Protein Expression and Purification 68 (2009) 110–120

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

HaloTag7: A genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification

Rachel Friedman Ohana *, Lance P. Encell, Kate Zhao, Dan Simpson, Michael R. Slater
Marjeta Urh, Keith V. Wood

Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711, USA

ARTICLE INFO

Article history:
Received 12 May 2009
and in revised form 16 May 2009
Available online 21 May 2009

Keywords:
HaloTag7
Protein expression
Protein solubility
Protein purification
Fusion tag
Protein labeling
Enhanced expression in *E. coli*
HaloTag

ABSTRACT

Over-expression and purification of soluble and functional proteins remain critical challenges for many aspects of biomolecular research. To address this, we have developed a novel protein tag, HaloTag7, engineered to enhance expression and solubility of recombinant proteins and to provide efficient protein purification coupled with tag removal. HaloTag7 was designed to bind rapidly and covalently with a unique synthetic linker to achieve an essentially irreversible attachment. The synthetic linker may be attached to a variety of entities such as fluorescent dyes and solid supports, permitting labeling of fusion proteins in cell lysates for expression screening, and efficient capture of fusion proteins onto a purification resin. The combination of covalent capture with rapid binding kinetics overcomes the equilibrium-based limitations associated with traditional affinity tags and enables efficient capture even at low expression levels. Following immobilization on the resin, the protein of interest is released by cleavage at an optimized TEV protease recognition site, leaving HaloTag7 bound to the resin and pure protein in solution. Evaluation of HaloTag7 for expression of 23 human proteins in *Escherichia coli* relative to MBP, GST and His_6Tag revealed that 74% of the proteins were produced in soluble form when fused to HaloTag7 compared to 52%, 39% and 22%, respectively, for the other tags. Using a subset of the test panel, more proteins fused to HaloTag7 were successfully purified than with the other tags, and these proteins were of higher yield and purity.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Recombinant DNA technologies have greatly expanded our access to the broad diversity of proteins represented in living organisms, and potentially to an even broader range of mutant proteins not found in nature. Accordingly, the expression and purification of recombinant proteins has become fundamental to many aspects of life science research. Yet, owing to complexities in protein structures and interactions within the host organism, success with these techniques often remains frustratingly elusive. Successful purification of functional proteins generally requires efficient expression of these proteins in soluble form followed by their separation from the highly complex crude lysate of the host. The most frequently used host for protein expression is *Escherichia coli* due to its ease of use, rapid cell growth, low cost of culturing and well documented protocols [1,2]. However, over-expression of heterologous proteins in *E. coli*, particularly human proteins, often yields inadequate levels of soluble protein [1–4].

One approach for overcoming this limitation is to optimize expression conditions such as temperature, growth media, induction parameters, promoters and *E. coli* expression strain [1,5]. Systematic screening of such variables can be simplified by using reporter fusion tags such as GFP[1] [6,7] or S-tag [8]. Another common strategy is to use solubility fusion tags for boosting expression of soluble protein, presumably by promoting proper folding of the fusion partner and suppressing proteolysis [1,9]. A variety of different solubility tags are available, yet not all are equally efficient as solubility enhancers. The most commonly used include GST [10,11], TRX [12], MBP [13,14] and NusA [15,16].

Once adequate expression of soluble protein is achieved, the next step is to purify the target protein from the biological mixture. Affinity tags are widely used to simplify the purification process and to provide a generic method that is straightforward and adaptable to all target proteins. Many affinity tags have been developed,

* Corresponding author. Fax: +1 608 277 2601.
E-mail address: rachel.ohana@promega.com (R.F. Ohana).

[1] *Abbreviations used:* amp, ampicilin; DTT, diothiothreitol; *Escherichia coli*, *E. coli*; EDTA, ethylenediaminetetraacetic acid; GFP, green fluorescent protein; GST, glutathione-S-transferase; kan, kanamycin; LB, Luria broth; MBP, maltose binding protein; NusA; N utilization substance A; MW, molecular weight; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; TRX, thioredoxin; TEV, tobacco etch virus; TMR, tetramethylrhodamine.

1046-5928/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2009.05.010

ranging in size from a few amino acids to entire proteins, that are capable of selective interaction with their corresponding ligands coupled to a chromatography matrix [17]. His_6Tag [18,19] is most widely used due to its small size and ability to frequently provide sufficient protein yield and purity for many applications. Other common affinity tags include GST and MBP, which are often favored for their additional ability to enhance protein solubility [3,4,20].

Although appending a fusion tag onto a protein of interest can improve expression and purification, the tag can also interfere with protein structure or function [13,21,22]. Consequently, it is commonly recommended to remove the tag after purification [18,21,22]. Tag removal can be performed by proteolytic cleavage at a defined sequence in the interconnecting polypeptide, i.e. linker separating the tag and the target protein [18,23–25]. This approach can be problematic due to non-specific or inefficient cleavage or loss of protein stability and solubility following tag removal [25–27]. Furthermore, this step often requires additional effort to separate the free target protein from the affinity tag and the protease.

While a variety of fusion tags is available to facilitate aspects of protein expression, solubility, detection, or purification, most tags are lacking or inefficient in some of these features. Many proteins are poorly expressed with available solubility tags, or are difficult to purify with existing affinity technologies due to low binding onto the purification matrix [25,26,28,29]. These shortcomings are addressed by a new tag, HaloTag7, designed to support efficient expression of soluble protein and bind rapidly and covalently to a unique synthetic ligand. HaloTag7 is a catalytically inactive derivative of DhaA, a bacterial haloalkane dehalogenase from *Rhodococcus* [30] present only among selected microbial groups. This 34 kDa monomeric protein was engineered through rational design and molecular evolution to rapidly form a covalent attachment to synthetic chloroalkane ligands [31,32], and to provide enhanced expression and solubility when fused to a protein partner.

The synthetic ligands comprise a chloroalkane linker attached to a variety of functional groups including fluorophores, affinity handles and solid supports. These features enable both fluorescent labeling of fusion proteins in cell lysate for expression screening and irreversible capture of fusion proteins onto a purification matrix. The rapid, specific, and covalent capture offered by HaloTag7 overcomes the inherent limitation of affinity tags by effectively eliminating protein loss associated with equilibrium-based binding. This feature is especially important for purification of poorly expressed proteins. Following immobilization on the purification matrix, the target protein can be released by cleavage at an optimized TEV protease recognition site contained within the interconnecting polypeptide separating HaloTag7 and the fusion partner. The HaloTag7-based protein purification method yields highly pure proteins in solution while the fusion tag remains covalently attached to the matrix, eliminating contamination by free tag or un-cleaved fusion protein.

To demonstrate the efficacy of HaloTag7 for protein expression and purification with *E. coli*, we compared its performance to the commonly used affinity tags, GST, MBP, and His_6Tag (see Table 1 for tags characteristics). We chose GST and MBP as they are used in a manner similar to HaloTag7; both promote expression of soluble protein in *E. coli* and both provide a means for protein purification. Although His_6Tag does not assist in protein expression or solubilization, this tag was also chosen because of its widespread use for protein purification. The relative performance of these tags was evaluated using a panel of cDNA clones encoding 23 human proteins that are difficult to express in *E. coli* [33]. The set of proteins ranges broadly in both size (~9–155 kDa) and function (e.g. kinases, membrane proteins and transcription factors). Our results showed that HaloTag7 delivered superior performance for protein expression, solubility, purification yield and purity. Furthermore, using two additional model proteins, we found that HaloTag7 produced proteins with higher specific activity.

Materials and methods

Bacterial strain and materials

Single Step *E. coli* KRX ([F′,*tra*D36, Δ*omp*P, *proA*$^+$B$^+$, *lacI*^{q}, Δ(*lacZ*)M15] Δ*omp*T, *end*A1, *rec*A1, *gyr*A96 (Nal^r), *thi*-1, *hsd*R17(r_k-,m_k+), e14$^-$ (McrA$^-$) *rel*A1, *sup*E44, Δ(*lac-pro*AB), Δ*rha*BAD)::T7 RNA polymerase) [34] (Promega, Madison, WI) was used for both cloning and expression. Precession Plus protein MW markers were from BioRad (Hercules, CA). All enzymes and other reagents were from Promega unless otherwise noted.

Expression vectors

Bacterial T7 promoter-based Flexi vectors pFN18K and pFN2K expressing HaloTag7 and GST, respectively, as N-terminal fusions were from Promega. pFN2K was modified by replacing the interconnecting polypeptide (linker) between the GST coding region and the barnase positive selection cassette with the interconnecting polypeptide from pFN18K that contained an optimized TEV recognition sequence [Encell, personal communication]. The MBP Flexi expression vector was constructed by replacing the GST coding region in the modified pFN2K with the MBP coding region from pMALC2 (New England Biolabs, Ipswich, MA). The His_6Tag Flexi expression vector was constructed by inserting the barnase positive selection cassette between the BamHI and MspI sites of pET14b (Novagen-EMD, Madison, WI).

Transfer of protein coding regions between flexi vectors

The protein coding sequences, already available as Flexi vector clones, were transferred into four expression vectors. Donor and acceptor vectors were combined, digested simultaneously with the Flexi enzyme blend (SgfI and PmeI) and ligated. Ligations were used to transform chemically competent KRX cells and transformants containing the acceptor vector were selected on LB plates supplemented with the appropriate antibiotic [35]. Resulting clones were verified by endonuclease restriction digestion (SgfI and PmeI) and sequencing.

Expression in E. coli KRX

Single colonies of KRX cells harboring T7 promoter-based Flexi expression vectors were grown overnight in 2 ml LB supplemented with kan (25 μg/ml) or amp (100 μg/ml) at 37 °C. Overnight cultures were diluted 1:100 into auto-induction media (LB, 0.05% glucose, 0.2% rhamnose and 25 μg/ml kan or 100 μg/ml amp) and grown for 18 h at 25 °C [36]. Fifty milliliters of cultures were used for medium scale protein purifications.

Expression and solubility analysis

Pellets from 1 ml of each expression culture were resuspended in 0.5 ml of 50 mM HEPES (pH 7.5), frozen for 20 min at −70 °C, thawed at room temperature and lysed by addition of 0.5 ml lysis buffer (1× FastBreak Cell Lysis Reagent (Promega), 0.2 mg/ml lysozyme (Sigma, St. Louis, MO) and 20 U RQ1-DNase (Promega)) for 30 min with slow rotation mixing. These crude lysates were used as total expression fractions, and soluble expression fractions were prepared from the crude material by high speed centrifugation (12,000 *g*) at 4 °C for 30 min. Total and soluble fractions were analyzed by 4–20% Tris–Glycine SDS–PAGE (BioRad) and proteins

Table 1
Characteristics of the protein tags used in this study.

Feature	HaloTag7	GST	MBP	His6Tag
Size (kDa)	34	26	40	<1
Uses	Purification Enhanced solubility	Purification Enhanced solubility	Purification Enhanced solubility	Purification
Purification method	Ligand specific covalent immobilization	Affinity	Affinity	Affinity
Matrix	HaloLink resin: chloroalkane ligand attached to agarose beads	Glutathione resin	Cross-linked amylose resin	Immobilized metal chelating chromatography
Reagent for elution of fusion protein	NA	Reduced glutathione	Maltose	Imidazole /low pH
Method for fusion detection	Coomassie blue staining Western blot Fluorescent labeling	Coomassie blue staining Western blot Enzymatic assay	Coomassie blue staining Western blot	Coomassie blue staining Western blot Ni^{2+}–NTA alkaline phosphatase

were detected by SimplyBlue staining (Invitrogen, Carlsbad, CA). Reference markers to estimate expression levels were prepared by adding known amounts of four different purified proteins: ubiquitin (8.5 kDa), HaloTag7 (34 kDa), GST:HaloTag7 (60 kDa) and GST:MBP (67 kDa) into 1 ml of KRX lysate. The reference markers designated for intermediate expression contained GST:HT7 at 50 μg/ml, and the other three proteins at a final concentration of 100 μg/ml each; the markers designated for high expression contained these marker proteins at 4-fold greater concentrations.

Protein purification

For small scale protein purifications, 1 ml of the induced cultures was harvested and the cell pellets were resuspended in 0.5 ml of the appropriate purification buffer (see below) without EDTA, freeze thawed and lysed using the detergent based lysis buffer described above. For medium scale protein purifications, pellets from 50 ml induced cultures were resuspended with 5 ml of the appropriate purification buffer and sonicated using a Misonix 3000 sonicator equipped with a microtip (2 min total time on; 5 s on/5 s off; power output of 3.5). For both purification scales, lysates were centrifuged at 12,000 *g* for 30 min at 4 °C and the supernatants were directly applied onto the appropriate pre-calibrated purification resin at a ratio of 50:1 original culture volume to settled resin.

Binding to the resin was conducted at room temperature for 1 h with constant end-over-end gentle rotation, followed by a wash of the resins with 10× resin volume of the appropriate purification buffer. For proteolytic release of the target protein from HaloTag7, GST and MBP tags, the resins were washed with an additional 10× resin volume of HaloTag7 purification buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT and 0.5 mM EDTA), which is suitable for proteolytic cleavage mediated by ProTEV (HQ-tagged ProTEV from Promega). Proteolytic cleavage was performed in 1× resin volume of the HaloTag7 purification buffer using a ratio of 230 U ProTEV for 1 ml settled resin for 1 h at room temperature.

Target proteins were recovered in a final 2× resin volume of HaloTag7 purification buffer and the HQ-tagged ProTEV was removed through selective binding onto HisLink resin (Promega). His_6Tag fusion proteins were eluted with 500 mM imidazole and subsequently dialyzed against HaloTag7 purification buffer to remove the imidazole. Protein yields were quantitated by Bradford assay (Thermo-Scientific, Pittsburgh, PA).

Resins and buffers for protein purification

Purification resins

HaloLink resin for HaloTag7-based purification and HisLink resin for His_6Tag-based purification were from Promega; Glutathione Sepharose 4 Fast Flow for GST-based purification was from GE Healthcare (Piscataway, NJ); Amylose resin for MBP-based purification was from New England Biolabs.

Purification buffers

Purification buffers used were as recommended by the supplier of the resins. HaloTag7 buffer: 50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT and 0.5 mM EDTA; GST buffer: 1× PBS, pH 7.3, and 1 mM DTT; MBP buffer: 20 mM Tris–HCl, pH 7.4, 200 mM NaCl and 2 mM EDTA; His_6Tag binding buffer: 100 mM HEPES, pH 7.5, 300 mM NaCl and 10 mM imidazole; His_6Tag wash buffer: 100 mM HEPES, pH 7.5, 300 mM NaCl and 20 mM imidazole; His_6Tag elution buffer: 100 mM HEPES, pH 7.5, 300 mM NaCl and 500 mM imidazole.

Determination of protein purity

Purified proteins were resolved by SDS–PAGE, stained with SimplyBlue (Invitrogen) and subsequently scanned using a Typhoon 9400 (GE Healthcare) with the following setting: E_{ex} = 633 nm, no emission filter. Bands were quantitated with ImageQuant (GE Healthcare), and purity (in %) was determined from the ratio of specific protein to total protein.

Gel filtration analysis of purified proteins

Gel filtration chromatography was performed at 24 °C on an Agilent 1200 HPLC, using a Superdex 200 5/150 GL column (GE Healthcare) with a flow rate of 0.25 ml/min and a running buffer of 50 mM Tris–HCl, pH 7.5, and 150 mM NaCl. Void volume was determined with Blue Dextran and a standard calibration curve was generated using: bovine serum albumin (66 kDa, Thermo-Scientific), ovalbumin (43 kDa, Sigma), carbonic anhydrase (29 kDa, Sigma), ribonuclease B (17 kDa, New England Biolabs), myoglobin (17 kDa, Sigma), ribonuclease A (14 kDa, Sigma) and cytochrome C (12 kDa, Sigma) [37]. The molecular weights of the purified proteins were calculated directly from the calibration curve.

Functional assays

PKA activity was determined using Kinase-Glo Plus (Promega) and Kempeptide (Promega), a synthetic peptide substrate [38]. Units of activity for 1 μg of purified proteins were calculated from a calibration curve generated with the purified catalytic subunit of bovine PKA (Promega).

Id activity was measured with respect to its interaction with MyoD protein. HaloTag7:MyoD fusion protein expressed in TNT SP6 High Yield Wheat Germ cell free extract (Promega) was captured on HaloLink slides (Promega) by incubation of 5 μl extract/well for 1 h at room temperature. Slides were washed with 1× PBS containing 0.05% IGEPAL (Sigma) and spin dried at 350 × *g*

for 3 min. The activity for 1 μg of purified Id proteins was determined by addition of the Id proteins to the wells containing HaloTag7:MyoD. Following a 1 h incubation at room temperature, slides were washed and spin dried. The captured Id was detected using sequential incubation with a goat anti-Id IgG (R&D Systems, Minneapolis, MN) and an anti-goat IgG-Alexa647 (Invitrogen/Molecular Probes, Eugene, OR). Slides were washed, spin dried and scanned on a GenePix 4000B (Axon Instruments, Sunnyvale, CA) with a 635 nm laser; captured Id was quantitated using GenePix Pro 6 [39].

Detection and quantitation of HaloTag7 fusion proteins

Lysates expressing HaloTag7 fusion proteins were labeled with 1 μM HaloTag TMR ligand (Promega) for 30 min, mixed with SDS loading buffer, boiled for 2 min at 95 °C and resolved by SDS–PAGE. Gels were scanned on a Typhoon 9400 (E_{ex} = 532 nm; E_{em} = 580 nm) and bands were quantitated using ImageQuant. Expression levels were calculated from a linear calibration curve generated by a serial dilution of purified and TMR ligand-labeled GST:HaloTag7 (60 kDa) and HaloTag7 (34 kDa).

Results

Assembly of test constructs

The test panel for the comparison of HaloTag7, GST, MBP and His_6-Tag contained full-length cDNAs encoding 23 human proteins that were previously shown to express poorly in *E. coli* in the absence of a tag [33]. The test panel sequences, summarized in Table 2, represent proteins of varying size (~9–155 kDa) and function (e.g. kinases, membrane proteins, and transcription factors). These 23 coding regions, previously available as Flexi vectors clones [33], were transferred to four different Flexi expression vectors that allowed for the production of each target protein in the context of HaloTag7, GST, MBP and His_6Tag as the N-terminal fusion partner.

The Flexi vector technology provides a rapid, directional and high fidelity method for transferring protein coding regions between compatible vectors using two rare-cutting restriction endonucleases, SgfI and PmeI [35]. Efficient recovery of the desired clones is achieved by the combination of two selection methods: antibiotic resistance and a lethal gene, barnase, residing in the acceptor vector. In a successful transfer reaction, the lethal barnase sequence is replaced by the coding sequence of interest.

Expression of human proteins in E. coli

The 23 human protein coding regions were expressed in *E. coli* KRX cells as HaloTag7, GST, MBP and His_6Tag N-terminal fusion proteins and analyzed by SDS–PAGE for total and soluble expression. A representative subset (six human proteins) of these results is shown in Fig. 1. To estimate the expression levels we prepared two reference expression markers by adding known amounts of purified proteins into soluble KRX lysates (as described in Materials and methods). The reference expression markers allowed for visual estimation of target protein expression levels based on the known protein concentrations in the standards. Figs. 2 and 3 summarize the total and soluble expression levels for all 23 recombinant proteins fused to each of the protein tags, as well as the impact of the different tags on solubility.

Our results show that HaloTag7 enhanced both expression of total target proteins, and more importantly the expression of soluble target proteins. In this panel of human proteins, 74% of the target proteins were successfully expressed and remained soluble when fused to HaloTag7, compared to only 52%, 39%, and 22% for MBP, GST, and His_6Tag, respectively. It is important to note that the six proteins not detected in a soluble form as HaloTag7 fusions were not solubilized by any of the other tags. In addition, we observed for each of the tags a decrease in soluble expression with increased protein size (Fig. 2). However, the impact of protein size was less pronounced for HaloTag7, as this tag showed the highest success rate (37%) in expressing target proteins $\geqslant$60 kDa compared to MBP (12.5%), GST (0%) and His_6Tag (0%).

Small scale protein purification

As the improved solubility provided by HaloTag7 was encouraging, we continued our comparison of the four tags with regard to protein purification, examining the impact of the tags on target protein yield, purity and structural integrity. From the 23 human proteins, we purified six proteins (Fig. 1) displaying a range of soluble expression levels (from undetectable to $\geqslant$100 μg/ml). The six

Table 2
List of the human coding regions used in this study.

Gene	Accession No.	Description	kDa
HSBP1	NM_001537	Heat shock factor binding protein 1	8.6
SMPX	BC005948	Small muscular protein	9.7
MAR1	BC014423	Melanoma antigen recognized by T-cells	13.3
Q9Y605	BC022797	T-cell activation protein	14.7
GFER	XM_034465	Growth factor, augmenter of liver regeneration	15.5
BTG1	NM_001731	B-cell translocation gene 1 anti-proliferative	19.3
TCPT	BC003352	T-cell tumor protein	19.7
CHP	NM_007236	Calcium binding protein P22	22.5
AURKC	NM_003160	Aurora kinase C	32.3
TEC	NM_003215	Tec protein tyrosine kinase	34.7
CREB1	NM_004379	cAMP responsive element binding protein 1	35.2
MAPK14	NM_001315	Mitogen-activated protein kinase 14	41.4
MAP2K4	NM_003010	Mitogen-activated protein kinase 4	44.4
HTR1A	NM_000524	5-hydroxytryptamine (serotonin) receptor 1A	46.2
MAPK8	NM_139049	Mitogen-activated protein kinase 8	48.4
NRBP1	NM_013392	Nuclear receptor binding protein	59.9
PRKCG	NM_002739	Protein kinase C, gamma	78.5
MCM5	NM_006739	Minichromosome maintenance deficient 5 Cell division cycle 46	82.4
GRM2	NM_000839	Glutamate receptor, metabotropic 2	95.7
GRM3	NM_000840	Glutamate receptor, metabotropic 3	99
EPHB1	NM_004441	EPH receptor B1	110.4
ERBB2(V1)	NM_004448	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2	138
IGF1R(V1)	NM_000875	*Homo sapiens* insulin-like growth factor 1 receptor	154.9

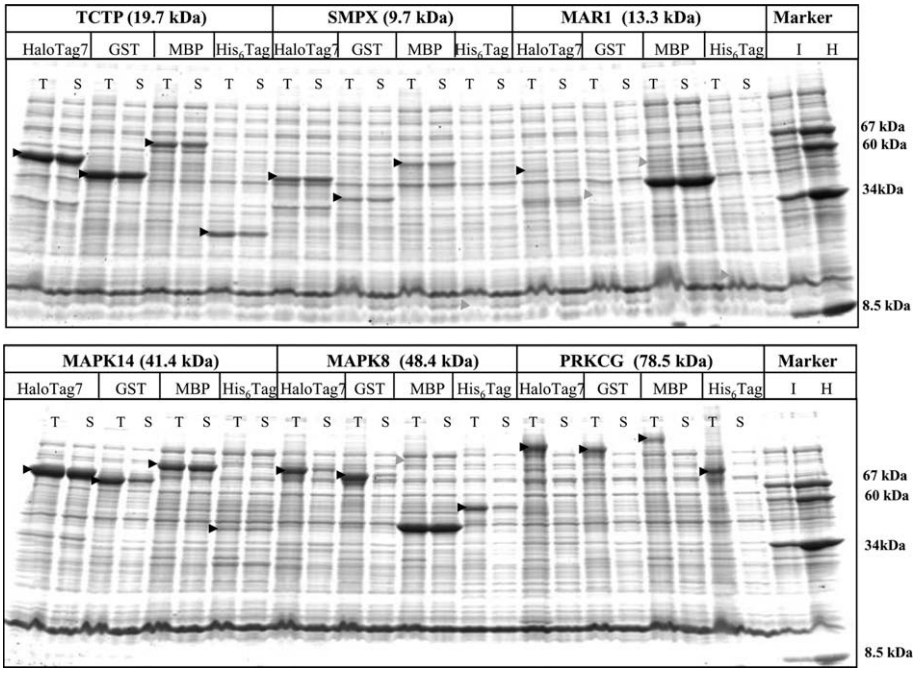


Fig. 1. Expression of six representative human proteins fused to four different protein tags. Ten microliters (1% of the *E. coli* culture) of total (T) and soluble (S) fractions were analyzed by SDS–PAGE gel and visualized by SimplyBlue staining. Reference expression markers for visual estimation of expression levels were prepared as described in Materials and methods: intermediate (I) expression 50–100 μg/ml; high (H) expression 200–400 μg/ml. (►) Correct size of fusion protein; (▷) correct size of fusion protein which was not observed.

		Total expression (μg/ml)				Soluble expression (μg/ml)			
Gene	kDa	HaloTag7	GST	MBP	His_6Tag	HaloTag7	GST	MBP	His_6Tag
HSBP1	8.6								
SMPX	9.7								
MAR 1	13.3								
Q9Y605	14.7								
GFER	15.5								
BTG1	19.3								
TCPT	19.7								
CHP	22.5								
AURKC	32.3								
TEC	34.7								
CREB1	35.2								
MAPK14	41.4								
MAP2K4	44.4								
HTR1A	46.2								
MAPK8	48.4								
NRBP1	59.9								
PRKCG	78.5								
MCM5	82.4								
GRM2	95.7								
GRM3	99.0								
EPHB1	110.4								
ERBB2(V1)	138.0								
IGF1R(V1)	154.9								

Fig. 2. Expression of 23 human proteins fused to four different protein tags in *E. coli* KRX cells. Total and soluble expression levels were assessed by SDS–PAGE analysis and the reference expression markers as described in Materials and methods. Total or soluble expression level of the fusion proteins: (■) ⩾100 μg/ml; (▩) ⩽100 μg/ml; (□) not detected on SDS–PAGE gel stained with SimplyBlue.

clones were expressed in *E. coli* KRX as HaloTag7, GST, MBP and His_6Tag N-terminal protein fusions and purified using the appropriate purification resin. HaloTag7-based protein purification differs from affinity-based methods because it utilizes covalent attachment to capture the fusion protein onto the purification resin. As a result, the fusion proteins cannot be eluted from the resin. However, the target protein can be released from the resin by TEV protease-mediated proteolytic cleavage.

Since the general view is that structural integrity of purified proteins should be determined after the solubility tag has been removed [13,21,22,25], we applied proteolytic cleavage to release the target proteins from the GST and MBP purification resins as well.

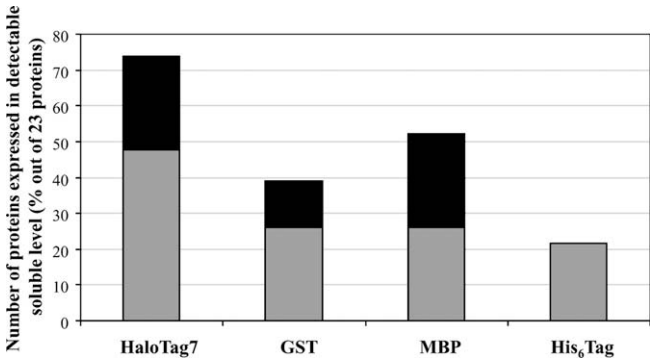


Fig. 3. Effect of tags on soluble and detectable expression of 23 human proteins. The number of soluble proteins (expressed as a % out of 23) is based on the data presented in Fig. 2. (■) expressed at ⩾100 μg/ml; (▩) ⩽100 μg/ml.

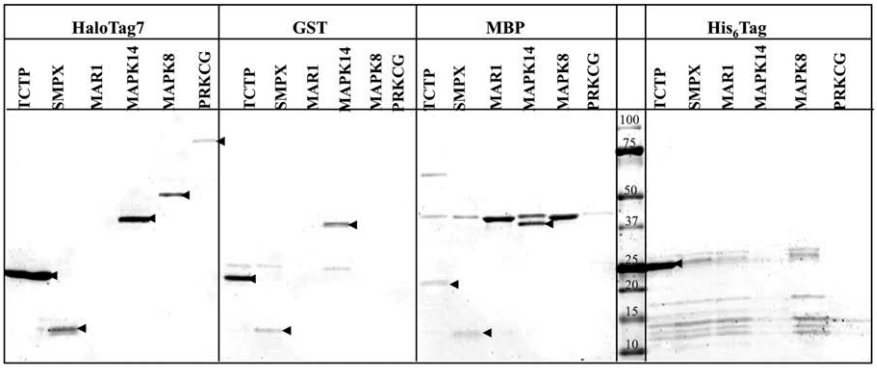
To normalize proteolytic cleavage efficiency, which can affect protein yield, we introduced the same interconnecting polypeptide (linker) containing an optimized TEV site to the HaloTag7, GST and MBP vectors. Cleavage at this optimized TEV site releases the target proteins with six vector-derived amino acids (SDNAIA) at the N-termini. Although a small tag such as His_6Tag may also interfere with protein structure or function [18,21,22], it is considered to be relatively unobtrusive and therefore commonly not removed [22,40]. Consequently, proteins expressed and purified using His_6Tag were eluted without removing the tag.

Our results, shown in Fig. 4, indicate that five out of the six human proteins were successfully purified using HaloTag7, compared to three purified using GST and MBP tags and only one purified using His_6Tag. Furthermore, the low-expressing 78 kDa PRKCG (undetectable soluble expression, Fig. 1) could only be purified using HaloTag7. In addition to providing the highest purification success rate, higher yields and purity (⩾95%) were achieved using HaloTag7 compared to the other tags. The lower purity observed for GST (62–84%) and MBP (30–60%) is primarily the result of contamination with free fusion tag released from the resin during proteolytic cleavage. All proteins purified using His_6Tag contained similar impurities, suggesting they were endogenous *E. coli* proteins with affinity to divalent metals. The low success rate we observed using His_6Tag-based protein purification is likely the result of low expression (three out of the six proteins expressed at undetectable levels) and/or inefficient capture due to the inaccessibility of the His_6Tag [3].

Biophysical characterization of purified proteins

Although the solubility of a protein may reflect on its folding nature, more precise biophysical methods are required to determine the stability and oligomeric state of the target protein [2,41]. We used gel filtration analysis to verify the expected MW of the purified proteins based on the theoretical values and consequently to determine their oligomeric state (Fig. 5). We observed good agreement between the expected and calculated MW for all purified proteins regardless of purification method, with the exception of SMPX (discussed below). Deviations of 10%, which were observed for most purified proteins, are typical for this approach, and were likely the result of variation in protein charge or tertiary structure affecting the migration of proteins through the gel filtration matrix [37,42]. A larger deviation (20%) was observed for the relatively smaller TCTP (19.7 kDa), which may be inherent to the wide separation range (10–600 kDa) of the column used. All proteins purified using MBP showed a predominate peak with an estimated MW of 33–36 kDa, corresponding to free MBP tag released from the resin during proteolytic cleavage. This contaminating peak reduced the resolution for target proteins of similar size, causing their peaks to be broader. These results indicate that all the proteins were purified as non-aggregating monomers with the exception of SMPX.

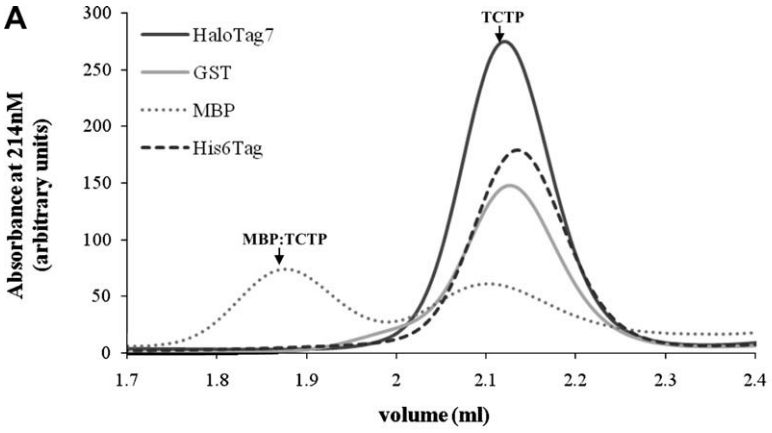
A



B

Tag	TCTP		SMPX		MAR1		MAPK14		MAPK8		PRKCG	
	μg/ml	purity	μg/ml	purity	μg/ml	purity	μg/ml	purity	μg/ml	purity	μg/ml	purity
HaloTag7	35	>95%	15	>95%	[a]0	[a]0	23	>95%	17	>95%	9	>95%
GST	8	84%	5	62%	[a]0	[a]0	7.5	79%	[a]0	[a]0%	[a]0	[a]0%
MBP	5.9	33%	4.6	43%	7	[b]0%	11.4	60%	10	[b]0%	5	[b]0%
His_6Tag	26	81%	6.5	[b]0%	7	[b]0%	1.7	[b]0%	8	[b]0%	3	[b]0%

Fig. 4. Effect of tags on protein purification yields and purity. (A) SDS–PAGE analysis of purified proteins. Six proteins representing different soluble expression levels were purified from 1 ml cell culture using four different tags (HaloTag7, GST, MBP, and His_6Tag) as described in Materials and methods. The purified proteins were analyzed by 4–20% Tris–Glycine SDS–PAGE and visualized by SimplyBlue staining. (►) Bands of correct size. (B) Analysis of protein yields and purity. Protein yields (μg from 1 ml culture) were determined by Bradford assay and protein purity (%) was determined by gel analysis as described in Materials and methods. [a]No protein was detected by Bradford analysis; [b]only contaminates were detected.



B

Recombinant protein	Expected MW(kDa)	Calculated MW (kDa)			
		HaloTag7	GST	MBP	His_6Tag
TCTP	20.7	25.6	25.5	[c]28.6 + [b]67.3	25.2
SMPX	10.7	9.9+32+67	34+66.6	[a]33.5	NA
MARI	14.3	NA	NA	[a]35.9	NA
MAPK14	42.4	42.7	42.9	[c]46.4	NA
MAPK8	49.4	46.5	NA	[a]36.1	NA
PRKCG	79.5	85.8	NA	[a]33.1	NA

Fig. 5. Gel filtration analysis of purified proteins. (A) Representative gel filtration chromatograms showing TCTP purified using four different protein tags. (B) Calculated MW of all purified proteins. Expected MW is the MW reported for these proteins + ~1 kDa for six extra residues (corresponding to His_6Tag or linker driven amino acids remaining after ProTEV proteolytic cleavage). Calculated MW is the MW calculated for these proteins from the calibration curve. [a]Contaminating MBP; [b]contaminating fusion; [c]broad peak containing MBP and the target protein.

SMPX, a small human muscular protein (9.7 kDa), showed the expected MW on SDS–PAGE. However, in the gel filtration analysis it was separated into multiple peaks. The protein purified using HaloTag7 had three peaks with calculated MW of 10 kDa, 32 kDa and 67 kDa, corresponding to a monomer, trimer and hexamer. The protein purified using GST showed only two peaks, corresponding to a trimer and hexamer. After prolonged incubation at 4 °C, we observed conversion of the 32 kDa peak to the 67 kDa peak (data not shown), suggesting that this protein may exist in a state of equilibrium between two or three different oligomerization species.

Functional analysis of purified proteins

Although the gel filtration analysis shows that the target proteins were not denatured or produced as aggregates, functional analysis is the ultimate validation of proper folding. Therefore, we added to the study two model proteins having distinct modes of functional analysis, mouse Id and the catalytic subunit of human protein kinase A (PKAc). Id, a negative regulator of myogenic differentiation, can be assayed with respect to its ability to interact with MyoD, another myogenic regulatory protein [43]. PKA is a member of the cAMP dependent protein kinase family and a key regulatory enzyme for many cellular processes [44]. Its catalytic subunit, PKAc, can be assayed for the ability to phosphorylate kempeptide, a synthetic peptide substrate [38].

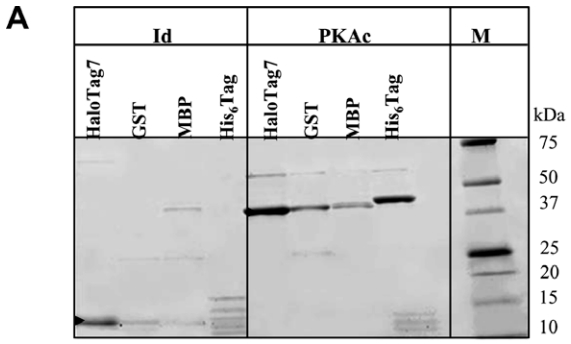
The two proteins were expressed in KRX cells as HaloTag7, GST, MBP and His_6Tag N-terminal fusions, purified from 50 ml culture and tested for yield, purity and specific activity (Fig. 6). In addition, by gel filtration analysis, both proteins purified by any of the four tags appeared as monomers (data not shown).

In evaluating specific activity, we took into account the variability in protein purity. We used SDS–PAGE analysis to estimate the fraction of the target protein in each sample and corrected the specific activity accordingly. Results showed that HaloTag7-based protein purifications provided not only the highest yield and purity, but also higher specific activity for both test proteins. Id and PKAc purified from HaloTag7 fusions were 50–60% and 15–30% more active than Id and PKAc purified from the other tags, respectively. Although it was purified successfully, His_6Tag:PKAc had a significantly lower specific activity than PKAc purified from fusions with larger tags. This might be the result of the presence of His_6Tag, which was not removed and has been previously shown in some cases to interfere with protein function [18,21,22]. The low activity could also be due to residual imidazole that was not efficiently removed during dialysis [45].

HaloTag7-based protein detection

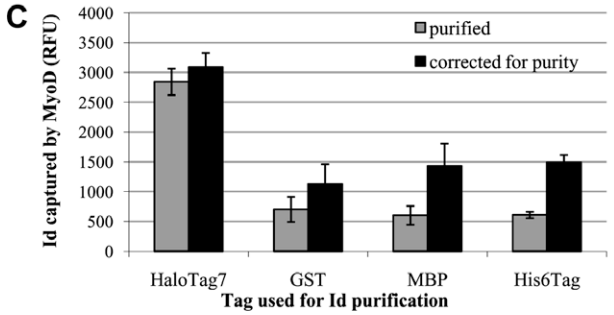
HaloTag7 was also designed to provide specific protein labeling through binding to fluorescent chloroalkane ligands. The stability of the covalent bond under conditions expected to disrupt protein structure allows for resolution by SDS–PAGE without significant loss of labeled protein, providing a simple and accurate method for detection of fusion proteins. The utility of this protein detection method was tested using the six HaloTag7 fusion proteins shown in Fig. 1, which displayed different soluble expression levels ranging from $\geqslant$100 μg/ml to undetectable levels.

Soluble fractions of bacterial lysates were labeled with the HaloTag TMR ligand, resolved by SDS–PAGE, and detected using a fluorescence imager (Fig. 7). All six proteins, including the low-expressing PRKCG (not detected by SimplyBlue staining), were de-



B

Tag	Recombinant protein purified from 50 ml culture			
	Id		PKAc	
	Yield (mg)	Purity (%)	Yield (mg)	Purity (%)
HaloTag7	0.6	92	1.7	91
GST	0.3	63	0.5	63
MBP	0.2	42	0.4	42
His_6Tag	0.4	41	1	70



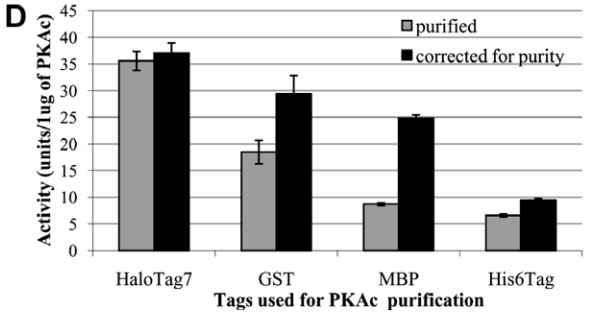
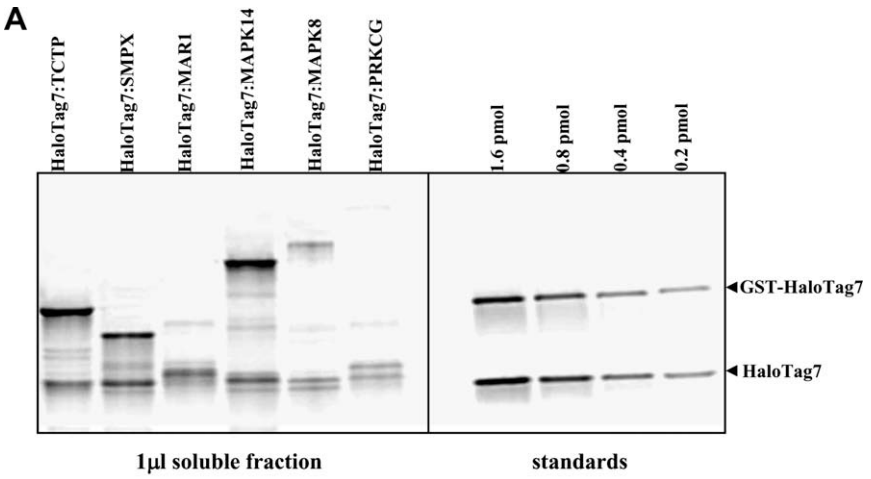


Fig. 6. Functional analysis of proteins purified using different tags. (A) Purification of Id and PKAc with four different tags as described in Materials and methods. The purified proteins were analyzed by SDS–PAGE gel and visualized by SimplyBlue staining. (►) Expected MW. (B) Analysis of protein yields and purity. Protein yields were determined by Bradford assay and protein purity (%) was determined by gel analysis as described in Materials and methods. (C) Specific activity of purified Id proteins. Activity of Id with respect to its ability to interact with MyoD was determined as described in Materials and methods. (D) Specific activity of purified PKAc proteins. Activity of PKAc was determined by its ability to phosphorylate Kemptide as described in Materials and methods.



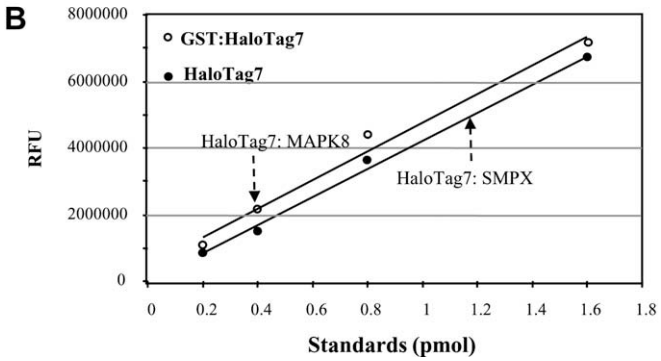


Fig. 7. HaloTag7-based protein detection and quantitation. (A) Fluorescence imaging of labeled HaloTag7 fusions. One microliter of soluble lysates (corresponding to 0.1% of cell culture) were labeled with 1 µM HaloTag TMR ligand for 30 min at room temperature and resolved by SDS–PAGE. A standard calibration curve was generated with serial dilutions of purified and TMR ligand-labeled GST:HaloTag7 and HaloTag7 proteins loaded onto the same gel. (B) Linear calibration curve generated from the labeled standards.

tected by this approach. The 1:1 stoichiometric labeling of HaloTag7 fusion proteins with the fluorescent ligand permits reproducible quantitation of the expressed fusion proteins within the lysate. In Fig. 7B we demonstrated how expression levels can be calculated from in-gel common standards comprised of serial dilutions of purified and labeled HaloTag7 proteins.

Discussion

Purification of functional recombinant proteins requires the ability to express adequate levels of soluble protein in an appropriate host and then efficiently isolate the protein to homogeneity. Although many protein fusion tags are available to assist in this process, none is ideal when applied to the diversity of proteins routinely studied. Moreover, available tags are often better suited to specific aspects of the overall process, such as expression, solubilization, protein capture, quantification, and so forth. Therefore, in selecting a tag it is common to compromise one feature in favor of another, or to combine two tags together to broaden functionality [25,26,28,29]. For example, a solubilization and a purification tag may be linked to exploit the advantages of both.

Unlike most fusion tags in common use, HaloTag7 was designed for optimal performance with respect to multiple features. Here we have demonstrated that HaloTag7 provides enhanced expression and solubility in *E. coli*, efficient protein purification with high yield and specific activity, and efficient protein labeling for screening and quantitation. Enhanced solubility coupled with covalent capture gives HaloTag7 advantages over conventional affinity fusion tags, as demonstrated by the successful purification of more proteins (including low expressers) with higher yield and purity compared to the other tested tags. Covalent capture coupled with efficient tag removal to release the target protein while HaloTag7 remains covalently bound to the resin provides higher purity by eliminating tag contaminates (either free tag or tag bound to the target protein). Furthermore, the usage of a single physiological buffer compatible with most applications throughout the HaloTag7 purification process provides a streamlined protocol and eliminates the need for dialysis.

Using full-length cDNAs encoding 23 human proteins of different size and function, we compared the efficacy of HaloTag7 as an expression and purification tag to the frequently used solubility enhancers MBP and GST. We also included His_6Tag, which is not utilized to enhance solubility but is the most frequently used purification tag, and thereby can serve as a reference point for solubility enhancement. Consistent with previous reports, the three larger tags (e.g. HaloTag7, MBP and GST) provided higher total and soluble expression compared to the smaller His_6Tag [3,4]. When we compared soluble protein expression for these larger tags we found that HaloTag7 solubilizes 74% of the 23 human proteins examined compared to only 52% for MBP and 39% in case of GST. These results indicate that for this test panel, HaloTag7 was superior to MBP, which is currently considered to be one of the best solubility tags available [3,4,13,20,27]. In general, solubility was inversely correlated to increased target protein size, an observation in agreement with a previous study [10]. However, the impact of size on solubility was less pronounced for HaloTag7 fusions, highlighting that HaloTag7 may be particularly useful for expression of larger proteins.

The precise mechanism of how protein tags enhance solubility is unclear. It has been suggested that they promote proper folding of their fusion partner because the N-terminal tag can rapidly reach a stable conformation either during, or shortly after synthesis [1]. Another proposed possibility is that expression tags function as general molecular chaperones in the context of a fusion protein or serve to attract chaperones [13,46]. HaloTag7 was engineered through molecular evolution for maximal soluble expression of fusion proteins in *E. coli* and thus its capabilities were developed through empirical selection. Subsequent analysis revealed that HaloTag7 had acquired increased stability and negative charge relative to the parental dehalogenase [Encell, personal communication]. We believe that the increased stability contributes to structural compatibility of the tag with different fusion partners, and fosters improved expression of properly folded proteins leading to higher protein solubility. The uniform negative surface charge of HaloTag7 ($pI \sim 5$) may reduce electrostatic attraction between the tag molecules and thus reduce the tendency for aggregation [47,48]. We suggest that the combination of enhanced stability and reduced tendency for aggregation may contribute to the enhanced performance of HaloTag7 as an expression solubility tag relative to GST and MBP.

Combining the improved expression of soluble fusion protein with efficient covalent capture should generally increase purification yields and purity. We tested this assertion with a subset of full-length cDNAs encoding six proteins that displayed a range of soluble expression levels. These proteins were purified using each of the four fusion tags, with the tags subsequently proteolytically removed from the purified proteins in all cases except for His_6Tag. Our data reveal that HaloTag7 delivered superior performance for both purification yields and purity. Of six proteins in the subset, five were purified when fused to HaloTag7 compared to three when fused to GST and MBP and only one when fused to His_6Tag.

These results suggest that the highly specific, rapid and covalent capture overcomes the equilibrium-based limitations associated with affinity tags, i.e. poor capture of proteins expressed at low levels and protein loss during washing of the purification resin. This proposal is supported by the case of the poorly expressed 78 kDa PRKCG, which could be purified only when fused to HaloTag7. The covalent capture and proteolytic release of target protein also provided higher purity ($\geqslant 95\%$), a result likely due to HaloTag7 being retained bound to the purification resin. The purification of only three proteins at lower yield and purity using GST and MBP was probably the result of low expression and lower affinity to the purification matrix, leading to inefficient capture and contamination with free tag released from the resin during proteolytic cleavage. The observed low affinity of MBP for the amylose resin has been well documented [3,29,49] and is likely the cause for the very low protein yield we observed for MBP-based protein purifications. The purification of only one protein using His_6Tag is probably the result of inefficient capture onto the resin due to low expression and/or inaccessibility of the tag. Low expression, along with the presence of native proteins that have affinity to divalent cations, is probably the cause for the inefficient capture and the presence of impurities [45,50]. The possibility of inaccessibility of His_6Tag due to fusion protein conformation is supported by a previous report indicating that capture of His_6Tag onto the Ni–NTA purification matrix under denaturing conditions is more efficient [3].

It has been well documented that certain proteins requiring a tag for purification will convert to an insoluble aggregate upon removal of the tag [2,25,26,41]. To examine whether this occurred with any of our target proteins, we used gel filtration chromatography to determine the apparent molecular weight and aggregation state of the purified proteins. We observed good agreement between calculated and theoretical molecular weights (with the exception of SMPX) regardless of the tag used, suggesting that the purified proteins were not denatured or aggregated, and therefore likely to be properly folded. Purified SMPX, regardless of the purification method, separated in gel filtration analysis into multiple peaks corresponding to monomers, trimers and hexamers. As we observed conversion of trimers into hexamers with time, we suggest that this protein is not denatured or aggregated but exists in a state of equilibrium between two or three oligomerization species.

Functional analysis was also carried out to provide further validation of proper folding. We observed higher specific activities for both PKAc and Id purified using HaloTag7, suggesting better folding provided by this tag. This result may suggest that the high intrinsic stability of HaloTag7 facilitates proper folding of the entire fusion protein. It is also noteworthy that one of the purified target proteins, PKAc, showed much lower specific activity when purified using His_6Tag. This may be caused by the presence of the His_6Tag and its potential interference with protein function [18,21,22]. This highlights that tags, regardless of size, may potentially interfere with protein structure or function and should be removed.

Finally, HaloTag7 also provides convenient fluorescence detection of recombinant proteins, an important and valuable feature for the early screening and optimization of cell culture and expression conditions. This labeling method is very sensitive, allowing detection of fusion proteins that cannot be detected with colorimetric staining. Unlike other detection methods, the labeling of HaloTag7 occurs at a 1:1 M ratio, enabling quantitation of the expression levels from an in-gel linear calibration curve generated with a serial dilution of a purified and labeled HaloTag7 protein.

It is important to note that the sensitivity of HaloTag7-based protein labeling allowed the detection of unintended truncations within the polypeptide linker region that were not detected by SimplyBlue staining (Fig. 7). Similar unintended truncation not detected by colorimetric staining, which can be detected by more sensitive detection methods such as Western blot, are likely for GST and MBP as well.

HaloTag7-based protein detection is rapid, reliable and quantitative. These features are useful for optimization of different steps in the production of recombinant proteins, including growth conditions for higher expression of soluble fusion protein with minimal truncations, and efficient fusion protein capture onto the purification matrix. Furthermore, this labeling method can be easily adapted to high throughput screening based on methods such as fluorescence polarization, facilitating systematic screening for optimized expression conditions.

The data presented here demonstrate that HaloTag7 is a comprehensive tag addressing all the steps necessary for efficient and successful purification including enhanced expression of soluble proteins in bacteria, convenient fusion detection, as well as immobilization and purification of recombinant proteins. Furthermore, HaloTag7 out-performed three commonly used tags (GST, MBP and His_6Tag) with respect to all of these features. The enhanced solubility compared to MBP, considered one of the best solubility enhancers, and the enhanced purification yields and purity compared to His_6Tag, the most widely used purification tag, suggest that HaloTag7 provides a considerable advantage over existing tags. In addition to purification applications, the ability of HaloTag7 to react with a variety of customized chloroalkane ligands containing different functional groups enables a broad spectrum of *in vivo* and *in vitro* applications using a single genetic construct, including live and fixed cellular imaging [31,32] and analysis of protein interactions [31].

Acknowledgments

We thank Nidhi Nath and Jim Hartnett for help with experiments and Robin Hurst for helpful discussion.

References

[1] F. Baneyx, Recombinant protein expression in *Escherichia coli*, Current Opinion in Biotechnology 10 (1999) 411–421.
[2] A.M. Edwards, C.H. Arrowsmith, D. Christendat, A. Dharamsi, J.D. Friesen, J.F. Greenblatt, M. Vedadi, Protein production: feeding the crystallographers and NMR spectroscopists, Nat. Struct. Mol. Biol. 7 (2000) 970–972.
[3] P. Braun, Y. Hu, B. Shen, A. Halleck, M. Koundinya, E. Harlow, J. LaBaer, Proteome-scale purification of human proteins from bacteria, Proc. Natl. Acad. Sci. USA 99 (2002) 2654–2659.
[4] M. Hammarstrom, N. Hellgren, S. van Den Berg, H. Berglund, T. Hard, Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*, Protein Sci. 11 (2002) 313–321.
[5] W. Schumann, L.C. Ferreira, Production of recombinant proteins in *Escherichia coli*, Genetics and Molecular Biology 27 (2004) 442–453.
[6] G.S. Waldo, Improving protein folding efficiency by directed evolution using the GFP folding reporter, Methods Mol. Biol. 230 (2003) 343–359.
[7] K. Omoya, Z. Kato, E. Matsukuma, A. Li, K. Hashimoto, Y. Yamamoto, H. Ohnishi, N. Kondo, Systematic optimization of active protein expression using GFP as a folding reporter, Protein Expr. Purif. 36 (2004) 327–332.
[8] B.R. Kelemen, T.A. Klink, M.A. Behlke, S.R. Eubanks, P.A. Leland, R.T. Raines, Hypersensitive substrate for ribonucleases, Nucleic Acids Res. 27 (1999) 3696–3701.
[9] E.R. LaVallie, J.M. McCoy, Gene fusion expression systems in *Escherichia coli*, Curr. Opin. Biotechnol. 6 (1995) 501–506.
[10] D.B. Smith, Generating fusions to glutathione S-transferase for protein studies, Methods Enzymol. 326 (2000) 254–270.
[11] 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.
[12] E.R. LaVallie, Z. Lu, E.A. Diblasio-Smith, L.A. Collins-Racie, J.M. McCoy, Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*, Methods Enzymol. 326 (2000) 322–340.
[13] 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.
[14] D. Sachdev, J.M. Chirgwin, Fusions to maltose-binding protein: control of folding and solubility in protein purification, Methods Enzymol. 326 (2000) 312–321.
[15] G.D. Davis, C. Elisee, D.M. Newham, R.G. Harrison, New fusion protein systems designed to give soluble expression in *Escherichia coli*, Biotechnol. Bioeng. 65 (1999) 382–388.
[16] V. De Marco, G. Stier, S. Blandin, A. De Marco, The solubility and stability of recombinant proteins are increased by their fusion to NusA, Biochem. Biophys. Res. Commun. 322 (2004) 766–771.
[17] J. Nilsson, S. Stahl, J. Lundeberg, M. Uhlen, P.A. Nygren, Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins, Protein Expr. Purif. 11 (1997) 1–16.
[18] M.H. Hefti, C.J. Van Vugt-Van der Toorn, R. Dixon, J. Vervoort, A novel purification method for histidine-tagged proteins containing a thrombin cleavage site, Anal. Biochem. 295 (2001) 180–185.
[19] E. Hochuli, W. Bannwarth, H. Dobeli, R. Gentz, D. Stuber, Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent, Nat. Biotech. 6 (1988) 1321–1325.
[20] Y.P. Shih, W.M. Kung, J.C. Chen, C.H. Yeh, A.H. Wang, T.F. Wang, High-throughput screening of soluble recombinant proteins, Protein Sci. 11 (2002) 1714–1719.
[21] J. Wu, M. Filutowicz, Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale, Acta Biochim. Pol. 46 (1999) 591–599.
[22] 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.
[23] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 60 (2003) 523–533.
[24] P.A. Walker, L.E. Leong, P.W. Ng, S.H. Tan, S. Waller, D. Murphy, A.G. Porter, Efficient and rapid affinity purification of proteins using recombinant fusion proteases, Biotechnology (NY) 12 (1994) 601–605.
[25] D.S. Waugh, Making the most of affinity tags, Trends Biotechnol. 23 (2005) 316–320.
[26] D. Esposito, D.K. Chatterjee, Enhancement of soluble protein expression through the use of fusion tags, Curr. Opin. Biotechnol. 17 (2006) 353–358.
[27] S. Nallamsetty, D.S. Waugh, Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners, Protein Expr. Purif. 45 (2006) 175–182.
[28] I. Hunt, From gene to protein: a review of new and enabling technologies for multi-parallel protein expression, Protein Expr. Purif. 40 (2005) 1–22.
[29] K.D. Pryor, B. Leiting, High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double-affinity fusion system, Protein Expr. Purif. 10 (1997) 309–319.
[30] D.B. Janssen, Evolving haloalkane dehalogenases, Curr. Opin. Chem. Biol. 8 (2004) 150–159.
[31] G.V. Los, L.P. Encell, M.G. McDougall, D.D. Hartzell, N. Karassina, C. Zimprich, M.G. Wood, R. Learish, R. Friedman Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D.H. Klaubert, R.F. Bulleit, K.V. Wood, HaloTag: a novel protein labeling technology for cell imaging and protein analysis, ACS Chem. Biol. 3 (2008) 373–382.
[32] G.V. Los, K. Wood, The HaloTag: a novel technology for cell imaging and protein analysis, Methods Mol. Biol. 356 (2007) 195–208.
[33] M. Slater, R. Hurst, B. Pferdehirt, D. White, A. Niles, N. Betz, E. Schenborn, Expression of soluble native human proteins in cell-free extracts, Promega Notes 91 (2005) 21–25.

[34] J. Hartnett, J. Gracyalny, M.R. Slater, The single step (KRX) competent cells: efficient cloning and high protein yields, Promega Notes 94 (2006) 27–30.
[35] M. Slater, J. Hartnett, N. Betz, J. English, E. Strauss, B. Pferdehirt, E. Schenborn, A new system for cloning and expressing protein-coding regions, Promega Notes 89 (2005) 11–15.
[36] T. Schagat, R. Friedman-Ohana, P. Otto, J. Hartnett, M. Slater, KRX autoinduction protocol: a convenient method for protein expression, Promega Notes 98 (2008) 16–18.
[37] P. Cutler, Protein purification protocols, Methods in Molecular Biology 244 (2004) 239–252.
[38] R. Somberg, B. Pferdehirt, K. Kuocho, Introducing the kinase-Glo luminescent kinase assay, Promega Notes 83 (2003) 14–17.
[39] N. Nath, R. Hurst, B. Hook, P. Meisenheimer, K.Q. Zhao, N. Nassif, R.F. Bulleit, D.R. Storts, Improving protein array performance: focus on washing and storage conditions, J. Proteome Res. 7 (2008) 4475–4482.
[40] S.A. Lesley, High-throughput proteomics: protein expression and purification in the postgenomic world, Protein Expr. Purif. 22 (2001) 159–164.
[41] L.D. Cabrita, W. Dai, S.P. Bottomley, A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production, BMC Biotechnol. 6 (2006) 12.
[42] P. Andrews, The gel-filtration behaviour of proteins related to their molecular weights over a wide range, Biochem. J. 96 (1965) 595–606.
[43] R. Benezra, R.L. Davis, D. Lockshon, D.L. Turner, H. Weintraub, The protein Id: a negative regulator of helix-loop-helix DNA binding proteins, Cell 61 (1990) 49–59.
[44] J.B. Shabb, Physiological substrates of cAMP-dependent protein kinase, Chem. Rev. 101 (2001) 2381–2411.
[45] A.D. Keefe, D.S. Wilson, B. Seelig, J.W. Szostak, One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag, Protein Expr. Purif. 23 (2001) 440–446.
[46] J.D. Fox, R.B. Kapust, D.S. Waugh, Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins, Protein Sci. 10 (2001) 622–630.
[47] G. Calloni, S. Zoffoli, M. Stefani, C.M. Dobson, F. Chiti, Investigating the effects of mutations on protein aggregation in the cell, J. Biol. Chem. 280 (2005) 10607–10613.
[48] M.S. Lawrence, K.J. Phillips, D.R. Liu, Supercharging proteins can impart unusual resilience, J. Am. Chem. Soc. 129 (2007) 10110–10112.
[49] D.R. Smyth, M.K. Mrozkiewicz, W.J. McGrath, P. Listwan, B. Kobe, Crystal structures of fusion proteins with large-affinity tags, Protein Sci. 12 (2003) 1313–1322.
[50] X. Zuo, M.R. Mattern, R. Tan, S. Li, J. Hall, D.E. Sterner, J. Shoo, H. Tran, P. Lim, S.G. Sarafianos, L. Kazi, S. Navas-Martin, S.R. Weiss, T.R. Butt, Expression and purification of SARS coronavirus proteins using SUMO-fusions, Protein Expr. Purif. 42 (2005) 100–110.