DOI 10.1002/biot.201100155 Biotechnol. J. 2012, 7, 620–634

Review

Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications

Carissa L. Young[1], Zachary T. Britton[1] and Anne S. Robinson[1,2]

[1] Department of Chemical Engineering, University of Delaware, Newark, DE, USA
[2] Department of Chemical and Biomolecular Engineering, Tulane University, New Orleans, LA, USA

Protein fusion tags are indispensible tools used to improve recombinant protein expression yields, enable protein purification, and accelerate the characterization of protein structure and function. Solubility-enhancing tags, genetically engineered epitopes, and recombinant endoproteases have resulted in a versatile array of combinatorial elements that facilitate protein detection and purification in microbial hosts. In this comprehensive review, we evaluate the most frequently used solubility-enhancing and affinity tags. Furthermore, we provide summaries of well-characterized purification strategies that have been used to increase product yields and have widespread application in many areas of biotechnology including drug discovery, therapeutics, and pharmacology. This review serves as an excellent literature reference for those working on protein fusion tags.

Received 22 July 2011
Revised 23 November 2011
Accepted 29 November 2011

Supporting information available online

Keywords: Affinity purification · Affinity tags · Proteases · Solubility enhancement

1 Introduction

Recent advances in genomics, proteomics, and bioinformatics have facilitated the use of recombinant DNA technology in order to evaluate any protein of interest, without prior knowledge of the protein's cellular location or function. The parallel use of affinity tags with recombinant DNA techniques, allows the facile modification of proteins of interest leading to efficient identification, production, and isolation from the host system. However, protein insolubility, conformation, stability, and structural flexibility, as well as low purification yields and host cell toxicity are challenges that must be resolved when microbial hosts are used to express recombinant proteins. To address these challenges of production and purification efficiency, fusion tags are incorporated to increase expression yields and influence solubility and native folding; novel tags in combination with affinity techniques increase purification yields; and proteases result in tag removal.

In recent years, numerous fusion tags have been developed for recombinant protein production. While contemporary reports have included several overviews of available affinity tags for protein detection or purification [1–4], solubility enhancement [5], tag removal [1], or applications [6, 7] – in this comprehensive review, we provide an extensive summary of various tags and established purification strategies. We describe several design aspects for these tags that should be considered for recombinant fusion protein expression, including amino acid composition and size; N- or C-terminal fusions, used individually or in tandem; as well as optimized tags, techniques, and buffer conditions that lead to purified protein. Additionally, we discuss the availability of protocols and expression vectors through commercial vendors or DNA

Correspondence: Dr. Anne S. Robinson, Department of Chemical and Biomolecular Engineering, Tulane University, 300 Boggs Laboratory, New Orleans, LA 70118–5674, USA
E-mail: asr@tulane.edu

Abbreviations: Arg, arginine; **DHFR,** dihydrofolate reductase; **DsbA,** protein disulfide isomerase I; **GFP,** green fluorescent protein; **GST,** glutathione S-transferase; **His,** Histidine; **IMAC,** immobilized metal-affinity chromatography; **KSI,** ketosteroid isomerase; **MBP,** maltose-binding protein; **Ni^{2+}-NTA,** Ni(II)-nitrilotriacetic acid; **NusA,** N-utilization substance A; **SUMO,** small ubiquitin related modifier; **TAP,** tandem affinity purification; **TEV,** tobacco etch virus; **TrxA,** thioredoxin A

repositories. Details of these plasmid-based designs and methods are available in the Supporting information.

2 Using protein tags to improve yields of recombinant proteins

2.1 Enhancing protein expression and/or solubility

2.1.1 Fusion proteins affecting solubility

A common method to overcome expression obstacles is to use a fusion protein strategy. In this strategy, a difficult-to-express protein or peptide is fused to one or more other proteins to stabilize expression in the soluble or insoluble fraction, with the idea that the fusion protein will drive the resulting expression. A number of fusion proteins have been used for this purpose, including, glutathione S-transferase (GST), maltose-binding protein (MBP), thioredoxin A (TrxA), small ubiquitin related modifier (SUMO), ketosteroid isomerase (KSI), and TrpΔLE, which are outlined briefly below and in Supporting information, Table S1.

While providing potential advantages in protein expression and solubility, the fusion protein strategy also presents a new set of issues, including the removal of the carrier protein and questions regarding whether or not the protein of interest retains native structure and activity. For example, carrier proteins may sterically hinder native complex formation. Ultimately, parallel analysis of many fusion proteins may be required to generate the yields and properties of the protein of interest for the intended application.

2.1.1.1. Proteins that increase solubility

2.1.1.1.1 Glutathione s-transferase

Glutathione S-transferase (GST) from *Schistosoma japonicum* is a 26 kDa protein that has been utilized for single-step purification of fusion proteins [8]. GST fusion proteins can be purified from crude lysate by affinity for immobilized glutathione and eluted under non-denaturing conditions with 10 mM reduced glutathione. GST fusion proteins can be detected using an enzymatic assay or immunoassay. In many cases, the GST-tag protects against intracellular proteolysis and stabilizes the recombinant protein in the soluble fraction as monomers or homodimers [9–13]. However, GST is considered to be a poor solubility enhancer as some GST fusion proteins are wholly or partly insoluble. Proteins fused to GST that are enriched in either hydrophobic regions or charged residues or are larger than 100 kDa may contribute to insoluble expression. Insoluble GST fusion proteins may be purified after solubilization with mild detergents [14]. GST expression vectors, such as the pGEX series from GE Healthcare, commonly include specific protease cleavage sites between the tag and partner proteins. Thus, the GST-tag can readily be removed from GST fusion proteins after or during purification, where affinity for immobilized glutathione greatly simplifies purification of the protein of interest from cleavage products. The GST-tag has been used as an N- or C-terminal tag in a variety of expression systems, including bacteria [8, 14], yeast [15–17], insect cells [18, 19], and mammalian cells [20].

In addition to advantages in expression and purification, GST-tagged fusion proteins have proved useful in studies on protein–DNA interactions [19, 21], protein–protein interactions [22, 23], and as antigens for vaccine studies [24].

2.1.1.1.2 Maltose-binding protein

Maltose-binding protein (MBP) is a 42 kDa protein encoded by the *mal*E gene of *Escherichia coli* K12 [25]. MBP fusion proteins have been utilized for single-step purification by affinity to cross-linked amylose [26]. MBP fusion proteins bound to immobilized amylose are eluted under non-denaturing conditions with 10 mM maltose. MBP fused to either the N- and C-terminus has been shown to increase the expression and folding of eukaryotic fusion proteins expressed in bacteria [27–31]. Although not completely understood [32], fusion of MBP onto a protein has been shown to enhance the solubility of the partner protein [31, 33–35]. There are a number of vectors for generating MBP fusion proteins that are commercially available, including the pMAL series from New England Biolabs and pIVEX series from Roche, which contain a specific protease cleavage site in the region between MBP and the multiple cloning site. In addition, there are also commercial vectors for the expression of MBP fusion proteins to the non-reducing environment of the *E. coli* periplasm (e.g., New England Biolabs pMAL), addition of which may improve the folding of disulfide bond-containing proteins. Although not generally required, MBP has been used in conjunction with a small affinity tag to improve purification purity [36, 37].

2.1.1.1.3 Thioredoxin A

Thioredoxins are universal oxido-reductaces that reduce disulfide bonds through thio-disulfide exchange. One of the *E. coli* thioredoxins, TrxA, is an 11.6 kDa protein that demonstrates high solubility in the *E. coli* cytoplasm and inherent thermal stability, which may be conferred to TrxA fusion proteins. For example, TrxA has been used as an N- or

C-terminal fusion protein [38, 39] to increase soluble protein expression of recombinant proteins [40]. In addition to enhancing the solubility of fusion proteins, TrxA has been shown to aid in the crystallization of proteins. Though in most instances TrxA must be removed prior to structural characterization, TrxA when joined with the linker sequence GSAM aided crystallization of the U2AF homology motif domain of splicing factor Puf60 [41]. Unlike GST or MBP, TrxA does not facilitate purification on its own; small affinity tags are commonly used in conjunction with TrxA to enable purification.

2.1.1.1.4 Small ubiquitin-related modifier (SUMO)

In its native system, small ubiquitin-like modifier (SUMO) protein, *Saccharomyces cerevisiae* Smt3, is a form of post-translational modification that plays roles in nuclear-cytosolic transport [42], apoptosis [43], protein activation [44], and stability [45], response to stress [46], and the cell cycle [47]. When used as an N-terminal carrier protein during prokaryotic expression, SUMO promotes folding and structural stability, which leads to enhanced functional production compared to untagged protein [48–54]. Unlike GST and MBP, SUMO does not itself serve as a means for purification of fusion proteins; however, the His_6 in series with the SUMO tag has been established to facilitate purification of fusion proteins. One unique advantage of the SUMO tag over other expression-enhancing carrier proteins is a specific SUMO protease (*S. cerevisiae* UlpI), which recognizes and removes the SUMO tag at a Gly–Gly motif.

The wild-type SUMO tag is an excellent carrier protein in prokaryotic expression systems but is efficiently removed in eukaryotes in vivo by naturally occurring SUMO proteases [55]. LifeSensors, Inc. has engineered a SUMO-based tag, SUMOstar, that has been used to enhance protein expression in yeast [56, 57], insect cells [58], and mammalian cells [59]. In addition to the SUMOstar carrier protein, LifeSensors, Inc. has also developed a SUMOstar-specific protease.

2.1.1.1.5 N-utilization substance A (NusA)

The transcription termination anti-termination factor, also known as N-utilization substance A (NusA), is a 55 kDa hydrophilic protein that promotes the soluble expression of even very hydrophobic fusion proteins in *E. coli* [60]. In *E. coli*, wild-type NusA promotes pauses in DNA transcription by RNA polymerase [61]. Because transcription and translation are coupled in prokaryotes, improvements in soluble expression of NusA fusion proteins over other fusion proteins may result from the biological activity of NusA. NusA slows translation and permits more time for critical folding events to occur [60]. A direct comparison of normally aggregation-prone proteins tagged with either MBP or NusA showed that both MBP and NusA enhanced solubility overall while not affecting the structure of the aggregation-prone proteins [32].

2.1.1.1.6 Protein disulfide isomerase I (DsbA)

Protein disulfide isomerase I (DsbA) is a 21.1 kDa protein that catalyzes the formation of disulfide bonds in *E. coli* [62, 63]. When fused to eukaryotic proteins, an inactive mutant of DsbA that lacked the periplasmic signal sequence was shown to promote the soluble expression of proteins in the *E. coli* cytoplasm [64]. Furthermore, a comparison of DsbA and thioredoxin fusion proteins showed that DsbA increased the solubility of fusion proteins by two- to three-fold over thioredoxin [64]. The vectors pET-39 and pET-40 available from EMD Biosciences enable expression of DsbA and DsbC, respectively, to the non-reducing environment of the *E. coli* periplasm, which may result in improved solubility and folding of the target protein [65].

This section has outlined a number of fusion proteins that increase the solubility of a protein of interest and it should be clear that there is no universal solution to achieve this end. Oftentimes, parallel analysis of different fusion proteins must be sought, as not only does the fusion protein affect properties of the protein of interest but the protein of interest also affects properties of the carrier protein. Furthermore, the use of carrier proteins, especially for those significantly enhancing solubility of the fusion, presents obstacles for maintaining the protein of interest in the soluble fraction after the carrier protein is removed.

2.1.1.1.7 Mistic

Recently, a novel membrane-associating protein from *B. subtilis* [72] referred to as Mistic has been shown to enhance expression levels of eukaryotic integral membrane proteins in *E. coli*, when used as a fusion protein linked to the N-terminus of cargo proteins [73, 74]. Mistic and its orthologs lack an identified signal sequence. Overexpression of these fusion partners does not result in cell toxicity, a common result due to overloading the protein translocation machinery. A truncated version of Mistic shorter by 25 residues – M110 – expresses primarily in the bacterial membrane [72]. Interestingly, shorter Mistic proteins maintain comparable levels of expression as M110 even though they are soluble in the presence of detergents and are localized mainly to the cytoplasm [75]. Mistic's mecha-

nism of action remains unclear; however, in some cases the use of Mistic and its orthologs have significantly increased levels of eukaryotic integral membrane protein expression in *E. coli* [73, 74, 76].

2.1.1.2 Proteins that decrease solubility

A number of carrier proteins have been utilized to drive expression to inclusion bodies, which protects them from intracellular degradation, provides a simple process for recovery, and commonly leads to the highest expression yields [66]. Furthermore, expression to inclusion bodies may be required in cases where protein production is toxic to the host cell.

Despite its advantages, expression to inclusion bodies requires solubilization with chaotropes or detergents and subsequent refolding steps that may ultimately affect biological activity. Therefore, use of fusion proteins that direct expression to inclusion bodies is generally limited to small easily refolded proteins and peptides. Two carrier proteins that direct expression to inclusion bodies include KSI and the gene product of *E. coli TrpE*.

2.1.1.2.1 Ketosteroid isomerase (KSI)

Ketosteroid isomerase is an extremely insoluble 13-kDa protein [67, 68] that efficiently forms inclusion bodies. Insoluble expression KSI fusion proteins are commonly employed for the production of recombinant peptides and used in conjunction with purification tags. Utilization of KSI as a carrier protein has increased protein production levels by two- to five-fold when compared to other methods [68].

2.1.1.2.2 TrpΔLE

The truncated gene product of *E. coli TrpE* (*TrpΔLE*), a 27-kDa protein, has also been used to direct expression of fusion proteins to inclusion bodies [69–71].

3 Facilitating affinity purification and emerging combinatorial tags

3.1 Selection of affinity tags and utility

Affinity tags are highly efficient tools for protein detection, characterization, and purification. An epitope is a short sequence of amino acids that typically serves as the antigenic determinant, or the region to which an antibody binds. Thus, epitope tagging is a technique in which a short sequence (i.e., epitope) is added to a protein of interest by recombinant DNA methods. Used in a variety of applications including western blot analysis, immunoprecipitation (IP), co-immunoprecipitation (co-IP), immunofluorescence (IF), and affinity purification, epitopes enable the determination of protein size, concentration, post-translation modifications, protein interactions, intracellular trafficking, and in some cases provide a means of obtaining pure product (via affinity purification) in order to attain high-resolution crystallographic or NMR structures. The impact of protein engineering and the extensive use of affinity tags on crystallization techniques and structure/function studies have been reviewed elsewhere [77–80]. Described below are numerous affinity tags commonly used in microbial systems, as well as relevant caveats for their use. Well-characterized purification schemes are briefly summarized and relevant protocols for novel purification strategies are appropriately referenced. Additional details are provided in Supporting information, Tables S2 and S3 with references cited therein.

3.1.1 c-myc, HA, and FLAG

Traditionally, *c*-myc, hemaglutinin antigen (HA), and FLAG epitopes have been used primarily as a means of protein detection. An epitope of the human *c*-myc proto – oncogene product (EQKLISEEDL) was one of the earliest affinity tags developed and is highly specific to the anti-*c*-myc antibody 9E10 [81]. Fused to either the N- or C- terminus [82], the *c*-myc tag is widely used in protein engineering approaches, including western blot analysis, IP, co-IP, IF, and flow-cytometry [83] in both bacteria and yeast. Although rarely used for protein purification, it is possible to couple the mAb9E10 to divinyl sulfone-activated agarose and purify *c*-myc-tagged proteins under physiological conditions; however, elution requires a low pH that may adversely affect protein activity. Despite these concerns, structural studies have been performed using *c*-myc fusion proteins purified in this way [84].

Similarly, as one of the most widely used epitope tags in cell biology, the HA tag is a nine amino acid sequence (YPYDVPDYA) derived from the human influenza virus hemaglutinin protein [85, 86]. The hemaglutinin epitope may be located within the DNA coding region of a protein, or at the N- or C-terminus [87]. However, the hemaglutinin epitope may influence trafficking, folding, and function of the target protein [88, 89].

The FLAG epitope is a short, hydrophilic octapeptide (DYKDDDDK) that can be used for antibody-based purification [90]. A unique aspect of FLAG utilization is the inherent enterokinase cleavage site located within the five C-terminal residues of the peptide sequence [91], which is ad-

vantageous if the recombinant protein of interest is intended for therapeutic use, since the epitope itself is immunogenic. The FLAG peptide sequence is recognized by the antibody M1, in either a calcium-dependent [92], or independent [93], manner; however, additional commercial mAb choices include M2 and M5, which comprise slightly different recognition sites and affinity. FLAG-tagged fusion proteins can be purified using an immobilized monoclonal antibody matrix under non-denaturing conditions and eluted by lowering the pH or adding chelating agents such as EDTA. However, antibody-based purification matrices are in general not as stable as either Ni^{2+}-NTA, used with polyhistidine-tagged recombinant proteins, or Strep-Tactin, which is used with Strep-tag II fusions, as discussed below.

To increase the signal in protein detection applications, a recombinant protein can be designed to include multiple epitope tags in series. Examples include multiple copies of the hemaglutinin [94], *c*-myc [95, 96], and FLAG [97] where typically three copies are used, although nine copies of the *c*-myc epitope in series have been incorporated in order to detect very dilute proteins in yeast [98].

3.1.2 1D4

The 1D4 epitope is a short hydrophilic sequence (TETSQVAPA) derived from the C-terminus of bovine rhodopsin [99, 100]. Combining this epitope and the high affinity rho-1D4 monoclonal antibody (available from www.flintbox.com) has established useful tools in antibody-based purification, localization studies, and western blot analysis of 1D4-tagged membrane proteins [101, 102]. Additionally, the 1D4 enrichment strategy offers a highly specific, non-denaturing method for purifying membrane proteins with yields and purities consistent with structural characterization and functional proteomics applications [101].

3.1.3 polyArg and polyHis

The first commercially available affinity tags were designed from naturally occurring epitopes and used primarily for purification. Yet, during the past decade novel engineered tags have been optimized for use in many cell types and for purification under various expression conditions (i.e., inclusion bodies, membrane localization).

The polyarginine (polyArg) tag was initially used by Sassenfeld and Brewer [103] and typically consists of five or six consecutive arginines at the C-terminal end of a recombinant protein. A protein containing the polyArg tag is purified by absorption to cation exchange resin SP-Sephadex, and eluted via a linear NaCl gradient at alkaline pH. Following purification, the C-terminus of polyArg-tagged proteins can be removed by carboxypeptidase B, although poor cleavage yields and non-specificity have been reported [104]. It is important to note that the polyArg tag may affect the tertiary structure of proteins whose C-terminus is hydrophobic [103], despite its use as a standard immobilization method for electron and scanning-probe microscopy applications [105]. This is perhaps not surprising because of the addition of the positive charges under neutral and acidic pH conditions. This may also contribute to the observation that the polyArg tag promotes protein internalization [106].

In general, the polyArg tail is not the optimal purification tag in the protein-engineering field and its use is infrequent. In contrast, another polyamino acid construct, the polyhistidine (polyHis) tags are the most widely used affinity tags for purifying recombinant proteins that lead to biophysical and structural studies. Advantages of the polyHis tag include its low immunogenicity and small size (0.84 kDa) – with composition ranging from 3 to 10 His tags in series. In addition, many proteins function with the polyHis tag positioned at either the N- or C-terminus, and purification methods can be carried out under both native and denaturing conditions [107, 108].

To purify recombinant polyHis-tagged proteins, immobilized metal-affinity chromatography (IMAC) is used to isolate the metal-binding peptides based on the interaction between the negatively charged His and transition metal ions (Ni^{2+}, Co^{2+}, Cu^{2+}, Zn^{2+}) immobilized on a matrix [109]. Developed by Hochuli et al. [110], Ni(II)-nitrilotriacetic acid (Ni^{2+}-NTA) exhibits high affinity for adjacent histidine residues; provides an inexpensive matrix that withstands multiple regeneration cycles under stringent conditions; and permits dissociation of bound recombinant proteins by an imidazole gradient (20 to 250 mM for polyHis-tagged fusion proteins; [111, 112]), changes in pH, or metal chelation. During purification, non-specific binding of host proteins can occur and may be overcome by shorter incubation steps (3 h or less), or a low concentration of imidazole, where typically 2–50 mM is recommended in column washes. It is important to note that the use of imidazole may result in protein aggregates and affect activity, competition studies, NMR experiments, or crystallographic trials [112]; therefore, dialysis is recommended in order to remove residual imidazole following purification. PolyHis-tagged proteins can be purified by IMAC under denaturing conditions, and then refolded (reviewed in [113]).

Since the first recombinant protein to be successfully purified with a Ni^{2+}-NTA matrix – Dihydrofolate reductase (DHFR) fused to a polyHis-tag expressed in *E. coli* [114] – novel matrices (i.e., TALON) have been developed for this affinity tag. When directly compared to Ni^{2+}-NTA, the use of TALON has resulted in decreased levels of non-specifically bound proteins and consequent yields consisting of increased purity [115, 116]. TALON consists of Co^{2+}-carboxylmethylaspartate (Co^{2+}-CMA) that is coupled to resin. Following Co^{2+}-CMA matrix development, a natural 19-residue histidine affinity tag [117] with optimal binding to TALON [115] has been reported. Overexpression with the histidine affinity-tag has been demonstrated in bacteria with N-terminal fusions to chloramphenicol acetyltransferase (CAT), DHFR, and green fluorescent protein (GFP) that were eluted by decreasing pH (5.0) or increasing imidazole (150 mM) concentration [115]. More recently, Dong et al. [118] have developed an artificial chaperone-assisted IMAC method for the refolding and purification of a polyHis-tagged protein from solubilized inclusion bodies with increased refolding efficiency.

IMAC technology combined with polyHis-tag recombinant protein expression continues to dominate generic single-step purification strategies leading to the biophysical manipulations of polypeptides; however, there are a few caveats to note. Addition of a polyHis-tag may affect the function of the targeted protein or result in weak binding to select matrices as a result of the change in protein charge; however, these aspects may be corrected by changing the location of the polyHis-tag (i.e., N- vs. C-terminus), incorporating a flexible polylinker, or increasing the number of histidine residues. As a consequence of the latter aspect, Lee and Kim [119] have developed a vector-based strategy using a single insert to optimize the molecular engineering of target genes with different polyHis-tag lengths. Furthermore, IMAC should not be employed when purifying proteins with a metal center since the metal itself could be adsorbed to the column; and purification should not be attempted under anaerobic conditions, since Ni^{2+}-NTA will be reduced, and therefore ineffective, under that environment. The presence of inherent cysteine- or histidine-rich regions in host proteins may result in non-specific binding, leading to yield contamination or decreased purification efficiency. Subsequently, a strategy has been designed to increase *E. coli* product yields by engineering a strain to be deficient in three prevalent host proteins, which under native conditions are a significant fraction of the eluate [120]; as a result of gene knockouts, purification of a recombinant polyHis-tagged protein resulted in increased yields [120].

3.1.4 Streptavidin binding tags

The original Strep-tag (WRHPQFGG) was selected from a random genetic library [121] based on its affinity to the streptavidin core – a proteolytically truncated version of the bacterial protein [122]. To permit greater flexibility and affinity to an engineered variant of streptavidin, referred to as Strep-Tactin [123, 124], Strep-tag II is a small affinity peptide (WSHPQFEK) that was simultaneously optimized. Strep-tag II is inert, largely resistant to cellular proteases, can be used with mild detergents, and optimal for the purification of recombinant proteins under physiological conditions. However, Strep-tag II is not well suited for purification strategies that use denaturing conditions. Dissociation of Strep-tag II from Strep-Tactin occurs with mild buffer conditions in the presence of D-biotin; or if resin regeneration is preferred, then D-desthiobiotin is used. An advantage of Strep-tag II, when compared to polyHis-tag purification, is its independence from metal ions during purification strategies. If metalloproteins are the recombinant protein of choice, then the Strep-tag is optimal [125].

Strep-tag II fusion proteins have been incorporated and retained for protein crystallography [126, 127] and employed in a variety of hosts, including bacteria [128, 129] and yeast [130–132]. Streptavidin-binding peptide (SBP) is the larger 38 amino acid peptide that was selected based on its increased affinity to streptavidin [133]. SBP-tagged proteins have been expressed and purified in bacteria [134] under mild elution conditions of biotin (2 mM).

3.1.5 Additional affinity tags used in microbial hosts

Alternative affinity purification tags include the calmodulin-binding peptide, chitin-binding domain, cellulose-binding domain, S-tag, and Softag3.

Calmodulin-binding peptide consists of 26 amino acids derived from the C-terminus of human skeletal muscle myosin light chain kinase [135], which tightly binds calmodulin in the presence of 0.2 mM $CaCl_2$ [136]. Nanomolar affinity ensures that stringent washing conditions result in few contaminating proteins and pure yields of recombinant proteins. Detergents (up to 0.1%) and reducing agents are compatible during recovery steps of calmodulin-binding peptide fusions [137]. Calmodulin-binding peptide is only suitable for expression and purification in bacterial hosts since no endogenous host proteins interact with calmodulin. In eukaryotes including yeast, this affinity tag may

disrupt calcium-signaling pathways and induce protein–protein interactions in a calcium-dependent manner [138].

Similarly, fusions tags consisting of the chitin-binding domain and inteins have been expressed and purified in bacterial hosts [139, 140]. The chitin-binding domain consists of 51 residues from *Bacillus circulans* [141], and may be fused to either the N- or C-terminus. In general, this affinity tag is combined in tandem with self-splicing inteins [142–144] and is commercially available via the IMPACT™ (Intein-Mediated Purification with an Affinity Chitin-binding Tag) system. A chitin matrix is employed for affinity purification of the recombinant protein, whereas self-cleavage of the thioester bond is induced by a thiol reagent (e.g., 1,4-dithiothreitol or β-mercaptoethanol) or pH adjustment combined with a temperature shift [144, 145]. To reduce non-specific binding, either non-ionic detergents or high salt concentrations are used.

Cellulose-binding domains are non-catalytic domains that vary in size (4–20 kDa) and affinity for their natural substrate [146]. These domains may be fused to the N- or C-terminus and bind to cellulose over a wide-range of pH (3.5–9.5). Cellulose matrices are an attractive method for purification since cellulose inherently maintains a low non-specific affinity. In general, cellulose-binding domain interactions with cellulose are so strong that they require the use of chaotropic agents, such as urea or guanidine hydrochloride, for dissociation. Thus, protein refolding is required following elution. However, ethylene glycol has been used as a milder elution condition when cellulose-binding domains of other families have been incorporated as an affinity tag and bind to alternative forms of cellulose matrices with lower affinities. Nahalka et al. [147] have recently described a purification strategy that induces cellulose-binding domain-tagged fusions to aggregate, which results in their selective isolation via differential centrifugation steps.

The S-tag is a 15-residue soluble affinity tag that binds to the S-protein, derived from RNaseA [148, 149]. This interactive complex (S-tag/S-protein) results in micromolar affinity – dependent upon pH, temperature, and ionic strength [150]. The S-tag itself enables colorimetric detection via western blot analysis or a rapidly detected assay dependent upon ribonucleolytic reconstitution and activity. When purified from the S-protein matrix, the S-tag is eluted under harsh conditions (i.e., pH = 2). The S-tag has enabled the purification of recombinant proteins in bacteria [151], and is typically used as a dual-tagging system that incorporates selective protease cleavage sites.

Softag3 is an 8-residue (TKDPSRVG) peptide that binds with high affinity to polyol-responsive mAb, and is eluted from antibodies under mild conditions consisting of a low molecular weight polyhydroxylated compound and non-chaotropic salt [152]. Although less frequently employed as a recombinant protein, Softag3 has been used for the purification of multi-subunit enzyme complexes [152, 153] and implemented to study protein interactions [154]. Softag3 is commercially available through neoclone®.

In contrast to affinity purification, the genetically engineered HaloTag® (Promega) has been designed to enhance expression and solubility of recombinant proteins in bacterial hosts, as well as to provide an efficient means of protein purification coupled with tag removal by TEV protease-mediated proteolytic cleavage [155]. Ohana et al. demonstrated the efficacy of HaloTag® for protein expression and purification by evaluating 23 human proteins recombinantly expressed in *E. coli*. Each heterologous protein was fused to the HaloTag®, MBP, GST, and His_6 at its C-terminus. Results indicated that an overwhelming 73% of proteins were produced in soluble form when fused to the HaloTag®, compared to 52, 39, and 22% of soluble proteins fused to MBP, GST, or His_6, respectively [155]. HaloTag® technology includes a synthetic linker which covalently binds to a purification matrix leading to increased purification yields and purity when compared to affinity purification schemes. The versatility associated with the HaloTag® includes a variety of ligands applicable for protein visualization, trafficking, and turnover [156, 157], as well as analysis of protein interactions [156].

3.2 Tandem affinity purification and dual-tagging methods

Unfortunately, no single affinity tag is ideal from every experimental perspective. Therefore, to increase the versatility of tags and downstream applications, novel dual-tagging methods have been developed for different needs. Two affinity tags expressed in tandem (i.e., tandem affinity purification (TAP); [158]) enable the use of multiple purification strategies. By incorporating a solubility-enhancing tag in combination with a purification tag (e.g., MBP combined with His_6; [159]), improvements in solubility and expression yields as well as methods for efficient purification are provided. A fusion can be included to evaluate recombinant protein expression (detected via an epitope or

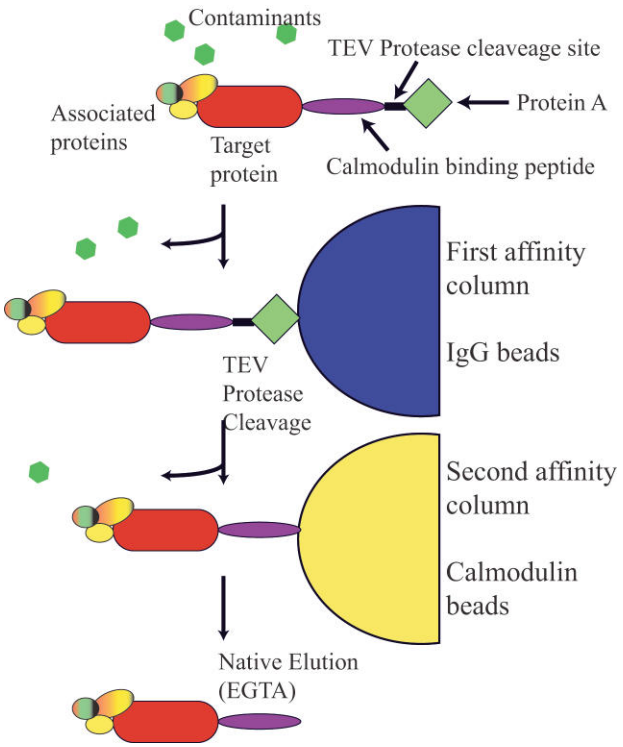

Figure 1. Illustrations of optimized tandem affinity purification (TAP) strategy, adapted from the following publications [158]. Overview of the TAP strategy for a dual-tagging construct consisting of a calmodulin-binding peptide, TEV protease cleavage site, followed by two IgG-binding units of protein A. The recombinant fusion protein is recovered from cell extracts using the protein A tag. Following repetitive wash cycles, the bound IgG material is then released by protease cleavage. Any residual contaminants are removed by a second purification step that uses the affinity of calmodulin-binding peptide for calmodulin beads. The bound material is released with EGTA resulting in efficient and high-purity recovery of recombinant proteins.

GFP) followed by a tag optimal for purification (e.g., GFP-His_6 [160] or GFP-His_{10} [161]).

Tandem affinity purification is an approach where the protein of interest is fused in-frame to two affinity tags, generally separated by a protease cleavage site. The original TAP tag consisted of a calmodulin binding peptide, a TEV cleavage site, followed by two immunoglobulin binding domains of *Staphylococcus aureus* protein A (ProtA) [158] and demonstrated maximum protein recovery compared to alternative affinity tags such as FLAG-, Strep-, His-, calmodulin-binding protein-, and cellulose-binding domain-tags [162, 163]. TAP tagging involves two sequential purification steps as illustrated in Fig. 1 and has been implemented in microbial systems. Combinatorial tagging has been a useful strategy to evaluate the genome scale mapping of protein interaction networks in yeast [96, 164–166] and may also provide a strategic means of purifying heterodimeric or higher-order protein complexes where each protein is selectively fused to a different tag by recombinant DNA methods.

Dual-tagging constructs, such as GFP-His_6, offer an additional method of detection using in vivo fluorescence microscopy instrumentation or in vitro techniques. Recently, the intrinsic fluorescence of GFP and its variants have been used to rapidly evaluate positive clones by in-gel fluorescence techniques [167, 168], in addition to the use of GFP fluorescent standards and established protocols [169] permitting the estimation of recombinant protein concentration via western blot analysis.

Multiple in-depth reviews of TAP tags and purification schemes are currently available [5, 170, 171]. Detailed elements of combinatorial cassettes used for dual-tagging and TAP strategies, as well as vectors available to the bacteria and yeast research communities, relevant methods, and protocols are provided in Supporting information, Tables S3 and S5.

4 Removal of affinity tags: Efficiency of proteases

4.1 Common proteases

Removal of carrier proteins and affinity tags is often necessary when the presence of the fusion tag affects the structure or biological function of the protein of interest. To achieve this end, specific sequences are included between the tag(s) and native protein and then cleaved with site-specific proteases. Common proteases include enterokinase [172], factor Xa [8], SUMO protease [48], tobacco etch virus (TEV) protease [173], thrombin [174, 175], and 3C [176], outlined in Supporting information, Table S4.

Proteases may cleave fusion proteins at unintended sites [151] and the buffer conditions that promote protease activity and specificity may not be suitable for fusion protein and product solubility [177]. It should be noted that protease cleavage efficiency might vary in an unpredictable manner with each fusion protein. Cleavage efficiency may be improved by applying an increased concentration of the protease or by prolonging the digestion. In some cases, the cleavage site is sterically hindered and improvements may result by the inclusion of extra amino acids that flank the cleavage site.

4.1.1 Enterokinase

Enterokinase is a protease that recognizes DDDDK^X and cleaves at the carboxyl lysine with variable efficiencies that depend on the amino acid

in the X position. For example, enterokinase cleavage efficiency varied from 61% for proline to 88% for alanine in the X position [178]. Although enterokinase has been shown to cleave at unintended sites [172, 179], it remains useful in some applications due to the internal enterokinase recognition site in the FLAG tag (DYKDDDK).

4.1.2 Factor Xa

Factor Xa is a protease that recognizes I(E/D) GR^X, where X can be any amino acid except arginine or proline [180], and cleaves after the carboxyl arginine. Ineffective and non-specific proteolysis has been reported for factor Xa [181, 182], although factor Xa has been used successfully [183]. Factor Xa consists of two disulfide-linked chains of 27 and 16 kDa; therefore, the activity of factor Xa may decrease in buffers with reducing agents.

4.1.3 SUMO protease

The SUMO protease (*S. cerevisiae* Ulp1) recognizes the tertiary structure of SUMO and cleaves at the N-terminus of the fused protein irrespective of the sequence with the only exception being proline [48]. The SUMO protease cleaves efficiently under a wide range of buffer conditions, pH (5.5–10.5) and temperatures (4–37°C). Additionally, SUMO protease has been shown to efficiently cleave fusion proteins in 2 M urea without yielding non-specific cleavage products [48], which may facilitate purification of SUMO-tagged proteins produced in inclusion bodies.

4.1.4 Tobacco etch virus (TEV) protease

Tobacco etch virus (TEV) protease recognizes ENLYFQ^S and cleaves between glutamine and serine [184–186]. TEV protease is highly specific, active on a variety of substrates, and cleaves efficiently at low temperatures [173]. One distinct advantage of TEV protease compared to other commonly used proteases is that "in house" methods have been optimized in order to express and purify the enzyme [187, 188]. As a direct result, stabilizing mutations and truncations of the enzyme have improved TEV protease activity and expression yields. Although wild-type TEV has been shown to cleave itself and result in greatly diminished activity [188], a mutation (S219V) not only confers resistance to unintended cleavage and autoinactivation but also increases activity of the enzyme by two-fold [189]. In addition to the S219V stabilizing mutation, an engineered TEV protease that lacks the C-terminal residues 238–242 further improved expression yields in *E. coli*, where yields were nearly 400 mg/L [189].

4.1.5 Thrombin

Thrombin is a protease that recognizes LVPR^GS and cleaves between arginine and glycine [190–192]. Although thrombin cleavage at the designated sequence is relatively specific, there have been reports of thrombin cleavage at alternative sites [193]. Additionally, impurities in commercial thrombin preparations, most notably plasmin, have resulted in non-specific cleavage products in the past [193, 194], but modern purification methods have improved thrombin purity [195]. However, thorough characterization of cleaved products must be performed following the use of any protease.

Despite these caveats, thrombin may play a specific role in the preparation of membrane proteins for structural characterization as thrombin was determined to be insensitive to an entire panel of detergents in a recent study by Vergis and Wiener [177]. In addition to advantages in detergent sensitivity, thrombin is efficiently removed from cleavage products by benzamidine sepharose.

4.1.6 3C and PreScission™

The 3C protease is derived from human rhinovirus (HRV 3C) and recognizes the exact residue sequence (ETLFQ^GP) as the native enzyme where it cleaves between glutamine and glycine [176]. This 22 kDa protease maintains optimal activity at 4°C and has been constructed in combination with His-, GST-, NusA, S-, StrepTag II-, and Trx-tags (EMD4Biosciences). A genetically engineered derivative of HRV 3C, PreScission™ (GE Healthcare) has been designed to cleave selectively between the glutamine and glycine of the LEVLFQ^GP recognition site [176, 196].

5 Conclusion

In this post-genomic era, the emergence of bioinformatics combined with proteomic research endeavors has accelerated the development of fusions tags. High-throughput protein expression and purification maintain a pivotal role in structural biology (reviewed in [6]). The polyHis-tag IMAC purification process has enabled numerous structural studies; in fact, more than 60% of the all protein structures that exist include a polyHis-tag [78]. This well-characterized purification strategy has been optimized recently for high-throughput on-chip purifications using protein microarrays in order to evaluate protein function [197]. Moreover, an automated IMAC purification strategy has been described [198] which is likely to have a significant, positive impact on the pharmaceutical industry

and drug design. We anticipate that further advances in affinity tags and complementary assays integrated with high-throughput technologies will provide additional insights into protein-interaction networks, accelerate the discovery of new drug targets and therapeutics, and increase our ability to investigate protein structure–function relationships that are presently inaccessible due to technical limitations.

The authors declare no conflict of interest.

6 References

[1] Arnau, J., Lauritzen, C., Petersen, G. E., Pedersen, J., Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr. Purif.* 2006, *48*, 1–13.

[2] Waugh, D. S., Making the most of affinity tags. *Trends Biotechnol.* 2005, *23*, 316–320.

[3] Brizzard, B., Epitope tagging. *Biotechniques* 2008, *44*, 693–695.

[4] Terpe, K., Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 2003, *60*, 523–533.

[5] Walls, D., Loughran, S. T., Tagging recombinant proteins to enhance solubility and aid purification. *Methods Mol. Biol.* 2011, *681*, 151–175.

[6] Koehn, J., Hunt, I., High-throughput protein production (HTPP): A review of enabling technologies to expedite protein production. *Methods Mol. Biol.* 2009, *498*, 1–18.

[7] Stevens, R. C., Design of high-throughput methods of protein production for structural biology. *Structure* 2000, *8*, R177–R185.

[8] Smith, D. B., Johnson, K. S., Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 1988, *67*, 31–40.

[9] Parker, M. W., Lo Bello, M., Federici, G., Crystallization of glutathione S-transferase from human placenta. *J. Mol. Biol.* 1990, *213*, 221–222.

[10] Lim, K., Ho, J. X., Keeling, K., Gilliland, G. L. et al., Three-dimensional structure of *Schistosoma japonicum* glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV. *Protein Sci.* 1994, *3*, 2233–2244.

[11] Ji, X., Zhang, P., Armstrong, R. N., Gilliland, G. L., 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* 1992, *31*, 10169–10184.

[12] Maru, Y., Afar, D. E., Witte, O. N., Shibuya, M., The dimerization property of glutathione S-transferase partially reactivates Bcr-Abl lacking the oligomerization domain. *J Biol. Chem.* 1996, *271*, 15353–15357.

[13] Kaplan, W., Husler, P., Klump, H., Erhardt, J. et al., Conformational stability of pGEX-expressed *Schistosoma japonicum* glutathione S-transferase: A detoxification enzyme and fusion-protein affinity tag. *Protein Sci.* 1997, *6*, 399–406.

[14] Frangioni, J. V., Neel, B. G., Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem.* 1993, *210*, 179–187.

Anne Skaja Robinson is the Chair of Chemical and Biomolecular Engineering at Tulane University and Catherine and Henry Boh Professor of Engineering as of January 2012. Prior to Tulane, Prof. Robinson was Full Professor and Associate Chair at the University of Delaware, where she started her academic career. She obtained her Bachelor's and Master's degrees in Chemical Engineering from the Johns Hopkins University. She obtained her PhD in Chemical Engineering from the University of Illinois at Urbana-Champaign, and was an NIH Postdoctoral Fellow in the Department of Biology at MIT before joining the faculty at UD. Her honors include a DuPont Young Professor Award, and a National Science Foundation Presidential Early Career Award for Science and Engineering (PECASE) Award, and more recently being inducted into AIMBE. She has several patents and over 55 publications in the field of biochemical engineering.

[15] Lu, Q., Bauer, J. C., Greener, A., Using Schizosaccharomyces pombe as a host for expression and purification of eukaryotic proteins. *Gene* 1997, *200*, 135–144.

[16] Mitchell, D. A., Marshall, T. K., Deschenes, R. J., Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. *Yeast* 1993, *9*, 715–722.

[17] Rodal, A. A., Duncan, M., Drubin, D., Purification of glutathione S-transferase fusion proteins from yeast. *Methods Enzymol.* 2002, *351*, 168–172.

[18] Parker, L. L., Atherton-Fessler, S., Piwnica-Worms, H., p107wee1 is a dual-specificity kinase that phosphorylates p34cdc2 on tyrosine 15. *Proc. Natl. Acad. Sci. USA* 1992, *89*, 2917–2921.

[19] Beekman, J. M., Cooney, A. J., Elliston, J. F., Tsai, S. Y., Tsai, M. J., A rapid one-step method to purify baculovirus-expressed human estrogen receptor to be used in the analysis of the oxytocin promoter. *Gene* 1994, *146*, 285–289.

[20] Rudert, F., Visser, E., Gradl, G., Grandison, P. et al., pLEF, a novel vector for expression of glutathione S-transferase fusion proteins in mammalian cells. *Gene* 1996, *169*, 281–282.

[21] Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L. et al., MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell* 1989, *58*, 823–831.

[22] Ron, D., Dressler, H., pGSTag—a versatile bacterial expression plasmid for enzymatic labeling of recombinant proteins. *Biotechniques* 1992, *13*, 866–869.

[23] Mayer, B. J., Jackson, P. K., Baltimore, D., The noncatalytic src homology region 2 segment of abl tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 627–631.

[24] McTigue, M. A., Williams, D. R., Tainer, J. A., 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.* 1995, *246*, 21–27.

[25] Duplay, P., Hofnung, M., Two regions of mature periplasmic maltose-binding protein of *Escherichia coli* involved in secretion. *J. Bacteriol.* 1988, *170*, 4445–4450.

[26] di Guan, C., Li, P., Riggs, P. D., Inouye, H., Vectors that facilitate the expression and purification of foreign peptides in

Escherichia coli by fusion to maltose-binding protein. *Gene* 1988, *67*, 21–30.

[27] Kapust, R. B., Waugh, D. S., *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 1999, *8*, 1668–1674.

[28] Sachdev, D., Chirgwin, J. M., Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. *J. Protein Chem.* 1999, *18*, 127–136.

[29] Sachdev, D., Chirgwin, J. M., Fusions to maltose-binding protein: Control of folding and solubility in protein purification. *Methods Enzymol.* 2000, *326*, 312–321.

[30] Riggs, P., Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol. Biotechnol.* 2000, *15*, 51–63.

[31] Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., McCafferty, J., Production of soluble mammalian proteins in *Escherichia coli*: Identification of protein features that correlate with successful expression. *BMC Biotechnol.* 2004, *4*, 32.

[32] Nallamsetty, S., Austin, B. P., Penrose, K. J., Waugh, D. S., Gateway vectors for the production of combinatorially tagged His6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. *Protein Sci.* 2005, *14*, 2964–2971.

[33] Kataeva, I., Chang, J., Xu, H., Luan, C. H. et al., Improving solubility of *Shewanella oneidensis* MR-1 and *Clostridium thermocellum* JW-20 proteins expressed into *Esherichia coli*. *J. Proteome Res.* 2005, *4*, 1942–1951.

[34] Busso, D., Delagoutte-Busso, B., Moras, D., Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Anal. Biochem.* 2005, *343*, 313–321.

[35] Braud, S., Moutiez, M., Belin, P., Abello, N. et al., Dual expression system suitable for high-throughput fluorescence-based screening and production of soluble proteins. *J. Proteome Res.* 2005, *4*, 2137–2147.

[36] Hamilton, S. R., O'Donnell, J. B., Jr., Hammet, A., Stapleton, D. et al., AMP-activated protein kinase: Detection with recombinant AMPK alpha1 subunit. *Biochem. Biophys. Res. Commun.* 2002, *293*, 892–898.

[37] Podmore, A. H., Reynolds, P. E., Purification and characterization of VanXY(C), a D,D-dipeptidase/D,D-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4174. *Eur. J. Biochem.* 2002, *269*, 2740–2746.

[38] Katti, S. K., LeMaster, D. M., Eklund, H., Crystal structure of thioredoxin from *Escherichia coli* at 1.68 A resolution. *J. Mol. Biol.* 1990, *212*, 167–184.

[39] LaVallie, E. R., Lu, Z., Diblasio-Smith, E. A., Collins-Racie, L. A., McCoy, J. M., Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*. *Methods Enzymol.* 2000, *326*, 322–340.

[40] LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L. et al., A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Biotechnology (N Y)* 1993, *11*, 187–193.

[41] Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M., Stier, G., Thioredoxin as a fusion tag for carrier-driven crystallization. *Protein Sci.* 2008, *17*, 2070–2079.

[42] Seeler, J. S., Bischof, O., Nacerddine, K., Dejean, A., SUMO, the three Rs and cancer. *Curr. Top. Microbiol. Immunol.* 2007, *313*, 49–71.

[43] Meinecke, I., Cinski, A., Baier, A., Peters, M. A. et al., Modification of nuclear PML protein by SUMO-1 regulates Fas-induced apoptosis in rheumatoid arthritis synovial fibroblasts. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 5073–5078.

[44] Rajan, S., Plant, L. D., Rabin, M. L., Butler, M. H., Goldstein, S. A., Sumoylation silences the plasma membrane leak K+ channel K2P1. *Cell* 2005, *121*, 37–47.

[45] Martin, S., Nishimune, A., Mellor, J. R., Henley, J. M., SUMOylation regulates kainate-receptor-mediated synaptic transmission. *Nature* 2007, *447*, 321–325.

[46] Mabb, A. M., Wuerzberger-Davis, S. M., Miyamoto, S., PIASy mediates NEMO sumoylation and NF-kappaB activation in response to genotoxic stress. *Nat. Cell Biol.* 2006, *8*, 986–993.

[47] Li, S. J., Hochstrasser, M., A new protease required for cell-cycle progression in yeast. *Nature* 1999, *398*, 246–251.

[48] Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M. et al., SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics* 2004, *5*, 75–86.

[49] Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P. et al., Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. *Protein Sci.* 2006, *15*, 182–189.

[50] Butt, T. R., Edavettal, S. C., Hall, J. P., Mattern, M. R., SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* 2005, *43*, 1–9.

[51] Zuo, X., Li, S., Hall, J., Mattern, M. R. et al., Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. *J. Struct. Funct. Genomics* 2005, *6*, 103–111.

[52] Zuo, X., Mattern, M. R., Tan, R., Li, S. et al., Expression and purification of SARS coronavirus proteins using SUMO-fusions. *Protein Expr. Purif.* 2005, *42*, 100–110.

[53] Guzzo, C. M., Yang, D. C., Systematic analysis of fusion and affinity tags using human aspartyl-tRNA synthetase expressed in *E. coli*. *Protein Expr. Purif.* 2007, *54*, 166–175.

[54] Dominy, J. E., Jr., Simmons, C. R., Hirschberger, L. L., Hwang, J. et al., Discovery and characterization of a second mammalian thiol dioxygenase, cysteamine dioxygenase. *J. Biol. Chem.* 2007, *282*, 25189–25198.

[55] Hughes, R. S., Dowd, P. F., Hector, R. E., Riedmuller, S. B. et al., Cost-effective high-throughput fully automated construction of a multiplex library of mutagenized open reading frames for an insecticidal peptide using a plasmid-based functional proteomic robotic workcell with improved vacuum system. *J. Assoc. Lab. Autom.* 2007, *12*, 202–212.

[56] Hughes, S. R., Dowd, P. F., Hector, R. E., Panavas, T. et al., Lycotoxin-1 insecticidal peptide optimized by amino acid scanning mutagenesis and expressed as a coproduct in an ethanologenic *Saccharomyces cerevisiae* strain. *J. Pept. Sci.* 2008, *14*, 1039–1050.

[57] Hughes, S. R., Sterner, D. E., Bischoff, K. M., Hector, R. E. et al., Engineered *Saccharomyces cerevisiae* strain for improved xylose utilization with a three-plasmid SUMO yeast expression system. *Plasmid* 2009, *61*, 22–38.

[58] Liu, L., Spurrier, J., Butt, T. R., Strickler, J. E., Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expr. Purif.* 2008, *62*, 21–28.

[59] Peroutka, R. J., Elshourbagy, N., Piech, T., Butt, T. R., Enhanced protein expression in mammalian cells using engineered SUMO fusions: Secreted phospholipase A2. *Protein Sci.* 2008, *17*, 1586–1595.

[60] Davis, G. D., Elisee, C., Newham, D. M., Harrison, R. G., New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol. Bioeng.* 1999, *65*, 382–388.

[61] Gusarov, I., Nudler, E., Control of intrinsic transcription termination by N and NusA: The basic mechanisms. *Cell* 2001, *107*, 437–449.

[62] Martin, J. L., Bardwell, J. C., Kuriyan, J., Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* 1993, *365*, 464–468.
[63] Bardwell, J. C., McGovern, K., Beckwith, J., Identification of a protein required for disulfide bond formation in vivo. *Cell* 1991, *67*, 581–589.
[64] Zhang, Y., Olsen, D. R., Nguyen, K. B., Olson, P. S. et al., Expression of eukaryotic proteins in soluble form in *Escherichia coli*. *Protein Expr. Purif.* 1998, *12*, 159–165.
[65] Collins-Racie, L. A., McColgan, J. M., Grant, K. L., DiBlasio-Smith, E. A. et al., Production of recombinant bovine enterokinase catalytic subunit in *Escherichia coli* using the novel secretory fusion partner DsbA. *Biotechnology (N Y)* 1995, *13*, 982–987.
[66] Lilie, H., Schwarz, E., Rudolph, R., Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.* 1998, *9*, 497–501.
[67] Kuliopulos, A., Shortle, D., Talalay, P., Isolation and sequencing of the gene encoding delta 5-3-ketosteroid isomerase of *Pseudomonas testosteroni*: Overexpression of the protein. *Proc. Natl. Acad. Sci. USA* 1987, *84*, 8893–8897.
[68] Kuliopulos, A., Walsh, C. T., Production, purification, and cleavage of tandem repeats of recombinant peptides. *J. Am. Chem. Soc.* 1994, *116*, 4599–4607.
[69] Yansura, D. G., Expression as trpE fusion. *Methods Enzymol.* 1990, *185*, 161–166.
[70] Miozzari, G. F., Yanofsky, C., Translation of the leader region of the *Escherichia coli* tryptophan operon. *J. Bacteriol.* 1978, *133*, 1457–1466.
[71] Kleid, D. G., Yansura, D., Small, B., Dowbenko, D. et al., Cloned viral protein vaccine for foot-and-mouth disease: Responses in cattle and swine. *Science* 1981, *214*, 1125–1129.
[72] Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S. et al., NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* 2005, *307*, 1317–1321.
[73] Roosild, T. P., Vega, M., Castronovo, S., Choe, S., Characterization of the family of Mistic homologues. *BMC Struct. Biol.* 2006, *6*, 10.
[74] Kefala, G., Kwiatkowski, W., Esquivies, L., Maslennikov, I., Choe, S., Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from *Escherichia coli*. *J. Struct. Funct. Genomics* 2007, *8*, 167–172.
[75] Dvir, H., Lundberg, M. E., Maji, S. K., Riek, R., Choe, S., Mistic: Cellular localization, solution behavior, polymerization, and fibril formation. *Protein Sci.* 2009, *18*, 1564–1570.
[76] Dvir, H., Choe, S., Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs. *Protein Expr. Purif.* 2009, *68*, 28–33.
[77] Xie, H., Guo, X. M., Chen, H., Making the most of fusion tags technology in structural characterization of membrane proteins. *Mol. Biotechnol.* 2009, *42*, 135–145.
[78] Derewenda, Z. S., The use of recombinant methods and molecular engineering in protein crystallization. *Methods* 2004, *34*, 354–363.
[79] Derewenda, Z. S., Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr. D Biol. Crystallogr.* 2010, *66*, 604–615.
[80] Moon, A. F., Mueller, G. A., Zhong, X., Pedersen, L. C., A synergistic approach to protein crystallization: Combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci.* 2010, *19*, 901–913.
[81] Evan, G. I., Lewis, G. K., Ramsay, G., Bishop, J. M., Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell Biol.* 1985, *5*, 3610–3616.
[82] Manstein, D. J., Schuster, H. P., Morandini, P., Hunt, D. M., Cloning vectors for the production of proteins in *Dictyostelium discoideum*. *Gene* 1995, *162*, 129–134.
[83] Kipriyanov, S. M., Kupriyanova, O. A., Little, M., Moldenhauer, G., Rapid detection of recombinant antibody fragments directed against cell-surface antigens by flow cytometry. *J. Immunol. Methods* 1996, *196*, 51–62.
[84] McKern, N. Ms., Lou, M., Frenkel, M. J., Verkuylen, A. et al., Crystallization of the first three domains of the human insulin-like growth factor-1 receptor. *Protein Sci.* 1997, *6*, 2663–2666.
[85] Wilson, I. A., Niman, H. L., Houghten, R. A., Cherenson, A. R. et al., The structure of an antigenic determinant in a protein. *Cell* 1984, *37*, 767–778.
[86] Field, J., Nikawa, J., Broek, D., MacDonald, B. et al., Purification of a RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.* 1988, *8*, 2159–2165.
[87] Surdej, P., Jacobs-Lorena, M., Strategy for epitope tagging the protein-coding region of any gene. *Biotechniques* 1994, *17*, 560–565.
[88] Brothers, S. P., Janovick, J. A., Conn, P. M., Unexpected effects of epitope and chimeric tags on gonadotropin-releasing hormone receptors: Implications for understanding the molecular etiology of hypogonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.* 2003, *88*, 6107–6112.
[89] Houle, T. D., Ram, M. L., McMurray, W. J., Cala, S. E., Different endoplasmic reticulum trafficking and processing pathways for calsequestrin (CSQ) and epitope-tagged CSQ. *Exp. Cell. Res.* 2006, *312*, 4150–4161.
[90] Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T. et al., A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio-Technology* 1988, *6*, 1204–1210.
[91] Maroux, S., Baratti, J., Desnuelle, P., Purification and specificity of porcine enterokinase. *J. Biol. Chem.* 1971, *246*, 5031–5039.
[92] Hopp, T. P., Gallis, B., Prickett, K. S., Metal-binding properties of a calcium-dependent monoclonal antibody. *Mol. Immunol.* 1996, *33*, 601–608.
[93] Einhauer, A., Schuster, M., Wasserbauer, E., Jungbauer, A., Expression and purification of homogenous proteins in *Saccharomyces cerevisiae* based on ubiquitin-FLAG fusion. *Protein Expr. Purif.* 2002, *24*, 497–504.
[94] Ross-Macdonald, P., Sheehan, A., Roeder, G. S., Snyder, M., A multipurpose transposon system for analyzing protein production, localization, and function in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 1997, *94*, 190–195.
[95] Sharrock, R. A., Clack, T., Heterodimerization of type II phytochromes in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 11500–11505.
[96] Graumann, J., Dunipace, L. A., Seol, J. H., McDonald, W. H. et al., Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Mol. Cell. Proteomics* 2004, *3*, 226–237.
[97] Hernan, R., Heuermann, K., Brizzard, B., Multiple epitope tagging of expressed proteins for enhanced detection. *Biotechniques* 2000, *28*, 789–793.
[98] Swanson, R., Locher, M., Hochstrasser, M., A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Mat alpha 2 repressor degradation. *Genes Dev.* 2001, *15*, 2660–2674.
[99] Wong, J. P., Reboul, E., Molday, R. S., Kast, J., A carboxy-terminal affinity tag for the purification and mass spectrometric characterization of integral membrane proteins. *J. Proteome Res.* 2009, *8*, 2388–2396.

[100] Hodges, R. S., Heaton, R. J., Parker, J. M., Molday, L., Molday, R. S., Antigen-antibody interaction. Synthetic peptides define linear antigenic determinants recognized by monoclonal antibodies directed to the cytoplasmic carboxyl terminus of rhodopsin. *J. Biol. Chem.* 1988, *263*, 11768–11775.

[101] Zhong, M., Molday, R. S., Binding of retinoids to ABCA4, the photoreceptor ABC transporter associated with Stargardt macular degeneration. *Methods Mol. Biol.* 2010, *652*, 163–176.

[102] Farrens, D. L., Dunham, T. D., Fay, J. F., Dews, I. C. et al., Design, expression, and characterization of a synthetic human cannabinoid receptor and cannabinoid receptor/G-protein fusion protein. *J. Pept. Res.* 2002, *60*, 336–347.

[103] Sassenfeld, H. M., Brewer, S. J., A Polypeptide fusion designed for the purification of recombinant proteins. *BioTechnology* 1984, *2*, 76–81.

[104] Nagai, K., Thogersen, H. C., Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. *Methods Enzymol.* 1987, *153*, 461–481.

[105] Nock, S., Spudich, J. A., Wagner, P., Reversible, site-specific immobilization of polyarginine-tagged fusion proteins on mica surfaces. *FEBS Lett* 1997, *414*, 233–238.

[106] Fuchs, S. M., Raines, R. T., Polyarginine as a multifunctional fusion tag. *Protein Sci.* 2005, *14*, 1538–1544.

[107] Chaga, G. S., Twenty-five years of immobilized metal ion affinity chromatography: past, present and future. *J. Biochem. Biophys. Methods* 2001, *49*, 313–334.

[108] Manjasetty, B. A., Turnbull, A. P., Panjikar, S., Bussow, K., Chance, M. R., Automated technologies and novel techniques to accelerate protein crystallography for structural genomics. *Proteomics* 2008, *8*, 612–625.

[109] Porath, J., Carlsson, J., Olsson, I., Belfrage, G., Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 1975, *258*, 598–599.

[110] Hochuli, E., Dobeli, H., Schacher, A., New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* 1987, *411*, 177–184.

[111] Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A. et al., Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 8972–8976.

[112] Hefti, M. H., Van Vugt-Van der Toorn, C. J., Dixon, R., Vervoort, J., A novel purification method for histidine-tagged proteins containing a thrombin cleavage site. *Anal. Biochem.* 2001, *295*, 180–185.

[113] Li, M., Su, Z. G., Janson, J. C., In vitro protein refolding by chromatographic procedures. *Protein Expr. Purif.* 2004, *33*, 1–10.

[114] Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R., Stuber, D., Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio-Technology* 1988, *6*, 1321–1325.

[115] Chaga, G., Bochkariov, D. E., Jokhadze, G. G., Hopp, J., Nelson, P., Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose. *J. Chromatogr. A* 1999, *864*, 247–256.

[116] Chaga, G., Hopp, J., Nelson, P., Immobilized metal ion affinity chromatography on Co2+-carboxymethylaspartate-agarose Superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast muscle. *Biotechnol. Appl. Biochem.* 1999, *29*, 19–24.

[117] Chatterjee, D. K., Esposito, D., Enhanced soluble protein expression using two new fusion tags. *Protein Expr. Purif.* 2006, *46*, 122–129.

[118] Dong, X. Y., Chen, L. J., Sun, Y., Refolding and purification of histidine-tagged protein by artificial chaperone-assisted metal affinity chromatography. *J. Chromatogr. A* 2009, *1216*, 5207–5213.

[119] Lee, J., Kim, S. H., High-throughput T7 LIC vector for introducing C-terminal poly-histidine tags with variable lengths without extra sequences. *Protein Expr. Purif.* 2009, *63*, 58–61.

[120] Liu, Z., Bartlow, P., Varakala, R., Beitle, R. et al., Use of proteomics for design of a tailored host cell for highly efficient protein purification. *J. Chromatogr. A* 2009, *1216*, 2433–2438.

[121] Schmidt, T. G., Skerra, A., The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment. *Protein Eng.* 1993, *6*, 109–122.

[122] Pahler, A., Hendrickson, W. A., Kolks, M. A., Argarana, C. E., Cantor, C. R., Characterization and crystallization of core streptavidin. *J. Biol. Chem.* 1987, *262*, 13933–13937.

[123] Korndorfer, I. P., Skerra, A., Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site. *Protein Sci.* 2002, *11*, 883–893.

[124] Voss, S., Skerra, A., Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein Eng.* 1997, *10*, 975–982.

[125] Skerra, A., Schmidt, T. G., Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins. *Methods Enzymol.* 2000, *326*, 271–304.

[126] Breustedt, D. A., Korndorfer, I. P., Redl, B., Skerra, A., The 1.8-A crystal structure of human tear lipocalin reveals an extended branched cavity with capacity for multiple ligands. *J. Biol. Chem.* 2005, *280*, 484–493.

[127] Korndorfer, I. P., Dommel, M. K., Skerra, A., Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat. Struct. Mol. Biol.* 2004, *11*, 1015–1020.

[128] Skerra, A., Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene* 1994, *151*, 131–135.

[129] Han, R., Zwiefka, A., Caswell, C. C., Xu, Y. et al., Assessment of prokaryotic collagen-like sequences derived from streptococcal Scl1 and Scl2 proteins as a source of recombinant GXY polymers. *Appl. Microbiol. Biotechnol.* 2006, *72*, 109–115.

[130] Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J., Tan, S., Comparison of affinity tags for protein purification. *Protein Expr. Purif.* 2005, *41*, 98–105.

[131] Prinz, B., Schultchen, J., Rydzewski, R., Holz, C. et al., 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* 2004, *5*, 29–44.

[132] Boettner, M., Prinz, B., Holz, C., Stahl, U., Lang, C., High-throughput screening for expression of heterologous proteins in the yeast Pichia pastoris. *J. Biotechnol.* 2002, *99*, 51–62.

[133] Wilson, D. S., Keefe, A. D., Szostak, J. W., The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 3750–3755.

[134] Keefe, A. D., Wilson, D. S., Seelig, B., Szostak, J. W., One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr. Purif.* 2001, *23*, 440–446.

[135] Stofko-Hahn, R. E., Carr, D. W., Scott, J. D., A single step purification for recombinant proteins. Characterization of a

microtubule associated protein (MAP 2) fragment which associates with the type II cAMP-dependent protein kinase. *FEBS Lett.* 1992, *302*, 274–278.
[136] Blumenthal, D. K., Takio, K., Edelman, A. M., Charbonneau, H. et al., Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase. *Proc. Natl. Acad. Sci. USA* 1985, *82*, 3187–3191.
[137] Vaillancourt, P., Zheng, C. F., Hoang, D. Q., Breister, L., Affinity purification of recombinant proteins fused to calmodulin or to calmodulin-binding peptides. *Methods Enzymol.* 2000, *326*, 340–362.
[138] Head, J. F., A better grip on calmodulin. *Curr. Biol.* 1992, *2*, 609–611.
[139] Szweda, P., Pladzyk, R., Kotlowski, R., Kur, J., Cloning, expression, and purification of the *Staphylococcus simulans* lysostaphin using the intein-chitin-binding domain (CBD) system. *Protein Expr. Purif.* 2001, *22*, 467–471.
[140] Wiese, A., Wilms, B., Syldatk, C., Mattes, R., Altenbuchner, J., Cloning, nucleotide sequence and expression of a hydantoinase and carbamoylase gene from *Arthrobacter aurescens* DSM 3745 in *Escherichia coli* and comparison with the corresponding genes from *Arthrobacter aurescens* DSM 3747. *Appl. Microbiol. Biotechnol.* 2001, *55*, 750–757.
[141] Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M. et al., The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J. Bacteriol.* 1994, *176*, 4465–4472.
[142] Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E. et al., Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* 1997, *192*, 271–281.
[143] Chong, S., Shao, Y., Paulus, H., Benner, J. et al., Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J. Biol. Chem.* 1996, *271*, 22159–22168.
[144] Xu, M. Q., Paulus, H., Chong, S., Fusions to self-splicing inteins for protein purification. *Methods Enzymol.* 2000, *326*, 376–418.
[145] Evans, T. C., Jr., Xu, M. Q., Intein-mediated protein ligation: Harnessing nature's escape artists. *Biopolymers* 1999, *51*, 333–342.
[146] Carrard, G., Koivula, A., Soderlund, H., Beguin, P., Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 10342–10347.
[147] Nahalka, J., Nidetzky, B., Fusion to a pull-down domain: A novel approach of producing Trigonopsis variabilisD-amino acid oxidase as insoluble enzyme aggregates. *Biotechnol. Bioeng.* 2007, *97*, 454–461.
[148] Karpeisky, M., Senchenko, V. N., Dianova, M. V., Kanevsky, V., Formation and properties of S-protein complex with S-peptide-containing fusion protein. *FEBS Lett.* 1994, *339*, 209–212.
[149] Kim, J. S., Raines, R. T., A misfolded but active dimer of bovine seminal ribonuclease. *Eur. J. Biochem.* 1994, *224*, 109–114.
[150] Connelly, P. R., Varadarajan, R., Sturtevant, J. M., Richards, F. M., Thermodynamics of protein–peptide interactions in the ribonuclease S system studied by titration calorimetry. *Biochemistry* 1990, *29*, 6108–6114.
[151] Lellouch, A. C., Geremia, R. A., Expression and study of recombinant ExoM, a beta1-4 glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* 1999, *181*, 1141–1148.
[152] Duellman, S. J., Thompson, N. E., Burgess, R. R., An epitope tag derived from human transcription factor IIB that reacts with a polyol-responsive monoclonal antibody. *Protein Expr. Purif.* 2004, *35*, 147–155.
[153] Burgess, R. R., Thompson, N. E., Advances in gentle immunoaffinity chromatography. *Curr. Opin. Biotechnol.* 2002, *13*, 304–308.
[154] Thompson, N. E., Arthur, T. M., Burgess, R. R., Development of an epitope tag for the gentle purification of proteins by immunoaffinity chromatography: Application to epitope-tagged green fluorescent protein. *Anal. Biochem.* 2003, *323*, 171–179.
[155] Ohana, R. F., Hurst, R., Vidugiriene, J., Slater, M. R. et al., HaloTag-based purification of functional human kinases from mammalian cells. *Protein Expr. Purif.* 2011, *76*, 154–164.
[156] Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D. et al., HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* 2008, *3*, 373–382.
[157] Los, G. V., Wood, K., The HaloTag: A novel technology for cell imaging and protein analysis. *Methods Mol. Biol.* 2007, *356*, 195–208.
[158] Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M. et al., A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 1999, *17*, 1030–1032.
[159] Tropea, J. E., Cherry, S., Nallamsetty, S., Bignon, C., Waugh, D. S., A generic method for the production of recombinant proteins in *Escherichia coli* using a dual hexahistidine-maltose-binding protein affinity tag. *Methods Mol. Biol.* 2007, *363*, 1–19.
[160] Liu, H., Naismith, J. H., A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr. Purif.* 2009, *63*, 102–111.
[161] O'Malley, M. A., Lazarova, T., Britton, Z. T., Robinson, A. S., High-level expression in *Saccharomyces cerevisiae* enables isolation and spectroscopic characterization of functional human adenosine A2a receptor. *J. Struct. Biol.* 2007, *159*, 166–178.
[162] Ford, C. F., Suominen, I., Glatz, C. E., Fusion tails for the recovery and purification of recombinant proteins. *Protein Expr. Purif.* 1991, *2*, 95–107.
[163] Sheibani, N., Prokaryotic gene fusion expression systems and their use in structural and functional studies of proteins. *Prep. Biochem. Biotechnol.* 1999, *29*, 77–90.
[164] Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D. et al., Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 2002, *415*, 180–183.
[165] Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W. et al., Global analysis of protein expression in yeast. *Nature* 2003, *425*, 737–741.
[166] Gavin, A. C., Bosche, M., Krause, R., Grandi, P. et al., Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002, *415*, 141–147.
[167] Newstead, S., Kim, H., von Heijne, G., Iwata, S., Drew, D., High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 13936–13941.
[168] Young, C. L., Yuraszeck, T., Robinson, A. S., Decreased secretion and unfolded protein response upregulation. *Methods Enzymol.* 2011, *491*, 235–260.
[169] Drew, D., Newstead, S., Sonoda, Y., Kim, H. et al., GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in *Saccharomyces cerevisiae*. *Nat. Protoc.* 2008, *3*, 784–798.
[170] Xu, X., Song, Y., Li, Y., Chang, J. et al., The tandem affinity purification method: An efficient system for protein com-

plex purification and protein interaction identification. *Protein Expr. Purif.* 2010, *72*, 149–156.
[171] Puig, O., Caspary, F., Rigaut, G., Rutz, B. et al., The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 2001, *24*, 218–229.
[172] Choi, S. I., Song, H. W., Moon, J. W., Seong, B. L., Recombinant enterokinase light chain with affinity tag: Expression from *Saccharomyces cerevisiae* and its utilities in fusion protein technology. *Biotechnol. Bioeng.* 2001, *75*, 718–724.
[173] Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A., Dougherty, W. G., Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal. Biochem.* 1994, *216*, 413–417.
[174] Sticha, K. R., Sieg, C. A., Bergstrom, C. P., Hanna, P. E., Wagner, C. R., Overexpression and large-scale purification of recombinant hamster polymorphic arylamine N-acetyltransferase as a dihydrofolate reductase fusion protein. *Protein Expr. Purif.* 1997, *10*, 141–153.
[175] Vothknecht, U. C., Kannangara, C. G., von Wettstein, D., Expression of catalytically active barley glutamyl tRNAGlu reductase in *Escherichia coli* as a fusion protein with glutathione S-transferase. *Proc. Natl. Acad. Sci. USA* 1996, *93*, 9287–9291.
[176] Cordingley, M. G., Callahan, P. L., Sardana, V. V., Garsky, V. M., Colonno, R. J., Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. *J. Biol. Chem.* 1990, *265*, 9062–9065.
[177] Vergis, J. M., Wiener, M. C., The variable detergent sensitivity of proteases that are utilized for recombinant protein affinity tag removal. *Protein Expr. Purif. 78*, 139–142.
[178] Hosfield, T., Lu, Q., Influence of the amino acid residue downstream of (Asp)4Lys on enterokinase cleavage of a fusion protein. *Anal. Biochem.* 1999, *269*, 10–16.
[179] Liew, O. W., Ching Chong, J. P., Yandle, T. G., Brennan, S. O., Preparation of recombinant thioredoxin fused N-terminal proCNP: Analysis of enterokinase cleavage products reveals new enterokinase cleavage sites. *Protein Expr. Purif.* 2005, *41*, 332–340.
[180] Nagai, K., Thogersen, H. C., Generation of beta-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature* 1984, *309*, 810–812.
[181] Ko, Y. H., Thomas, P. J., Pedersen, P. L., The cystic fibrosis transmembrane conductance regulator. Nucleotide binding to a synthetic peptide segment from the second predicted nucleotide binding fold. *J. Biol. Chem.* 1994, *269*, 14584–14588.
[182] Jenny, R. J., Mann, K. G., Lundblad, R. L., A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expr. Purif.* 2003, *31*, 1–11.
[183] Pryor, K. D., Leiting, B., High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expr. Purif.* 1997, *10*, 309–319.
[184] Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F., Fletterick, R. J., Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. *Virology* 1989, *172*, 302–310.
[185] Dougherty, W. G., Carrington, J. C., Cary, S. M., Parks, T. D., Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *Embo J.* 1988, *7*, 1281–1287.
[186] Carrington, J. C., Dougherty, W. G., A viral cleavage site cassette: Identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. USA* 1988, *85*, 3391–3395.
[187] Tropea, J. E., Cherry, S., Waugh, D. S., Expression and purification of soluble His(6)-tagged TEV protease. *Methods Mol. Biol.* 2009, *498*, 297–307.
[188] Blommel, P. G., Fox, B. G., A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* 2007, *55*, 53–68.
[189] Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E. et al., Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* 2001, *14*, 993–1000.
[190] Chang, J. Y., Alkan, S. S., Hilschmann, N., Braun, D. G., Thrombin specificity. Selective cleavage of antibody light chains at the joints of variable with joining regions and joining with constant regions. *Eur. J. Biochem.* 1985, *151*, 225–230.
[191] Chang, J. Y., Thrombin specificity. Requirement for apolar amino acids adjacent to the thrombin cleavage site of polypeptide substrate. *Eur. J. Biochem.* 1985, *151*, 217–224.
[192] Haun, R. S., Moss, J., Ligation-independent cloning of glutathione S-transferase fusion genes for expression in *Escherichia coli*. *Gene* 1992, *112*, 37–43.
[193] Koehl, C., Abecassis, J., Determination of oxalic acid in urine by atomic absorption spectrophotometry. *Clin. Chim. Acta* 1976, *70*, 71–77.
[194] Donaldson, V. H., Kleniewski, J., The role of plasmin in kinin-release by preparations of human thrombin. *Thromb Res.* 1979, *16*, 401–406.
[195] Litwiller, R. D., Jenny, R. J., Mann, K. G., Identification and isolation of vitamin K-dependent proteins by HPLC. *Anal. Biochem.* 1986, *158*, 355–360.
[196] Walker, P. A., Leong, L. E., Ng, P. W., Tan, S. H. et al., Efficient and rapid affinity purification of proteins using recombinant fusion proteases. *Biotechnology (N Y)* 1994, *12*, 601–605.
[197] Kwon, K., Grose, C., Pieper, R., Pandya, G. A. et al., High quality protein microarray using in situ protein purification. *BMC Biotechnol.* 2009, *9*, 72.
[198] Steen, J., Uhlen, M., Hober, S., Ottosson, J., High-throughput protein purification using an automated set-up for high-yield affinity chromatography. *Protein Expr. Purif.* 2006, *46*, 173–178.

BiologicsCorp

Your protein manufacturer

BiologicsCorp provides high-quality cost-effective recombinant protein expression services in a timely manner. Based on our in-house codon optimization technology and protein expression platform, we provide one-stop services including gene optimization and synthesis, protein expression and purification, protein refolding, endotoxin/RNase control and removal, and fermentation. Other services available include custom antibody production, and biomolecular interaction analysis.

Protein Expression Service Highlights

Guaranteed quantity and purity: Get your desired quantity and purity or your money back.

Flexible solutions: Multiple approaches for assured success.

Guaranteed soluble proteins: You only get billed if solubility is achieved.

Maximum activity: We prioritize supernatant proteins.

Protein Expression

Bacterial Expression System	
Guaranteed *E.coli* Protein Expression	$1,550 for 3-5mg ~85% purity soluble protein No protein, no billing
SupernateIN™ Protein Expression	Protein with maximum biological activity High soluble expression level
Tag-free Protein Production	Close to native protein Soluble protein with over 95% purity
Mammalian Expression System	
Transient Transfection	High yield: >200mg/L High transfection efficiency
Stable Cell Line Establishment	Scalable and reproducible production 30% below industry average
Recombinant Antibody Production	High yield: up to 1g/L expression level High-specificity: bind to corresponding antigen specifically with high-affinity
Other Protein Services	
Endotoxin/RNase Removal	Endotoxin level below 1 EU/μg >90% RNase removal from protein
Fermentation	Fed-batch fermentation Protein production up to 2000 L
Biomolecular Interaction Analysis	Label-free and real-time analysis K_a, K_d and K_D for small molecule and protein binding interactions

E-mail: Sales@BiologicsCorp.com
Telephone: +1 (317) 703-0614
Fax: +1 (855) 427-1516
Website: http://www.biologicscorp.com
Address: Biologics International Corp
5323 East 82nd Street, Unit 109
Indianapolis, IN 46250, USA