Appl Microbiol Biotechnol (2003) 60:523–533
DOI 10.1007/s00253-002-1158-6

MINI-REVIEW

K. Terpe

Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems

Received: 8 July 2002 / Revised: 25 September 2002 / Accepted: 27 September 2002 / Published online: 7 November 2002

Abstract In response to the rapidly growing field of proteomics, the use of recombinant proteins has increased greatly in recent years. Recombinant hybrids containing a polypeptide fusion partner, termed affinity tag, to facilitate the purification of the target polypeptides are widely used. Many different proteins, domains, or peptides can be fused with the target protein. The advantages of using fusion proteins to facilitate purification and detection of recombinant proteins are well-recognized. Nevertheless, it is difficult to choose the right purification system for a specific protein of interest. This review gives an overview of the most frequently used and interesting systems: Arg-tag, calmodulin-binding peptide, cellulose-binding domain, DsbA, c-myc-tag, glutathione S-transferase, FLAG-tag, HAT-tag, His-tag, maltose-binding protein, NusA, S-tag, SBP-tag, Strep-tag, and thioredoxin.

Introduction

The production of recombinant proteins in a highly purified and well-characterized form has become a major task for the protein chemist working in the pharmaceutical industry. In recent years, several epitope peptides and proteins have been developed to over-produce recombinant proteins. These affinity-tag systems share the following features: (a) one-step adsorption purification; (b) a minimal effect on tertiary structure and biological activity; (c) easy and specific removal to produce the native protein; (d) simple and accurate assay of the recombinant protein during purification; (e) applicability to a number of different proteins. Nevertheless, each affinity tag is purified under its specific buffer conditions, which could affect the protein of interest (Table 1). Thus, several different strategies have been developed to produce recombinant proteins on a large scale. One approach is to use a very small peptide tag that should not interfere with the fused protein. The most commonly used small peptide tags are poly-Arg-, FLAG-, poly-His-, c-myc-, S-, and Strep II-tag. For some applications, small tags may not need to be removed. The tags are not as immunogenic as large tags and can often be used directly as an antigen in antibody production. The effect on tertiary structure and biological activity of fusion proteins with small tags depends on the location and on the amino acids composition of the tag (Bucher et al. 2002). Another approach is to use large peptides or proteins as the fusion partner. The use of a large partner can increase the solubility of the target protein. The disadvantage is that the tag must be removed for several applications e.g. crystallization or antibody production.

In general, it is difficult to decide on the best fusion system for a specific protein of interest. This depends on the target protein itself (e.g. stability, hydrophobicity), the expression system, and the application of the purified protein. This review provides an overview on the most frequently used and interesting tag-protein fusion systems (Table 2).

K. Terpe (✉)
Technical Consultant of the IBA GmbH,
Protein expression/purification and nucleic acids, 37079 Göttingen, Germany
e-mail: terpe@iba-go.de
Tel.: +49-551-50672121
Fax: +49-551-50672181

Polyarginine-tag (Arg-tag)

The Arg-tag was first described in 1984 (Sassenfeld and Brewer 1984) and usually consists of five or six arginines. It has been successfully applied as C-terminal tag in bacteria, resulting inrecombinant protein with up to 95% purity and a 44% yield. Arginine is the most basic amino acid. Arg_5-tagged proteins can be purified by cation exchange resin SP-Sephadex, and most of the contaminating proteins do not bind. After binding, the tagged proteins are eluted with a linear NaCl gradient at alkaline pH. Polyarginine might affect the tertiary structure of proteins whose C-terminal region is hydrophobic (Sassenfeld and Brewer 1984). The Arg-tagged maltodextrin-binding protein of *Pyrococcus furiosus* has been crystal-

Table 1 Matrices and elution conditions of affinity tags

Affinity tag	Matrix	Elution condition
Poly-Arg	Cation-exchange resin	NaCl linear gradient from 0 to 400 mM at alkaline pH>8.0
Poly-His	Ni^{2+}-NTA, Co^{2+}-CMA (Talon)	Imidazole 20–250 mM or low pH
FLAG	Anti-FLAG monoclonal antibody	pH 3.0 or 2–5 mM EDTA
Strep-tag II	Strep-Tactin (modified streptavidin)	2.5 mM desthiobiotin
c-myc	Monoclonal antibody	Low pH
S	S-fragment of RNaseA	3 M guanidine thiocyanate, 0.2 M citrate pH 2, 3 M magnesium chloride
HAT (natural histidine affinity tag)	Co^{2+}-CMA (Talon)	150 mM imidazole or low pH
Calmodulin-binding peptide	Calmodulin	EGTA or EGTA with 1 M NaCl
Cellulose-binding domain	Cellulose	Family I: guanidine HCl or urea>4 M Family II/III: ethylene glycol
SBP	Streptavidin	2 mM Biotin
Chitin-binding domain	Chitin	Fused with intein: 30–50 mM dithiothreitol, β-mercaptoethanol or cysteine
Glutathione S-transferase	Glutathione	5–10 mM reduced glutathione
Maltose-binding protein	Cross-linked amylose	10 mM maltose

Table 2 Sequence and size of affinity tags

Tag	Residues	Sequence	Size (kDa)
Poly-Arg	5–6 (usually 5)	RRRRR	0.80
Poly-His	2–10 (usually 6)	HHHHHH	0.84
FLAG	8	DYKDDDDK	1.01
Strep-tag II	8	WSHPQFEK	1.06
c-myc	11	EQKLISEEDL	1.20
S-	15	KETAAAKFERQHMDS	1.75
HAT-	19	KDHLIHNVHKEFHAHAHNK	2.31
3x FLAG	22	DYKDHDGDYKDHDIDYKDDDDK	2.73
Calmodulin-binding peptide	26	KRRWKKNFIAVSAANRFKKISSSGAL	2.96
Cellulose-binding domains	27–189	Domains	3.00–20.00
SBP	38	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP	4.03
Chitin-binding domain	51	TNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ	5.59
Glutathione S-transferase	211	Protein	26.00
Maltose-binding protein	396	Protein	40.00

lized (Bucher et al. 2002). The crystals were visually indistinguishable from crystals of the native protein; however, the crystals did differ in mosaicity and diffraction. C-terminal series of arginine residues can be removed by carboxypeptidase B treatment. This enzymatic process has been successfully used in several instances, but often has been limited by poor cleavage yields or by unwanted cleavage occurred within the desired protein sequence (Nagai and Thogerson 1987). The Arg-tag can be used to immobilize functional proteins on flat surfaces; this is important for studying interactions with ligands. GFP with an Arg_6-tag on one of its termini can be reversibly and specifically bound via this sequence onto a mica surface, which has been established as a standard substrate for electron and scanning probe microscopy applications (Nock et al. 1997). While the Arg-tag is not used very often, in combination with a second tag it can be an interesting tool for protein purification.

Polyhistidine-tag (His-tag)

A widely employed method utilizes immobilized metal-affinity chromatography to purify recombinant proteins containing a short affinity-tag consisting of polyhistidine residues. Immobilized metal-affinity chromatography (IMAC; described by Porath et al. 1975) is based on the interaction between a transition metal ion (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) immobilized on a matrix and specific amino-acid side chains. Histidine is the amino acid that exhibits the strongest interaction with immobilized metal ion matrices, as electron donor groups on the histidine imidazole ring readily form coordination bonds with the immobilized transition metal. Peptides containing sequences of consecutive histidine residues are efficiently retained on IMAC. Following washing of the matrix material, peptides containing polyhistidine sequences can be easily eluted by either adjusting the pH of the column buffer or by adding free imidazole (Table 1). The method to purify proteins with histidine residues was first

Table 3 Affinity of polyhistidine dihydrofolate reductase (DHFR) for the Ni^{2+}-NTA adsorbent in 6 M guanidine hydrochloride (GuHCl) and 0.05 M phosphate buffer (Hochuli et al. 1988)

	Phosphate		GuHCl	
	Retained (%)	Eluted (%)	Retained (%)	Eluted (%)
Polyhistidine dihydrofolate reductase				
$(His)_2$-DHFR	30	10	<5	–
$(His)_3$-DHFR	90	75	<10	–
$(His)_4$-DHFR	>90	30	10	10
$(His)_5$-DHFR	>90	20	50	50
$(His)_6$-DHFR	>90	10	>90	90
DHFR-$(His)_2$	>90	90	<5	–
DHFR-$(His)_3$	>90	80	<10	–
DHFR-$(His)_4$	>90	50	10	10
DHFR-$(His)_5$	>90	40	50	50
DHFR-$(His)_6$	>90	30	>90	90

described in 1987 (Hochuli et al. 1987). Hochuli has developed a nitrilotriacetic acid (NTA) adsorbent for metal-chelate affinity chromatography. The NTA resin forms a quadridentate chelate and is especially suitable for metal ions with coordination numbers of six, since two valencies remain for the reversible binding of biopolymers. Dihydrofolate reductase with a poly-His-tag was successfully purified with Ni^{2+}-NTA matrices in 1988 (Hochuli et al. 1988). The purification efficiency of this system was dependent on the length of the poly-histidine and the solvent system (Table 3). While the system worked efficiently with His_6-tagged proteins under denaturing conditions, His_3-tagged proteins were efficiently purified under physiological conditions. However, His_6-tagged proteins can be bound to Ni^{2+}-NTA matrices under native conditions in low- or high-salt buffers. After binding, the target protein can be eluted by an imidazole gradient from 0.8 to 250 mM. Washing with a low concentration of imidazole (e.g. 0.8 mM) reduces non-specific binding of host proteins with histidines. Elution of His_6-tagged proteins is effective within a range of 20–250 mM imidazole (Hefti et al. 2001; Janknecht et al. 1991). A disadvantage of using imidazole is that it can influence NMR experiments, competition studies, and crystallographic trials, and the presence of imidazole often results in protein aggregates (Hefti et al. 2001). Another material that has been developed to purify His-tagged proteins is TALON. It consists of a Co^{2+}-carboxylmethylaspartate (Co^{2+}-CMA), which is coupled to a solid-support resin. TALON allows the elution of tagged proteins under mild conditions, and it has been reported to exhibit less non-specific protein binding than the Ni^{2+}-NTA resin, resulting in higher elution product purity (Chaga et al. 1999a, b). A final preparation of enzymes exhibited a purity higher than 95% as ascertained by SDS-PAGE. Purification with Co^{2+}-CMA allowed the development of a natural 19-amino-acid poly-histidine affinity tag (HAT-tag; for the sequence, see Table 2). Chloramphenicol acetyltransferase, dihydrofolate reductase, and green fluorescent protein with N-terminal HAT-tags were purified under mild conditions in one step with a purity over 95%. Adsorption of weakly bound unspecific proteins was eliminated by using 5 mM imidazole in the equilibration and loading buffer, and 150 mM imidazole was used to elute the HAT-tagged proteins. Elution of tagged proteins was also possible by decreasing the pH to 5.0. Urea turned out to have a much stronger negative effect on the binding of HAT-tagged proteins than guanidinium HCl. However, over-expression with HAT-tag has only been tested in bacteria.

Poly-histidine affinity tags are commonly placed on either the N- or the C-terminus of recombinant proteins. Optimal placement of the tag is protein-specific. Purification using poly-histidine tags has been carried out successfully using a number of expression systems including bacteria (Chen and Hai 1994; Rank et al. 2001), yeast (Borsing et al. 1997; Kaslow and Shiloach 1994), mammalian cells (Janknecht et al. 1991; Janknecht and Nordheim 1992), and baculovirus-infected insect cells (Kuusinen et al. 1995; Schmidt et al. 1998). More than 100 structures of His-tagged proteins have been deposited in the Protein Data Bank. Proteins with a His-tag may vary slightly as far as their mosaicity and diffraction compared to the native protein (Hakansson et al. 2000). In principle, it cannot be excluded that the affinity tag may interfere with protein activity (Wu and Filutowicz 1999), although the relatively small size and charge of the polyhistidine affinity tag ensure that protein activity is rarely affected. Moving the affinity tag to the opposite terminus (Halliwell et al. 2001) or carrying out the purification under denaturing conditions often solves this problem. Purification of protein with a metal center is not recommended because the metal can be absorbed by the NTA. Purification under anaerobic conditions is also not recommended because Ni^{2+}-NTA is reduced. Nevertheless, purification of proteins with His-tag is the most commonly used method.

FLAG-tag

The FLAG-tag system utilizes a short, hydrophilic 8-amino-acid peptide (Table 1) that is fused to the protein of interest (Hopp et al. 1988). The FLAG peptide binds to the antibody M1. Whether binding is calcium-dependent manner (Hope et al. 1996) or -independent (Einhauer and Jungbauer 2000) remains controversial. Kinetic studies for binding of FLAG-GFP, evaluated by BIACORE

analysis, were identical in the presence and absence of Ca^{2+} ions. Additional targets are the monoclonal antibodies M2 and M5, each with different recognition and binding characteristics. The FLAG-tag can be located at the C- or N-terminus of the protein. The system has been used in a variety of cell types, including examples from bacterial (Blanar and Rutter 1992; Su et al. 1992), yeast (Einhauer et al. 2002; Schuster et al. 2000), and mammalian cells (Kunz et al. 1992; Zhang et al. 1991). The purification condition of the system is non-denaturing and thus allows active fusion proteins to be purified. The complex can be dissociated by chelating agents such as EDTA or by transiently reducing the pH (Table 1). A disadvantage of the system is that the monoclonal-antibody purification matrix is not as stable as others, e.g. Ni^{2+}-NTA or Strep-Tactin. The purity of isolated proteins is in the range of 90% (Schuster et al. 2000). In general, small tags can be detected with specific monoclonal antibodies. To improve the detection of the FLAG-tag the 3x FLAG system has been developed. This three-tandem FLAG epitope is hydrophilic, 22-amino-acids long (Table 2) and can detect up to 10 fmol of expressed fusion protein. The FLAG-tagged maltodextrin-binding protein of *Pyrococcus furiosus* has been crystallized (Bucher et al. 2002) and the quality of the crystals was very similar to that of crystals of untagged protein. Finally, the FLAG-tag can be removed by treatment with enterokinase, which is specific for the five C-terminal amino acids of the peptide sequence (Maroux et al. 1971).

Strep-tag

The Strep-tag is an amino acid peptide that was developed as an affinity tool for the purification of corresponding fusion proteins on streptavidin columns (Schmidt and Skerra 1993). Streptavidin mutants with a specific mutation at position 44, 45, and 47 have a higher affinity for the octapeptide Strep-tag II than for the native form (for the sequence, see Table 2; Schmidt et al. 1996; Voss and Skerra 1997; Korndörfer and Skerra 2001). This streptavidin variant is called Strep-Tactin. Strep-tagged proteins are bound under physiological buffer conditions in the biotin binding pocket, and can be eluted gently with biotin derivatives. Elution with 2.5 mM desthiobiotin is recommended. The matrix can be regenerated with 4-hydroxy azobenzene-2-carboxylic acid, which is yellow in solution and red when bound on the matrix. The binding conditions are very specific. Biotinylated proteins such as the carboxyl carrier protein of *Escherichia coli* are also bound on Strep-Tactin, but biotin or biotinylated proteins can be blocked with avidin. The purification conditions are highly variable. Chelating agents, mild detergents, reduction detergents, and salt up to 1 M can be added to the buffer. Denaturing purification conditions, such 6 M urea, destroy the Strep-tag/Strep-Tactin interaction but not Strep-Tactin. The interaction between the tag and Strep-Tactin is close to the range of 1 μM (Voss and Skerra 1997). Fusion proteins can be specifically detected by Strep-Tactin conjugates or by antibodies. The tag can be engineered to either the C- or N-terminus of a protein. Recombinant Strep-tag-hybrids are produced in bacteria (Fontaine et al. 2002), yeast (Murphy and Lagarias 1997), mammalian systems (Sárdy et al. 2002; Smyth et al. 2000), plants (Drucker et al. 2002) and baculovirus-infected insect cells. This method is recommended for purifying active fusion proteins with a small tag under anaerobic conditions (Hans and Buckel 2000; Juda et al. 2001), and for metal-containing enzymes. Integration of tagged proteins into the membrane is also possible (Groß et al. 2002). Membrane protein subunits with no tag could be co-purified. A special application of the tag is that it can be used for eukaryotic surface display (Ernst et al. 2000). The compatibility of Strep-Tactin binding biotin and Strep-tag was used to observe the rotating c-subunit oligomer of EF_0EF_1-F-ATPase (Pänke et al. 2000). The use of Strep-tag has widely increased during the last years. Recombinant proteins with the tag can be used for NMR and crystallization (Ostermeier et al. 1997). The Strep-tag system is of relevance for studies on protein-protein interaction and special applications in which large or charged tags are not functional.

c-myc-tag

The murin anti-c-myc antibody 9E10 was developed in 1985 (Evan et al. 1985) and is used as an immunochemical reagent in cell biology and in protein engineering. The antibody epitope of eleven amino acids (Table 2) can be expressed in a different protein context and still confers recognition by the 9E10 immunoglobulin (Munro and Pelham 1986). The c-myc-tag has been successfully used in Western-blot technology, immunoprecipitation, and flow cytometry (Kipriyanov 1996). It is therefore useful for monitoring expression of recombinant proteins in bacteria (Dreher et al. 1991; Vaughan et al. 1996), yeast (Sequi-Real et al. 1995; Weiss et al. 1998), insect cells (Schioth et al. 1996), and mammalian cells (McKern 1997; Moorby and Gherardi 1999). The successful co-immunopurification of interacting proteins expressed in *Agrobacterium*-transformed *Arabidobsis* cells was also reported (Ferrando et al. 2001). c-myc-tagged proteins can be affinity-purified by coupling Mab 9E10 to divinyl sulphone-activated agarose. The washing conditions are physiological followed by elution at low pH, which could exert a negative effect on protein activity. Purified c-myc-tagged proteins have been crystallized (McKern et al. 1997). The c-myc-tag can be placed at the N- or C-terminus (Manstein et al. 1995). It is a widely used detection system but is rarely applied for purifications.

S-tag

The S-tag sequence is a fusion-peptide tag that allows detection by a rapid, sensitive homogeneous assay or by colorimetric detection in Western blots. The system is

based on the strong interaction between the 15-amino-acid S-tag (Table 2) and the 103-amino-acid S-protein, both of which are derived from RNaseA (Karpeisky et al. 1994; Kim and Raines 1994). The S-protein/S-tag complex has a k_d of ~0.1 μM which depends on pH, temperature, and ionic strength (Connelly et al. 1990). The tag is composed of four cationic, three anionic, three uncharged polar, and five non-polar residues. This composition makes the S-tag soluble. The S-tag rapid assay is based on the reconstitution of ribonucleolytic activity. Tagged proteins can be bound on S-protein matrices. The elution conditions are very harsh, e.g. buffer with pH 2 (Table 1); however, it is recommended to cleave the tag with protease to get functional proteins. The system is functional to purify recombinant proteins from bacteria (Lellouch and Geremia 1999), mammalian cells, and baculovirus-infected insect cell extracts. The system is often used together with a second tag. The discovery of a hypersensitive fluorogenic substrate for RNase A makes the system interesting for detection in combination with high-throughput screening (Kelemen et al. 1999).

Calmodulin-binding peptide

Purification of fusion proteins containing calmodulin-binding peptide was first described in 1992 (Stofko-Hahn et al. 1992). The peptide has 26 residues (for the sequence, see Table 2) derived from the C-terminus of skeletal-muscle myosin light-chain kinase, which binds calmodulin with nanomolar affinity in the presence of 0.2 mM $CaCl_2$ (Blumenthal et al. 1985) The tight binding allows more stringent washing conditions, ensuring that few contaminating proteins will be co-purified with the fusion protein. A second elution step with EGTA and 1 M NaCl is useful if the protein does not elute completely at the first step. The system has a high specificity to purify recombinant proteins in *E. coli* because there are no endogenous proteins that interact with calmodulin. Recovery of fusion proteins is 80–90%. Reducing agents and detergents in amounts up to 0.1% are compatible with the system (Vaillancourt et al. 2000). Purification in eukaryotic cells is not recommended because many endogenous proteins interact with calmodulin in a calcium-dependent manner (Head 1992). A calmodulin-binding peptide thrombin fusion tag is an excellent target for isotopic labeling with $\gamma[^{32}]$ATP using protein kinase A (Vaillancourt et al. 2000). His-tagged protein kinase can be removed by Ni^{2+}-NTA chromatography. This allows studies of protein interaction or screening of bacteriophage expression libraries. The calmodulin-binding peptide can be placed at the N- or C-terminus. The N-terminal location may reduce the efficiency of translation, while calmodulin-binding peptide at the C-terminus can result in high expression levels (Zheng et al. 1997).

Cellulose-binding domain

More than 13 different families of proteins with cellulose-binding domains (CBDs) have been classified. CBDs can vary in size from 4 to 20 kDa; they occur at different positions within polypeptides: N-terminal, C-terminal and internal. Some CBDs bind irreversibly to cellulose and can be used for immobilization of active enzymes (Xu et al. 2002); others bind reversibly and are more useful for separation and purification. CBDs of family I bind reversibly to crystalline cellulose and are a useful tag for affinity chromatography. Hydrogen bond formation and van der Waals interaction are the main driving forces for binding (Tomme et al. 1998). The advantage of cellulose is that it is inert, has low non-specific affinity, is available in many different forms, and has been approved for many pharmaceutical and human uses. CBDs bind to cellulose at a moderately wide pH range, from 3.5 to 9.5. The tag can be placed at the N- or C-terminus of the target protein. The affinity of the tag is so strong that an immobilized fusion protein can only be released with buffers containing urea or guanidine hydrochloride. This denaturating elution conditions make refolding of the fused target protein necessary. Fused proteins with CBDs of families II and III can be eluted gently from cellulose with ethylene glycol (McCormick and Berg 1997). This low-polarity solvent presumably disrupts the hydrophobic interaction at the binding site. Ethylene glycol can be removed easily by dialysis. Recombinant CBD-hybrids have been produced in bacteria, yeast, mammalian cells, and baculovirus-infected insect cells (Tomme et al. 1998).

SBP-tag

The SBP-tag is a new streptavidin-binding peptide and has a length of 38 amino acids (for the sequence, see Table 2; Wilson et al. 2001). The dissociation constant of the tag to streptavidin is 2.5 nM. SBP-tagged proteins can be purified with immobilized streptavidin. The elution conditions are very mild, using 2 mM biotin. Proteins with C-terminal SBP-tagged proteins were expressed in bacteria and successfully purified (Keefe et al. 2001). Little is known regarding further applications, but the tag seems to be an interesting tool to immobilize proteins on streptavidin-coated chips.

Chitin-binding domain

The chitin-binding domain from *Bacillus circulans* consists of 51 amino acids (Watanabe et al. 1994). The affinity tag is commonly available in combination with self-splicing inteins. The intein from the *Saccharomyces cerevisiae* VMA1 gene, which consists of 454 amino acids, is often used (Chong et al. 1996, 1997). Other, shorter inteins have also been employed (Xu et al. 2000). Self-cleavage of the thioester bond can be induced by thiol reagents, such as 1,4-dithiothreitol or β-mercapto-

ethanol (Table 2). The C- or N-terminal amino acid residue of the target protein has an effect on in vivo and in vitro cleavage (Xu et al. 2000). A high salt concentration or the use of non-ionic detergents can be employed to reduce non-specific binding, thus increasing purity. The uncleaved fusion precursor and the intein tag remain bound to the chitin resin during target protein elution and can be stripped from the resin by 1% SDS or 6 M guanidine HCl. Proteins with C- or N-terminal chitin-binding domains fused with inteins have been expressed in bacterial systems (Cantor and Chong 2001; Sweda et al. 2001; Wiese et al. 2001).

Glutathione S-transferase-tag

Single-step purification of polypetides as fusions with glutathione S-transferase (GST) was first described in 1988 (Smith and Johnson 1988). A 26-kDa GST of *Schistosoma japonicum* (Taylor et al. 1994) was cloned in an *E. coli* expression vector. Fusion proteins could be purified from crude lysate by affinity chromatography on immobilized glutathione. Bound fusion proteins can be eluted with 10 mM reduced glutathione under non-denaturing conditions. In the majority of cases, fusion proteins are soluble in aqueous solutions and form dimers. The GST-tag can be easily detected using an enzyme assay or an immunoassay. The tag can help to protect against intracellular protease cleavage and stabilize the recombinant protein. In some cases GST fusion proteins are totally or partly soluble. It remains unclear which factors are responsible for insolubility, but in several instances insolubility of GST fusion proteins was associated with the presence of hydrophobic regions. Other insoluble fusion proteins either contain many charged residues or are larger than 100 kDa. In some cases insoluble fusion proteins can be purified by affinity chromatography if they are solubilized in 1% Triton X-100, 1% Tween, 10 mM dithiothreitol, 0.03% SDS or 1.5% sarcosyl buffer (Frangioni and Neel 1993). Sarcosyl inhibits co-aggregation of proteins with bacterial outer membrane components. Purification of other insoluble proteins must be done by conventional methods. It is recommended to cleave the GST-tag from fusion proteins by a site-specific protease such as thrombin or factor X_a. The PreScission protease contains the human rhinovirus 3C protease including the GST-tag; the GST carrier and the protease can be removed after proteolysis by affinity chromatography on gluthatione-agarose. The GST-tag can be placed at the N- or C-terminus and can be used in bacteria (Smith and Johnson 1988), yeast (Lu et al. 1997), mammalian cells (Rudert et al. 1996), and baculovirus-infected insect cells (Beekman et al. 1994). GST fusion proteins have become a basic tool for the molecular biologist. They are also commonly used in studies on protein-DNA interactions (Beekman et al. 1994; Lassar et al. 1989), protein-protein interactions (Mayer et al. 1991; Ron and Dressler 1992) and as antigens for immunology or vaccination studies (McTigue et al. 1995).

Maltose-binding protein

The 40-kDa maltose-binding protein (MBP) is encoded by the *mal*E gene of *E. coli* K12 (Duplay et al. 1988). Vectors that facilitate the expression and purification of foreign peptides in *E. coli* by fusion to MPB were first described in 1988 (Di Guan et al. 1988). Fused proteins can be purified by one-step affinity chromatography on cross-linked amylose. Bound fusion proteins can be eluted with 10 mM maltose in physiological buffer. Binding affinity is in the micro-molar range. Some fusion proteins do not bind efficiently in the presence of 0.2% Triton X-100 or 0.25% Tween 20, while other fusions are unaffected. Buffer conditions are compatible from pH 7.0–8.5, and up to 1 M salt. Denaturing agents cannot be used. MBP can increase the solubility of over-expressed fusion proteins in bacteria, especially eukaryotic proteins (Sachdev and Chirgwin 1999). A spacer sequence coding for ten asparagine residues between the MBP and the protein of interest increases the chances that a particular fusion will bind tightly to the amylose resin. The MBP-tag can be easily detected using an immunoassay. It is necessary to cleave the tag with a site-specific protease. The MBP can be fused at the N- or C-terminus of the protein if the proteins are expressed in bacteria (Sachdev and Chirgwin 2000). N-terminal location can reduce the efficiency of translation. The MBP system is widely used in combination with a small affinity tag (Hamilton et al. 2002; Podmore and Reynolds 2002).

NusA, TrxA and DsbA

One disadvantage when heterologous proteins are produced in *E. coli* is that proteins frequently aggregates as insoluble folding intermediates, known as inclusion bodies. In order to recover an active protein, it must be solubilized with denaturing agents such as 8 M urea or 6 M guanidine hydrochloride. One possibility to avoid inclusion bodies is to use large affinity tags such as GST or MBP. Hydrophilic tags, such as transcription termination anti-termination factor (NusA), *E. coli* thioredoxin (TrxA), or protein disulfide isomerase I (DsbA) can increase solubility. A disadvantage is, however, that proteins with these tags cannot be purified with a specific affinity matrix. The fusion construct must be used in combination with a small affinity tag for purification. Especially, the NusA protein increases the solubility of fusion proteins (Davis et al. 1999). Usually, *E coli* NusA protein promotes hairpin folding and termination (Gusarov and Nudler 2001). Some insoluble proteins expressed in *E. coli* remained soluble when tagged N-terminal with NusA. NusA has often been used in combination with the His-tag (Harrisson 2000). Thioredoxin can be fused to the amino or carboxyl terminus of the protein of interest (Katti et al. 1990; LaVallie et al. 2000), but typically the *trxA* sequence is placed at the 5' end. DsbA increases the solubility of the target protein in the cytoplasm and periplasm of *E. coli*. It is recommended to cleave fusion

proteins with NusA, TrxA or DsbA by a site-specific protease; the cleavage site can be used as linker peptide.

Other tag-systems

There are also other tag systems in use, which are not described in detail in this review:

Staphylococcal protein A gene fusion vectors were developed to purify recombinant proteins by IgG affinity chromatography (Uhlén et al. 1983; Nilsson et al. 1985). This protein is well-suited for affinity purification due to its specific binding to the Fc part of immunoglobulins of many species including human. Analogously to protein A, protein G from *Streptococcus* strain G148 can be used in the same manner because it binds the Fc portion of IgG (Goward et al. 1990). Biotinylation of proteins using small peptide tags are commonly used for detection, immobilization, and purification (Cronan 1990). Different tags, such the AviTag, PinPoint X_a protein purification system, and Bio-tag (Schatz 1993; Tucker and Grisshammer 1996), have been described. The bacteriophage T7 and V5 epitopes are interesting tags for sensitive detection. Other epitope tags for detection are: ECS (enterokinase cleavage site), HA (hemaglutinin A), and Glu-Glu.

Cleavage of the tag

The presence of affinity tags may affect important characteristics or functions of the protein to be studied. Removal of the tag from a protein of interest can be accomplished with a site-specific protease, and cleavage should not reduce protein activity. Removal of the protease after cleavage is easier using a recombinant protease with an affinity tag or using a biotinylated protease. A biotinylated protease can be directly purified during affinity chromatography using Strep-tag/Strep-Tactin chromatography, or in a second step with streptavidin. Cleavage of the tag without using a protease is also possible by introducing a self-splicing intein (Xu et al. 2000). The most commonly used proteases are: enterokinase, tobacco etch virus (TEV), thrombin, and factor X_a. Recovery of the target protein depends on the cleavage efficiency.

Enterokinase is often the protease of choice for N-terminal fusions, since it specifically recognizes a five-amino-acid polypeptide (D-D-D-D-K-X_1) and cleaves at the carboxyl site of lysine. Sporadic cleavage at other residues was observed to occur at low levels, depending on the conformation of the protein substrate (Choi et al. 2001). The molecular weight of the light-chain of enterokinase is 26.3 kDa. One unit is defined as the amount of enterokinase that will cleave 95% of 50 μg of a fusion protein in 8 hat 23 °C. Biochemical analyses have shown that the cleavage efficiency depends on the amino acid residue X_1 downstream of the D_4K recognition site (Table 4; Hosfield and Lu 1999). In contrast to other tags, the FLAG-tag (DYKDDDK) has an internal recognition site of the enterokinase.

Table 4 Cleavage (%) of enterokinase through densitometry (Hosfield and Lu 1999) based on the amino acid residue X_1. The sequence....-GSDYKDDDDK-X_1-ADQLTEEQIA-... of a GST-calmodulin fusion protein was tested using 5 mg protein digested with 0.2 Uof enterokinase for 16 h at 37 °C

Amino acid in position X_1	Cleavage of enterokinase (%)
Alanine	88
Methionine	86
Lysine	85
Leucine	85
Asparagine	85
Phenylalanine	85
Isoleucine	84
Aspartic acid	84
Glutamic acid	80
Glutamine	79
Valine	79
Arginine	78
Threonine	78
Tyrosine	78
Histidine	76
Serine	76
Cysteine	74
Glycine	74
Tryptophan	67
Proline	61

TEV protease is a site-specific protease that has a seven-amino-acid recognition site. The sequence is E-X-X-Y-X-Q-S, and cleavage occurs between the conserved glutamine and serine (Dougherty et al. 1989). X can be various amino acid residues but not all are tolerated. The optimal sequence for cleavage is E-N-L-Y-F-Q-S (Carrington and Dougherty 1988; Doughery et al. 1988). Best results will be obtained when the TEV protease recognition site is placed between two domains. When cleavage is not optimal, insertion of short linker sequence introducing structural flexibility can improve efficiency. The high specificity, its activity on a variety of substrates, and the efficient cleavage at low temperature makes TEV protease an ideal tool for removing tags from fusion proteins (Parks et al. 1994). The efficiency of cleavage is dependent on both the tag and the protein fused to the carboxyl terminus of the TEV cleavage site.

Thrombin is a protease widely used to cleave tags. Cleavage can be carried out at temperatures between 20 and 37 °C for 0.3–16 h. In contrast to enterokinase and factor Xa, thrombin cleavage results in the retention of two amino acids on the C-terminal side of the cleavage point of the target protein. The optimal cleavage site for α-thrombin has the structures of X_4-X_3-P-R[K]-X_1'-X_2', where X_4 and X_3 are hydrophobic amino acid and X_1', X_2' are non-acidic amino acids (Chang 1985; Chang et al. 1985; Haun and Moos 1992). Some frequently used recognition sites are L-V-P-R-G-S, L-V-P-R-G-F, and M-Y-P-R-G-N. Cleavage between X_4-X_3-P-R-G-X_2' is more efficient than cleavage between X_4-X_3-P-K-L-X_2'. Other short recognition sites are X_2-R[K]-X_1', where X_2 or X_1' are glycine. Examples are A-R-G and G-K-A, where

cleavage occurs after the second residue. Five glycine residues between the thrombin cleavage site and the N-terminal tag enhance the cleavage (Guan and Dixon 1991). Using this "glycine kinker", less enzyme is necessary to effect complete digestion, and inappropriate cleavage, where it occurs, may be avoided. Effective digestion was carried out with pure Tris buffer, pH 8. NaCl in the buffer has an inhibitory effect (Haun and Moos 1992). Thrombin can be removed from the cleaved product by affinity purification on *p*-amino agarose, gel filtration with a superose-12 FPLC column (Yu et al. 1995) or benzamidine sepharose.

A factor X_a recognition site between the tag and a protein of interest can be a useful tool to completely remove N-terminal affinity tags. Factor X_a cleaves at the carboxyl side of the four-amino-acid peptide I-E[D]-G-R-X_1 (Nagai and Thogerson 1984), where X_1 can be any amino acid except arginine and proline. Cleavage can be carried out at temperatures ranging from 4 to 25 °C. The predominant form of factor X_a has a molecular weight of approximately 43 kDa, consisting of two disulfide-linked chains of approximately 27 kDa and 16 kDa. On SDS-PAGE, the reduced chains have apparent molecular weights of 30 kDa and 20 kDa. Cleavage of the tag by a site-specific protease such as factor X_a has sometimes been ineffective, and non-specific digestion has been reported using factor X_a (Ko et al. 1994). The reasons can be insolubility of fusion proteins or the presence of denaturing reagents. Cleavage can also be increased by introducing a polyglycine region of five amino acids (Rodriguez and Carrasco 1995). Dansyl-glu-gly-arg-chloromethyl ketone irreversibly inactivates 95% of factor X_a activity in 1 min at room temperature. Although factor X_a has been less popular because cleavage requires longer incubation time and is less effective, there are several examples of its successful use (Pryor and Leiting 1997).

Conclusion

Affinity tags are important in protein purification. They can be helpful for stabilizing proteins or enhancing their solubility. Affinity chromatography usually results in 90–99% purity. The choice of the purification system depends on the protein itself and the further applications. Sometimes the fused protein cannot be purified because the tag is not surface-exposed. Using denaturing conditions or placing the tag at the other terminus can solve this problem. In many cases, a second affinity tag is used to increase the purity after a second affinity chromatography step (Pryor and Leiting 1997; Schioth et al. 1996); alternatively, one tag can be used for purification and the other for detection (Vaughan et al. 1996; Lu et al. 1997). If two different tags are placed at opposite termini, full-length products will be generated after two affinity chromatography steps (Ostermeier et al. 1995; Sun and Budde 1995). Multi-tagging is also possible, each tag being suitable for a special application. Multi-tagging also allows consecutive purification steps, resulting in high purity. These highly purified proteins allow protein-protein interactions to be measured. Associated proteins can be identified using mass spectroscopy (Honey et al. 2001). A special multi-tag is the tandem affinity purification tag (TAP; Rigaut et al. 1999; Puig et al. 2001). It consists of a protein of interest, a calmodulin-binding peptide, a TEV protease cleavage site, and protein A for immobilization. The TAP tag allows the rapid purification of specific complexes. The applications of the procedure are similar to those of the yeast two-hybrid screen (Fromont-Racine et al. 1997). The Tap-tag is a tool for proteome exploration (Gavin et al. 2002). The method has been tested in yeast but should be applicable to other cells or organisms. Many tags with high affinity to their binding partner are also useful tools to immobilize peptides or proteins on surfaces. Immobilization of biologically active proteins is important for research and industry. Furthermore, the importance of affinity-tag technology will increase for use in peptide/protein chip design, high-throughput purification, peptide/protein libraries, large-scale production systems, and drug delivery strategies.

Acknowledgements The author thanks Prof. A. Steinbüchel to support this review and Prof. F. Mayer for his advice during preparation of the manuscript.

References

Beekman JM, Cooney AJ, Elliston JF, Tsai SY, Tsai MJ (1994) A rapid one-step method to purify baculovirus-expressed human estrogen receptor to be used in the analysis of the oxytocin promoter. Gene 146:285–289

Blanar MA, Rutter WJ (1992) Interaction cloning: identification of a helix-loop-helix zipper protein that interacts with c-*fos*. Science 256:1014–1018

Blumenthal DK, Takio K, Edelman A, Charbonneau H, Titani K, Walsh KA, Krebs EG (1985) Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase. Proc Natl Acad Sci USA 82:3187–3191

Borsing L, Berger EG, Malissard M (1997) Expression and purification of His-tagged beta-1,4 galactosyltransferase in yeast and in COS cells. Biochem Biophys Res Commun 240:586–590

Bucher MH, Evdokimov AG, Waugh DS (2002) Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein. Biol Cryst 58:392–397

Cantor EJ, Chong S (2001) Intein-mediated rapid purification of Cre recombinase. Protein Expr Purif 22:135–140

Carrington JC, Dougherty WG (1988) A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proc Natl Acad Sci USA 85:3391–3395

Chaga G, Hopp J, Nelson P (1999a) Immobilized metal ion affinity chromatography on Co^{2+}-carboxymethylaspartate-agarose superflow, as demonstrated by one-step purification of lactate dehydrogenase from chicken breast. Biotechnol Appl Biochem 29:19–24

Chaga G, Bochkariov DE, Jokhadze GG, Hopp J, Nelson P (1999b) Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose. J Chromatogr A 864:257–256

Chang JY (1985) Thrombin specificity. Requirement for apolar amino acids adjacent to the thrombin cleavage site of polypeptide substrate. Eur J Biochem 151:217–224

Chang JY, Alkan SS, Hilschmann N, Braun DG (1985) Thrombin specificity. Selective cleavage of antibody light chains at the joints of variable with joining regions and joining with constant regions. Eur J Biochem 151:225–230

Chen BP, Hai T (1994) Expression vectors for affinity purification and radiolabeling of proteins using *Escherichia coli* as host. Gene 139:73–75

Choi SI, Song HW, Moon JW, Seong BL (2001) Recombinant enterokinase light chain with affinity tag: expression from *Saccharomyces cerevisiae* and its utilities in fusion protein technology. Biotechnol Bioeng 75:718–724

Chong S, Shao Y, Paulus H, Benner J, Perler FB, Xu M-Q (1996) Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. J Biol Chem 271:22159–22168

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 M-Q (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 192:271–281

Connelly PR, Varadarajan R, Sturtevant JM, Richards FM (1990) Thermodynamics of protein-peptide interactions in the ribonuclease S system studied by titration calorimetry. Biochemistry 29:6108–6114

Cronan JE Jr (1990) Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins. J Biol Chem 265:10327–10333

Davis GD, Elisee C, Newham DM, Harrison RG (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. Biotechnol Bioeng 65:382–388

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

Dougherty WG, Carrington JC, Cary SM, Parks TD (1988) Biochemical and mutational analysis of a plant virus polyprotein cleavage site. EMBO J 7:1281–1287

Dougherty WG, Parks TD, Cary SM, Bazan JF, Fletterick RJ (1989) Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. Virology 172:302–310

Dreher ML, Gherardi E, Skerra A, Milstein C (1991) Colony assays for antibody fragments expressed in bacteria. J Immunol Methods 139:197–205

Drucker M, German-Retana S, Esperandieu P, LeGall O, Blanc S (2002) Purification of viral protein from infected plant tissues using the Strep-tag. Botechnol Intl June:16–18

Duplay P, Bedouelle H, Fowler A, Zabin I, Saurin W, Hofnung M (1988) Sequence of male gene and of its product, the maltose-binding protein of *Escherichia coli* K12. J Biol Chem 259:10606–10613

Einhauer A, Jungbauer A (2000) Kinetics and thermodynamical properties of the monoclonal antibody M1 directed against the FLAG peptide. 20th International symposium on the separation of proteins, peptides, and polynucleotides (ISPPP). Lublijana, Slovenia, November 5–8, 2000

Einhauer A, Schuster M, Wasserbauer E, Jungbauer A (2002) Expression and purification of homogenous proteins in *Saccharomyces cerevisiae* based on ubiquitin-FLAG fusion. Protein Expr Purif 24:497–504

Ernst WJ, Spenger A, Toellner L, Katinger H, Grabherr RM (2000) Expanding baculovirus surface display. Eur J Biochem 267:4033–4039

Evan GI, Lewis GK, Ramsay G, Bishop JM (1985) Isolation of monoclonal antibodies specific for human c-*myc* proto-oncogene product. Mol Cell Biol 5:3610–3616

Ferrando A, Koncz-Lalman Z, Farras R, Tiburcio A, Schell J, Koncz C (2001) Detection of *in vivo* protein interactions between Snf1-related kinase subunits with intron-tagged epitope-labeling in plant cells. Nucleic Acids Res 29:3685–3693

Fontaine L, Meynial-Salles I, Girbal L, Yang X, Croux C, Soucaille P (2002) Molecular characterization and transcriptional analysis of adhE2, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. J Bacteriol 184:821–830

Frangioni JV, Neel BG (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal Biochem 210:179–187

Fromont-Racine M, Rain J-C, Legrain R, Legrain P (1997) Toward a functional analysis of the yeast genome exhausted two-hybrid screens. Nat Genet 16:277–282

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415:141–147

Goward CR, Murphy JP, Atkinson T, Barstow DA (1990) Expression and purification of a truncated recombinant streptococcal protein G. Biochem J 267:171–177

Groß R, Pisa R, Simon J, Kröger A (2002) Isolierung der trimeren Hydrogenase aus *Wolinella succinogenes* durch StrepTactin-Affinitätschromatographie. Biospektrum 1:101

Guan K, Dixon JE (1991) Eukaryotic proteins expressed in *E. coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. Anal Biochem 192:262–267

Gusarov I, Nudler E (2001) Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell 107:437–449

Hakansson K, Broder D, Wang AH, Miller CG (2000) Crystallization of peptidase T from *Salmonella typhimurium*. Acta Crystallogr D Biol Crystallogr 56:924–926

Halliwell CM, Morgan G, Ou CP, Cass AE (2001) Introduction of a (poly)histidine tag in L-lactate dehydrogenase produces a mixture of active and inactive molecules. Anal Biochem 295:257–261

Hamilton SR, O'Donnell JB Jr, Hammet A, Stapleton D, Habinowski SA, Means AR, Kemp BE, Witters LA (2002) AMP-activated protein kinase: detection with recombinant AMPK alpha1 subunit. Biochem Biophys Res Commun 293:892–898

Hans M, Buckel W (2000) Purification of recombinant component A of 2-hydroxylglutaryl-CoA dehydratase from *Acidaminococcus fermentans* using Strep-Tactin affinity-chromatography. Biotechnol Intl September: 12

Harrison RG (2000) Expression of soluble heterologous proteins via fusion with NusA protein. inNovation 11:4–7

Haun RS, Moss J (1992) Ligation-independent cloning of glutathione fusion genes for expression in *Escherichia coli*. Gene 112:37–43

Head JF (1992) A better grip on calmodulin. Curr Biol 2:609–611

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

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

Hochuli E, Bannwarth W, Döbeli H, Gentz R, Stüber D (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology 6:1321–1325

Honey S, Schneider BL, Schieltz DM, Yates JR, Futcher B (2001) A novel multiple affinity purification tag and its use in identification of proteins associated a cyclin-CDK complex. Nucleic Acids Res 29:e24 (http://nar.oupjournals.org/cgi/content/full/29/4/e24)

Hope TP, Gallis B, Prikett KS (1996) Metal-binding properties of a calcium dependent monoclonal antibody. Mol Immunol 33:601–608

Hopp TP, Pricket KS, Price VL, Libby RT, March CJ, Ceretti DP, Urdal DL, Conlon PJ (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology 6:1204–1210

Hosfield T, Lu Q (1999) Influence of the amino acid residue downstream of (Asp)4Lys on enterokinase cleavage of a fusion protein. Anal Biochem 269:10–16

Janknecht R, Nordheim A (1992) Affinity purification of histidine-tagged proteins transiently produced in HeLa cells. Gene 121:321–324

Janknecht R, de Martynoff G, Lou J, Hipskind R, Nordheim A, Stunnenberg HG (1991) Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccina virus. Proc Natl Acad Sci USA 88:8972–8976

Juda GA, Bollinger JA, Dooley DM (2001) Construction, overexpression, and purification of *Arthrobacter globiliformis* amine oxidase-*Strep*-Tag II fusion protein. Protein Expr Purif 22:455–461

Karpeisky MY, Senchenko VN, Dianova MV, Kanevsky V (1994) Formation and properties of S-protein complex with S-peptide-containing fusion protein. FEBS Lett 339:209–212

Kaslow DC, Shiloach J (1994) Production, purification and immunogenicity of a malarian transmission-blocking vaccine candidate: TBV25H expressed in yeast and purified using nickel-NTA agarose. Bio/Technology 12:494–499

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

Keefe AD, Wilson DS, Seelig B, Szostak JW (2001) One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. Protein Expr Purif 23:440–446

Kelemen BR, Klink TA, Behlke MA, Eubanks SR, Leland PA, Raines RT (1999) Hypersensitive substrate for ribonucleases. Nucleic Acids Res 27:3696–3701

Kim JS, Raines RT (1994) A misfolded but active dimer of bovine seminal ribonuclease. Eur J Biochem 224:109–114

Kipriyanov SM, Kupriyanova OA, Little M, Moldenhauer G (1996) Rapid detection of recombinant antibody fragments directed against cell-surface antigens by flow cytometry. J Immunol Methods 196:51–62

Ko YH, Thomas PJ, Delannoy MR, Pedersen PL (1993) The cystic fibrosis transmembrane conductance regulator. Overexpression, purification, and characterization of wild type and delta F508 mutant forms of the first nucleotide binding fold in fusion with the maltose-binding protein. J Biol Chem 268:24330–24338

Korndörfer IP, Skerra A (2001) 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 11:883–893

Kunz D, Gerad NP, Gerad C (1992) The human leukocyte platelet-activating factor receptor. J Biol Chem 267:9101–9106

Kuusinen A, Arvola M, Oker-Blom C, Keinämen K (1995) Purification of recombinant GluR-D glutamate receptor produced in Sf21 insect cells. Eur J Biochem 233:720–726

Lassar AB, Buskin JN, Lockshon D, Davis RL, Apone S, Hauschka SD, Weintraub H (1989) MyoD is a sequence-specific DNA binding protein requiring a region of *myc* homology to bind to the muscle creatine kinase enhancer. Cell 58:823–831

LaVallie ER, Lu Z, Diblasimo-Smith EA, Collins-Racie LA, McCoy JM (2000) Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*. Methods Enzymol 326:322–340

Lellouch AC, Geremia RA (1999) Expression and study of recombinant ExoM, a beta1–4 glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*. J Bacteriol 181:1141–1148

Lu Q, Bauer JC, Greener A (1997) Using *Schizosaccharomyces pombe* as host for expression and purification of eukaryotic proteins. Gene 200:135–144

Manstein DJ, Schuster H-P, Morandini P, Hunt DM (1995) Cloning vectors for the production of proteins in *Dictyostelium discoideum*. Gene 162:129–134

Maroux S, Baratti J, Desnuelle P (1971) Purification and specificity of procine enterokinase. J Biol Chem 246:5031–5039

Mayer BJ, Jackson PK, Baltimore D (1991) The noncatalytic *src* homology 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. Proc Natl Acad Sci USA 88:627–631

McCormick M, Berg J (1997) Purification and S·Tag detection of CBD fusion proteins. inNovations 7:12–15

McKern NM, Lou M, Frenkel MJ, Verkuylen A, Bentley JD, Lovrecz GO, Ivancic N, Elleman TC, Garrett TP, Cosgrove LJ, Ward CW (1997) Crystallization of the first three domain of the human insuline-like growth factor-1 receptor. Protein Sci 6:2663–2666

McTigue MA, Williams DR, Tainer JA (1995) Chrystal 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

Moorby CD, Gherardi E (1999) Expression of a Cx43 deletion mutant in 3T3 A31 fibroblasts prevents PDGF-induced inhibition of cell communication and suppress cell growth. Exp Cell Res 249:367–376

Munro S, Pelham HRB (1986) An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell 46:291–300

Murphy JT, Lagarias JC (1997) Purification and characterization of recombinant affinity peptide-tagged oat phytochrom A. Photochem Photobiol 65:750–758

Nagai K, Thogersen HC (1984) Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature 309:810–812

Nagai K, Thogersen HC (1987) Synthesis and sequence specific proteolysis of hybrid proteins produced in *Escherichia coli*. Methods Enzymol 153:461–481

Nilsson B, Abrahmsén L, Uhlén M. (1985) Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors. EMBO J 4:1075–1080

Nock S, Spudich JA, Wagner P (1997) Reversible, site-specific immobilization of polyarginine-tagged fusion proteins on mica surfaces. FEBS Lett 414:233–238

Ostermeier C, Essen L-O, Michel H (1995) Crystals of an antibody $F_v$fragment against an integral membrane protein diffracting to 1.28 Å resolution. Proteins 21:74–77

Ostermeier C, Harrenga A, Ermler U, Michel H (1997) Structure at 2.7 Å resolution of *Paracoccus denitrificans* two-subunit cytochrom c oxidase complexed with an antibody $F_v$fragment. Proc Natl Acad Sci USA 94:10547–10553

Pänke O, Karin G, Junge W, Engelbrecht S (2000) F-ATPase: specific observation of the rotating c subunit oligomer of EF_0EF_1. FEBS Lett 472:34–38

Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG (1994) Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. Anal Biochem 216:413–417

Podmore AH, Reynolds PE (2002) Purification and characterization of VanXYC, a d,d-dipeptidase/d,d-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4174. Eur J Biochem 269:2740–2746

Porath J, Carlsson J, Olsson I, Belfrage G (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. Nature 258:598–599

Pryor KD, Leiting B (1997) High-level expression of soluble protein in *Escherichia coli* using His_6-tag and maltose-binding-protein double-affinity system. Prot Expr Purif 10:309–319

Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods 24:218–229

Rank KB, Mildner AM, Leone JW, Koeplinger KA, Chou KC, Tomasselli AG, Heinrikson RL, Sharma SK (2001). [W206R]-procaspase 3: an inactivateable substrate for caspase 8. Protein Expr Purif 22:258–266

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol 17:1030–1032

Rodriguez PL, Carrasco L (1995) Improved factor Xa cleavage of fusion proteins containing maltose binding protein. Biotechniques 238:241–243

Ron D, Dressler H (1992) pGSTag – a versatile bacterial expression Plasmid for enzymatic labelling of recombinant proteins. Biotechniques 13:866–869

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

Sachdev D, Chirgwin JM (1999) Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. J Protein Chem 18:127–136

Sachdev D, Chirgwin JM (2000) Fusions to maltose-binding protein: control of folding and solubility in protein purification. Methods Enzymol 326:312–321

Sárdy M, Kárpáti S, Merkl B, Paulsson M, Smyth N (2002) Epidermal transglutaminase (TGase 3) is the autoantigen of dermatitis herpetiformis. J Exp Med 195:747–757

Sassenfeld HM, Brewer SJ (1984) A polypeptide fusion designed for purification of recombinant proteins. Bio/Technology 2:76–81

Schatz PJ (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. Bio/Technology 11:1138–1143

Schioth HB, Kuusinen A, Muceniece R, Szardenings M, Keinanen K, Wikberg JE (1996) Expression of functional melanocorcin 1 receptor in insect cells. Bichem Biophys Res Commun 221:807–814

Schmidt M, Tuominen N, Johansson T, Weiss SA, Keinamen K, Oker-Blom C (1998) Baculovirus-mediated large-scale expression and purification of a polyhistidine-tagged *Rubrella* virus capsid protein. Protein Expr Purif 12:323–330

Schmidt TGM, Skerra A (1993) 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 6:109–122

Schmidt TGM, Koepke J, Frank R, Skerra A (1996) Molecular interaction between the *Strep*-tag affinity peptide and its cognate target, Streptavidin. J Mol Biol 255:753–766

Schuster M, Wasserbauer E, Einhauer A, Ordner C, Jungbauer A, Hammerschmidt F, Werner G (2000) Protein expression strategies for identification of novel target proteins. J Biomol Screen 5:89–97

Sequi-Real B, Martinez M, Sandoval IV (1995) Yeast aminopeptidase I is post-translationally sorted from the cytosol to the vacuole by a mechanism mediated by its bipartite N-terminal extension. EMBO J 14:5476–5484

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

Smyth N, Odenthal U, Merkl B, Paulsson M (2000) Eucaryotic expression and purification of recombinant extracellular matrix proteins carrying the Strep II tag. In: Streuli CH, Grant ME (eds) Extracellular matrix protocols. Humana, Totowa, New Jersey, pp 49–57

Stofko-Hahn RE, Carr DW, Scott JD (1992) A single step purification for recombinant proteins. FEBS Lett 302:274–278

Su XA, Prestwood AK, McGraw RA (1992) Production of recombinant porcine tumor necrosis factor alpha in a novel *E. coli* expression system. BioTechniques 13:756–762

Sun G, Budde RJA (1995) A modified pGEX expression system that eliminates degradation products and thrombin from the recombinant protein. Anal Biochem 231:458–460

Sweda P, Pladzyk R, Kotlowski R, Kur J (2001) Cloning, expression, and purification of the *Staphylococcus simulans* lysostaphin using the intein-chitin-binding domain (CBD) system. Protein Expr Purif 22:467–471

Taylor MG, Bushara H, Capron A, Xu S, Butterworth A (1994) Laboratory and field evaluation of defined antigen vaccines against *Schistosoma bovis* and *S. japonicum* in animals and of defined antigens for the immunodiagnosis of human *S. japonicum* infection. In: Jepsen S, Hagan P, Klein R, Taylor D (eds). Science and technology for development-health-second programme (1987–1991). Summaries of the final reports of the research contracts. Parasitology European Commission DGXII, A:189–200

Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, Gilkes NR, Haynes CA, Warren RA, Kilburn DG (1998) Characterization and affinity applications of cellulose-binding domains. J Chromatogr B 715:283–296

Tucker J, Grisshammer R (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. Biochem J 317:891–899

Uhlén M, Nilsson B, Guss B, Lindberg M, Gatenbeck S, Philipson L (1983) Gene fusion vectors based on the gene for staphylococcal protein A. Gene 23:369–378

Vaillancourt P, Zheng C-F, Hoang DQ, Breister L (2000) Affinity purification of recombinant proteins fused to calmodulin or calmodulin-binding peptides. Methods Enzymol 326:340–362

Vaughan, TJ Williams AJ, Pritchard K, Osbourn JK, Pope AR, Earnshaw JC, McCafferty J, Hodits RA, Wilton J, Johnson KS (1996) Human antibodies with sub-nanomolar affinities isolated from large non-immunized phage display library. Nat Biotechnol 14:309–314

Voss S, Skerra A (1997) 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 10:975–982

Watanabe T, Ito Y, Yamada T, Hashimoto M, Sekine S, Tanaka H (1994) The role of the C-terminal domain and type III domains of chitinase A1 from *bacillus circulans* WL-12 in chitin degradation. J Bacteriol 176:4465–4472

Weiss HM, Haase W, Michel H, Reihlander H (1998) Comparative biochemical and pharmacological characterization of mouse 5HT5A 5-hydroxytryptamine receptor and the human beta2-adrenergic receptor produced in the methylotrophic yeast *Pichia pastoris*. Biochem J 330:1137–1147

Wiese A, Wilms B, Syldatk C, Mattes R, Altenbuchner J (2001) Cloning, nucleotide sequences and expression of a hydantoinase and carbamoylase gene from *Arthrobacter aurescens* DSM 3745 in *Escherichia coli* and comparison with corresponding genes from *Arthrobacter aurescens* DSM 3747. Appl Microbiol Biotechnol 55:750–757

Wilson DS, Keefe AD, Szostak JW (2001) The use of mRNA display to select high-affinity protein-binding peptides. PNAS 98:3750–3755

Wu J, Filutowicz M (1999) Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale. Acta Biochim Pol 46:591–599

Xu M-Q, Paulus H, Chong S (2000) Fusions to self-splicing inteins for protein purification. Methods Enzymol 236:376–418

Xu Z, Bae W, Mulchandani A, Mehra RK, Chen W (2002) Heavy metal removal by novel CBD-EC20 sorbents immobilized on cellulose. Biomacromolecules 3:462–465

Yu L, Deng K-P, Yu C-A (1995) Cloning, gene sequencing, and expression of the small molecular mass ubiquinone-binding protein of mitochondrial ubiquinol-cytochrom c reductase. J Biol Chem 270:25634–25638

Zhang XK, Wills KN, Hunsmann M, Hermann T, Pfahl M (1991) Novel pathway for thyroid hormone receptor action through interaction with *jun* and *fos* oncogene activities. Mol Biol 11:6016–6025

Zheng C-F, Simcox T, Xu L, Vaillancourt P (1997) A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins. Gene 186:55–60