# REVIEWS

# To fuse or not to fuse: What is your purpose?

**Mark R. Bell, Mark J. Engleka, Asim Malik, and James E. Strickler***

LifeSensors, Inc., Malvern, Pennsylvania 19083

**Abstract:** Since the dawn of time, or at least the dawn of recombinant DNA technology (which for many of today's scientists is the same thing), investigators have been cloning and expressing heterologous proteins in a variety of different cells for a variety of different reasons. These range from cell biological studies looking at protein-protein interactions, post-translational modifications, and regulation, to laboratory-scale production in support of biochemical, biophysical, and structural studies, to large scale production of potential biotherapeutics. In parallel, fusion-tag technology has grown-up to facilitate microscale purification (pull-downs), protein visualization (epitope tags), enhanced expression and solubility (protein partners, e.g., GST, MBP, TRX, and SUMO), and generic purification (e.g., His-tags, streptag, and FLAG™-tag). Frequently, these latter two goals are combined in a single fusion partner. In this review, we examine the most commonly used fusion methodologies from the perspective of the ultimate use of the tagged protein. That is, what are the most commonly used fusion partners for pull-downs, for structural studies, for production of active proteins, or for large-scale purification? What are the advantages and limitations of each? This review is not meant to be exhaustive and the approach undoubtedly reflects the experiences and interests of the authors. For the sake of brevity, we have largely ignored epitope tags although they receive wide use in cell biology for immunoprecipitation.

**Keywords:** fusion tag; protein expression; GST; MBP; SUMO; TRX; His6

## Introduction

Advances in genomics, proteomics, and bioinformatics over the last thirty years have dramatically increased the use of recombinant DNA as a way to study proteins of interest for a variety of applications. Combined with affinity tagging, recombinant DNA techniques allow for the identification, modification, production, isolation, and purification of proteins from a range of host systems, including *E. coli*, yeast, plant, insect, and mammalian cell lines. However, production of recombinant proteins routinely encoun-

ters problems, including the formation of inclusion bodies, incorrect protein conformation, toxicity to the host cell, or low protein yield. These issues are most often addressed by changing expression hosts or through fusion of the protein of interest (POI) to a carrier protein (fusion tag). Located at either the N- or C- terminus of the protein of interest, fusion tags can improve protein solubility, achieve native protein folding, and increase total yield by improving expression and decreasing degradation. Fusion tags may be used *in tandem* with affinity tags and other markers to improve detection, allow for protein secretion, and achieve greater total yield.

There are several published reviews on both affinity and fusion tags from the past several

*Correspondence to: James E. Strickler, LifeSensors, Inc., 271 Great Valley Parkway, Malvern, PA 19355. E-mail: strickler@lifesensors.com

years.[1,2] While these reviews do an excellent job of describing the many tags and tag removal systems currently available, it can be difficult to determine which tags are the best candidates for specific applications. For example, it is estimated that 20–40% of eukaryotic proteins cannot be expressed in soluble form in prokaryotic hosts.[3] Given, the variability of protein structures, many tags may have similar issues. This review, therefore, focuses on tags that are utilized in specific protein applications, including protein-protein interaction "pull-down" assays, structure determination, for example, X-ray crystallography, control, and maintenance of protein functionality, and large scale manufacturing. While this list is by no means exhaustive, we hope to provide insight on the prominent tags used for these applications.

**Protein-Protein Interaction "Pull-Down" Assay**

Proteins do not work in isolation, but instead interact in complex networks. When studying any one protein, the isolation of other proteins in its complex can have several uses, including the purification of one or more of the binding partners, or the identification of unknown binding partners. Protein complex immunoprecipitation (Co-IP) is a technique in which an antibody is bound to a known target protein, allowing this protein and other proteins that are bound to it to be precipitated, or "pulled-down," out of solution and analyzed.

While effective, one of the problems with this technique is the difficulty in generating specific antibodies to the target protein. A solution is to clone the DNA of the target protein into an expression vector containing a fusion tag at either terminus of the protein. Depending on the tag, either affinity chromatography or an antibody can be used for capture of the complex. The use of affinity chromatography drastically speeds up the process of protein isolation and identification, and allows the same purification process to be used repeatedly. Additionally, this system can be used to increase the expression of the target protein beyond endogenous levels, potentially allowing more complete pull-down of the protein complex and providing greater amounts of specific bound proteins.[4,5]

The most common fusion tag used in pull-down assays is glutathione-S-transferase (GST). A 26 kDa protein from the parasitic helminth *Schistosomajaponicum*, GST binds with high affinity to glutathione.[6] When used as a fusion tag, GST can increase protein yield by allowing efficient initiation of translation.[7] GST has been used in a wide range of cell types, including *E. coli*,[8–10] yeast,[11,12] plant,[13,14] insect,[15,16] and mammalian cells.[17,18] For purification, the GST-protein fusion is bound to glutathione immobilized to a solid support such as agarose beads or magnetic particles. The fusion construct is eluted by the addition of 10 m*M* reduced glutathione. The target and associated proteins can be analyzed by standard methods such as SDS-PAGE or western blotting.[19] GST has been used as both an N- and C-terminal tag, and in many commercially available systems, a protease cleavage site is encoded between GST and the target protein, allowing removal of GST after purification.

One of the problems that can occur with GST-based pull down assays is the solubility of the binding protein. Specifically, proteins that are either highly hydrophobic or larger than 100 kDa tend to form insoluble aggregates and inclusion bodies when tagged with GST, rendering them inactive. To correct for this, detergents such as Triton X and CHAPS are often used in the purification process to enhance the solubility of the fusion complex. If the detergents disrupt the biological activity of the binding protein, a high salt buffer can also be used to encourage solubility.[20]

Another issue with GST is its propensity to dimerize. Native GST exists as a homodimer, and when fused with a target protein that can also oligomerize, the resulting fusion can form large complexes that are not easily eluted from the bound glutathione resin.[21] Strategies to prevent dimerization include modification of salt or pH, or the addition of a strong reducing agent such as dithiothreitol (DTT). Another tactic is to promote elution by the addition of extra free glutathione, or by switching the elution agent to S-butylglutathione, which has a 25-fold higher affinity for GST than does glutathione.[22] As in the case of insolubility above, several of these solutions may disrupt the functionality of the tagged protein, and some researchers have suggested that GST is not suitable for pull down assays of proteins that are known to oligomerize.[7,23]

While the most common, GST is not the only tag that is used for pull-down assays. Technically, any affinity tag that can be fused to the target protein will work, and as the most widely used affinity tag in general, it comes as little surprise that poly-histidine tags (usually hexahistidine or His$_6$) are also popular for pull-down assays. Although a His$_6$ tag does not offer increased expression or solubility levels, it is small (0.84 kDa), immunogenitically inactive, and does not dimerize. Like GST, the His$_6$ tag may be attached to the target protein at either the N- or C-terminus, and most proteins are functional with the tag attached. Purification of His$_6$-tagged fusions involves immobilized metal-affinity chromatography (IMAC), as the negatively charged histidine binds to the positively charged metal ions, most commonly $Ni^{2+}$.[24] The fusion construct can be eluted with an imidazole gradient (either stepwise or linear).[25] One disadvantage of imidazole elution is the observation that high imidazole concentrations have been found to remove metal ions from a variety

of proteins leaving them inactive and possibly altering the nature of their protein-protein interactions.[26] Beyond both GST and His[6,] a number of additional tags are used for pull-down assays, including Strep-tag,[27] and fragment crystallization (Fc)-fusions,[28] although these are seen in much lower overall numbers.

## Tags for Structural Studies

Generation of good quality protein for structural studies, such as NMR spectroscopy, X-ray crystallography, and cryoelectron microscopy, places stringent demands on protein production. These approaches require multimilligram quantities of protein at high purity and production techniques that minimize the use of detergents, chaotropes, and reducing agents that might alter the final structure of the protein or prevent the formation of good quality diffracting crystals. While incorporating fusion tags can overcome some of these challenges by increasing yield, enhancing folding, and streamlining purification, they can also create new obstacles. Multidomain fusion proteins joined by a flexible linker may be less likely to form well-ordered, diffracting crystals or be too large for NMR studies. Strategies that require tag removal introduce challenges including optimization of cleavage conditions, added costs of proteases for tag removal, and failure to recover soluble or structurally intact protein after tag removal. On balance, however, the advantages of using fusion tags for producing proteins for structural studies outweigh the disadvantages, and fusion tags have been widely adopted as the method of choice for protein production for structural study. Over 75% of proteins produced for crystallization are expressed as fusion constructs.[29]

By far, the His[6]-tag is the most common fusion partner, being used in 60% of crystallographic studies.[30] The His[6]-tag's small size does not generally interfere with crystallization and its utility in nickel affinity chromatography facilitates simple and cost-effective purification at the multimilligram scale.[31] His[6]-tags do have certain disadvantages, however, as they are variably gluconoylated in $E.$ $coli$,[32] which can lead to heterogeneity that is not conducive to crystal formation. While His[6]-tags have clearly been effective in structural studies, the Structural Genomics Center has estimated that up to 50% of all prokaryotic proteins are insoluble when expressed in $E.$ $coli$ with a His-tag,[33,34] and additional studies suggest that this number is higher for eukaryotic proteins.[35–37]

Many large fusion tags, such as maltose binding protein (MBP),[38] GST,[39] thioredoxin,[40] and small ubiquitin-like modifier (SUMO),[41] enhance solubility and, either directly or in combination with small affinity tags, simplify purification. Therefore, expression of target proteins as fusions with these tags is often used as a rescue strategy for difficult to express proteins (cf.[42]). In most cases, tags are proteolytically removed prior to crystallization by engineering endoprotease cleavage sites between the fusion tag and the protein of interest. The high homogeneity necessary for crystallization trials requires minimal non-specific or variable cleavage during tag removal. Thrombin, enterokinase (enteropeptidase) and FactorXa, commonly used for tag removal, have historically shown spurious cleavage at sites distinct from the engineered site.[43] In contrast, the viral proteases, tobacco etch virus (TEV) or human rhinovirus 3C protease, have a lower turnover, which results in fewer catalytic events at sub-optimal cleavage sites and thus greater specificity.[43] However, the lower turnover requires that large quantities of protease are required for tag removal and increases costs of protein production particularly at the multi-milligram scale. While limited to removal of SUMO tags, SUMO proteases provide both high specificity and high efficiency.[44] SUMO proteases specifically recognize the tertiary structure of the SUMO tag, a structure not found elsewhere in the proteome, rather than a linear peptide sequence, and cleave precisely at the C-terminus of the tag. Furthermore, SUMO protease has a $k_{cat}$ ~25-fold higher than TEV making the SUMO tag/SUMO protease system an efficient and cost-effective option for structural work.[45]

A growing number of cases are being documented where crystallization is performed with the intact fusion protein (reviewed in Smyth[3] and Moon[46]). Successful crystals have been described for fusions with MBP,[47–49] GST,[50–53] Thioredoxin,[54,55] Lysozyme,[56–58] and SUMO.[59] Leaving the fusion tag in place presents a number of advantages in so-called "carrier-driven" crystallization,[60–62] particularly, when the protein of interest is small enough to occupy available space between the neighboring tag molecules in the crystal lattice. High resolution structures of MBP,[63–65] GST,[66] TRX,[67,68] and SUMO[69,70] have been determined. The tags provide surfaces favorable to crystal lattice formation and by extension, the conditions for crystallization of the tag may translate to conditions for successful crystallization of the fusion protein. In addition, the structures of the tags can be used as search models to solve the crystallographic phase problem by molecular replacement methods. Examples of protein structures determined with or without tag removal are shown in Figure 1.

GST has a number of properties that make it a good candidate for carrier-driven crystallization.[39] GST's hydrophilic surface can improve the solubility of the protein of interest, GST fusions can be readily purified via glutathione resins, and finally, the structure of recombinant GST is known.[66] Several structures of small peptides and protein regulatory domains have been determined as GST fusions,
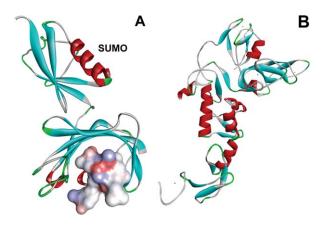
**Figure 1.** Structures of two proteins expressed as SUMO-fusions in E. coli. **Panel A** shows the structure of a peptidyl-prolylcis-trans isomerase from *Burkholderiapsuedomallei* bound to 8-deethyl-8-[but-3-enyl]-ascomycin (solid surface). The protein was crystallized and the structure determined to 1.9Å resolution with SUMO still attached (labeled). (PDB ID 3UQB, Fox III, Abendroth, Staker, and Stewart, Seattle Structural Genomics Center for Infectious Disease, deposited Nov. 2011). **Panel B** shows the structure of the human Ub E3 ligase Parkin. SUMO was enzymatically removed prior to crystallization. The structure was determined at 2.0Å resolution (PDB ID 4I1H, Riley et al.[42]).

including gp41 from HIV,[60] the C-terminal fibrinogen gamma chain,[53] the ankryn-binding domain of α-Na/K ATPase,[52] acute myelogenous leukemia-1 nuclear matrix targeting sequence [amidomethyl-luciferin (AML)-1 NMTS],[51] and DNA replication-related element-binding factor.[50] A further set of proteins has been crystallized but no structures have been reported, presumably because the fused fragment is disordered. The success of carrier-driven GST appears to be limited to date to protein fragments less than 100 amino acids.[62] The use of GST presents added drawbacks in that it does not improve solubility in all cases[71] and it forms dimers,[72,73] which can lead to aggregation of certain targets.

MBP has also been used successfully in carrier-driven crystallization.[3,46] Like GST, MBP confers increased solubility to its fusion partner[71] and its crystal structure has been solved.[63–65] In addition, the C-terminus of MBP forms a solvent accessible α-helix, which provides a rigid support for linking the protein of interest.[3] Although MBP provides a strategy for affinity purification by amylose affinity chromatography, MBP-fusion proteins often fail to bind to the amylose resin in practice.[31,74,75] Furthermore, MBP adopts different conformations in the presence or absence of bound maltose,[63–65] so partial occupancy of maltose can lead to heterogeneity in the final product, which can be detrimental to crystal formation. The addition of excess maltose may be required under certain purification schemes and crystallization trials. To circumvent these issues,

MBP is frequently used in conjunction with a His$_6$-tag for purification. To date, MBP has shown the highest degree of success as a carrier protein for protein structure determination with over 25 structures determined as MBP fusions.[46] Design of a short, rigid linker between the fusion tag and the protein of interest is critical to the success of carrier-driven crystallization. A variety of linkers and N-terminal truncations of the target protein should be tried. Linkers too long will result in excessive flexibility, whereas linkers too short may influence the target protein structure.

Additional tags have been incorporated into successful crystals, including thioredoxin,[54,55] lysozyme,[56,58] and SUMO,[76] but to a lesser extent than MBP or GST. Like MBP, the C-terminus of thioredoxin ends in a rigid, solvent accessible α-helix.[55] Much of thioredoxin's polar surface can form crystal contacts and thioredoxin readily crystallizes. Despite these features, only two crystals and one solved structure have been reported using thioredoxin as a fusion partner.[54,55] Rather than as a traditional solubility or affinity tag, lysozyme was used to replace an unstructured loop in the β2-adrenergic G-protein coupled receptor.[56,58] Lysozyme provided an increased polar surface conducive to lattice formation. Multiple structures of SUMO fusions have been placed in the NCBI structure database. The SUMO fusion tag provides an interesting option for carrier-driven crystallization. As discussed above, SUMO improves yield and solubility across a wide range of targets[44,45,77–81] and its structure has been solved to high resolution.[69,70] Most importantly, the SUMO star system is amenable to expression in eukaryotic hosts, allowing production of proteins for structural study that require expression in a eukaryotic host to produce proper folding or desired eukaryotic post-translational modifications.[41,78,82]

## Tags for Functional Activity

The need to generate functionally active proteins is a necessity of many studies, but is especially important when the protein in question is a potential therapeutic. Conformational characteristics, including proper folding and solubility, are an essential component of functionally active proteins, and these can be improved by the presence of fusion tags.[83] However, the generation of a native N-terminus is also critical for functional activity, particularly among cytokines, small peptides, and cytotoxic proteins, presenting additional challenges for tag use.

Cytokines, including interleukins (e.g., IL-6, IL-8), interferons (e.g., hIFN-γ), colony stimulating factors (G-CSF, M-CSF, and GM-CSF), and hematopoietic factors such as erythropoeitin and thrombopoeitin are of interest for their immunomodulatory effects and therapeutic potential. A range of tags have been used to express cytokines, with varying

levels of success. GST and thioredoxin have performed inconsistently to solubilize these types of proteins.[1,84–86] MBP and NusA are found to display solubility enhancing properties and increases expression levels of the target protein, perhaps due to their size.[84,87] In fact, NusA has been used to successfully express and purify cytokine homologs with the IL8-like fold.[85] Removal of these tags requires endoproteases that leave additional amino acid residues on the N-terminus of the processed protein,[1,85] which can affect the structure and/or function of the protein.[88] A tag that has seen more consistent success is SUMO. The conformation-specific activity of SUMO protease ensures that cleavage occurs precisely at the N-terminus of the target protein, and SUMO has been shown to produce an array of functionally active cytokines. IL-1β and IL-8 were expressed at high levels using SUMO, and both were biologically active.[88,89] Recently, IFN-γ was generated in a functionally active form following SUMO expression and purification, although it was expressed as an inclusion body previously.[90] TNF-α was also generated as a mature and active product for use in drug development assays.[91] Finally, the activity of chemokines can be drastically altered by truncation or extension at the N-terminus (see for instance[92,93]). In an extensive study, Lu *et al*. used a SUMO-tag to express and purify 15 chemokines in active form.[94] In their system, SUMO fusion did not help with solubility, but was essential for producing the proteins with the authentic N-terminus. They used both the standard His$_6$-SUMO-tag and, in an interesting variation on the theme, they used a tandem fusion; His$_6$-thioredoxin-SUMO, to promote refolding of the chemokines in the absence of a redox buffer.

Antimicrobial peptides are studied for their roles in physiology and potential as therapeutics. The main peptide families of interest, defensins, and cathelicidins, are synthesized as precursor-proteins that are proteolytically cleaved to produce mature peptides.[95,96] These precursors protect host cells from the cytotoxic effects of the mature peptides,[97,98] and fusion tags are used to mimic these precursors, preventing cytotoxicity, and successfully generating functional peptides.

Similar to expression work with cytokines, tags used for the production of antimicrobial peptides have encountered mixed results (see review by Li[99]). Once again GST performs inconsistently, as in a number of cases the tag failed to protect against proteolytic cleavage of the precursor peptide.[100–102] This may be due to the large size of GST (28 kDa), especially in relation to the small antimicrobial peptides. MBP has been used for the production of Human β-defensin 25 (hBD25) and Human β-defensin 28 (hBD28), but both required refolding steps to recover fusion protein from aggregates in purification.[98,103]

Additional tags have been more consistently successful in expressing antimicrobial peptides. Thioredoxin has been used to achieve high-yields of precursor peptides in the cytoplasm, perhaps because of its smaller size (11.8 kDa; 40). SUMO, also small in size (11.2 kDa), has also been used successfully in generating defensins and cathelicidins. In the production of Human β-Defensin-4, 166 mg per 1 L fermentation was obtained after purification.[104] LL-37, regarded as the only cathelicidin-derived antimicrobial peptide found in humans,[105] has been produced in conjunction with thioredoxin in a dual-tag expression.[104,106] Other functionally active proteins produced with SUMO include the antibacterial peptide CM4 (ABP-CM4),[107] the PnTx3–4 toxin isolated from spider venom,[108] and the antitumor-analgesic peptide purified from scorpion venom.[108,109]

### Self-Cleaving Affinity Tags

The use of self-splicing tags (inteins) for the purification of recombinant proteins was first described in 1997[110,111] by a group working at New England Biolabs. Their work gave rise to NEB's IMPACT (Intein-mediated purification with an affinity chitin-binding tag) system, which remains the most frequently used intein purification methodology. The technology revolves around the use of a yeast-derived protein splicing element coupled to a bacterial chitin-binding affinity purification tag. The expressed protein is captured on a chitin column, washed free of host derived proteins, and then released by induction of cleavage of the intein. Mechanistically, the intein spontaneously undergoes an S-N acyl shift at its N-terminal cysteine residue to form a thioester bond with the protein of interest. This thioester is then readily cleaved by a number of small molecules such as 2-mercaptoethanol, 2-mercaptoethansulfonic acid, or DTT. The protein of interest is released as a thioester of the small molecule which, depending on the intended use, can be used to carry-out C-terminal modification or simply hydrolyzed to generate a C-terminal carboxylate. While this technology has worked well for a large number of proteins from small anti-bacterial toxins[112–114] to antibody fragments[115,116] to human choline acetyltransferase[117] it has not become a widely used method for general protein purification. Rather it has found its widest use to generate proteins that can subsequently be used for C-terminal modification via the thioester generated during cleavage. This includes expressed protein ligation or native chemical ligation. In this technique, two disparate peptide segments are joined together through the ligation of a C-terminal peptide containing an N-terminal cysteine residue to the N-terminal peptide with a C-terminal thioester. Attack of the sulfhydryl of the cysteine on the thioester results in a thioester linkage between the two

peptides. In a reverse of the intein reaction, the thioester undergoes an N to S migration generating a standard amide bond. This methodology is frequently used to introduce non-natural amino acids into a protein when one of the two partners is produced synthetically rather than biosynthetically. One interesting use was the joining of two peptides, one of which was expressed as a heavy isotope labeled protein ($^{13}$C and $^{15}$N) to produce a more highly refined NMR structure of two protein domains, which identified an extended interaction interface between the two proteins previously thought to behave independently.[118] In another iteration of the methodology, the thioester is used to introduce small molecules such as fluorophores at the C-terminus of the protein of interest to generate enzyme substrates. For instance, a small cottage industry has grown up around the C-terminal modification of ubiquitin (Ub) and ubiquitin-like proteins (Ubls) as substrates/inhibitors of Ubl-deconjugating enzymes (deubiquitylases, desumoylases, deNEDDylases, etc). These include inhibitors such as Ub-aldehyde[119] and Ub-vinylsulfone[120] among others[121–123] and substrates such as Ub-amidomethylcoumarin (AMC),[124] Ub-rhodamine 110,[125] and Ub-AML.[126]

There are two principle disadvantages to intein technology. First, the binding capacity of chitin-agarose is very low. Typically, the capacity of chitin-agarose is 1–2 mg of recombinant protein/mL of matrix. Contrast this to Ni$^{+2}$-IMAC or IEX-sepharose columns with capacities of 40–80 mg/mL of resin. As a consequence, one must use relatively large columns to insure complete capture of the expressed protein. While this is generally not a major concern at laboratory-scale, it decreases the desirability of this methodology at manufacturing scales. Second, the cleavage reaction is very slow. On column cleavage is usually carried out for 16 h at room temperature[110] or up to 5 days at 4°C.[114,127] Long incubation times and elevated temperatures can lead to increased nonspecific proteolysis by contaminating proteases or increased risk of protein denaturation. Finally, the use of high levels of reducing agents (30 to 100 m$M$) can be detrimental to the recovery of disulfide linked proteins leading to the subsequent need for refolding.

To overcome some of these disadvantages, the Wood laboratory at Princeton has been actively engaged in modification of the technology. Their major focus has been on replacing the chitin-binding domain with other types of purification entities to avoid the low capacity of chitin-agarose as well as replace column chromatography with mechanical methods of primary separation (see review by Banki and Wood[128]). Briefly, these new methods either require a precipitation step mediated by the controlled aggregation of elastin-like polypeptide tags attached to the intein[129–132] or adsorption of phasin-tagged inteins to polyhydroxybutyrate nanogranules produced in specially engineered strains of *E. coli*.[133–136] None of these methods have been tried extensively outside of the originator's laboratory. In a separate attempt to improve purification capacity of the intein technology, Wang *et al*.[137] added either a His$_6$-Ub-tag or a His$_6$-SUMO-tag at the N-terminus of the intein with a C-terminal target protein. They could then use Ni$^{+2}$-IMAC for purification and then induce intein cleavage to release the protein of interest without the use of a protease. As an added bonus, the SUMO- and Ub-tags enhanced expression levels compared with the intein alone.

None of these approaches address, the issue of slow cleavage rate. To tackle this problem a number of groups have moved away from inteins to use proteases themselves as part of the fusion tag. These include a His$_6$-tagged cysteine protease domain from *Vibrio cholerae*[138] and a His$_6$-tagged catalytic sortase domain from *Staphylococcus aureus*.[139] In both cases, enzymatic cleavage is initiated on-column through the addition of a small molecule, inositol hexakisphosphate or Ca$^{+2}$, respectively. Again, these methodologies have not gained wide acceptance outside of the originators' laboratories. Perhaps the most well-developed of these approaches is that developed by Bryan's group[140] and marketed by Bio-Rad. as the Profinity system. In this instance, the affinity tag is the prodomain of the bacterial enzyme subtilisin. An engineered form of subtilisin, which binds with high affinity to this prodomain, is coupled to agarose for affinity purification. As with the two protease tags described above, cleavage is initiated by a small molecule, in this case fluoride. Although less potent, the enzyme can be activated by other halide ions limiting their use in buffers. Another drawback of this self-cleaving system is the requirement for separate expression, purification, and conjugation of the mutant subtilisin, which could make the system cost-prohibitive at large scale. With the other two protease-based systems, the enzyme is expressed as part of the fusion protein insuring that sufficient protease is always present.

## Large Scale Manufacturing

Large scale protein manufacturing is often quite different from small scale research production. While protein production for research and development might typically involve a few liters of culture or less to obtain the desired amount of purified product, commercial production often requires bioreactors capable of holding several thousand liters each. The challenges of high quantity and low cost bioprocessing mean that affinity and fusion tags are not commonly used due in part to the extra time and cost associated with tag removal and the subsequent purification steps at such large scales. Therefore,

when a tag system is used, it must present clear advantages in production or to the therapeutic itself.

Fc-fusion proteins are the best-known example of a fusion tag being used in large scale manufacturing. First created in 1989 as a potential AIDS therapeutic,[141] several commercially available drugs are based on Fc-fusions. These include belatacept and alefacept for organ rejection, abatacept and etanercept for rheumatoid arthritis, and a fibercept for macular degeneration, among others.[142] Fc-fusion proteins consist of a protein of interest linked to a Fc domain of an immunoglobulin, which is the segment of the heavy chain furthest from the antigen binding site. At ~250 amino acids in size, the Fc domain can be attached to either terminus of the protein of interest. Currently all commercial therapeutics use the Fc domain from human IgG1, although other options, such as IgG3, IgA, and IgM are currently being explored.[143]

As an affinity tag, the Fc domain binds with high affinity to Protein A, a surface protein originally isolated from *Staphylococcus aureus* that is often bound to a stationary phase resin. Elution can be accomplished via a pH gradient or by commercial detergents, and the combination of high affinity and simple elution allows the Fc system to be cost-effective at scales large enough for manufacturing. This purification process is only an ancillary benefit; the addition of an Fc domain to a therapeutic compound provides the fusion with several beneficial pharmacological properties. Most prominently, the addition of an Fc domain increases the serum half-life of a therapeutic, increasing its activity. This is accomplished by the Fc domain binding with the neonatal Fc-receptor (FcRn), which plays a role in protein recycling by preventing lysosomal degradation.[144–146] The Fc domain can also interact with Fc-receptors present on certain cells in the immune system, such as B-lymphocytes and natural killer cells, an ability which is important to certain types of oncology therapeutics.[147,148] Finally, the addition of an Fc domain can improve the solubility of the overall fusion due to the fact that it folds independently of its fusion partner.[149]

Since Fc-domains are essentially a functional section of the final therapeutic, no cleavage step is necessary. In order for afusion system that does undergo tag removal to be considered for larger scale protein expression, additional benefits must be present to offset the additional purification costs. One such system is SUMO. As previously mentioned, SUMO can increase both the yield and solubility of its fusion partner and cleavage of the tag is performed by a specific SUMO protease, which leaves no additional amino acids behind that can disrupt therapeutic function. When combined with a $His_6$ or other affinity tag on both the SUMO tag and the protease, a simple purification scheme based on

IMAC chromatography can be used, and current SUMO tags can be used in several expression systems, including *E. coli*, *Pichia pastoris*, and mammalian cells such as HEK and CHO.[45] Recent examples include the production of the antiviral protein cyanovirin-N (CVN), a microbicide with anti-HIV effects. The expression of CVN with a $His_6$-SUMO tag lead to a soluble protein level in *E. coli* of over 30% of the total soluble protein using a 30 L bioreactor.[150] Additionally, SUMO was used in the production of subunit antigens to anthrax and swine flu in a test sponsored by DARPA that required both rapid antigen production and cost-per-dose thresholds. At the 30 L scale, the maximum product output was 14 g/L, and the cost-per-dose was well under the requirements.

## Acknowledgements

## References

1. Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Exp Purif 48:1–13.
2. Young CL, Britton ZT, Robinson AS (2012) Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol J 7:620–634.
3. Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P, Kobe B (2003) Crystal structures of fusion proteins with large-affinity tags. Protein Sci 12:1313–1322.
4. Ron D, Dressler H (1992) pGSTag-a versatile bacterial expression plasmid for enzymatic labeling of recombinant proteins. Biotechniques 13:866–869.
5. Malhotra A (2009) Tagging for protein expression. Methods Enzymol 463:239–258.
6. Smith DB, Johnson KS (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene 67:31–40.
7. Waugh DS (2005) Making the most of affinity tags. Trends Biotechnol 23:316–320.
8. Frangioni JV, Neel BG (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal Biochem 210:179–187.
9. Schrodel A, de Marco A (2005) Characterization of the aggregates formed during recombinant protein expression in bacteria. BMC Biochem 6:10.
10. Fang M, Wang W, Wang Y, Ru B (2004) Bacterial expression and purification of biologically active human TFF3. Peptides 25:785–792.
11. Mitchell DA, Marshall TK, Deschenes RJ (1993) Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. Yeast 9:715–722.
12. Rodal AA, Duncan M, Drubin D (2002) Purification of glutathione S-transferase fusion proteins from yeast. Methods Enzymol 351:168–172.

13. Zhang N, Qiao Z, Liang Z, Mei B, Xu Z, Song R (2012) Zea mays Taxilin protein negatively regulates opaque-2 transcriptional activity by causing a change in its sub-cellular distribution. PLoS One 7:e43822.

14. Song CP, Galbraith DW (2006) AtSAP18, an ortho-logue of human SAP18, is involved in the regulation of salt stress and mediates transcriptional repression in Arabidopsis. Plant Mol Biol 60:241–257.

15. Wong A, Albright SN, Wolfner MF (2006) Evidence for structural constraint on ovulin, a rapidly evolving Drosophila melanogaster seminal protein. Proc Natl Acad Sci U S A 103:18644–18649.

16. Bao X, Zhang W, Krencik R, Deng H, Wang Y, Girton J, Johansen J, Johansen KM (2005) The JIL-1 kinase interacts with lamin Dm0 and regulates nuclear lamina morphology of Drosophila nurse cells. J Cell Sci 118:5079–5087.

17. Thomson RB, Wang T, Thomson BR, Tarrats L, Girardi A, Mentone S, Soleimani M, Kocher O, Aronson PS (2005) Role of PDZK1 in membrane expression of renal brush border ion exchangers. Proc Natl Acad Sci U S A 102:13331–13336.

18. Rudert F, Visser E, Gradl G, Grandison P, Shemshedini L, Wang Y, Grierson A, Watson J (1996) pLEF, a novel vector for expression of glutathione S-transferase fusion proteins in mammalian cells. Gene 169:281–282.

19. Harper S, Speicher DW (2011) Purification of proteins fused to glutathione S-transferase. Methods Mol Biol 681:259–280.

20. Deceglie S, Lionetti C, Roberti M, Cantatore P, Loguercio Polosa P (2012) A modified method for the purification of active large enzymes using the glutathione S-transferase expression system. Anal Biochem 421:805–807.

21. Maru Y, Afar DE, Witte ON, Shibuya M (1996) The dimerization property of glutathione S-transferase partially reactivates Bcr-Abl lacking the oligomerization domain. J Biol Chem 271:15353–15357.

22. Vinckier NK, Chworos A, Parsons SM (2011) Improved isolation of proteins tagged with glutathione S-transferase. Protein Expr Purif 75:161–164.

23. Tolia NH, Joshua-Tor L (2006) Strategies for protein coexpression in Escherichia coli. Nat Methods 3:55–64.

24. Hochuli E, Dobeli H, Schacher A (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J Chromatogr 411:177–184.

25. Hefti MH, Van Vugt-Van der Toorn CJ, Dixon R, Vervoort J (2001) A novel purification method for histidine-tagged proteins containing a thrombin cleavage site. Anal Biochem 295:180–185.

26. Huang A, de Jong RN, Folkers GE, Boelens R (2010) NMR characterization of foldedness for the production of E3 RING domains. J Struct Biol 172:120–127.

27. Lamberti A, Sanges C, Chambery A, Migliaccio N, Rosso F, Di Maro A, Papale F, Marra M, Parente A, Caraglia M, Abbruzzese A, Arcari P (2011) Analysis of interaction partners for eukaryotic translation elongation factor 1A M-domain by functional proteomics. Biochimie 93:1738–1746.

28. Noberini R, Rubio de la Torre E, Pasquale EB (2012) Profiling Eph receptor expression in cells and tissues: a targeted mass spectrometry approach. Cell Adh Migr 6:102–112.

29. Uhlen M, Forsberg G, Moks T, Hartmanis M, Nilsson B (1992) Fusion proteins in biotechnology. Curr Opin Biotechnol 3:363–369.

30. Derewenda ZS (2004) The use of recombinant methods and molecular engineering in protein crystallization. Methods 34:354–363.

31. Bucher MH, Evdokimov AG, Waugh DS (2002) Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrin-binding protein. Acta Cryst D58:392–397.

32. Geoghegan KF, Dixon HB, Rosner PJ, Hoth LR, Lanzetti AJ, Borzilleri KA, Marr ES, Pezzullo LH, Martin LB, LeMotte PK, McColl AS, Kamath AV, Stroh JG (1999) Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. Anal Biochem 267:169–184.

33. Stevens RC (2000) Design of high-throughput methods of protein production for structural biology. Structure 8:R177–185.

34. Edwards AM, Arrowsmith CH, Christendat D, Dharamsi A, Friesen JD, Greenblatt JF, Vedadi M (2000) Protein production: feeding the crystallographers and NMR spectroscopists. Nat Struct Biol 7:970–972.

35. Shih YP, Kung WM, Chen JC, Yeh CH, Wang AH, Wang TF (2002) High-throughput screening of soluble recombinant proteins. Protein Sci 11:1714–1719.

36. Hammarstrom M, Hellgren N, van Den Berg S, Berglund H, Hard T (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. Protein Sci 11:313–321.

37. Braun P, Hu Y, Shen B, Halleck A, Koundinya M, Harlow E, LaBaer J (2002) Proteome-scale purification of human proteins from bacteria. Proc Natl Acad Sci U S A 99:2654–2659.

38. di Guan C, Li P, Riggs PD, Inouye H (1988) Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene 67:21–30.

39. Smith DB, Johnson KS (1988) Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 67:31–40.

40. LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. Biotechnology (N Y) 11:187–193.

41. Panavas T, Sanders C, Butt TR (2009) SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. Methods Mol Biol 497:303–317.

42. Riley BE, Lougheed JC, Callaway K, Velasquez M, Brecht E, Nguyen L, Shaler T, Walker D, Yang Y, Regnstrom K, Diep L, Zhang Z, Chiou S, Bova M, Artis DR, Yao N, Baker J, Yednock T, Johnston JA (2013) Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases. Nat Commun 4.

43. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. Protein Expr Purif 80:283–293.

44. Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J Struct Funct Genomics 5:75–86.

45. Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR (2006) Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. Protein Sci 15:182–189.

46. Moon AF, Mueller GA, Zhong X, Pedersen LC (2010) A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. Protein Sci 19:901–913.

47. Liu Y, Manna A, Li R, Martin WE, Murphy RC, Cheung AL, Zhang G (2001) Crystal structure of the SarR protein from Staphylococcus aureus. Proc Natl Acad Sci U S A 98:6877–6882.

48. Ke A, Wolberger C (2003) Insights into binding cooperativity of MATa1/MATalpha2 from the crystal structure of a MATa1 homeodomain-maltose binding protein chimera. Protein Sci 12:306–312.

49. Kobe B, Center RJ, Kemp BE, Poumbourios P (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proc Natl Acad Sci U S A 96:4319–4324.

50. Kuge M, Fujii Y, Shimizu T, Hirose F, Matsukage A, Hakoshima T (1997) Use of a fusion protein to obtain crystals suitable for X-ray analysis: crystallization of a GST-fused protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF. Protein Sci 6:1783–1786.

51. Tang L, Guo B, Javed A, Choi JY, Hiebert S, Lian JB, van Wijnen AJ, Stein JL, Stein GS, Zhou GW (1999) Crystal structure of the nuclear matrix targeting signal of the transcription factor acute myelogenous leukemia-1/polyoma enhancer-binding protein 2alphaB/core binding factor alpha2. J Biol Chem 274:33580–33586.

52. Zhang Z, Devarajan P, Dorfman AL, Morrow JS (1998) Structure of the ankyrin-binding domain of alpha-Na,K-ATPase. J Biol Chem 273:18681–18684.

53. Ware S, Donahue JP, Hawiger J, Anderson WF (1999) Structure of the fibrinogen gamma-chain integrin binding and factor XIIIa cross-linking sites obtained through carrier protein driven crystallization. Protein Sci 8:2663–2671.

54. Stoll VS, Manohar AV, Gillon W, MacFarlane EL, Hynes RC, Pai EF (1998) A thioredoxin fusion protein of VanH, a D-lactate dehydrogenase from Enterococcus faecium: cloning, expression, purification, kinetic analysis, and crystallization. Protein Sci 7:1147–1155.

55. Corsini L, Hothorn M, Scheffzek K, Sattler M, Stier G (2008) Thioredoxin as a fusion tag for carrier-driven crystallization. Protein Sci 17:2070–2079.

56. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK (2007) GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. Science 318:1266–1273.

57. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322:1211–1217.

58. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318:1258–1265.

59. Madej T, Addess KJ, Fong JH, Geer LY, Geer RC, Lanczycki CJ, Liu C, Lu S, Marchler-Bauer A, Panchenko AR, Chen J, Thiessen PA, Wang Y, Zhang D, Bryant SH (2012) MMDB: 3D structures and macromolecular interactions. Nucleic Acids Res 40:D461–464.

60. Lim K, Ho JX, Keeling K, Gilliland GL, Ji X, Ruker F, Carter DC (1994) Three-dimensional structure of Schistosoma japonicum glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. Protein Sci 3:2233–2244.

61. Donahue JP, Patel H, Anderson WF, Hawiger J (1994) Three-dimensional structure of the platelet integrin recognition segment of the fibrinogen gamma chain obtained by carrier protein-driven crystallization. Proc Natl Acad Sci U S A 91:12178–12182.

62. Zhan Y, Song X, Zhou GW (2001) Structural analysis of regulatory protein domains using GST-fusion proteins. Gene 281:1–9.

63. Quiocho FA, Spurlino JC, Rodseth LE (1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure 5:997–1015.

64. Spurlino JC, Lu GY, Quiocho FA (1991) The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. J Biol Chem 266:5202–5219.

65. Sharff AJ, Rodseth LE, Spurlino JC, Quiocho FA (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. Biochemistry 31:10657–10663.

66. McTigue MA, Williams DR, Tainer JA (1995) Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from Schistosoma japonica and its complex with the leading antischistosomal drug praziquantel. J Mol Biol 246:21–27.

67. Katti SK, LeMaster DM, Eklund H (1990) Crystal structure of thioredoxin from Escherichia coli at 1.68 A resolution. J Mol Biol 212:167–184.

68. Jeng MF, Campbell AP, Begley T, Holmgren A, Case DA, Wright PE, Dyson HJ (1994) High-resolution solution structures of oxidized and reduced Escherichia coli thioredoxin. Structure 2:853–868.

69. Sheng W, Liao X (2002) Solution structure of a yeast ubiquitin-like protein Smt3: the role of structurally less defined sequences in protein-protein recognitions. Protein Sci 11:1482–1491.

70. Mossessova E, Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Mol Cell 5:865–876.

71. Kapust RB, Waugh DS (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 8:1668–1674.

72. Parker MW, Lo Bello M, Federici G (1990) Crystallization of glutathione S-transferase from human placenta. J Mol Biol 213:221–222.

73. Ji X, Zhang P, Armstrong RN, Gilliland GL (1992) The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-A resolution. Biochemistry 31:10169–10184.

74. Baneyx F (1999) Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol 10:411–421.

75. Routzahn KM, Waugh DS (2002) Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. J Struct Funct Genomics 2:83–92.

76. Assenberg R, Delmas O, Graham SC, Verma A, Berrow N, Stuart DI, Owens RJ, Bourhy H, Grimes

JM (2008) Expression, purification and crystallization of a lyssavirus matrix (M) protein. Acta Cryst F64: 258–262.

77. Liu X, Chen Y, Wu X, Li H, Jiang C, Tian H, Tang L, Wang D, Yu T, Li X (2012) SUMO fusion system facilitates soluble expression and high production of bioactive human fibroblast growth factor 23 (FGF23). Appl Microbiol Biotechnol 96:103–111.

78. Peroutka RJ, Elshourbagy N, Piech T, Butt TR (2008) Enhanced protein expression in mammalian cells using engineered SUMO fusions: secreted phospholipase A2. Protein Sci 17:1586–1595.

79. Butt TR, Edavettal SC, Hall JP, Mattern MR (2005) SUMO fusion technology for difficult-to-express proteins. Protein Expr Purif 43:1–9.

80. Zuo X, Li S, Hall J, Mattern MR, Tran H, Shoo J, Tan R, Weiss SR, Butt TR (2005) Enhanced expression and purification of membrane proteins by SUMO fusion in Escherichia coli. J Struct Funct Genomics 6: 103–111.

81. Zuo X, Mattern MR, Tan R, Li S, Hall J, Sterner DE, Shoo J, Tran H, Lim P, Serafinaos S, Kazi L, Navas-Martin, Weiss SR, Butt TR (2005) Expression and purification of SARS coronavirus proteins using SUMO fusions. Protein Expr Purif 42:100–110.

82. Liu L, Spurrier J, Butt TR, Strickler JE (2008) Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. Protein Expr Purif 62:21–28.

83. Vazquez E, Corchero JL, Villaverde A (2011) Post-production protein stability: trouble beyond the cell factory. Microb Cell Fact 10:60.

84. Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307:249–264.

85. Nausch H, Huckauf J, Koslowski R, Meyer U, Broer I, Mikschofsky H (2013) Recombinant production of human interleukin 6 in Escherichia coli. PLoS One 8: e54933.

86. Kim TW, Chung BH, Chang YK (2005) Production of soluble human interleukin-6 in cytoplasm by fed-batch culture of recombinant E. coli. Biotechnol Prog 21:524–531.

87. Tomczak A, Sontheimer J, Drechsel D, Hausdorf R, Gentzel M, Shevchenko A, Eichler S, Fahmy K, Buchholz F, Pisabarro MT (2012) 3D profile-based approach to proteome-wide discovery of novel human chemokines. PLoS One 7:e36151.

88. Kirkpatrick RB, Grooms M, Wang F, Fenderson H, Feild J, Pratta MA, Volker C, Scott G, Johanson K (2006) Bacterial production of biologically active canine interleukin-1beta by seamless SUMO tagging and removal. Protein Expr Purif 50:102–110.

89. Cui X, Han Y, Pan Y, Xu X, Ren W, Zhang S (2011) Molecular cloning, expression and functional analysis of interleukin-8 (IL-8) in South African clawed frog (Xenopus laevis). Dev Comp Immunol 35:1159–1165.

90. Zhu F, Wang Q, Pu H, Gu S, Luo L, Yin Z (2013) Optimization of soluble human interferon-gamma production in Escherichia coli using SUMO fusion partner. World J Microbiol Biotechnol 29:319–325.

91. Hoffmann A, Muller MQ, Gloser M, Sinz A, Rudolph R, Pfeifer S (2010) Recombinant production of bioactive human TNF-alpha by SUMO-fusion system-high yields from shake-flask culture. Protein Expr Purif 72:238–243.

92. Nibbs RJ, Salcedo TW, Campbell JD, Yao XT, Li Y, Nardelli B, Olsen HS, Morris TS, Proudfoot AE, Patel VP, Graham GJ (2000) C-C chemokine receptor 3 antagonism by the beta-chemokine macrophage inflammatory protein 4, a property strongly enhanced by an amino-terminal alanine-methionine swap. J Immunol 164:1488–1497.

93. Proudfoot AE, Buser R, Borlat F, Alouani S, Soler D, Offord RE, Schroder JM, Power CA, Wells TN (1999) Amino-terminally modified RANTES analogues demonstrate differential effects on RANTES receptors. J Biol Chem 274:32478–32485.

94. Lu Q, Burns MC, McDevitt PJ, Graham TL, Sukman AJ, Fornwald JA, Tang X, Gallagher KT, Hunsberger GE, Foley JJ, Schmidt DB, Kerrigan JJ, Lewis TS, Ames RS, Johanson KO (2009) Optimized procedures for producing biologically active chemokines. Prot Exp Purif 65:251–260.

95. Ganz T (2005) Defensins and other antimicrobial peptides: a historical perspective and an update. Comb Chem High Throughput Screen 8:209–217.

96. Daher KA, Lehrer RI, Ganz T, Kronenberg M (1988) Isolation and characterization of human defensin cDNA clones. Proc Natl Acad Sci U S A 85: 7327–7331.

97. Tongaonkar P, Golji AE, Tran P, Ouellette AJ, Selsted ME (2012) High fidelity processing and activation of the human alpha-defensin HNP1 precursor by neutrophil elastase and proteinase 3. PLoS One 7: e32469.

98. Li X, Leong SS (2011) A chromatography-focused bioprocess that eliminates soluble aggregation for bioactive production of a new antimicrobial peptide candidate. J Chromatogr A 1218:3654–3659.

99. Li Y (2011) Recombinant production of antimicrobial peptides in Escherichia coli: a review. Protein Expr Purif 80:260–267.

100. Chen YQ, Zhang SQ, Li BC, Qiu W, Jiao B, Zhang J, Diao ZY (2008) Expression of a cytotoxic cationic antibacterial peptide in Escherichia coli using two fusion partners. Protein Expr Purif 57:303–311.

101. Si LG, Liu XC, Lu YY, Wang GY, Li WM (2007) Soluble expression of active human beta-defensin-3 in Escherichia coli and its effects on the growth of host cells. Chin Med J (Engl) 120:708–713.

102. Skosyrev VS, Kulesskiy EA, Yakhnin AV, Temirov YV, Vinokurov LM (2003) Expression of the recombinant antibacterial peptide sarcotoxin IA in Escherichia coli cells. Protein Expr Purif 28:350–356.

103. Tay DK, Rajagopalan G, Li X, Chen Y, Lua LH, Leong SS (2011) A new bioproduction route for a novel antimicrobial peptide. Biotechnol Bioeng 108:572–581.

104. Li JF, Zhang J, Zhang Z, Ma HW, Zhang JX, Zhang SQ (2010) Production of bioactive human beta-defensin-4 in Escherichia coli using SUMO fusion partner. Protein J 29:314–319.

105. Durr UH, Sudheendra US, Ramamoorthy A (2006) LL-37, the only human member of the cathelicidin family of antimicrobial peptides. Biochim Biophys Acta 1758:1408–1425.

106. Bommarius B, Jenssen H, Elliott M, Kindrachuk J, Pasupuleti M, Gieren H, Jaeger KE, Hancock RE, Kalman D (2010) Cost-effective expression and purification of antimicrobial and host defense peptides in Escherichia coli. Peptides 31:1957–1965.

107. Li JF, Zhang J, Song R, Zhang JX, Shen Y, Zhang SQ (2009) Production of a cytotoxic cationic antibacterial peptide in Escherichia coli using SUMO fusion partner. Appl Microbiol Biotechnol 84:383–388.

108. Souza IA, Cino EA, Choy WY, Cordeiro MN, Richardson M, Chavez-Olortegui C, Gomez MV, Prado MA, Prado

VF (2012) Expression of a recombinant Phoneutria toxin active in calcium channels. Toxicon 60:907–918.

109. Cao P, Yu J, Lu W, Cai X, Wang Z, Gu Z, Zhang J, Ye T, Wang M (2010) Expression and purification of an antitumor-analgesic peptide from the venom of Mesobuthus martensii Karsch by small ubiquitin-related modifier fusion in Escherichia coli. Biotechnol Prog 26:1240–1244.

110. Chong S, Mersha FB, Comb DG, Scott ME, Landry D, Vence LM, Perler FB, Benner J, Kucera RB, Hirvonen CA, Pelletier JJ, Paulus H, Xu MQ (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 192:277–281.

111. Chong S, Montello GE, Zhang A, Cantor EJ, Liao W, Xu M-Q, Benner J (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. Nucleic Acids Res 26:5109–5115.

112. Morassutti C, De Amicis F, Bandiera A, Marchetti S (2005) Expression of SMAP-29 cathelicidin-like peptide in bacterial cells by intein-mediated system. Protein Expr Purif 39:160–168.

113. Morassutti C, De Amicis F, Skerlavaj B, Zanetti M, Marchetti S (2002) Production of a recombinant antimicrobial peptide in transgenic plants using a modified VMA intein expression system. FEBS Lett 519: 141–146.

114. Wang H, Meng XL, Xu JP, Wang J, Ma CW (2012) Production, purification, and characterization of the cecropin from Plutella xylostella, pxCECA1, using an intein-induced self-cleavable system in Escherichia coli. Appl Microbiol Biotechnol 94:1031–1039.

115. Wu WY, Miller KD, Coolbaugh M, Wood DW (2011) Intein-mediated one-step purification of Escherichia coli secreted human antibody fragments. Protein Expr Purif 76:221–228.

116. Reulen SW, van Baal I, Raats JM, Merkx M (2009) Efficient, chemoselective synthesis of immunomicelles using single-domain antibodies with a C-terminal thioester. BMC Biotechnol 9:66.

117. Kim AR, Doherty-Kirby A, Lajoie G, Rylett RJ, Shilton BH (2005) Two methods for large-scale purification of recombinant human choline acetyltransferase. Protein Expr Purif 40:107–117.

118. Vitali F, Henning A, Oberstrass FC, Hargous Y, Auweter SD, Erat M, Allain FH-T (2006) Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. EMBO J 25: 150–162.

119. Melandri F, Grenier L, Plamondon L, Huskey WP, Stein RL (1996) Kinetic studies on the inhibition of isopeptidase T by ubiquitin aldehyde. Biochemistry 35:12893–12900.

120. Borodovsky A, Kessler BM, Casagrande R, Overkleeft HS, Wilkinson KD, Ploegh HL (2001) A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. EMBO J 20:5187–5196.

121. Borodovsky A, Ovaa H, Kolli N, Gan-Erdene T, Wilkinson KD, Ploegh HL, Kessler BM (2002) Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. Chem Biol 9:1149–1159.

122. Hemelaar J, Galardy PJ, Borodovsky A, Kessler BM, Ploegh HL, Ovaa H (2003) Chemistry-based functional proteomics: mechanism-based activity-profiling tools for ubiquitin and ubiquitin-like specific proteases. J Proteome Res 3:268–276.

123. Hemelaar J, Galardy PJ, Borodovsky A, Kessler BM, Ploegh HL, Ovaa H (2004) Chemistry-based functional proteomics: mechanism-based activity-profiling tools for ubiquitin and ubiquitin-like specific proteases. J Proteome Res 3:268–276.

124. Dang LC, Melandri FD, Stein RL (1998) Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. Biochemistry 37:1868–1879.

125. Hassiepen U, Eidhoff U, Meder G, Bulber JF, Hein A, Bodendorf U, Lorthiois E, Martoglio B (2007) A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate. Anal Biochem 371:201–207.

126. Orcutt SJ, Wu J, Eddins MJ, Leach CA, Strickler JE (2012) Bioluminescence assay platform for selective and sensitive detection of Ub/Ubl proteases. Biochim Biophys Acta 1823:2079–2086.

127. Humphries HE, Christodoulides M, Heckels JE (2002) 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 26:243–248.

128. Banki MR, Wood DW (2005) Inteins and affinity resin substitutes for protein purification and scale up. Microb Cell Fact 4:32.

129. Banki MR, Feng L, Wood DW (2005) Simple bioseparations using self-cleaving elastin-like polypeptide tags. Nat Methods 2:659–661.

130. Fong BA, Gillies AR, Ghazi I, LeRoy G, Lee KC, Westblade LF, Wood DW (2010) Purification of Escherichia coli RNA polymerase using a self-cleaving elastin-like polypeptide tag. Protein Sci 19: 1243–1252.

131. Fong BA, Wu WY, Wood DW (2009) Optimization of ELP-intein mediated protein purification by salt substitution. Protein Expr Purif 66:198–202.

132. Wu WY, Mee C, Califano F, Banki MR, Wood DW (2006) Recombinant protein purification by self-cleaving aggregation tag. Nat Protoc 1:2257–2262.

133. Banki MR, Gerngross TU, Wood DW (2005) Novel and economical purification of recombinant proteins: intein-mediated protein purification using in vivo polyhydroxybutyrate (PHB) matrix association. Protein Sci 14:1387–1395.

134. Gillies AR, Mahmoud RB, Wood DW (2009) PHB-intein-mediated protein purification strategy. Methods Mol Biol 498:173–183.

135. Wang Z, Wu H, Chen J, Zhang J, Yao Y, Chen GQ (2008) A novel self-cleaving phasin tag for purification of recombinant proteins based on hydrophobic polyhydroxyalkanoate nanoparticles. Lab Chip 8: 1957–1962.

136. Barnard GC, McCool JD, Wood DW, Gerngross TU (2005) Integrated recombinant protein expression and purification platform based on Ralstonia eutropha. Appl Environ Microbiol 71:5735–5742.

137. Wang Z, Li N, Wang Y, Wu Y, Mu T, Zheng Y, Huang L, Fang X (2012) Ubiquitin-intein and SUMO2-intein fusion systems for enhanced protein production and purification. Protein Expr Purif 82: 174–178.

138. Shen A, Lupardus PJ, Morell M, Ponder EL, Sadaghiani AM, Garcia KC, Bogyo M (2009) Simplified, enhanced protein purification using an inducible, autoprocessing enzyme tag. PLoS One 4:e8119.

139. Mao H (2004) A self-cleavable sortase fusion for one-step purification of free recombinant proteins. Protein Expr Purif 37:253–263.

140. Ruan B, Fisher KE, Alexander PA, Doroshko V, Bryan PN (2004) Engineering subtilisin into a fluoride-triggered processing protease useful for one-step protein purification. Biochemistry 43:14539–14546.

141. Capon DJ, Chamow SM, Mordenti J, Marsters SA, Gregory T, Mitsuya H, Byrn RA, Lucas C, Wurm FM, Groopman JE, Broder S, Smith DH (1989) Designing CD4 immunoadhesins for AIDS therapy. Nature 337: 525–531.

142. Strohl WR, Knight DM (2009) Discovery and development of biopharmaceuticals: current issues. Curr Opin Biotechnol 20:668–672.

143. Czajkowsky DM, Hu J, Shao Z, Pleass RJ (2012) Fc-fusion proteins: new developments and future perspectives. EMBO Mol Med 4:1015–1028.

144. Roopenian DC, Akilesh S (2007) FcRn: the neonatal Fc receptor comes of age. Nat Rev Immunol 7: 715–725.

145. Huang C (2009) Receptor-Fc fusion therapeutics, traps, and MIMETIBODY technology. Curr Opin Biotechnol 20:692–699.

146. Jazayeri JA, Carroll GJ (2008) Fc-based cytokines: prospects for engineering superior therapeutics. Bio Drugs 22:11–26.

147. Beck A, Reichert JM (2011) Therapeutic Fc-fusion proteins and peptides as successful alternatives to antibodies. MAbs 3:415–416.

148. Nimmerjahn F, Ravetch JV (2008) Fcgamma receptors as regulators of immune responses. Nat Rev Immunol 8:34–47.

149. Carter PJ (2011) Introduction to current and future protein therapeutics: a protein engineering perspective. Exp Cell Res 317:1261–1269.

150. Xiong S, Fan J, Kitazato K (2010) The antiviral protein cyanovirin-N: the current state of its production and applications. Appl Microbiol Biotechnol 86:805–812.