Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 41 (2005) 98–105
Protein Expression & Purification
www.elsevier.com/locate/yprep

Comparison of affinity tags for protein purification

Jordan J. Lichty [a,b], Joshua L. Malecki [a], Heather D. Agnew [a,b,1], Daniel J. Michelson-Horowitz [a,b], Song Tan [a,*]

[a] *Center for Gene Regulation, Department of Biochemistry and Molecular Biology, 108 Althouse Laboratory, The Pennsylvania State University, University Park, PA 16802-1014, USA*
[b] *Schreyer Honors College, The Pennsylvania State University, University Park, PA 16802-3905, USA*

Received 27 October 2004, and in revised form 19 January 2005

Abstract

Affinity tags are highly efficient tools for purifying proteins from crude extracts. To facilitate the selection of affinity tags for purification projects, we have compared the efficiency of eight elutable affinity tags to purify proteins from *Escherichia coli*, yeast, *Drosophila*, and HeLa extracts. Our results show that the HIS, CBP, CYD (covalent yet dissociable NorpD peptide), Strep II, FLAG, HPC (heavy chain of protein C) peptide tags, and the GST and MBP protein fusion tag systems differ substantially in purity, yield, and cost. We find that the HIS tag provides good yields of tagged protein from inexpensive, high capacity resins but with only moderate purity from *E. coli* extracts and relatively poor purification from yeast, *Drosophila*, and HeLa extracts. The CBP tag produced moderate purity protein from *E. coli*, yeast, and *Drosophila* extracts, but better purity from HeLa extracts. Epitope-based tags such as FLAG and HPC produced the highest purity protein for all extracts but require expensive, low capacity resin. Our results suggest that the Strep II tag may provide an acceptable compromise of excellent purification with good yields at a moderate cost.

Keywords: Affinity tags; Affinity purification; Immunopurification; Dihydrofolate reductase; Proteomics

Protein and peptide affinity tags have become highly popular tools for purifying recombinant proteins and native protein complexes for excellent reasons [1]. They can provide hundred- or even thousand-fold purification from crude extracts without prior steps to remove nucleic acid or other cellular material. Second, the mild elution conditions employed make affinity tags useful for purifying individual proteins and especially protein complexes. Importantly, affinity tags allow diverse proteins to be purified using generalized protocols in contrast to highly customized procedures associated with conventional chromatography, a compelling consideration for proteomics or structural genomics ventures.

Most of the available protein and peptide affinity tags were developed within the last 20 years and can be categorized into three classes depending on the nature of the affinity tag and its target. The first class of epitope affinity tags uses peptide or protein fusion to bind to small molecule ligands linked to a solid support. For example, the hexahistidine tag binds to immobilized metal [2] while glutathione *S*-transferase protein fusions bind to glutathione attached to chromatography resin [3]. In the second class of affinity tags, a peptide tag binds to a protein-binding partner immobilized on chromatography resin. For example, the calmodulin-binding peptide binds specifically to calmodulin allowing proteins fused to the peptide to be purified over calmodulin resin [4]. The third class of epitope affinity tags can be considered a subset of the second class where the protein-binding partner attached to the resin is an antibody which recognizes a specific peptide epitope. Examples include the

* Corresponding author. Fax: +1 814 863 7024.
E-mail address: sxt30@psu.edu (S. Tan).
[1] Present address: Department of Chemistry, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA.

1046-5928/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2005.01.019

FLAG peptide, which can be used with one of several anti-FLAG antibody resins [5].

This ample choice of affinity tags for protein purification can make it difficult to decide on the most appropriate tag for a particular project. Although many, if not most, recombinant proteins are now expressed with and purified using affinity tags, the few studies comparing affinity tags for protein purification have only compared two or three tags [6,7]. As a result, affinity tags are typically chosen based on anecdotal information. To address this lack of a systematic comparison of affinity tags, we have assessed the purity, yield, and cost of using two protein fusions and six short peptide affinity tags to purify two proteins expressed in *Escherichia coli*. We have also assessed the ability of the affinity resins to purify tagged proteins from yeast, *Drosophila*, and HeLa extracts. Our results, which show the affinity tags differ substantially in effectiveness of purification, provide a more rigorous basis for selecting affinity tags.

Methods

Purification of tagged DHFR and Gcn5 from E. coli extracts

The coding region of the DHFR gene from pET22b-DHFR (gift of Craig Cameron) was subcloned into pST39 plasmids [8] with the appropriate dual affinity tags. Similarly, the coding region of yeast Gcn5(99–439) was subcloned into dual affinity tagged versions of pST50Trc1, a T7-based *E. coli* expression vector [9]. 0.05–5 ml of soluble extracts of tagged proteins expressed from pST39 or pST50Trc1 plasmids in BL21(DE3)pLysS cells was incubated with 0.1–0.5 ml of appropriate resins, then washed twice with 20 column volumes of incubation buffer, and then eluted according to manufacturer's recommendations. Control experiments using different amounts of each resin were performed to ensure that the capacity of each resin was not exceeded. Buffers used were: *HIS tag incubation and washing*: 50 mM sodium phosphate, pH 7.0, 100 mM NaCl, 1 mM benzamidine, 5 mM of 2-mercaptoethanol, and 5 mM imidazole; *elution*: same buffer with 100 mM imidazole. *CBP tag incubation and washing*: 50 mM Tris–Cl, pH 8.0, 150 mM NaCl, 1 mM benzamidine, 10 mM of 2-mercaptoethanol, 1 mM imidazole, 1 mM $Mg(Ac)_2$, and 2 mM $CaCl_2$; *elution*: 50 mM Tris–Cl, pH 8.0, 150 mM NaCl, 2 mM EGTA, and 10 mM of 2-mercaptoethanol. *CYD tag incubation and washing*: 20 mM Tris–Cl, pH 8.0, 100 mM NaCl, and 1 mM benzamidine; *elution*: same buffer with 100 mM DTT. *Strep II tag incubation and washing*: 50 mM Tris–Cl, pH 7.0, 150 mM NaCl, 1 mM benzamidine, and 10 mM of 2-mercaptoethanol; *elution*: same buffer with 2.5 mM desthiobiotin. *FLAG tag incubation and washing*: 100 mM Tris–Cl, pH 8.0, 150 mM NaCl, and 1 mM benzamidine; *elution*: same buffer with 100 μg/ml 3× FLAG peptide. *HPC tag incubation and washing*: 20 mM Tris–Cl, pH 7.5, 100 mM NaCl, 1 mM benzamidine, and 2 mM $CaCl_2$; *elution*: 20 mM Tris–Cl, pH 7.5, 100 mM NaCl, 1 mM benzamidine, and 10 mM EDTA. *GST tag incubation and washing*: 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 150 mM NaCl, and 1 mM benzamidine; *elution*: 50 mM Tris–Cl, pH 7.0, 150 mM NaCl, 10 mM reduced glutathione, and 1 mM benzamidine.

Elutions were performed in gravity feed disposable columns (Bio-Spin columns, Bio-Rad, catalog #732-6008). Resins used were Talon Superflow metal affinity (Clontech, catalog #8908-02), calmodulin affinity (Stratagene, catalog #214303), Strep-Tactin Superflow (IBA, catalog #2-1206-010), FLAG M2 monoclonal antibody matrix (Sigma, catalog #A2220), anti-protein C monoclonal antibody matrix (gift from Charles Esmon, University of Oklahoma; also available from Roche, catalog #1-815-024), and glutathione–Sepharose 4B (GE Healthcare, catalog #17-0756-01). InaD protein was expressed and purified, and the InaD resin was prepared essentially as described [10]. The extracts were incubated with the resin for 30 min (Talon, Strep-Tactin, glutathione–Sepharose), 1 h (anti-FLAG, anti-protein C, CYD) or 2 h (calmodulin).

Purification of tagged DHFR from yeast, Drosophila, and HeLa extracts

CBP-HIS-DHFR, STR-HIS-DHFR, and FLAG-HIS-DHFR were expressed from pST39 plasmids in BL21(DE3)pLysS cells. CBP-HIS-DFHR and FLAG-HIS-DHFR were purified over Talon resin followed by cation-exchange Source S (GE Healthcare, catalog #17-0944-01) HPLC. STR-HIS-DHFR was purified using Talon and Strep-Tactin resins in succession. Whole cell BY4742 yeast and Oregon R *Drosophila* extracts (generous gifts of Decha Sermwittayawong and David Gilmour, respectively) were dialyzed against incubation buffers appropriate for each resin. HeLa extracts were prepared from HeLa-S9 cells (National Cell Culture Center) by lysing the cells in 20 mM Tris–Cl, pH 7.5, 150 mM NaCl, 1% Nonidet P40, 5 mM EDTA, and dialyzing the cleared extract against incubation buffers appropriate for each resin. Tagged DHFR was then added to each extract to a final concentration of 2 μg/ml before purification over appropriate resin following the procedure described above for *E. coli* extracts.

DHFR assay

The tagged DHFR proteins in crude extracts, flow-through, and eluted fractions were quantified by a colorimetric NADPH assay [11] in three independent purification experiments. Appropriate dilutions of samples

were incubated in reaction buffer (100 mM imidazole–Cl, pH 7.0, 100 μM NADPH, 100 μM DHF, 6.5 mM of 2-mercaptoethanol, and 50 μg/ml BSA) at 30 °C for 5 min before sample absorbance was monitored at 320 nm over 10 min using a Hewlett-Packard 8452A Diode Array Spectrophotometer. We calculate the yield which compares the total DHFR activity in elution fractions to the total activity in the crude extract and the recovery which additionally includes the flowthrough fractions. We estimate 5% or less of the DHFR activity could be accounted for in the wash fractions and less than 10% of DHFR activity was left on the resins used.

ELISA quantitation of Gcn5

The amount of Gcn5 in different samples was determined using an indirect ELISA procedure [12]. Samples were incubated overnight in 96-well Nunc MaxiSorp immunomicrotiter plates at 4 °C before washing with 158 mM Tris, 1.5 M NaCl, 0.5% Tween 20, pH 8.0, and blocking at 4 °C for 3 h with 1% BSA. Plates were then incubated with polyclonal rabbit anti-Gcn5 antibodies prepared against recombinant full-length yeast Gcn5 protein for 1 h, washed, and then incubated with anti-rabbit HRP antibodies for an additional 1 h. After washing, the plates were incubated with color substrate (1 mg/ml *o*-phenylenediamine and 100 mM citrate buffer, pH 4.5) before absorbance was measured at 450 nm using a MaxSpec 190 plate reader.

Results

Experimental design

We examined the macromolecular glutathione *S*-transferase (GST) [3] and maltose-binding protein (MBP) [13] tags as well as the peptide hexahistidine (HIS) [2], calmodulin-binding peptide (CBP) [4], covalent yet dissociable (CYD) [10], Strep II (STR) [14], FLAG [5], and heavy chain of protein C (HPC) [15] tags. We restricted our study to affinity tags that could be eluted under native conditions to ensure that our results would be applicable to both analytical (microgram) and preparative (milligram or greater) scale purifications. Thus, we did not evaluate protein tags such as the protein A IgG-binding domain [16] or the chitin-binding domain [17] which require denaturing conditions or cleavage with specific protease or protein intein to separate the desired target protein from the IgG or chitin resin, respectively.

The tags we investigated employ different modes of interaction with the appropriate affinity resin, and the corresponding different elution methods. GST, MBP, and HIS tags bind to small molecules (immobilized glutathione, amylose, immobilized cobalt or nickel ions) and can be eluted from the resin by competition with the small molecule or an analog (reduced glutathione for GST, maltose for MBP, and imidazole for HIS). The calmodulin-binding peptide (CBP) tag forms an α-helix which is engulfed by the EF-hand motifs of calmodulin but is released by EDTA, which chelates the calcium ions necessary for proper folding of the EF-hand motif [18]. The covalent yet dissociable (CYD) tag is a five-residue C-terminal tag from the NorpA protein that binds to a cleft in an InaD protein primarily through an intermolecular disulfide bond [10]. This covalent disulfide bond can be subsequently broken by reducing agents such as DTT. The eight-residue Strep II tag binds reversibly to the modified biotin-binding pocket of an engineered streptavidin called Strep-Tactin and can be eluted with desthiobiotin, a biotin analog [14]. Both the FLAG and HPC peptides are epitope tags which bind to immobilized monoclonal antibodies. Thus, peptide containing the FLAG sequence elute the bound FLAG tagged protein from the anti-FLAG resin, while EDTA or EGTA elutes HPC tagged proteins from the anti-protein C monoclonal antibody resin by disrupting that calcium-dependent epitope–antibody interaction. The affinity tags used in our study are summarized in Table 1.

Table 1
Comparison of affinity tag technologies

Tag	Size (aa)	Resin	Eluting agent	Source	Capacity	Cost	Cost/10 mg
MBP	396	Amylose	Maltose	Biolabs	3 mg/ml	$105/10 ml	$12
HIS	6	Talon	Imidazole	Clontech	5–14 mg/ml	$220/25 ml	$18
		Ni–NTA	Imidazole	Qiagen	5–10 mg/ml	$257/25 ml	$21
GST	218	GSH–Sepharose	Glutathione	Amersham	10 mg/ml	$396/25 ml	$36
CBP	28	Calmodulin affinity	EGTA	Stratagene	2 mg/ml	$227/10 ml	$114
STR (Strep II)	8	Strep-Tactin–Sepharose	Desthiobiotin	IBA	50–100 nmol/ml	$1100/25 ml	$293
FLAG	8	Anti-FLAG M2 MAb agarose	FLAG peptide	Sigma	0.6 mg/ml	$1568/25 ml	$1045
HPC	12	Anti-Protein C MAb matrix	EDTA	Roche	2–10 nmol/ml	$299/1 ml	$4983
CYD	5	InaD	DTT	N/A	>0.2 mg/ml	N/A	N/A

Capacity for the fusion protein and retail cost information were obtained from manufacturer's catalogs and websites, except for the InaD resin for which the capacity was estimated from our experiments. The unit cost was calculated for purification of 10 mg of a 30 kDa protein fused to the appropriate tag, and the values for the MBP and GST fusions take into account the large size of those affinity tag. Since the affinity resins listed are reusable, actual costs may be less.

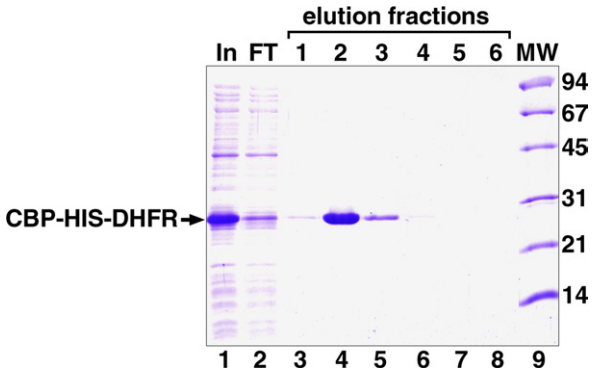

Fig. 1. Sample purification of tagged DHFR. Soluble extract of CBP-HIS-DHFR expressed in *E. coli* (Input in lane 1) was incubated with calmodulin affinity resin, washed with buffer, and the tagged protein eluted with EGTA (lanes 3–8). Some tagged DHFR can be detected in the flowthrough (FT) fraction (lane 2). Samples were separated by SDS–PAGE and stained with Coomassie blue.

We selected *E. coli* dihydrofolate reductase (DHFR) as the target protein to be purified because it is a well-behaved monomeric enzyme that can be assayed using a simple spectrophotometric assay [11]. DHFR was also used in one of the original characterizations of the hexahistidine tag [19]. Since the hexahistidine tag is probably the most popular affinity tag in use, we expressed each tag in tandem with the hexahistidine tag to permit a direct comparison between the tag of interest with the hexahistidine tag. For example, we expressed the fusion CBP-HIS-DHFR protein to assay both the calmodulin-binding peptide (CBP) and hexahistidine (HIS) tags. We assayed the purity of the eluted protein samples by silver staining of SDS–PAGE gels to maximize the detection of minor contaminants, while yields were calculated from DHFR determination by the NADPH assay [11]. We originally planned to quantitate the purity of each sample by normalizing the amount of active DHFR with the total protein content, but we found that experimental errors associated with Bradford and similar total protein assays were greater than the subtle difference in purity between our samples. Since silver does not stain all proteins equivalently, we have elected not to estimate the purity by densitometry of the silver stained gels.

Each of the various tagged DHFR proteins we examined was expressed solubly at similar levels (data not shown). Fig. 1 shows an example of the purification of tagged DHFR from *E. coli* extracts via the CBP tag. The majority of the CBP-HIS-DHFR fusion protein bound to the calmodulin resin in Fig. 1 (compare lanes 1 and 2) and could be eluted by EGTA (Fig. 1, lanes 3–8). The purity of the eluted CBP-HIS-DHFR protein in Fig. 1 appears very high by Coomassie blue staining, but contaminants are clearly visible with silver staining (compare Fig. 1, lane 4 with Fig. 2, lane 4). About 55% of the active DHFR present in the crude extract could be retrieved in the elution fractions, while approximately 20% of the DHFR was found in the flow through fraction (Fig. 1, lane 2 and Table 2). For each of the tagged DHFR proteins we have examined, the flowthrough and elution fractions account for 80–90% of the DHFR activity present in the crude extract, with the remaining 10–20% in washes of the resin before elution, left on the resin after elution or perhaps from loss of enzyme activity during the purification process.

Purification of tagged DHFR from E. coli extracts

We first compared CBP-HIS, STR-HIS, and GST-HIS tagged proteins purified via the HIS tag by metal affinity chromatography. A large number of proteins contaminate DHFR purified over Talon cobalt metal affinity resin (Fig. 2, lanes 1–3). Essentially the same contaminants are found when CBP-HIS-DHFR and STR-HIS-DHFR proteins are purified on Talon resin, suggesting the contaminants do not result from proteins which interact with the CBP or Strep II tags (Fig. 2, lanes 1 and 2). However, when extracts containing the GST-HIS-DHFR protein are purified over the same resin, significantly more contaminants are detected. GST degradation products may account for some of the lower molecular weight contaminants, but larger molecular weight contaminants likely correspond to *E. coli* proteins that interact with GST (Fig. 2, compare lane 3 with lanes 1 and 2). The purity of protein isolated via the HIS tag was moderate at best, but the HIS tag/Talon combination did provide the highest yield of almost 70% for the tag/resins tested (Table 2).

In general, the other affinity tags produced higher purity DHFR than the HIS tag. Our experimental design allows us to directly compare the purification of tagged DHFR over different resins. For example, STR-HIS-DHFR purified via the HIS tag over the Talon resin is clearly not as pure as the same protein purified via the Strep II (STR) tag over the Strep-Tactin resin (Fig. 2, lanes 2 and 6). CBP-HIS-DHFR purified via the CBP tag appears to be equivalent to or slightly purer than that protein purified via the HIS tag (Fig. 2, lanes 1 and 4). The epitope FLAG and HPC tags produced the most pure proteins, with few if any contaminating bands visible besides bands around 67 kDa which we believe to be keratin. Yields of the FLAG and HPC in the elution fractions were acceptable at 50–60% of the activity in the crude extract (Table 2). The Strep II tag produces nearly comparable purity protein to the FLAG or HPC tags with only a few contaminating bands coeluting with the desired tagged protein, and similar yields of approximately 60%. In contrast, the CBP, CYD, and GST tagged DHFR proteins all copurify with significantly more contaminants (Fig. 2, lanes 4, 5, and 9). About 55% of CBP and CYD tagged DHFR but only 40% of GST tagged DHFR was retrieved in the elution fractions (Table 2). We were unable to assay the MBP tag because

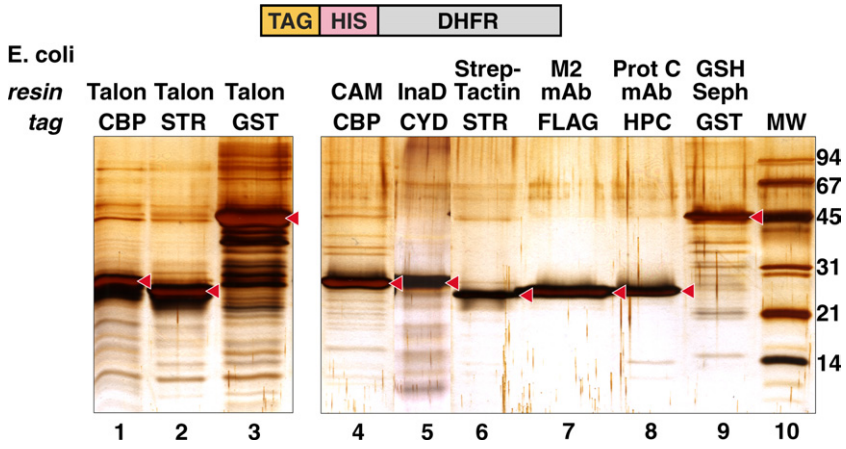

Fig. 2. Comparison of affinity tags to purify tagged DHFR proteins expressed in *E. coli*. As shown in the schematic above the gels, each of the tagged recombinant proteins contain a specific N-terminal tag (identified by TAG label just above each lane), followed by the hexahistidine tag and DHFR. The CYD tagged polypeptide was C-terminally tagged as CBP-HIS-DHFR-CYD. Eluted fractions were fractionated by SDS–PAGE and the gels were silver stained. Each tagged protein is highlighted by a red arrowhead. The identity of the tag used for each purification is shown immediately above each lane and the identity of the affinity resin just above that. Molecular weight markers are shown in lane 10 with sizes provided to the right. Abbreviations used for affinity tags: CBP, calmodulin-binding peptide; STR, Strep II; GST, glutathione *S*-transferase; CYD, covalent yet dissociable NorpA peptide; and HPC, heavy chain of protein C. Abbreviations used for affinity resins: CAM, calmodulin; InaD, PDZ domain of InaD protein; M2 mAb, anti-FLAG M2 monoclonal antibody; Prot C mAb, anti-protein C (clone HPC4) monoclonal antibody; GSH Seph, GSH–Sepharose.

Table 2
Purification yields for active tagged DHFR proteins from *E. coli* extracts determined by NADPH assay (affinity tag used to purify the protein highlighted in bold)

Affinity tag	Resin	Yield (elution fractions) (%)	Recovery (flowthrough and fractions) (%)
CBP-**HIS**	Talon	67.7 ± 4.3	82.4 ± 2.4
CBP-HIS	Calmodulin	55.3 ± 7.6	88.6 ± 5.4
STR-HIS	Strep-Tactin	61.5 ± 6.1	87.6 ± 5.7
FLAG-HIS	M2 MAb	56.1 ± 2.3	84.5 ± 2.5
HPC-HIS	Prot C MAb	62.5 ± 1.2	80.3 ± 3.9
CYD (C-term)	InaD PDZ	57.6 ± 4.0	81.1 ± 4.6
GST-HIS	GSH–Seph	40.8 ± 5.0	77.8 ± 4.6

even the minor fraction of soluble MBP tagged DHFR protein expressed in *E. coli* exhibited very weak DHFR activity (data not shown) despite MBPs previously demonstrated ability to improve the solubility of tagged proteins [20].

Purification of tagged Gcn5 from E. coli extracts

Many recombinant polypeptides expressed in *E. coli* are not as highly expressed or well behaved as DHFR. We therefore used a similar experimental design to assess the ability of four of the peptide tags (HIS, CBP, STR, and FLAG) to purify truncated yeast Gcn5 histone acetyltransferase, a polypeptide we find to be weakly expressed and prone to aggregation. Gcn5 is found in several multicomponent gene regulatory complexes in yeast [21] and the absence of stabilizing partner proteins likely accounts for the propensity of the solitary polypeptide to aggregate. We observed mixed results using the Talon resin to purify tagged Gcn5. While the Talon column was able to partially purify an STR-HIS tagged Gcn5 from crude *E. coli* extracts (Fig. 3, lane 2), a similarly produced CBP-HIS tagged Gcn5 was not purified in appreciable quantities by the Talon resin (Western blotting with anti-Gcn5 antibodies indicate that the 43 kDa band marked by the open arrowhead in Fig. 3, lane 1 is an *E. coli* contaminant that comigrates with the CBP-HIS tagged Gcn5 polypeptide). The calmodulin column was able to partially purify the CBP-HIS tagged Gcn5 polypeptide via its CBP tag but the purity was low (Fig. 3, lane 3). In contrast, the anti-FLAG M2 antibody and Strep-Tactin columns were able to purify the FLAG and Strep II tagged Gcn5 polypeptides, respectively, with higher purity than via the HIS or CBP tags (Fig. 3, lanes 4 and 5). However, the purity was still lower than when compared to tagged DHFR polypeptides. The yield of purification for the tagged Gcn5 polypeptides was much lower than for the tagged DHFR polypeptides (Table 3) with a maximum of 19% (STR-HIS via HIS tag) of the tagged protein found in the elution fractions. These low yields are not the result of overloading the resins because our DHFR purifications established

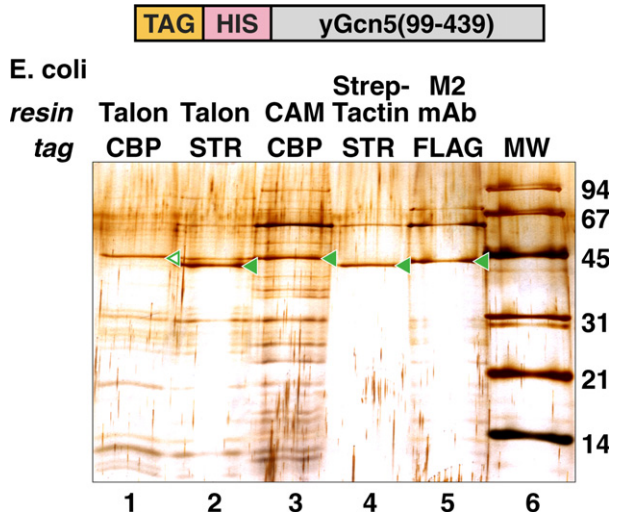

Fig. 3. Comparison of affinity tags to purify tagged Gcn5(99–439) proteins expressed in *E. coli*. Similar comments as for Fig. 2 except that green arrowheads are used. The 43 kDa band marked by the open arrowhead in lane 1 is an *E. coli* contaminant that coincidentally comigrates with the CBP-HIS tagged Gcn5 polypeptide. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

minimum binding capacities for the resins and the crude extracts used for the Gcn5 experiments contained much less Gcn5 than equivalent DHFR extracts (data not shown). Instead, the low yields likely reflect aggregation of the Gcn5 protein since 60–80% of Gcn5 in the soluble extract failed to bind to the resins (difference between yield and recovery in Table 3).

Purification of tagged DHFR from yeast, Drosophila, and HeLa extracts

Since *E. coli* is only one of several popular recombinant expression hosts, we have also analyzed the ability of the HIS, CBP, Strep II, and FLAG tags to purify tagged DHFR from other extracts. *E. coli* recombinant CBP-HIS-DHFR, STR-HIS-DHFR and FLAG-HIS-DHFR were each purified essentially to homogeneity by affinity and/or conventional chromatography, and then spiked at similar concentrations into yeast, *Drosophila*, and HeLa extracts. For the most part, our results mirror what we observed with the affinity purification from *E. coli* extracts: Strep II and FLAG tags produced highest purity DHFR protein for all three extracts, with relatively few proteins from the extracts copurifying (Fig. 4, lanes 4 and 5 in panels A, B, and C). In contrast, significant contaminants copurified using the HIS tag with all extracts (Fig. 4, lane 2 in panels A, B, and C). Interestingly, purification via the CBP tag over the calmodulin resin was significantly better for HeLa extracts than for yeast or *Drosophila* extracts which both produce similar levels of contamination as for *E. coli* extracts (Fig. 4, lanes 3 in panels A, B, and C).

Discussion

Several protein and peptide affinity tags are now available to facilitate the isolation of proteins expressed in a heterologous host like *E. coli* or the purification of native complexes via a tagged subunit. In general, an ideal affinity tag: (i) would allow the efficient purification of tagged proteins in high yield, (ii) can be used with any protein without affecting its function, (iii) can be placed at any position (N-terminal, middle, C-terminal) in the protein, (iv) can be used to purify protein expressed in any host strain or any expression system, (v) can be used to detect the recombinant protein, and (vi) binds to and elutes from a resin that is inexpensive, can be regenerated, and possesses good flow characteristics. Fortunately, there are several commercially available affinity tags that fulfill many, if not all, of these criteria.

We have examined two protein and six peptide affinity tags which match many of the above requirements for their effectiveness to purify recombinant proteins expressed in *E. coli*. We have also analyzed the ability of a subset of these tags to purify proteins from three eukaryotic cell extracts. Our results document the disparate purity that can be obtained. We find that epitope peptide tags such as the FLAG peptide produce the highest purity protein for both well-behaved and ill-behaved polypeptides in *E. coli* extracts, as well as from yeast, *Drosophila*, and HeLa extracts. Tags such as the CBP peptide found in the TAP tag commonly used to isolate "native" protein complexes [22] and the popular HIS tag produced proteins with many more contaminants. The Strep II (STR) tag was surprisingly effective, consistently producing protein almost as pure as epitope-based systems.

We obtained reasonable yields of generally high purity DHFR by affinity tag purification. In contrast, the truncated yeast Gcn5 polypeptide was isolated with relatively poor purity and in low amount or essentially

Table 3
Purification yields for active tagged yGcn5(99–439) from *E. coli* extracts determined by ELISA assays using anti-yGcn5 antibodies

Affinity tag	Resin	Yield (elution fractions) (%)	Recovery (flowthrough and fractions) (%)
CBP-HIS	Talon	3.6 ± 0.6	62.9 ± 10.9
STR-HIS	Talon	18.5 ± 6.4	82.6 ± 11.1
CBP-HIS	Calmodulin	7.2 ± 1.2	87.1 ± 3.8
STR-HIS	Strep-Tactin	7.0 ± 0.8	73.0 ± 1.9

no amount. Thus, even though purification via affinity tags allows the possibility of general purification protocols that do not depend on the nature of the protein being purified, nonideal behavior of a protein such as aggregation can still affect affinity tag purification experiments. It is worth noting that solubility of a recombinant protein is a minimum requirement for functional, properly folded nonmembrane proteins. Soluble extracts can contain small aggregates of the tagged protein which did not pellet out upon centrifugation. Since such nonspecific aggregates may mask the affinity tag, the tagged protein will be soluble and yet unable to bind to the affinity resin, as we observed with the tagged Gcn5 polypeptides.

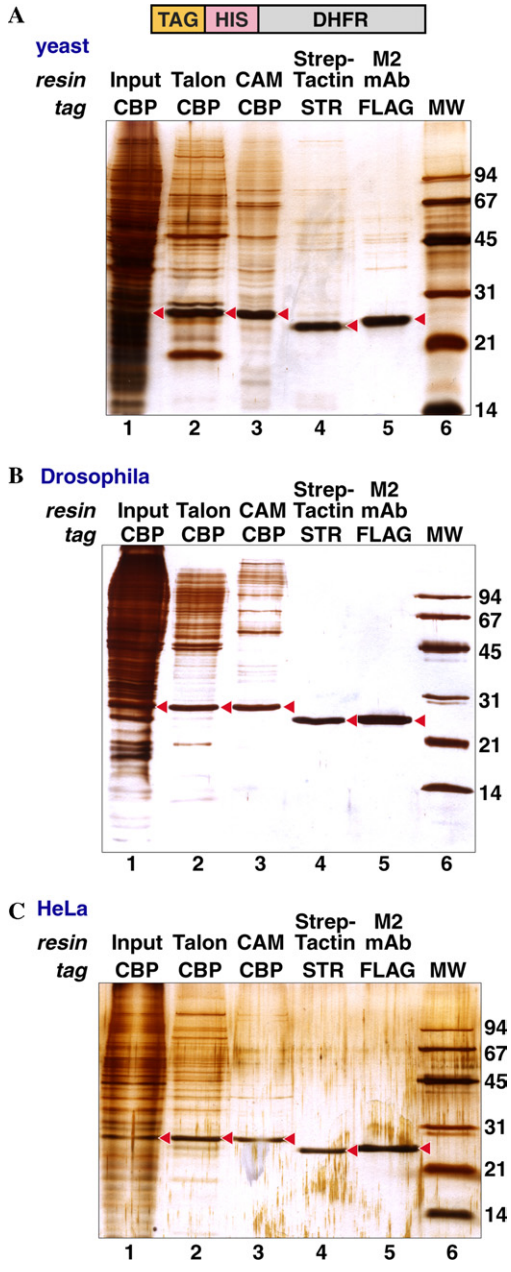

Fig. 4. Comparison of affinity tags to purify tagged DHFR proteins from yeast, *Drosophila*, and HeLa extracts. As shown in the schematic above the gel in (A), each of the tagged recombinant proteins contain a specific N-terminal tag (identified by TAG label just above each lane), followed by the hexahistidine tag and DHFR. The tagged DHFR proteins were expressed in *E. coli* and purified essentially to homogeneity before spiking into yeast, *Drosophila* or HeLa extracts. Eluted fractions were fractionated by SDS–PAGE and the gels were silver stained. Each tagged protein is highlighted by a red arrowhead. Molecular weight markers are shown in lane 10 with sizes provided to the right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

In addition to purity and yield, cost can be a critical parameter when selecting affinity tags and resins, particularly for preparative purifications. To compare the price of different affinity resins, we have calculated the cost to purify 10 mg of a 30 kDa polypeptide using retail cost and resin capacity information provided by the distributors (Table 1). The MBP, Talon, Ni–NTA, and GST columns were the least expensive, with the resins costing $12–36 to purify 10 mg of the tagged polypeptide. The calmodulin (for CBP tag) and Strep-Tactin (for Strep II tag) were noticeably more expensive at $114–293 to purify 10 mg of tagged polypeptide. The monoclonal-based resins for the FLAG and HPC tags are particularly expensive at $1000–5000 for the same amount of tagged polypeptide because of their high unit resin cost and their very limited capacities (0.6 mg FLAG tagged protein/ml resin vs 5–10 mg HIS tagged protein/ml resin).

The choice of an affinity tag clearly depends on the requirements of the particular experiment. Experiments requiring large quantities of partially purified material in high yields at a low cost may find the HIS and GST attractive, while those experiments requiring small quantities of the highest purity may find the benefits of the FLAG and HPC tags to outweigh their costs and limited capacity. The Strep II tag appears to be an excellent candidate for affinity purification in general since it is a short tag that produces high purity material in good yields at a moderate cost. One caveat for using in purifying proteins via the Strep II tag for protein crystallization is that the Strep II tag, but not the HIS or FLAG tags, was shown to interfere with the crystallization of one particular enzyme [23]. Since other proteins have been successfully crystallized with the Strep II tag [24–26], it is not clear if this tag adversely affects protein crystallization in general. Such incongruent results highlight the need for more systematic studies of how affinity tags affect the crystallization of proteins.

For our own preparative needs, we find that a combination HIS and Strep II tag allows rapid capture of the tagged protein or protein complex from crude extracts by the Talon column, followed by polishing over the Strep-Tactin column. To facilitate the use of the tags examined in this study, we have prepared a suite of expression vectors which incorporate the HIS, CBP, STR, FLAG, HPC, and CYD tags as cleavable and

noncleavable single or double tags for expression of individual polypeptides and polycistronic expression of protein complexes in *E. coli* [9].

We hope that our systematic comparison of affinity tags will help others to make informed choices of affinity tags for protein purification. Comparing diverse tags using a common target protein has significant benefits and we would like to recommend that new affinity tags be analyzed using DHFR in the future to facilitate direct comparisons with existing tags. The information provided in this study may also assist in the development of new, smaller tandem affinity purification (TAP) tags as alternatives to the popular protein A–CBP combination [22].

Acknowledgments

We are grateful to Charles Esmon for providing samples of the anti-protein C monoclonal antibody resin, to John Sondek for supplying the pPROEX-HTAInaD expression plasmid and sharing unpublished information, and to Craig Cameron for providing the pET22b-DHFR plasmid. We also thank Decha Sermwittayawong and David Gilmour for the yeast and *Drosophila* extracts. S.T. is a Pew Scholar in the Biomedical Sciences. This work was supported in part by NIH Grant GM060489 to S.T.

References

[1] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 60 (2003) 523–533.
[2] E. Hochuli, H. Dobeli, A. Schacher, New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues, J. Chromatogr. 411 (1987) 177–184.
[3] D.B. Smith, K.S. Johnson, Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-*S*-transferase, Gene 67 (1988) 31–40.
[4] R.E. Stofko-Hahn, D.W. Carr, J.D. Scott, 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. 302 (1992) 274–278.
[5] B.L. Brizzard, R.G. Chubet, D.L. Vizard, Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution, Biotechniques 16 (1994) 730–735.
[6] B.S. Wojczyk, M. Czerwinski, M.M. Stwora-Wojczyk, D.L. Siegel, W.R. Abrams, W.H. Wunner, S.L. Spitalnik, Purification of a secreted form of recombinant rabies virus glycoprotein: comparison of two affinity tags, Protein Expr. Purif. 7 (1996) 183–193.
[7] J. Fang, D. Ewald, Expression cloned cDNA for 10-deacetylbaccatin III-10-*O*-acetyltransferase in *Escherichia coli*: a comparative study of three fusion systems, Protein Expr. Purif. 35 (2004) 17–24.
[8] S. Tan, A modular polycistronic expression system for overexpressing protein complexes in *Escherichia coli*, Protein Expr. Purif. 21 (2001) 224–234.
[9] S. Tan, R.C. Kern, W. Selleck, The pST44 polycistronic expression system for producing protein complexes in *Escherichia coli*, Protein Expr. Purif. 40 (2005) 385–395.
[10] M.E. Kimple, J. Sondek, Affinity tag for protein purification and detection based on the disulfide-linked complex of InaD and NorpA, Biotechniques 33 (2002) 578–588.
[11] S.R. Stone, J.F. Morrison, Kinetic mechanism of the reaction catalyzed by dihydrofolate reductase from *Escherichia coli*, Biochemistry 21 (1982) 3757–3765.
[12] J.R. Crowther, in: J.R. Crowther (Ed.), Methods in Molecular Biology, vol. 42, Humana Press, Totowa, 1995, pp. 131–160.
[13] C. di Guan, P. Li, P.D. Riggs, H. Inouye, Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein, Gene 67 (1988) 21–30.
[14] S. Voss, A. Skerra, 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 (1997) 975–982.
[15] D.J. Stearns, S. Kurosawa, P.J. Sims, N.L. Esmon, C.T. Esmon, The interaction of a Ca^{2+}-dependent monoclonal antibody with the protein C activation peptide region. Evidence for obligatory Ca2+ binding to both antigen and antibody, J. Biol. Chem. 263 (1988) 826–832.
[16] B. Nilsson, L. Abrahmsen, M. Uhlen, Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors, EMBO J. 4 (1985) 1075–1080.
[17] S. Chong, F.B. Mersha, D.G. Comb, et al., Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, Gene 192 (1997) 271–281.
[18] W.E. Meador, A.R. Means, F.A. Quiocho, Target enzyme recognition by calmodulin: 2.4 a structure of a calmodulin-peptide complex, Science 257 (1992) 1251–1255.
[19] E. Hochuli, W. Bannwarth, H. Doebeli, R. Gentz, D. Steuber, Genetic approach to facilitate protein purification of recombinant proteins with a novel metal chelate adsorbent, Biotechnology 6 (1988) 1321–1325.
[20] R.B. Kapust, D.S. Waugh, *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused, Protein Sci. 8 (1999) 1668–1674.
[21] M.J. Carrozza, R.T. Utley, J.L. Workman, J. Cote, The diverse functions of histone acetyltransferase complexes, Trends Genet. 19 (2003) 321–329.
[22] O. Puig, F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, B. Seraphin, The tandem affinity purification (TAP) method: a general procedure of protein complex purification, Methods 24 (2001) 218–229.
[23] M.H. Bucher, A.G. Evdokimov, D.S. Waugh, Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 392–397.
[24] S.A. Roberts, G.F. Wildner, G. Grass, A. Weichsel, A. Ambrus, C. Rensing, W.R. Montfort, A labile regulatory copper ion lies near the T1 copper site in the multicopper oxidase CueO, J. Biol. Chem. 278 (2003) 31958–31963.
[25] I.P. Korndorfer, S. Schlehuber, A. Skerra, Structural mechanism of specific ligand recognition by a lipocalin tailored for the complexation of digoxigenin, J. Mol. Biol. 330 (2003) 385–396.
[26] I.P. Korndorfer, M.K. Dommel, A. Skerra, Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture, Nat. Struct. Mol. Biol. 11 (2004) 1015–1020.