- [9] C. Scheich, V. Sievert, K. Bussow, An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography, *BMC Biotechnol.* 3 (2003) 12.
- [10] Z.S. Derewenda, The use of recombinant methods and molecular engineering in protein crystallization, *Methods* 34 (2004) 354–363.
- [11] S.S. Rajan, H. Lackland, S. Stein, D.T. Denhardt, Presence of an N-terminal polyhistidine tag facilitates stable expression of an otherwise unstable N-terminal domain of mouse tissue inhibitor of metalloproteinase-1 in *Escherichia coli*, *Protein Expr. Purif.* 13 (1998) 67–72.
- [12] Q.M. Sun, L.L. Chen, L. Cao, L. Fang, C. Chen, Z.C. Hua, An improved strategy for high-level production of human vasostatin 120–180, *Biotechnol. Prog.* 21 (2005) 1048–1052.
- [13] W. Tang, Z.Y. Sun, R. Pannell, V. Gurewich, J.N. Liu, An efficient system for production of recombinant urokinase-type plasminogen activator, *Protein Expr. Purif.* 11 (1997) 279–283.
- [14] S. Shi, J. Xue, K. Fan, G. Kou, H. Wang, Y. Guo, Preparation and characterization of recombinant protein ScFv(CD11c)-TRP2 for tumor therapy from inclusion bodies in *Escherichia coli*, *Protein Expr. Purif.* (2005) (in press), doi:10.1016/j.pep.2005.08.002.
- [15] A. Mayer, S.K. Sharma, B. Tolner, N.P. Minton, D. Purdy, P. Amlot, G. Tharakan, R.H.J. Begent, K.A. Chester, Modifying an immunogenic epitope on a therapeutic protein: a step towards an improved system for antibody-directed enzyme prodrug therapy (ADEPT), *Br. J. Cancer* 90 (2004) 2402–2410.
- [16] H. Chen, Z. Xu, N. Xu, P. Cen, Efficient production of a soluble fusion protein containing human beta-defensin-2 in *E. coli* cell-free system, *J. Biotechnol.* 115 (2005) 307–315.
- [17] M.R. Dyson, S.P. Shadbolt, K.J. Vincent, R.L. Perera, J. McCafferty, Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression, *BMC Biotechnol.* 4 (2004) 32.
- [18] M. Hammarström, N. Hellgren, S. van Den Berg, H. Berglund, T. Hard, Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*, *Protein Sci.* 11 (2002) 313–321.
- [19] S. Nallamsetty, D.S. Waugh, Solubility-enhancing proteins MBP and NusA play a passive role during the folding of their fusion partners, *Protein Expr. Purif.* (2005).
- [20] X. Wang, M. Campoli, E. Ko, W. Luo, S. Ferrone, Enhancement of scFv fragment reactivity with target antigens in binding assays following mixing with anti-tag monoclonal antibodies, *J. Immunol. Methods* 294 (2004) 23–35.
- [21] A. Chant, C.M. Kraemer-Pecore, R. Watkin, G.G. Kneale, Attachment of a histidine tag to the minimal zinc finger protein of the *Aspergillus nidulans* gene regulatory protein AreA causes a conformational change at the DNA-binding site, *Protein Expr. Purif.* 39 (2005) 152–159.
- [22] A. Goel, D. Colcher, J.S. Koo, B.J. Booth, G. Pavlinkova, S.K. Batra, Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct, *Biochim. Biophys. Acta* 1523 (2000) 13–20.
- [23] K.M. Kim, E.C. Yi, D. Baker, K.Y. Zhang, Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase, *Acta Crystallogr. D Biol. Crystallogr.* 57 (2001) 759–762.
- [24] S. Cadel, C. Gouzy-Darmon, S. Petres, C. Piesse, V.L. Pham, M.C. Beinfeld, P. Cohen, T. Foulon, Expression and purification of rat recombinant aminopeptidase B secreted from baculovirus-infected insect cells, *Protein Expr. Purif.* 36 (2004) 19–30.
- [25] I. Fonda, M. Kenig, V. Gaberc-Porekar, P. Pristovaek, V. Menart, Attachment of histidine tags to recombinant tumor necrosis factor- $\alpha$  drastically changes its properties, *Sci. World J.* 15 (2002) 1312–1325.
- [26] D.R. Smyth, M.K. Mrozkiewicz, W.J. McGrath, P. Listwan, B. Kobe, Crystal structures of fusion proteins with large-affinity tags, *Protein Sci.* 12 (2003) 1313–1322.
- [27] E.G.E. de Vries, M.N. de Hooge, J.A. Gietema, S. de Jong, Correspondence re: C.G. Ferreira et al., Apoptosis: Target of Cancer Therapy, *Clin. Cancer Res.*, 9 (2003) 912.
- [28] I. Hunt, From gene to protein: a review of new and enabling technologies for multi-parallel protein expression, *Protein Expr. Purif.* 40 (2005) 1–22.
- [29] J.J. Lichty, J.L. Malecki, H.D. Agnew, D.J. Michelson-Horowitz, S. Tan, Comparison of affinity tags for protein purification, *Protein Expr. Purif.* 41 (2005) 98–105.
- [30] J. Nilsson, S. Ståhl, J. Lundberg, M. Uhlén, P.Å. Nygren, Affinity fusion strategies for detection, purification and immobilization of recombinant proteins, *Protein Expr. Purif.* 11 (1997) 1–16.
- [31] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, *Appl. Microbiol. Biotechnol.* 60 (2003) 523–533.
- [32] D.S. Waugh, Making the most of affinity tags, *Trends Biotechnol.* 23 (2005) 316–320.
- [33] G. Fassina, M. Ruvo, G. Palombo, A. Verdoliva, M. Marino, Novel ligands for the affinity-chromatographic purification of antibodies, *J. Biochem. Biophys. Methods* 49 (2001) 481–490.
- [34] U.M. Kent, Purification of antibodies using protein A-Sepharose and FPLC, *Methods Mol. Biol.* 115 (1999) 29–33.

- [35] J. Porath, Immobilized metal ion affinity chromatography, *Protein Expr. Purif.* 3 (1992) 263–281.
- [36] D.S. Hage, Affinity chromatography: a review of clinical applications, *Clin. Chem.* 45 (1999) 593–615.
- [37] N. Abdullah, H.A. Chase, Removal of poly-histidine fusion tags from recombinant proteins purified by expanded bed adsorption, *Biotechnol. Bioeng.* 92 (2005) 501–513.
- [38] G. Chaga, D.E. Bochkariov, G.G. Jokhadze, J. Hopp, P. Nelson, Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose, *J. Chromatogr. A* 864 (1999) 247–256.
- [39] D.F. Westra, G.W. Welling, D.G. Koedijk, A.J. Scheffer, T.H. The, S. Welling-Wester, Immobilised metal-ion affinity chromatography purification of histidine-tagged recombinant proteins: a wash step with a low concentration of EDTA, *J. Chromatogr. B Biomed. Sci. Appl.* 760 (2001) 129–136.
- [40] K. Chester, B. Pedley, B. Tolner, J. Violet, A. Mayer, S. Sharma, G. Boxer, A. Green, S. Nagl, R. Begent, Engineering antibodies for clinical applications in cancer, *Tumor Biol.* 25 (2004) 91–98.
- [41] Y.P. Tan, T.C. Ling, W.S. Tan, K. Yusoff, B.T. Tey, Purification of recombinant nucleocapsid protein of Newcastle disease virus from unclarified feedstock using expanded bed adsorption chromatography, *Protein Expr. Purif.* 2005 (in press), doi:10.1016/j.pep.2005.06.015.
- [42] A. Einhauser, A. Jungbauer, The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins, *J. Biochem. Biophys. Methods* 49 (2001) 455–465.
- [43] A. Knappik, A. Pluckthun, An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments, *Biotechniques* 17 (1994) 754–761.
- [44] M. Muller, M. Yong, X.H. Peng, B. Petre, S. Arora, S.V. Ambudkar, Evidence for the role of glycosylation in accessibility of the extracellular domains of human MRP1 (ABCC1), *Biochemistry* 41 (2002) 10123–10132.
- [45] J.W. Sloodstra, D. Kuperus, A. Pluckthun, R.H. Meloen, Identification of new tag sequences with differential and selective recognition properties for the anti-FLAG monoclonal antibodies M1, M2 and M5, *Mol. Divers.* 2 (1997) 156–164.
- [46] R.R. Burgess, N.E. Thompson, Advances in gentle immunoaffinity chromatography, *Curr. Opin. Biotechnol.* 13 (2002) 304–308.
- [47] S.J. Duellman, N.E. Thompson, R.R. Burgess, An epitope tag derived from human transcription factor IIB that reacts with a polyol-responsive monoclonal antibody, *Protein Expr. Purif.* 35 (2004) 147–155.
- [48] N.E. Thompson, T.M. Arthur, R.R. Burgess, Development of a new epitope tag for the gentle purification of proteins by immunoaffinity chromatography, *Anal. Biochem.* 323 (2003) 171–179.
- [49] B. Cass, P.L. Pham, A. Kamen, Y. Durocher, Purification of recombinant proteins from mammalian cell culture using a generic double-affinity chromatography scheme, *Protein Expr. Purif.* 40 (2005) 77–85.
- [50] T. Lamla, W. Stiege, V.A. Erdmann, An improved protein bioreactor: efficient product isolation during in vitro protein biosynthesis via affinity tag, *Mol. Cell Proteomics* 1 (2002) 466–471.
- [51] A. Skerra, T.G.M. Schmidt, Use of the Strep-tag and streptavidin for detection and purification of recombinant proteins, *Methods Enzymol.* 326 (2000) 271–304.
- [52] C.P. Witte, L.D. Noel, J. Gielbert, J.E. Parker, T. Romeis, Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope, *Plant Mol. Biol.* 55 (2004) 135–147.
- [53] J.D. Fox, D.S. Waugh, Maltose-binding protein as a solubility enhancer, *Methods Mol. Biol.* 205 (2003) 99–117.
- [54] D.B. Smith, K.S. Johnson, Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase, *Gene* 67 (1988) 31–40.
- [55] T.R. Butt, S.C. Edavettal, J.P. Hall, M.R. Mattern, SUMO fusion technology for difficult-to-express proteins, *Protein Expr. Purif.* 43 (2005) 1–9.
- [56] D.K. Chatterjee, D. Esposito, Enhanced soluble protein expression using two new fusion tags, *Protein Expr. Purif.* (2005) (in press), doi:10.1016/j.pep.2005.07.028.
- [57] R. Eliseev, A. Alexandrov, T. Gunter, High-yield expression and purification of p18 form of Bax as an MBP-fusion protein, *Protein Expr. Purif.* 35 (2004) 206–209.
- [58] K.M. Routzahn, D.S. Waugh, Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins, *J. Struct. Funct. Genomics* 2 (2002) 83–92.
- [59] Y. Tsunoda, N. Sakai, K. Kikuchi, S. Katoh, K. Akagi, J. Miura-Ohnuma, Y. Tashiro, K. Murata, N. Shibuya, E. Katoh, Improving expression and solubility of rice proteins produced as fusion proteins in *Escherichia coli*, *Protein Expr. Purif.* 42 (2005) 268–277.
- [60] A. Feher, P. Boross, T. Sperka, S. Oroszlan, J. Tozser, Expression of the murine leukemia virus protease in fusion with maltose-binding protein in *Escherichia coli*, *Protein Expr. Purif.* 35 (2004) 62–68.
- [61] J.Q. Guo, Q.M. Li, J.Y. Zhou, G.P. Zhang, Y.Y. Yang, G.X. Xing, D. Zhao, S.Y. You, C.Y. Zhang, Efficient recovery of the functional IP10-scFv fusion protein from inclusion bodies with an on-column refolding system, *Protein Expr. Purif.* 45 (2006) 168–174.
- [62] J. Pedersen, C. Lauritzen, M.T. Madsen, S.W. Dahl, Removal of N-terminal polyhistidine tags from recombinant proteins using engineered aminopeptidases, *Protein Expr. Purif.* 15 (1999) 389–400.
- [63] M.R. Banki, L. Feng, D.W. Wood, Simple bioseparations using self-cleaving elastin-like polypeptide tags, *Nat. Methods* (2005).
- [64] D.E. Meyer, K. Trabbic-Carlson, A. Chilkoti, Protein purification by fusion with an environmentally responsive elastin-like polypeptide: effect of polypeptide length on the purification of thioredoxin, *Biotechnol. Prog.* 17 (2001) 720–728.
- [65] K. Trabbic-Carlson, L. Liu, B. Kim, A. Chilkoti, Expression and purification of recombinant proteins from *Escherichia coli*: comparison of an elastin-like polypeptide fusion with an oligohistidine fusion, *Protein Sci.* 13 (2004) 3274–3284.
- [66] X. Ge, D.S. Yang, K. Trabbic-Carlson, B. Kim, A. Chilkoti, C.D. Filipe, Self-cleavable stimulus responsive tags for protein purification without chromatography, *J. Am. Chem. Soc.* 127 (2005) 11228–11229.
- [67] X.Q. Liu, Protein-splicing intein: genetic mobility, origin, and evolution, *Annu. Rev. Genet.* 34 (2000) 61–76.
- [68] H.E. Humphries, M. Christodoulides, J.E. Heckels, Expression of the class 1 outer-membrane protein of *Neisseria meningitidis* in *Escherichia coli* and purification using a self-cleavable affinity tag, *Protein Expr. Purif.* 26 (2002) 243–248.
- [69] Z. Sun, J. Chen, H. Yao, L. Liu, J. Wang, J. Zhang, J.N. Liu, Use of Ssp dnaB derived mini-intein as a fusion partner for production of recombinant human brain natriuretic peptide in *Escherichia coli*, *Protein Expr. Purif.* 43 (2005) 26–32.
- [70] R. Paul, F.U. Bosch, K.P. Schafer, Overexpression and purification of *Helicobacter pylori* flavodoxin and induction of a specific antiserum in rabbits, *Protein Expr. Purif.* 22 (2001) 399–405.
- [71] H. Mao, A self-cleavable sortase fusion for one-step purification of free recombinant proteins, *Protein Expr. Purif.* 37 (2004) 253–263.
- [72] R.D. Finn, I. Kapelioukh, M.J. Paine, Rainbow tags: a visual tag system for recombinant protein expression and purification, *Biotechniques* 38 (2005) 387–392.
- [73] R.I. Paramban, R.C. Bugos, W.W. Su, Engineering green fluorescent protein as a dual functional tag, *Biotechnol. Bioeng.* 86 (2004) 687–697.
- [74] A. Mitra, K.S. Chakrabarti, M.S. Shahul Hameed, K.V. Srinivas, G. Senthil Kumar, S.P. Sarma, High level expression of peptides and proteins using cytochrome *b5* as a fusion host, *Protein Expr. Purif.* 41 (2005) 84–97.
- [75] M.R. Banki, T.U. Gerngross, D.W. Woods, Novel and economical purification of recombinant proteins: intein mediated protein purification using in vivo polyhydroxybutyrate (PHB) matrix association, *Protein Sci.* 14 (2005) 1387–1395.
- [76] A.-C. Gingras, R. Aebersold, B. Raught, Advances in protein complex analysis using mass spectrometry, *J. Physiol.* 563 (2005) 11–21.

- [77] S. Graslund, M. Eklund, R. Falk, M. Uhlen, P.A. Nygren, S. Stahl, A novel affinity gene fusion system allowing protein A-based recovery of non-immunoglobulin gene products, *J. Biotechnol.* 99 (2002) 41–50.
- [78] V. Rubio, Y. Shen, Y. Saijo, Y. Liu, G. Gusmaroli, S.P. Dinesh-Kumar, X.W. Deng, An alternative tandem affinity purification strategy applied to *Arabidopsis* protein complex isolation, *Plant J.* 41 (2005) 767–778.
- [79] G.-H. Wei, D.-P. Liu, C.-C. Liang, Charting gene regulatory networks: strategies, challenges and perspectives, *Biochem. J.* 381 (2004) 1–12.
- [80] P.O. Olins, S.C. Lee, Recent advances in heterologous gene expression in *Escherichia coli*, *Curr. Opin. Biotechnol.* 4 (1993) 520–525.
- [81] H. Wang, D.J. O'Mahony, D.J. McConnell, S.Z. Qi, Optimization of the synthesis of porcine somatotropin in *Escherichia coli*, *Appl. Microbiol. Biotechnol.* 39 (1993) 324–328.
- [82] C. Jain, Overexpression and purification of tagged *Escherichia coli* proteins using a chromosomal knock-in strategy, *Protein Expr. Purif.* (2005).
- [83] W.D. Fairlie, A.D. Uboldi, D.P. De Souza, G.J. Hemmings, N.A. Nicola, M. Baca, A fusion protein system for the recombinant production of short disulfide-containing peptides, *Protein Expr. Purif.* 26 (2002) 171–178.
- [84] C. Rais-Beghdadi, M.A. Roggero, N. Fasel, C.D. Reymond, Purification of recombinant proteins by chemical removal of the affinity tag, *Appl. Biochem. Biotechnol.* 74 (1998) 95–103.
- [85] S. Chatterjee, J. Schoepe, S. Lohmer, D. Schomburg, High level expression and single-step purification of hexahistidine-tagged l-2-hydroxyisocaproate dehydrogenase making use of a versatile expression vector set, *Protein Expr. Purif.* 39 (2005) 137–143.
- [86] A.R. Kim, A. Doherty-Kirby, G. Lajoie, R.J. Rylett, B.H. Shilton, Two methods for large-scale purification of recombinant human choline acetyltransferase, *Protein Expr. Purif.* 40 (2005) 107–117.
- [87] P.K. Purbey, P.C. Jayakumar, P.D. Deepalakshmi, M.S. Patole, S. Galande, GST fusion vector with caspase-6 cleavage site for removal of fusion tag during column purification, *Biotechniques* 38 (2005) 360–364.
- [88] R.J. Jenny, K.G. Mann, R.L. Lundblad, A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa, *Protein Expr. Purif.* 31 (2003) 1–11.
- [89] O.W. Liew, J.P. Ching Chong, T.G. Yandle, S.O. Brennan, Preparation of recombinant thioredoxin fused N-terminal proCNP: analysis of enterokinase cleavage products reveals new enterokinase cleavage sites, *Protein Expr. Purif.* 41 (2005) 332–340.
- [90] H. Zhang, Q. Yuan, Y. Zhu, R. Ma, Expression and preparation of recombinant hepcidin in *Escherichia coli*, *Protein Expr. Purif.* 41 (2005) 409–416.
- [91] S.I. Choi, H.W. Song, J.W. Moon, B.L. Seong, Recombinant enterokinase light chain with affinity tag: expression from *Saccharomyces cerevisiae* and its utilities in fusion protein technology, *Biotechnol. Bioeng.* 75 (2001) 718–724.
- [92] M.H. Hefti, C.J.G. Van Vugt-Van der Toorn, R. Dixon, J. Vervoort, A novel purification method for histidine-tagged proteins containing a thrombin cleavage site, *Anal. Biochem.* 295 (2001) 180–185.
- [93] P. Turner, O. Holst, E.N. Karlsson, Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration, *Protein Expr. Purif.* 39 (2005) 54–60.
- [94] R.H. Lorentsen, C.H. Fynbo, H.C. Thogersen, M. Etzerodt, T.L. Holtet, Expression, refolding, and purification of recombinant human granzyme B, *Protein Expr. Purif.* 39 (2005) 18–26.
- [95] S.Y. Kwon, Y.J. Choi, T.H. Kang, K.H. Lee, S.S. Cha, G.H. Kim, H.S. Lee, K.T. Kim, K.J. Kim, Highly efficient protein expression and purification using bacterial hemoglobin fusion vector, *Plasmid* 53 (2005) 274–282.
- [96] L.E. Leong, The use of recombinant fusion proteases in the affinity purification of recombinant proteins, *Mol. Biotechnol.* 12 (1999) 269–274.
- [97] C. Dian, S. Eshaghi, T. Urbig, S. McSweeney, A. Heijbel, G. Salbert, D. Birse, Strategies for the purification and on-column cleavage of glutathione S-transferase fusion target proteins, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 769 (2002) 133–144.
- [98] A.L. Taylor, A. Haze-Filderman, A. Blumenfeld, B. Shay, L. Dafni, E. Rosenfeld, Y. Leiser, E. Fermon, Y. Gruenbaum-Cohen, D. Deutsch, High yield of biologically active recombinant human amelogenin using the baculovirus expression system, *Protein Expr. Purif.* 45 (2006) 43–53.
- [99] M. Kenig, S. Peternel, V. Gaberc-Porekar, V. Menart, Influence of the protein oligomericity on final yield after affinity tag removal of recombinant proteins, *J. Chromatogr.* (2005) (in press).
- [100] S. Blumberg, D. Ben Meir, Method for producing human growth hormone, EP 0 489 711 A2 (1986).
- [101] F.W. Wagner, S.H. Wilkes, J.M. Prescott, Specificity of *Aeromonas* aminopeptidase toward amino acid amides and dipeptides, *J. Biol. Chem.* 247 (1972) 1208–1210.
- [102] D. Turk, V. Janjic, I. Stern, M. Podobnik, D. Lamba, S.W. Dahl, C. Lauritzen, J. Pedersen, V. Turk, B. Turk, 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.* 20 (2001) 6570–6582.
- [103] S.W. Dahl, C. Slaughter, C. Lauritzen, R.C. Bateman, I. Connerton, J. Pedersen, Carica papaya glutamine cyclotransferase belongs to a novel plant enzyme subfamily: cloning and characterization of the recombinant enzyme, *Protein Expr. Purif.* 20 (2000) 27–36.
- [104] C. Lauritzen, J. Pedersen, M.T. Madsen, J. Justesen, P.M. Martensen, S.W. Dahl, Active recombinant rat dipeptidyl aminopeptidase I (cathepsin C) produced using the baculovirus expression system, *Protein Expr. Purif.* 14 (1998) 434–442.
- [105] M.R. Jensen, C. Lauritzen, S.W. Dahl, J. Pedersen, J.J. Led, Binding ability of a HHP-tagged protein towards Ni<sup>2+</sup> studied by paramagnetic NMR relaxation: the possibility of obtaining long-range structure information, *J. Biomol. NMR* 29 (2004) 175–185.
- [106] J.C. Coval, T.B. Cao, J.R. Magdaroag, L.A. Gross, P.A. Jennings, Temperature, media, and point of induction affect the N-terminal processing of interleukin-1 $\beta$ , *Protein Expr. Purif.* 41 (2005) 45–52.
- [107] H. Dalbøge, S. Bayne, J. Pedersen, In vivo processing of N-terminal methionine in *E. coli*, *FEBS Lett.* 266 (1990) 1–3.
- [108] A.K. Pavlou, J.M. Reichert, Recombinant protein therapeutics-success rates, market trends and values to 2010, *Nat. Biotechnol.* 22 (2004) 1513–1519.
- [109] V.R. Ratnala, H.G. Swarts, J. VanOostrum, R. Leurs, H.J. DeGroot, R.A. Bakker, W.J. DeGrip, Large-scale overproduction, functional purification and ligand affinities of the His-tagged human histamine H1 receptor, *Eur. J. Biochem.* 271 (2004) 2636–2646.
- [110] M.R. Junttila, S. Saarinen, T. Schmidt, J. Kast, J. Westermarck, Single-step Strep-tag purification for the isolation and identification of protein complexes from mammalian cells, *Proteomics* 5 (2005) 1199–1203.
- [111] S.S. Sharma, S. Chong, S.W. Harcum, Simulation of large-scale production of a soluble recombinant protein expressed in *Escherichia coli* using an intein-mediated purification system, *Appl. Biochem. Biotechnol.* 126 (2005) 93–118.
- [112] R. Tanaka, K. Kosugi, M. Mizukami, M. Ishibashi, H. Tokunaga, M. Tokunaga, Expression and purification of thioredoxin (TrxA) and thioredoxin reductase (TrxB) from *Brevibacillus choshinensis*, *Protein Expr. Purif.* 37 (2004) 385–391.
- [113] M. Kavoosi, J. Meijer, E. Kwan, A.L. Creagh, D.G. Kilburn, C.A. Haynes, Inexpensive one-step purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima*, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 807 (2004) 87–94.
- [114] P.J. Simpson, S.J. Jamieson, M. Abou-Hachem, E.N. Karlsson, H.J. Gilbert, O. Holst, M.P. Williamson, The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase, *Biochemistry* 41 (2002) 5712–5719.
- [115] de Marco, E. Casatta, S. Savaresi, A. Geerlof, Recombinant proteins fused to thermostable partners can be purified by heat incubation, *J. Biotechnol.* 107 (2004) 125–133.