ELSEVIER

Journal of Chromatography B, 769 (2002) 133–144

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Strategies for the purification and on-column cleavage of glutathione-*S*-transferase fusion target proteins

Cyril Dian[a,b], Said Eshaghi[a], Thomas Urbig[a], Seán McSweeney[b], Anna Heijbel[c], Gilles Salbert[d], Darcy Birse[a,*]

[a]*Department of Biochemistry and Biophysics, Arrhenius Laboratories for Natural Sciences, Stockholm University, SE-106 91 Stockholm, Sweden*

[b]*European Synchrotron Radiation Facility (ESRF), BF-220 CEDEX, 38043, Grenoble, France*

[c]*Amersham Biosciences, AB 2000, SE-751 25 Uppsala, Sweden*

[d]*Equipe Information et Programmation Cellulaire, University of Rennes I, UMR 6026 CNRS, Rennes, France*

Received 12 September 2001; received in revised form 18 December 2001; accepted 19 December 2001

Abstract

In this report, we describe a flexible, efficient and rapid protein purification strategy for the isolation and cleavage of glutathione-*S*-transferase (GST) fusion proteins. The purification and on-column cleavage strategy was developed to work for the purification of difficult proteins and for target proteins where efficient fusion-tag cleavage is essential for downstream processes, such as structural and functional studies. To test and demonstrate the flexibility of this method, seven diverse unrelated target proteins were assayed. A purification technique is described that can be applied to a wide range of both soluble and membrane inserted recombinant target proteins of differing function, structure and chemical nature. This strategy is performed in a single chromatographic step applying an on-column cleavage method, yielding ''native'' proteins in the 200 μg to 40 mg/l scale of 95–98% purity. © 2002 Published by Elsevier Science B.V.

Keywords: Enzymes; Glutathione-*S*-transferase; Fusion target proteins

1. Introduction

Molecular biology methods can manipulate, at the gene level, a target protein by fusing it to another protein or peptide fragment to facilitate rapid ''capture'' of the protein of interest using an affinity chromatography step (see recent reviews in Refs. [1–3] and references therein). At the protein expression level, fusion proteins can have the advantage of providing a more favourable gene construct organization permitting higher levels of soluble protein to be expressed [4], and possibly reducing the propensity to drive the protein folding process towards creating inclusion bodies [5]. An inherent problem with the fusion protein/peptide system is that the ''tag'' is often difficult to remove. Specific proteases required to perform the cleavage reaction necessary to separate the fusion tag from the target protein have inherent difficulties manifesting themselves as: (i) non-specific proteolytic attack of the target protein; (ii) the need for elevated temperatures for efficient cleavage, often resulting in the denaturation

*Corresponding author. Tel.: +46-8-16-4392; fax: +46-8-15-3679.

E-mail address: birse@dbb.su.se (D. Birse).

1570-0232/02/$ – see front matter
PII: S1570-0232(01)00637-7

or aggregation of the target protein; (iii) incomplete proteolytic processing resulting in partially cleaved target protein, thereby significantly reducing the yield and/or introducing heterogeneity to the purified protein; (iv) additional purification steps are necessary to separate the cleaved target protein from the fusion tag, deactivate and remove the processing protease and exchange or desalt buffer components. These problems are addressed in reviews [1–13].

This report describes a flexible, efficient and rapid protein purification strategy for the isolation, purification, and cleavage of glutathione-*S*-transferase (GST) fusion proteins. The groundwork for this purification method has been outlined previously [14,15] and reviewed relative to other fusion protein technologies [1,2] such as polyhistidine tags [16,17], FLAG-tags [18–20], thioredoxin [21–23], Protein A [24–26], *Strep*-tag [27,28] and Maltose-binding protein [29–31].

However, the wide range of diverse applications and effectiveness of the technique was not demonstrated in terms of: (i) on-column cleavage efficiency (>70%); (ii) suitable levels of protein enrichment (>95–98%); (iii) large-scale applicability; (iv) problematic GST-fusion proteins (poorly expressing, poorly soluble or membrane proteins); and (v) advances to accommodate high-throughput (HTP) protein purification applications. To determine if this strategy is compatible to perform with a wide range of proteins, seven unrelated target proteins with diverse properties and functions such as: DNA binding; RNA binding; signal transduction; nuclear receptor; membrane associated; and integral membrane proteins were investigated. The purification approach described in this article is performed in a single chromatographic step applying an on-column cleavage strategy, yielding ''native'' proteins with a high level of enrichment with a yield in the 95–98% purity range that can be applied in a fully automated high-throughput experimental design.

2. Experimental

2.1. Cloning

The genes encoding: GST::Pur-α (NCBI Accession: XP_006652); GST::Translin (NCBI Accession: AAH04615); GST::TLP40 (and constructs; NCBI Accession: O49939); GST::COUP-TFI (NCBI Accession: P10589); GST::Lep (NCBI Accession: NP_417794); GST::ecoKch (NCBI Accession: BAA35689); and GST::ProW (NCBI Accession: BAA07636) fusion proteins were independently subcloned into pGEX-6P (Amersham Biosciences, Uppsala, Sweden) expression vectors according to Ref. [32].

2.2. Cell growth

The pGEX-6P vector containing the fusion target gene was transformed into E. coli host strain BL21(DE3) (Stratagene, La Jolla, CA, USA) overexpressing cell line. Comprehensive expression tests were performed selecting for high expression levels of soluble, or targeted membrane proteins. The optimum conditions were determined and applied to large-scale preparations. Inoculated cultures (1/1000) were grown in LB medium in TUNAIR flasks (Shelton Scientific, Shelton, CT, USA) at 25 °C (COUP-TFI, Translin and ProW), 30 °C (Pur-α, TLP40 and ecoKch) and 37 °C (Lep) supplemented with 100 μg/ml Carbenicillin (Sigma, WI, USA) on a floor model shaker/incubator gyrating at 220 rpm. The cell culture was grown until $OD_{600} \cong 1.0$ whereby GST::target protein synthesis was induced with 1.0 m*M* IPTG (final concentration; Sigma, WI, USA) and grown further for: 3 h (Lep); 5 h (Pur-α, TLP40); 7 h (ecoKch and ProW) and 9 h (COUP-TFI and Translin). Cells were harvested by centrifugation at 5000 *g* for 30 min at 4 °C. The cell pellet was resuspended, washed with PBS buffer (4.3 m*M* Na_2HPO_4, 1.4 m*M* KH_2PO_4, 2.7 m*M* KCl, 137 m*M* NaCl at pH 7.3) and centrifuged at 4000 *g* for 10 min at 4 °C. Cell pellets were then frozen and stored at −80 °C.

2.3. Cell lysis

2.3.1. Soluble target proteins

The cell pellet containing the soluble expressed fusion GST::target protein (GST::Pur-α; GST::Translin; GST::TLP40 (and constructs); GST::COUP-TFI was resuspended in Lysis buffer (PBS) supplemented with: 1 mg/ml lysozyme (Sigma, WI, USA); Complete Protease Inhibitor Cocktail (accord-

ing to the manufacturer's recommendations; Roche, Manheim, Germany); 10 m*M* $MgCl_2$; and 10 U/ml DNase I (Sigma, WI, USA). The cells were lysed by repeated (3×) freezing (−170 °C) in $N_{2(l)}$ and thawing in a water bath (30 °C). The lysate was then clarified by centrifugation at 70 000 *g* for 30 min at 4 °C to remove cellular debris followed by ultracentrifugation of the supernatant at 300 000 *g* for 60 min at 4 °C to remove membrane components, large nucleic acids and aggregates.

2.3.2. Isolation of membrane fraction

For the preparation of GST-fused membrane proteins (GST::Lep; GST::ecoK; GST::ProW), isolated enriched membranes were selected as being most suitable for the starting material of the purification process. Cells were lysed as described above for soluble GST::target proteins. The remaining whole cells and cell debris were removed by centrifugation at 10 000 *g* for 20 min at 4 °C followed by an ultracentrifugation step of the supernatant at 300 000 *g* for 60 min at 4 °C to pellet the membrane fraction. The supernatant was discarded and the membrane pellet was frozen at −80 °C. To solubilize the membranes, PBS buffer containing dodecylmaltoside (DDM; Roche, Manheim, Germany) at a final concentration of 0.5% (w/v), was added and stirred in an ice bath for 30 min. After solubilization, the extract was clarified by ultracentrifugation at 300 000 *g* for 60 min at 4 °C. The supernatant containing the solubilized membrane components was decanted and used in subsequent purification steps.

2.4. Chromatographic steps

2.4.1. GST::target protein binding

The supernatant containing the GST::target protein fraction from the ultracentrifugation step (soluble or membrane containing fractions) was loaded on a GSTrap™ FF column (5 ml; Amersham Biosciences, Uppsala, Sweden), pre-equilibrated with PBS as binding buffer, at a flow rate of ~1 ml/min. For all chromatographic steps, an ÄKTA™ Explorer (Amersham Biosciences, Uppsala, Sweden) was used enclosed in a refrigeration unit cooled to 4 °C to ensure protein stability and reduce protein degradation. Chromatographic profiles continuously monitor absorbance (260 and 280 nm) and conductivity (mS/cm). The bound material was washed with PBS buffer until the absorbance baseline had returned. Once the baseline was stable, the buffer was exchanged with Cleavage buffer (50 m*M* Tris–HCl at pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA and 1 m*M* DTT). Cleavage buffer equilibration was continued until a stable absorbance baseline was achieved. At this stage, buffer flow was arrested.

For the purification and on-column cleavage of membrane proteins, 0.03% (w/v) DDM was included in all buffers throughout the entire purification process.

2.4.2. On-column cleavage reaction

This method uses the technologies of a GST-fusion protein linked with an infrequently biologically occurring proteolytic cleavage site, in conjunction with glutathione affinity columns and a highly specific engineered protease to purify desired target proteins in high yields with high levels of enrichment. PreScission™ protease (Amersham Bioscience, Uppsala, Sweden) is a genetically engineered fusion protein consisting of GST fused to a modified human rhinovirus 3C protease [14,15]. The proteolytic cleavage site: Leu-Glu-Val-Leu-Phe-Gln-$^{\nabla}$-Gly-Pro (where $^{\nabla}$ indicates the proteolytic cleavage site) has zero identical amino acid sequences based on searches of the SWISSPROT annotated protein sequence database and percentage occurrences of 1.6×10^{-3} for proteins containing the underlined sequence.

PreScission protease (2 U enzyme/100 μg of bound fusion GST::target protein) was diluted in Cleavage buffer equal to 90% of the volume of the GSTrap FF column and injected into the column at an increased flow rate of ~5–7 ml/min. Following injection, the column was placed in a closed flow status and the system was incubated on-line for 12–16 h at 4 °C. The on-column cleavage process of PreScission protease can have reduced efficiency in the presence of detergents due to micelle formation whereby proteolytic incubation times should be increased accordingly on protein/detergent-to-protein/detergent basis to reach maximum cleavage efficiency.

2.4.3. Elution of ''native'' target protein

An auxiliary GSTrap FF column (1 ml) pre-equilibrated with Cleavage buffer was connected downstream of the primary cleavage reaction column inline with the fraction collector. Cleaved ''native'' target protein elution occurs immediately upon flow start-up with Cleavage buffer at ~1 ml/min. Following cleaved target protein elution and the return of the absorbance baseline, the GST-affinity peak was eluted with Elution buffer (50 m*M* Tris–HCl at pH 8.0 and 10 m*M* reduced glutathione) in a full step gradient (100% elution buffer).

2.4.4. Column regeneration/equilibration

Following the elution of the target protein and the GST-affinity proteins, the column can be equilibrated for subsequent purification runs. Column equilibration is completed by flushing the column with three column volumes of Milli-Q water followed by three column volumes of PBS Binding buffer. This regeneration stage is important for the throughput of the protein production process and allows for multiple runs to be completed in series.

2.5. Analytical methods

2.5.1. SDS–PAGE analysis

The purification stages and chromatographic profiles were evaluated by SDS–PAGE using 3.5–12% stacking polyacrylamide gel in the Fling and Gregerson buffer system [33] stained with Coomassie™ Brilliant Blue R-250 (Roche, Manheim, Germany) and silver staining.

2.5.2. Mass spectrometry

The on-column cleavage process of GST::target proteins cleaved by PreScission protease was monitored by MALDI mass spectrometry. After tryptic digest of samples, peptides were extracted and analysed by matrix-assisted laser desorption/ionization mass spectrometry using Voyager Biospectrometry Workstation with Delayed Extraction Technology, PerSeptive Biosystems, Inc. Data obtained were analysed using Moverz software (Proteometrics, LLC).

2.5.3. Protein concentration analysis

Protein quantification was evaluated and determined using Coomassie® Protein Assay Reagent (Pierce, IL, USA) based on the Bradford Coomassie dye binding colourimetric method [34]. Measurements were normalized using bovine serum albumin as a protein standard measuring absorbance at 595 nm.

3. Results

The aim of this work is to isolate highly pure and homogenous recombinant proteins in a rapid and efficient manner suitable for structural, functional and drug discovery studies. To facilitate this, the gene encoding the desired target protein is cloned into a pGEX-6P expression vector allowing for the production of a GST-fusion protein with a highly specific proteolytic cleavage site. To best illustrate the versatility of this method, the procedure is presented in a flow scheme depicted in Fig. 1. To facilitate the description and effectiveness of the strategy and results, selected typical chromatograms (Fig. 2) and SDS–PAGE data (Fig. 3) are presented to best illustrate the observations and experimental measurements. Specific target protein purification data and observations are presented and summarized in Table 1.

3.1. Selected target proteins

Seven target proteins were selected to demonstrate the GST-fusion purification and on-column cleavage strategy.

3.1.1. Pur-α

Pur-α is a sequence specific, single-stranded DNA and RNA binding protein that binds to purine-rich promoter regions with a consensus $(GGN)_n$ sequence [35–41]. Previously, Pur-α has been shown to be a difficult protein to purify to homogeneity with problems of yield, purity, degradation and decreased function/activity following purification.

3.1.2. Translin

Translin is a recombination hot-spot binding-protein specifically recognizing DNA consensus se-

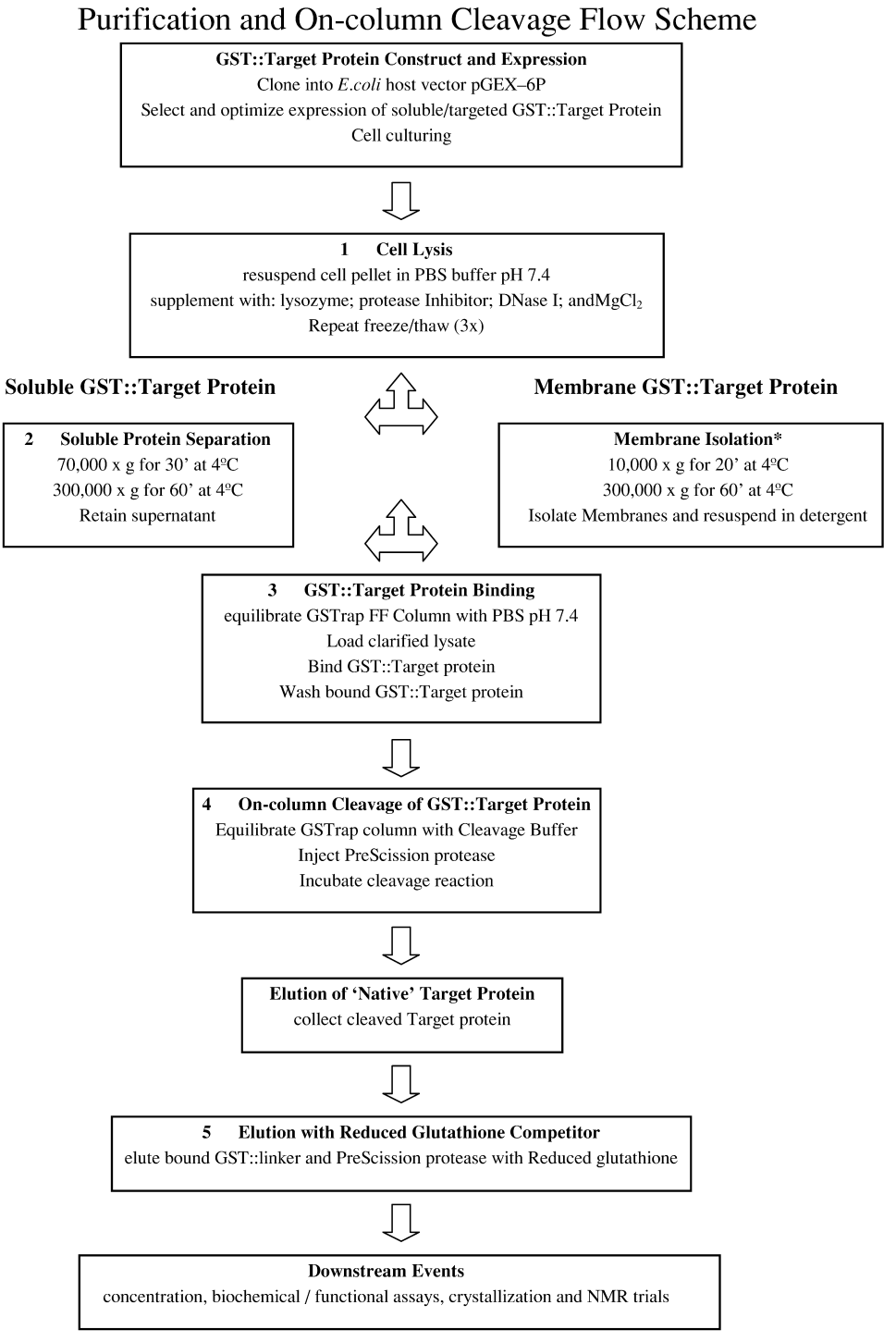

Fig. 1. Purification and on-column cleavage strategy. A general flow scheme illustrating the purification and on-column cleavage strategy for GST::target proteins producing ''native'' proteins in the >95% purity range. The asterisk indicates that detergents are present throughout all subsequent purification, cleavage reaction and ''native'' target protein elution steps.

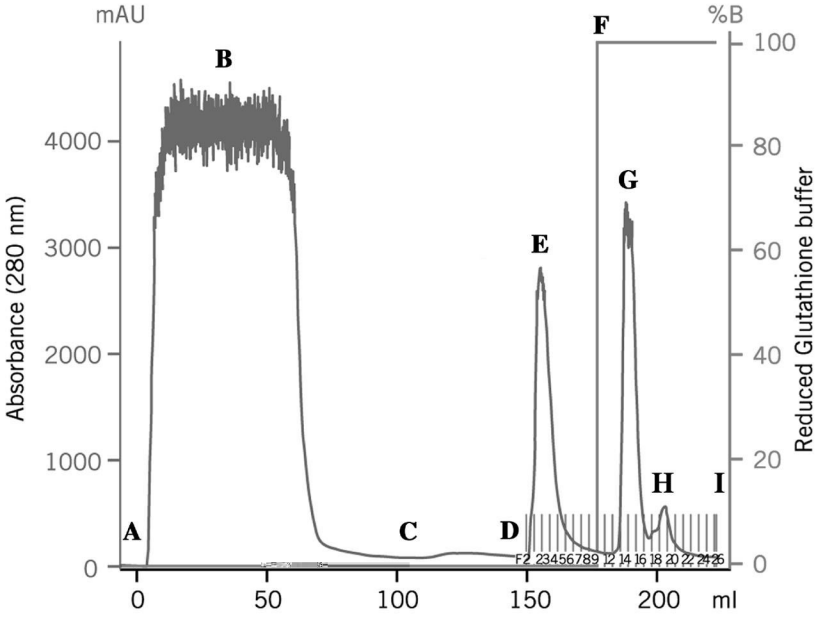

Fig. 2. Purification and on-column cleavage profile. A chromatogram depicting the purification and on-column cleavage profile of GST::TLP40 fusion protein. The purification strategy used a GSTrap affinity column and PreScission protease on an ÄKTA Explorer (Amersham Biosciences, Uppsala, Sweden) chromatographic system. Point A indicates the loading buffer (PBS at pH 7.4) and sample application; peak B, flow-through of unbound proteins in lysate; point C, following return of absorbance baseline (280 nm) and exchange to Cleavage buffer; point D, incubation of on-column cleavage reaction; peak E, eluted ''native'' TLP40; point F, exchange buffer to reduced glutathione buffer; peak G, eluted GST::linker and PreScission protease (Amersham Biosciences, Uppsala, Sweden); peak H, eluted impurities; point I, regenerated column.

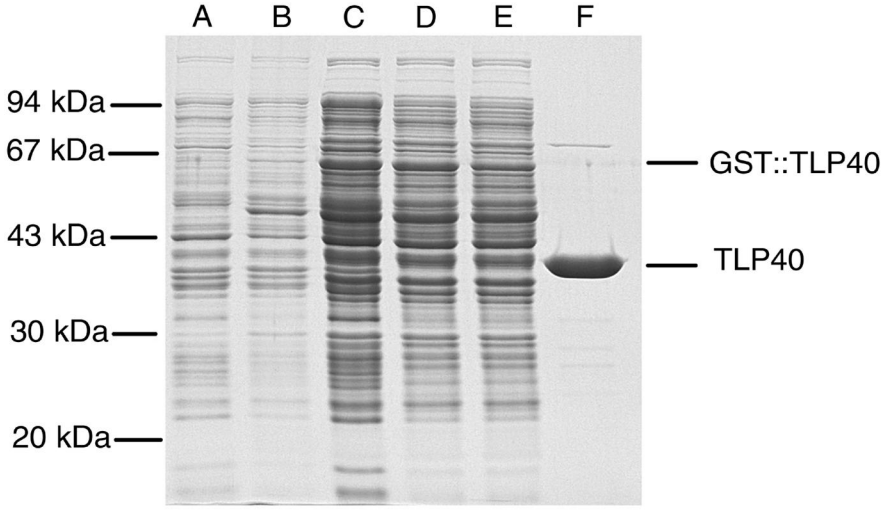

Fig. 3. SDS–PAGE profile of purification procedure. SDS–PAGE analysis profiling the purification and on-column cleavage procedure. Samples were prepared on a 3.5–12% stacking polyacrylamide gel (Fling and Gregerson buffer system [33]). Lane A, total protein extract on non-induced culture; lane B, total protein extract of induced culture; lane C, supernatant after 70 000 *g* centrifugation step; lane D, supernatant after 300 000 *g* centrifugation step; lane E, flow-through from GSTrap column; lane F, ''native'' TLP40 cleavage product eluted after purification and on-column digestion with PreScission protease.

Table 1
Summary of protein production and purity

Protein	Source	Cellular location	Function	Construct	Mr (kDa)	Yield (mg)	Recovery (% total)	Purity (%)
Pur-α	*Mus muscullus*	Nucleus/cytoplasm	-transcription regulation (DNA binding) -HIV-1 transactivation (RNA binding) -microtubule associated	Full-length	39	30	72	98
Translin	*Mus muscullus*	Nucleus/cytoplasm	-recombination regulation (DNA binding) -translational repressor (mRNA binding) -microtubule associated	Full-length	26	8	26	96
TLP40	*Spinacia oleraceae*	Lumen of thylakoid membrane	-Photosystem II regulation -signal transduction -protein folding	Full-length	40	32	81	98
				ΔLZD	35	15	41	95
				PBD	8	40	84	98
				CD	20	12	37	96
COUP-TFI	*Homo sapiens*	Nucleus	-nuclear hormone receptor -transcription factor (DNA binding)	aa 57–423	46	2	18	96
Lep	*Escherichia coli*	Inner membrane	-leader peptidase -transmembrane protein -signal peptide cleavage	Full-length	28	7	24	97
ecoKch	*Escherichia coli*	Inner membrane	-K^+ channel (homotetramer)	Full-length	46	0.4	9	95
ProW	*Escherichia coli*	Inner membrane	-member of the ATP-binding cassette (ABC) super-family of transporters -structural homology to GPCRs	Full-length	38	0.2	7	95

A chart summarizing protein production and purity yield for purification and on-column cleavage strategy of assayed GST::target proteins. All constructs were sub-cloned into the pGEX-6P family of vectors and over-expressed in E. coli BL21(DE3) (Stratagene, La Jolla, CA, USA) cell lines grown at 25, 30 or 37 °C (see Material and methods) shaking at 220 rpm in TUNAIR flasks (Shelton Scientific, Shelton, CT, USA).

quences at break-point junctions in chromosomal translocations [42] and functions as a translational repressor linking mRNAs to microtubules [43]. Translin poorly over-expresses in bacterial hosts, is difficult to purify to homogeneity and self-organizes in larger multimeric complexes.

3.1.3. TLP40

The photosynthetic regulatory protein TLP40, a cyclophilin-like protein in the lumenal compartment of the thylakoid membrane, is associated with the photosynthetic membrane of chloroplasts [44–46]. Using other fusion methods (polyhistidine tag), TLP40 protein has a history of forming inclusion bodies, aggregating resulting in inactive or denatured protein.

3.1.4. COUP-TFI

COUP-TFI is a member of the nuclear orphan receptor super-family involved in neurogenetic events [47–52]. Recombinant COUP-TFI is poorly over-expressed in bacteria, primarily insoluble, aggregates and is subject to rapid proteolytic breakdown rendering a highly unstable product.

3.1.5. Leader peptidase

Leader peptidase (Lep) is a polytopic integral membrane protein responsible for the processing of signal peptides [53,54]. Lep is a model membrane protein that is well suited to study membrane insertion and topology in E. coli hosts.

3.1.6. ProW

ProW is an E. coli bacterial homologue of G-protein coupled receptors (GPCRs) [55,56]. ProW is a difficult to over-express and purify membrane protein with seven transmembrane segments.

3.1.7. ecoKch

The potassium (K^+) channel ecoKch is a membrane protein controlling the flux of K^+ ions across the plasma membrane [57,58]. K^+ channels present a difficult protein expression and purification problem. The membrane proteins must be expressed and inserted with the correct topology in the membrane

and organized as macromolecular assemblies to be functional.

3.2. Target protein preparation

The desired gene product is assayed for over-expression levels as a fusion GST::target protein expressed primarily in the soluble, or membrane targeted fraction comparing normalized absorbance units ($OD_{600} \cong 0.4$ U; Fig. 3). For membrane proteins, the preparation of the crude membrane fractions prior to solubilization increases the target protein enrichment in cases where there is a low yield of over-expressed membrane protein. The supernatant following ultracentrifugation, enriched in fusion target proteins, is loaded on a GSTrap FF 5 ml affinity column, containing Glutathione Sepharose™ 4 Fast Flow (FF) media.

The clarified lysate is loaded using PBS buffer at pH 7.4 at a flow rate of ~1 ml/min. Following binding of GST::target protein to the GSTrap FF column and the return of the absorbance baseline, stringent washing with Cleavage buffer is performed. Cleavage buffer allows for the optimum proteolytic activity of the highly sequence specific PreScission protease. Following equilibration, the column flow is stopped and PreScission protease is injected onto the column at a flow rate of ~5–7 ml/min. This step is performed to obtain a uniform distribution of the protease throughout the column increasing the probability and efficiency of fusion protein cleavage. The even dispersion of proteolytic enzyme throughout the column is key to having both effective and efficient cleavage. Following incubation with PreScission protease, an auxiliary GSTrap FF 1 ml column is mounted downstream of the primary column. This column performs several roles in the target protein elution step: first, it buffers any start-up flow pressure peaks that may alter the baseline established prior to the cleavage reaction following proteolytic incubation; second, it creates a small void volume allowing for precise fraction collecting; and third, it captures any unbound GST-linked proteins that may contaminate the eluting target protein. The cleaved, ''native'' target protein is eluted immediately afterwards and the fractions are collected.

3.3. Purification and on-column cleavage strategy

A typical chromatographic profile of the purification and on-column cleavage scheme summarized from multiple purification runs to emphasize the robustness and reproducibility of the method, considering both models and difficult protein targets, is depicted in Fig. 2: fusion protein capture; on-column cleavage reaction; elution of ''native'' protein; and GST-reduced glutathione elution. As represented in Figs. 2 and 3, the over-expressed, GST::target protein is bound, efficiently cleaved on-column and eluted as a homogenous product. The purity of the eluted proteins is evaluated by SDS–PAGE, Coomassie-stained (Fig. 3). An example target protein chosen to best demonstrate this method is GST::TLP40 illustrated in Figs. 2 and 3. TLP40 is selected to represent the purification and on-column cleavage strategy due to its common purification characteristics and behaviour that best illustrates the isolated target proteins studied to date. After selecting for conditions which produce high levels of expressed GST::TLP40 protein, comparing non-induced and induced samples of total protein extract (Fig. 3—Lanes A and B) and monitoring soluble fusion protein levels during growth after induction, the GST::TLP40 containing lysate is clarified by subsequent centrifugation steps at 70 000 *g* (Fig. 3—Lane C) and at 300 000 *g* (Fig. 3—Lane D). Following the ultracentrifugation step, the clarified lysate containing GST::TLP40 fusion protein is loaded on a GSTrap FF column with PBS Binding buffer (Fig. 2—Point A), and the unbound fraction passes through the column (Fig. 2—Peak B and Fig. 3—Lane E). PBS Binding buffer is used to wash the GSTrap FF column until the absorbance baseline returns (Fig. 2—Point C). At this point the buffer is exchanged to Cleavage buffer. The Cleavage buffer acts primarily to equilibrate the GSTrap FF column prior to proteolytic cleavage by PreScission protease of the bound fusion protein and also acts to further wash the column removing unspecifically bound contaminants. The buffer exchanges the phosphate saline buffer (PBS) with a Tris-based buffer as well as introducing reducing agents (DTT). This buffer exchange plays an important role in downstream processes sensitive to phosphate salts, such as crys-

tallization condition screening and metal-dependent biochemical assays. After the absorbance baseline returns, PreScission protease is loaded on the GSTrap FF column and the system is in a ''closed'' position incubating for ~12 h (Fig. 2—Point D). Following incubation, the cleaved ''native'' TLP40 protein is eluted (Fig. 2—Peak E and Fig. 3—Lane F). After the absorbance baseline has returned, a 100% step gradient of reduced glutathione containing buffer, Elution buffer, is introduced (Fig. 2—Lane F) and acts as a competitor for GST binding sites. The bound GST::linker (proteolytic cleavage reaction product) and PreScission protease elute from the column (Fig. 2—Peak G) along with other contaminants having affinity for glutathione (Fig. 2—Peak H). The column is regenerated after the absorbance baseline has returned (Fig. 2—Point I) and can be re-equilibrated with Binding buffer for subsequent purification and on-column cleavage runs. This method can be scaled from small-scale (<1-ml cultures) to larger-scale (>20 l-cultures) in a linear manner and is directly proportional to materials and yield reproducing consistent levels of protein production. The overall purification and on-column cleavage strategy produces highly pure, enriched ''native'' TLP40 target protein with a yield of ~32 mg/l culture in the purity range of 98% pure protein completed in a single chromatographic step.

3.4. Discussion

This study presents an effective and flexible strategy for the purification of recombinant ''native'' target proteins. The technique is based on the properties of the pGEX-6P vector, an engineered plasmid that greatly facilitates the over-expression and purification of recombinant GST-fusion proteins. In this study we have described advances to previously existing recombinant protein purification methods. These modifications are highly flexible and have shown that they can be easily applied to suit a wide range of diverse target proteins, differing in function and overall chemical characteristics, and integrated into HTP protein purification applications.

This improved purification strategy takes the advantages of a fusion protein to enhance the solubility of the over-expressed protein and facilitate its purification. However, the disadvantages in having a foreign protein or peptide fusion incorporated in the end product are excluded in this method. This is of major interest in studies dealing with protein interactions probing substrates or ligands (drug target screening), lipid bilayers (membrane assays), nucleic acids (genetic machinery), carbohydrates (cell communication) or other proteins (macromolecular assemblies). Also of significant importance is the reduction of additional protein/peptide components fused to the target protein allowing for native protein conformation affecting biological function and complex formation. The contribution of additional amino acid residues (polyhistidines or other fusion tags) can contribute to increasing the overall flexibility of the purified protein disrupting or decreasing the effectiveness of structure determination experiments such as NMR spectroscopy studies resulting in noisy spectra due to highly mobile termini regions or X-ray crystallographic studies through disordered crystal packing contacts.

A significant advantage of the utilization of this strategy is the rapidity and simplicity of the procedure, starting from cell lysate, centrifugation and a single chromatographic step using GSTrap FF columns, resulting in a high level of enrichment of >95% pure target protein in its ''native'' form performed entirely at 4 °C. This is an important aspect when working with unstable target proteins, which can be highly labile and rapidly lose biological activity over longer purification steps. For the test cases described in this study, the soluble ''native'' target proteins were determined to be functionally active in nucleic acid binding, protein–protein interactions and/or enzymatic activity, and where the membrane target proteins were observed to be inserted with the correct topology in the isolated membrane compared to endogenous protein. These observations confer the functional and structural integrity of the purified ''native'' target proteins using this purification strategy. In addition, the integrity of the amino-terminal of the proteolytically processed target proteins was effectively cleaved as determined by MALDI mass spectrometry analysis.

This purification strategy can also be easily applied directly to accommodate HTP methodologies with minimal effort in an effective and efficient

manner. With its single chromatographic step, this process of producing highly enriched ''native'' target proteins becomes an attractive alternative compared to multi-dimensional chromatographic methods that are restrictive in terms of cost and efficiency. Today's generation of chromatographic systems allows for the purification, cleavage and elution of recombinant proteins in a fully automated process [59]. Currently, complete bioprocessing systems are commercially available, and are often considered standard equipment in many laboratories. The next generation of systems will directly couple protein expression and purification [60] to crystallization screening [61], functional (microchip assays) workstations [62] and structure determination [63,64].

In practice, the production of purified proteins on a HTP scale is defined by the quantity, quality and the number of the products required for a given application. For structural studies the production of target proteins is commonly the rate-limiting step. The protein yield must be significantly high, often in the tens of milligrams range, and the sample purity must be greater than 95%. Applying today's technology of HTP protein expression screening, molecular biology methods generating multitudes of protein variants and the rapid advances of modern chromatographic systems, it is rational to increase the rate by which target proteins can be purified. To satisfy these requirements, there must be a balance between protein yield, protein purity and the throughput of samples purified. A common element of HTP purification methods is affinity purification. A popular choice is metal-chelating affinity resins used in isolating polyhistidine-fused target proteins. This method is also simple, efficient and rapid. The major drawbacks of this method include: reduced purity levels; the requirement of further additional chromatographic steps including buffer exchange and/or desalting; and the final product remains ''non-native'' as a fused polyhistidine fusion protein (or if cleaved, see Introduction). A superior affinity purification method, in terms of yield, purity, throughput and ''native'' protein production exploits the GST-fusion protein method integrated with GSTrap FF chromatographic columns and proteolytic cleavage using PreScission protease. In a single chromatographic step, soluble, or membrane-targeted, ''native'' target proteins can be purified to high levels of purity (>95%), with a high yield (tens of milligrams/preparation) in a ''native'' form. The purification strategy has been applied in series to run on a currently existing chromatographic system (ÄKTA Explorer, Amersham Biosciences, Uppsala, Sweden) processing routinely, in an automated manner, multiple target proteins. In addition, the HTP approach has the invaluable aspect to purify seven target proteins per chromatographic platform per day. This throughput of protein production far exceeds current protein production methods in terms of scale, yield, purity and the suitability for downstream processing of the purified ''native'' target protein. Due to the scalability of the method, the magnitude of protein production can be directly controlled and proteins can be processed in a fully-automated manner. This method can also be applied on a smaller-scale, in terms of final product yield, in multi-well titre plates.

The application of the protein purification and on-column cleavage strategy described in this study is relevant to current structure determination projects that are highly dependent on reproducible, systematic production of pure protein in amounts to satisfy the needs of structural biologists. Furthermore, proteins purified following this strategy can immediately be incorporated into crystallization condition screening and is of suitable purity to successfully produce protein crystals (unpublished data D.B.). The purification process described in this paper may be easily adapted to suit a wide range of target proteins as well as accommodate novel purification processes, including HTP protein production and robotic automation. The success of structural genomics depends on the development of automated or semi-automated, simple, robust and inexpensive methods for protein purification.

Acknowledgements

The authors would like to thank Rosemarie Knaust and Marianne Carlsson for valuable discussions concerning experimental techniques. D.B. would like to thank Pär Nordlund for support and enthusiasm respecting method development. S.E. would like to thank Bertil Andersson for support. S.E. is supported by the Swedish Natural Science Research Council. T.U. was supported by grant BIO4-CT97-2357 from

the European Community. The authors wish to thank Gunnar von Heijne (Lep and ecoKch), Susanne Witt (ProW), Reinhold G. Hermann and Bertil Andersson (TLP40) for kindly providing clones and Catherine Martin for technical assistance.

References

[1] S. Markrides, Microbiol. Rev. 60 (1996) 512.
[2] J. Nilsson, S. Ståhl, J. Lundeberg, M. Uhlén, P.-Å. Nygren, Protein Expr. Purif. 11 (1997) 1.
[3] F. Baneyx, Curr. Opin. Biotechnol. 10 (1999) 411.
[4] R.B. Kapust, D.S. Waugh, Protein Sci. 8 (1999) 1668.
[5] H. Lilie, E. Schwarz, R. Rudolph, Curr. Opin. Biotechnol. 9 (1998) 497.
[6] K. Guan, J.E. Dixon, Anal. Biochem. 192 (1991) 262.
[7] K. Miyashita, M. Kusumi, R. Utsumi, T. Komano, N. Satoh, Biosci. Biotechnol. Biochem. 14 (1992) 746.
[8] E.R. LaVallie, J.M. McCoy, Curr. Opin. Biotechnol. 6 (1995) 501.
[9] S. Chong, Y. Shao, H. Paulus, J. Benner, F.B. Perler, M.-Q. Xu, J. Biol. Chem. 271 (1996) 22159.
[10] P. Jonasson, J. Nilsson, E. Samuelsson, T. Moks, S. Ståhl, M. Uhlén, Eur. J. Biochem. 236 (1996) 656.
[11] J. Nilsson, P. Jonasson, E. Samuelsson, S. Ståhl, M. Uhlén, P.-Å. Nygren, J. Biotechnol. 48 (1996) 241.
[12] M.-Q. Xu, F.B. Perler, EMBO J. 15 (1996) 5146.
[13] T. Gräslund, J. Nilsson, M. Lindberg, M. Uhlén, P.-Å. Nygren, Protein Expr. Purif. 9 (1997) 125.
[14] P.A. Walker, L.E. Leong, P.W. Ng, S.H. Tan, S. Waller, D. Murphy, A.G. Porter, Bio/Technology 12 (1994) 601.
[15] M.G. Cordingley, P.L. Callahan, V.V. Sardana, V.M. Garsky, R.J. Colonno, J. Biol. Chem. 265 (1990) 9062.
[16] E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz, D. Stüber, Bio/Technology 6 (1988) 1321.
[17] E. Hochuli, H. Döbeli, A. Schacher, J. Chromatogr. 411 (1987) 177.
[18] B.L. Brizzard, R.G. Chubet, D.L. Vizzard, Biotechniques 16 (1994) 730.
[19] T.P. Hopp, K.S. Prickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal, P.J. Conlon, Bio/Technology 6 (1988) 1204.
[20] A. Knappik, A. Plückthun, Biotechniques 17 (1994) 747.
[21] E.R. LaVallie, E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendel, J.M. McCoy, Bio/Technology 11 (1993) 187.
[22] Z. Lu, E.A. DiBlasio-Smith, K.L. Grant, N.W. Warne, E.R. LaVallie, L.A. Collins-Racie, M.T. Follettie, M.J. Williams, J.M. McCoy, J. Biol. Chem. 271 (1996) 5059.
[23] D.L. Wilkinson, N.T. Ma, C. Haught, R.G. Harrison, Biotechnol. Prog. 11 (1995) 265.
[24] J. Nilsson, P. Nilsson, Y. Williams, L. Pettersson, M. Uhlén, P.-Å. Nygren, Eur. J. Biochem. 224 (1994) 103.
[25] E. Samuelsson, T. Moks, B. Nilsson, M. Uhlén, Biochemistry 33 (1994) 4207.
[26] M. Uhlén, B. Nilsson, B. Guss, M. Lindberg, S. Gatenbeck, L. Philipson, Gene 23 (1983) 369.
[27] J. Nilsson, M. Larsson, S. Ståhl, P.-Å. Nygren, M. Uhlén, J. Mol. Recognit. 9 (1996) 585.
[28] T.G.M. Schmidt, A. Skerra, Protein Eng. 6 (1993) 109.
[29] H. Bedouelle, P. Duplay, Eur. J. Biochem. 171 (1988) 541.
[30] C. di Guan, P. Li, P.D. Riggs, H. Inouye, Gene 67 (1988) 21.
[31] C.V. Maina, P.D. Riggs, A.G. Grandea III, B.E. Slatko, L.S. Moran, J.A. Tagliamonte, L.A. McReynolds, C. diGuan, Gene 74 (1988) 365.
[32] F.M. Ausubel et al., Short Protocols in Molecular Biology, Wiley, New York, 1992.
[33] S.P. Fling, D.S. Gregerson, Anal. Biochem. 155 (1986) 83.
[34] M.M. Bradford, Anal. Biochem. 72 (1976) 248.
[35] A.D. Bergemann, E.M. Johnson, Mol. Cell. Biol. 12 (1992) 1257.
[36] S. Haas, J. Gordon, K. Khalili, J. Cell Biol. 130 (1993) 3103.
[37] C.P. Krachmarov, L.C. Chepenik, S. Barr-Vagell, K. Khalili, E.M. Johnson, Proc. Natl. Acad. Sci. USA 93 (1996) 14112.
[38] Q. Du, E. Tomkinson, P.D. Gardner, J. Biol. Chem. 272 (1998) 14990.
[39] C.H. Kuo, E. Nishikawa, H. Ichikawa, T. Sadakata, S.-Y. Niu, N. Miki, Biochem. Biophys. Res. Commun. 255 (1999) 406.
[40] L.G. Chepenik, A.P. Tretiakova, C.P. Krachmarov, E.M. Johnson, K. Khalili, Gene 210 (1998) 37.
[41] A.D. Frankel, C.A. Smith, Cell 92 (1998) 149.
[42] J.R. Han, G.K. Yiu, N.B. Hecht, Proc. Natl. Acad. Sci. USA 92 (1995) 9550.
[43] K. Aoki, K. Suzuki, T. Sugano, T. Tasaka, K. Nakahara, O. Kuge, A. Omori, M. Kasai, Nat. Genet. 10 (1995) 167.
[44] H. Fulgosi, A.V. Vener, L. Altschmied, R.G. Herrmann, B. Andersson, EMBO J. 17 (1998) 1577.
[45] A.V. Vener, A. Rokka, H. Fulgosi, B. Andersson, R.G. Herrmann, Biochemistry 38 (1999) 14955.
[46] A. Rokka, E-M. Aro, R.G. Herrmann, B. Andersson, A.V. Vener, Plant Physiol. 123 (2000) 1525.
[47] S.Y. Tsai, M.J. Tsai, Endocr. Rev. 18 (1997) 229.
[48] H. Connor, H. Nornes, T. Neuman, J. Biol. Chem. 270 (1995) 15066.
[49] Y. Qiu, F.A. Peirera, F.J. De Mayo, J.P. Lydon, S.Y. Tsai, M.J. Tsai, Genes Dev. 11 (1997) 1925.
[50] F. Adam, T. Sourisseau, R. Métivier, Y. Le Page, C. Desbois, D. Michel, G. Salbert, Mol. Endocrinol. 14 (2000) 1918.
[51] D.G. Kieback, T. Levi, P. Kohlberger, U. Friedrich, M.F. Press, H.E. Rosenthal, V.J. Mobus, I.B. Runnebaum, X.W. Tong, M.J. Tsai, Anticancer Res. 16 (1996) 3371.
[52] H. Shibata, T. Ando, T. Suzuki, I. Kurihara, M. Hayashi, I. Saito, H. Kawabe, M. Tsujioka, M. Mural, T. Saruta, Endocr. Res. 24 (1998) 881.
[53] C. Zwizinski, W. Wickner, J. Biol. Chem. 255 (1980) 7973.
[54] M. Paetzel, R.E. Dalbey, N.C. Strynadka, Nature 396 (1998) 186.
[55] P. Whitley, T. Zander, M. Ehrmann, M. Haardt, E. Bremer, G. von Heijne, EMBO J. 13 (1994) 4653.
[56] I. Nilsson, S. Witt, H. Kiefer, I. Mingarro, G. von Heijne, J. Biol. Chem. 275 (2000) 6207.

[57] R. Milkman, Proc. Natl. Acad. Sci. USA 91 (1994) 3510.
[58] M. Johansson, G. von Heijne, J. Biol. Chem. 271 (1996) 25912.
[59] A. Abbott, Nature 408 (2000) 130.
[60] A.M. Edwards, C.H. Arrowsmith, D. Christendat, A. Dharamsi, J.D. Friesen, J.F. Greenblatt, M. Vedadi, Nat. Struct. Biol. 7 (2000) 970.
[61] R.C. Stevens, Curr. Opin. Struct. Biol. 10 (2000) 558.
[62] G.C. Kennedy, EXS 89 (2000) 1.
[63] E. Abola, P. Kuhn, T. Earnest, R.C. Stevens, Nat. Struct. Biol. 7 (2000) 973.
[64] S.W. Muchmore, J. Olson, R. Jones, J. Pan, M. Blum, J. Greer, S.M. Merrick, P. Magdalinos, V.L. Nienaber, Structure Fold. Des. 8 (2000) R243.