Active membrane transport and receptor proteins from bacteria


Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK

****Corresponding author:
Peter JF Henderson
Astbury Centre for Structural Molecular Biology,
School of Biochemistry and Microbiology,
University of Leeds,
Leeds LS2 9JT, UK
Phone: 44-113-343-3175;
Fax: 44-113-1407
Email: p.j.f.henderson@leeds.ac.uk

*Department of Bacteriology and Immunology
Veterinary and Agrochemical Research Centre
VAR-CODA-CERVA
Groeselenberg 99
B-1180 Ukkel
Belgium

**Aston Academy of Life Sciences
Aston University
Aston Triangle
Birmingham B4 7ET
UK
Key words: membrane transport proteins, two-component systems, pathogens, *Helicobacter pylori*

Abbreviations: dodecyl-β-D-maltoside, DDM; Nickel nitrilo-triacetate, NiNTA.
Abstract
A general strategy for the expression in *Escherichia coli* of bacterial membrane transport and receptor genes is described. Expression is amplified so that the encoded proteins comprise 5-35% of *E. coli* inner membrane protein. Depending upon their topology, proteins are produced with RGSH₆ or a Strep-tag at the C-terminus. These enable purification in mg quantities for crystallisation and NMR studies. Examples of one nutrient uptake and one multidrug extrusion protein from *Helicobacter pylori* are described. This strategy is successful for membrane proteins from *Helicobacter pylori*, *Escherichia coli*, *Enterococcus faecalis*, *Bacillus subtilis*, *Staphylococcus aureus*, *Microbacterium liquefaciens*, *Brucella abortus*, *Brucella melitensis*, *Campylobacter jejuni*, *Neisseria meningitides*, *Streptomyces coelicolor* and *Rhodobacter sphaeroides*. 

©2005 Biochemical Society
Introduction

The lipid cell membrane of bacteria is inherently impermeable to nutrients required for metabolism. Uptake of nutrients (and secretion of wastes) therefore depends on the presence of transport proteins, activities of which are typically coupled to metabolic energy to drive transport against the prevailing electrochemical gradient of solute. Examples are ATP-dependent primary active transport, sugar-H⁺ or antibiotic/H⁺ secondary active transport and phosphotransferase [1] (Figure 1). In addition the bacterial membrane contains proteins that sense environmental conditions and, through the ‘two-component’ sensor/response system (TCS), facilitate an appropriate response of the cells [2] (Figure 1). In most cases the low abundance of these membrane proteins and their hydrophobic nature make them difficult to isolate in amounts required for elucidation of their 3D structure. Determination of structure is a major bottleneck in understanding the molecular mechanisms of membrane transport and sensor proteins, which comprise 5-15% of genomic potential in all organisms.

In this article a strategy is described that enables the amplified expression and purification of bacterial membrane transport and receptor proteins in amounts required for structural studies; the strategy is successful for many such proteins (Table 1). Many of these are prokaryote proteins homologous to those found in numerous organisms from cyanobacteria, eubacteria, protozoan parasites, fungi, plants and mammals [3]; the convenience of structure-activity studies in bacteria then illuminate the molecular mechanism of transporters in numerous organisms including man. Other transport systems and TCS’s are unique to bacteria and critical for growth of pathogenic organisms during infection; the availability of the purified active protein may then be useful for discovery of novel antibacterials.

Only two examples of amplified expression are illustrated here, an α-ketoglutarate transport protein (KgtP) and a putative multidrug resistance protein (Mdr), both originating from Helicobacter pylori. Tomb et al. [4] determined the complete genome sequence of the H. pylori strain 26695, later compared with that of the pathogenic strain J99 [5]. One of the genes found was JHP0334. This was thought to encode an α-ketoglutarate membrane transport protein, because its predicted amino acid sequence aligned with 40-44% identity to the known ‘KgtP’ protein from Escherichia coli [3, 6]. Homologues of KgtP also occur in Campylobacter jejuni, Salmonella typhimurium, Pseudomonas putida, Salmonella typhi and Brucella suis. The JHP0334 gene product is predicted to comprise 437 amino acids. Similarly, the JHP1092 gene
product is thought to encode an efflux protein for drugs and/or antibiotics of 386 amino acids. Both proteins belong to the widespread Major Facilitator Superfamily of transport proteins [7] and their hydropathy profiles suggest they are integral membrane proteins arranged in 12 transmembrane \( \alpha \)-helices. Their transport is likely to be energised by the transmembrane electrochemical potential of \( H^+ \), but this is not proven; where the nature of the cation linkage is thought to be established it is indicated in Table 1.

**Materials and methods**

The materials and methods used in this work are described in [8-12]. Genomic DNA from *H. pylori* strains J99 and 26695 was kindly provided by D. E. Berg (St Louis), from *Campylobacter jejuni* and *Neisseria meningitidis* by M.C.J. Maiden (Oxford), from *Brucella abortus* by R.C. Essenberg (Ohio) and from *Brucella melitensis* by K Walravens (Brussels). *Bacillus subtilis* was kindly provided by A. Moir (Sheffield). n-Dodecyl-beta-D-maltoside (DDM) was from Melford Labs, Ipswich, Suffolk. Nickel nitrilo-triacetic acid (NiNTA) and an antibody against the RGSH\(_6\) epitope were from QIAgen, Ltd., Dorking, Surrey.

*E. coli* strains XL1-Blue Stratagene\(^\text{TM}\) (*recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac [F'proAB lacI\(^\text{q}\) ZAM15, \( \text{Tn10 (Tet}^\text{R}) \)]) and BL21 Novagen\(^\text{TM}\) (F\(^-\) ompT hsdS\(_B\)(r\(^B\) m\(^B\)) gal dcm (DE3)) were used as hosts for transformation work. *E. coli* BL21(DE3) was also used for small- and large-scale isolation of plasmids, and for over-expression and purification of transport proteins.

Maintenance and growth of these *E. coli* strains was achieved by culturing the bacteria in Luria Broth (LB) [9] liquid medium, or in minimal salts medium [9] containing 20mM glycerol, or on plates containing 1.5% agar. Carbenicillin (at least 100 \( \mu \)g/ml) was used throughout all stages of growth to maintain plasmid integrity.

For small-scale investigation of protein expression, 50ml cultures in 250ml flasks were used. Total membranes were prepared from sphaeroplasts by the water lysis method [9]. Inner membrane vesicles were prepared from 500ml cultures in 2l baffled conical flasks or from 25l fermentor cultures. After harvesting the cells were disrupted by explosive decompression using a
French pressure cell. The inner and outer cell membranes were separated by sucrose density centrifugation, followed by washing in buffer to remove the sucrose and EDTA [9].

For both small-scale and inner membrane production of E.coli strains, growth was allowed to continue until the cell density reached an A₆₈₀ of approximately 0.6 when the expression of the cloned gene was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5mM. Growth was continued for 3-4hrs following the induction of the tac promoter, thus producing optimal yields of protein.

Results

Introduction of the gene encoding a putative symport or antiport protein into the plasmid pTTQ18

The genes putatively encoding the α-ketoglutarate or the Mdr transport proteins were amplified from the H. pylori genomic DNA, using appropriate mutagenic oligonucleotides. These were designed to introduce an EcoRI site at the 5' end and a PstI site at the 3' end to promote subsequent ligation with the 4.56kb pTTQ18/RGSH₆ fragment. The PCR product was isolated from an agarose gel and then digested with EcoRI and PstI.

In order to clone each of the genes into the pTTQ18 plasmid vector [13], the plasmid pNorAH₆ (pTTQ18 containing the gene norAH₆ [14], Figure 2) was isolated from E. coli strain BLR and digested with the restriction endonucleases EcoRI and PstI to yield two DNA fragments of 4.56kb and 1.2kb. The larger fragment (pTTQ18 with the RGSH₆ coding DNA sequence) was isolated from an agarose gel.

Ligation reactions were performed using the EcoRI-PstI digested gene and pTTQ18-RGSH₆ fragments at various vector:insert molar ratios. After the ligated product was transformed into E. coli XL1-blue, recombinant clones were selected on LB plates containing carbenicillin. Plasmid DNA was prepared from carbenicillin-resistant colonies and subjected to restriction analysis with HindIII, NcoI (not shown) uniquely-cutting restriction enzymes and automated DNA sequencing of the 5’ end. to confirm the correct size of the plasmid and the presence of each gene [11].
size and DNA sequence of the inserts in the new plasmids (Figure 2) were confirmed. The plasmid was then transformed into E. coli strain BL21(DE3) for expression studies.

This procedure can be applied to any gene that does not include EcoRI or PstI restriction sites. If these sites are present in the coding region then EcoRI and PstI can still be introduced as flanking sites and partial digestion used to obtain a fragment uncut at the internal site(s), or a two-step procedure adopted as in [12]. Alternatively, other flanking restriction sites can be chosen compatible with the multi-cloning site in pTTQ18 [9, 13], for example NdeI at the 5’ and HindIII at the 3’end of the gene. If the C-terminus is predicted to lie in the periplasm then a Strep-tag may well be successful if the RGSH$_6$ tag is not.

Detection of expressed histidine-tagged transport protein in E. coli membrane preparations

E. coli BL21(DE3) cells harbouring each plasmid were cultured in LB broth plus 20mm glycerol and expression trials performed with concentrations of IPTG 0.0-1.0 mM [11]. 0.5mM IPTG is sufficient for maximal expression of the putative KgpP(His)$_6$ and JHP1092 Mdr proteins.

After preparing membrane samples [9, 10, 11], SDS-PAGE analysis and staining with Coomassie brilliant blue, an IPTG-inducible protein is observed migrating at approximately 35 kDa for KgpP(His)$_6$ and at approximately 31 kDa for JHP1092(His)$_6$ (Figures 3A and 3B). It is typical for membrane transport proteins to migrate at 65-75% of their true molecular weight, possibly as a result of their hydrophobicity, high binding of SDS or the retention of secondary structure retarding passage through the gel mixture [9]. This is also the case for membrane sensor proteins of TCS, though the extent of anomalous migration behaviour is not as great, possibly due to fewer TM regions in these proteins compared with membrane transport proteins. The predicted molecular weight of KgpP(His)$_6$ is 50,245.3 Da, so the migrating protein would be expected at 32-37 kDa, and for JHP1092(His)$_6$ it is 44,994 kDa so the expected apparent mass is 29-34 kDa. Scanning densitometry analysis showed that the induced proteins were expressed at 20% of total membrane proteins, indicating that over-expression has occurred, whereas the protein at the same position in the uninduced sample is expressed at 3% of inner membrane protein. Their identities in the membranes were further confirmed by Western blotting [10, 11].

Solubilisation and purification of histidine-tagged transport proteins

©2005 Biochemical Society
Initial purification of KgtP(His)₆ or JHP1092(His)₆ using Immobilised Metal Affinity Chromatography (IMAC) to exploit the C-terminal His-tag (Methods) show proteins migrating at 35 or 31 kDa, respectively in the eluted fractions following Coomassie blue or silver-staining (Figure 3A & 3B), similar to the IPTG-inducible proteins migrating in the inner membrane fractions (Figure 3A & 3B).

In order to reinforce identification and integrity of the isolated proteins the gels were subjected to Western blot analysis using an antibody to the C-terminal RGS(His)₆ tag. Both purified proteins reacted with the antibody (Figure 3A & 3B). N-Terminal sequencing further confirmed their identification and, importantly, the integrity of the amino acid sequences of the purified proteins (MNSH MNPQIQ for KgtP(His)₆ and MNSH MLRKNILAY for JHP1092(His)₆, the first four amino acids are from the LacZα peptide). Some minor contaminants (hardly visible on silver staining) in KgtP(His)₆ preparations of lower Mr were positive with the antibody (Figure 3), suggesting that some breakdown of the purified protein may have occurred, while bands of higher Mr may be oligomers of each protein (Figure 3A & 3B).

Thus, the uncertainty of the identity of the protein resulting from the anomalous migration on SDS-PAGE gels is overcome by the combined detection of the predicted N-terminal sequence and Western blot analysis for the C-terminal RGS(His)₆ tag, confirming that the proteins migrating at 31-35 kDa are the ones desired and that they have not been degraded during isolation. These are generic tests of integrity applicable to any -RGS(His)₆-tagged proteins.

The conditions for solubilisation and purification, i.e. DDM concentration for solubilisation and imidazole concentration for washes and elution, vary depending on the characteristics of each of the transport proteins. However, the generic conditions described here for one symporter and one antiporter have proved feasible and provided an initial index of yield and purity.

**Retention of structure and activity is confirmed by circular dichroism spectroscopy and activity assays of reconstituted protein.**

CD spectroscopy is a useful technique for the detection of secondary structure within proteins, although the quantification of the proportions of different structural elements is severely limited for membrane proteins [15, 16]. The CD spectra obtained for the purified
reconstituted (His)$_6$-tagged JHP0334 and JHP1092 proteins [10, 11] reveal a predominantly $\alpha$-helix content, confirming retention of the secondary structure during purification of the protein. Furthermore, the purified reconstituted KgtP(His)$_6$ protein catalysed $\alpha$-ketoglutarate counterflow into liposomes [10, 11].

In the case of TCS proteins activity of the protein expressed in membranes and also after purification can be established by direct assays of phosphorylation of the protein [12].

**Wider application of the strategy for amplified expression and purification of membrane proteins**

The same strategy, with minor modifications in growth and purification conditions, has been used for overexpression of other membrane transport and TCS proteins from *H. pylori* [e.g. 17], *E. coli* [e.g. 12, 18] and other bacteria, both Gram-negative and Gram-positive (Table 1). Out of 40 attempts, 34 have so far been successful for amplification of expression, i.e. the induced cloned protein comprised at least 15% of the inner membrane preparations of the *E. coli* host strain, and in every case (so far 16) where the (His)$_6$ tag was added at the C-terminus the protein has been purified successfully by NiNTA chromatography. In some cases where a parallel construct with the (His)$_6$-tag at the N-terminus, instead of the C-terminus, was made, the level of expression was substantially reduced.

For each protein growth conditions should be optimised in 1-25 litre cultures of *E. coli* host strains, testing both minimal and complex media at temperatures between 25-37°C to maximise expression [9]. The concentration of IPTG required is tested between 0.1-1.0mM, and the period of growth before induction varied to obtain as high a cell density as possible commensurate with optimal protein expression (growth is often diminished, or abolished, after induction). Similarly, the period of exposure to IPTG (2-24 hours) is investigated in order to promote maximum expression. In many cases a 25l fermentor can conveniently be used without compromising expression, but for some proteins the level of expression is always higher in batch cultures of 500-800ml in 2l baffled flasks. Further examination of the parameters regulating growth and protein production in these conditions may enhance our understanding of expression and enable us to reproduce the complex growth behaviour in flasks in the more controlled environment of a fermentor.
Conclusions

The prime purpose of this review is to illustrate a generic procedure for obtaining sufficient quantities of correctly folded transport and receptor protein(s) from a variety of microorganisms, including pathogens, for structural studies. This has been achieved for 12 organisms, both Gram positive and Gram negative, and for 34 proteins (Table 1). It is possible to proceed from identification of a gene encoding a membrane protein in a bacterial genome to the production of mg of purified protein in a few weeks, and the application of higher throughput methods for cloning and purification will hopefully reduce the time required. The yield and purity of protein may well be increased by further optimisation of conditions, especially for detergent extraction of protein from the membrane. The yields are easily enough for trials to form 2D ordered arrays for electron crystallography [19, 20], and sufficient to start 3d crystallisation trials for Xray crystallography (Shibayama, Byrne, Iwata, Suzuki, Rutherford, O’Reilly and Henderson, unpublished) and for NMR studies [21,22].

In addition, the purified protein can be examined by a variety of biophysical techniques – mass spectrometry for precise Mr and sequence determination [23, 24], FTIR, fluorimetry, calorimetry and EPR that elucidate the structure-activity relationship, especially when performed in conjunction with directed mutagenesis and genetic recombination.

The activity and bioenergetics of each cloned transport protein can be confirmed and/or investigated in IPTG-induced E. coli hosts, by measuring transport of radioisotope-labelled substrates [e.g. 10,11] or phosphorylation kinetics of TCS proteins [e.g. 12]. If there is no significant transport activity in an uninduced strain, but substantial activity in the IPTG-induced one, then any E. coli wild-type host and vesicles prepared therefrom can be used to investigate cation dependence, substrate specificity, susceptibility to inhibitors, etc. Nevertheless, it is
preferable to use an *E. coli* host itself attenuated in the relevant transport system. Expression of genes encoded in the pTTQ18 vector is independent of the *E. coli* host in our experience so far, and movement between host strains by transformation is easily accomplished. By this means we have established, for example, that the PutP(His)$_6$ protein from *H. pylori* catalyses proline-$\text{Na}^+$ symport (11).

In recent landmark papers the 3D structures of the lactose and $\alpha$-glycerophosphate transport proteins (Fig. 1) of *E. coli* have been determined by X-ray crystallography [24, 25]. Future such elucidation of the structures of transport and other membrane proteins from the additional organisms described here may uncover means of preventing or treating bacterial infections.

**Acknowledgements**

This work was funded by BBSRC, EU, Ajinomoto Co., SmithKline Beecham (SB, now GlaxoSmithKline, GSK) and SWEGENE and by equipment grants from the Wellcome Trust and BBSRC. MS is grateful to the Iranian Government for financial support. BBSRC provided PhD studentships for JC, GP and SM, and MRC provided studentships for CKH and SLP. We thank Dr JG Keen for N-terminal amino acid sequencing of the purified proteins. We also acknowledge the constructive interest of Dr M Gwynne (GSK), Dr F Paul (GSK), Dr G Badman (GSK) and Dr Watanabe (Ajinomoto Co.). PJFH is grateful for a Leverhulme Research Fellowship and KS for a JSPS/RS Travel Fellowship.
References

Figure legends

**Figure 1.** Scheme for transport mechanisms and sensor kinases in bacteria. The oval represents the cytoplasmic membrane of a bacterium containing the enzymes of ATP synthesis and respiration on the left and in a clockwise direction: two symporters; two antiporters; a phosphotransferase; two ABC transporters for efflux and influx; and a TCS sensor kinase.

**Figure 2.** Cloning strategy for membrane proteins using plasmid pTTQ18. Details are given in refs [8-14].

**Figure 3.** (A) Purification (and identification) of the *H. pylori* KgtP(His)$_6$ ketoglutarate transport protein. Membrane preparations were made from IPTG-induced *E. coli* BL21(DE3) (pTTQ18$kgtP$). Samples of the original preparation ('Membrane'), the proteins solubilised in 1.5% DDM ('Soluble') and insoluble ('Pellet') material were examined by SDS-PAGE (Coomassie-stained). A sample of the material that failed to adhere to Ni$^{2+}$ NTA-agarose ('Flowthrough') and the Ni$^{2+}$ NTA-bound protein subsequently eluted by 200mM imidazole ('Protein') were processed in the same gel. The eluted protein, 0.1µg and 0.5 µg, was also tested for reaction with anti-RGSH$_6$ antibody by ‘Western blotting’.

(B) Purification (and identification) of the *H. pylori* JHP1092(His)$_6$ putative multidrug resistance protein. Samples from membranes of IPTG-induced *E. coli* BL21(DE3) (pTTQ18$jhp1092$) were prepared and are labeled as in A. The purified protein, 1µg and 2 µg was also tested for reaction with an anti RGSH$_6$ antibody by ‘Western blotting’.

**Table 1.**

Examples of 34 overexpressed proteins and their 12 organisms of origin

<table>
<thead>
<tr>
<th>Protein</th>
<th>Organism</th>
<th>Major substrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>KgtP (JHP0334)</td>
<td><em>Helicobacter pylori</em></td>
<td>Ketoglutarate</td>
</tr>
<tr>
<td>“ProP”</td>
<td><em>Helicobacter pylori</em></td>
<td>?</td>
</tr>
<tr>
<td>GluP</td>
<td><em>Helicobacter pylori</em></td>
<td>Glucose</td>
</tr>
<tr>
<td>NupC</td>
<td><em>Helicobacter pylori</em></td>
<td>?</td>
</tr>
<tr>
<td>PutP</td>
<td><em>Helicobacter pylori</em></td>
<td>Proline-Na$^+$</td>
</tr>
<tr>
<td>NixA</td>
<td><em>Helicobacter pylori</em></td>
<td>Nickel</td>
</tr>
</tbody>
</table>

©2005 Biochemical Society
<table>
<thead>
<tr>
<th>HP1092</th>
<th>Helicobacter pylori</th>
<th>Multidrugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP1181</td>
<td>Helicobacter pylori</td>
<td>Multidrugs</td>
</tr>
<tr>
<td>AraE</td>
<td>Escherichia coli</td>
<td>Arabinose-H⁺</td>
</tr>
<tr>
<td>XylE</td>
<td>Escherichia coli</td>
<td>Xylose-H⁺</td>
</tr>
<tr>
<td>GalP</td>
<td>Escherichia coli</td>
<td>Galactose-H⁺</td>
</tr>
<tr>
<td>ProP</td>
<td>Escherichia coli</td>
<td>Proline-H⁺</td>
</tr>
<tr>
<td>PutP</td>
<td>Escherichia coli</td>
<td>Proline-Na⁺</td>
</tr>
<tr>
<td>Bcr</td>
<td>Escherichia coli</td>
<td>Bicyclomycin</td>
</tr>
<tr>
<td>FucP</td>
<td>Escherichia coli</td>
<td>Fucose-H⁺</td>
</tr>
<tr>
<td>GusB</td>
<td>Escherichia coli</td>
<td>Glucuronide-H⁺</td>
</tr>
</tbody>
</table>

| “ProP” (Cj0250c) | Campylobacter jejuni | ? |
| FucP (Cj0486)    | Campylobacter jejuni | Fucose |
| Nma2100          | Neisseria meningitidis | Sugar? |
| GluP (Nma0714)   | Neisseria meningitidis | Glucose |

| “Bcr” (Nma2040) | Neisseria meningitidis | Multidrugs? |
| GluP            | Brucella abortus       | Glucose |
| KgtP            | Brucella melitensis    | Ketoglutarate |
| LmrB            | Brucella melitensis    | Lincomycin |
| NorA            | Staphylococcus aureus  | Multidrugs |
| Mj 1560         | Methanococcus janaschii | Multidrugs? |
| Mhp1            | Microbacterium liquefaciens | Hydantoins |
| AraE            | Bacillus subtilis      | Arabinose |
| Bmr             | Bacillus subtilis      | Multidrugs |
| Blt             | Bacillus subtilis      | Multidrugs |
| VicK            | Enterococcus faecalis  | Signal? |
| VanSₐ           | Enterococcus faecalis  | Signal? |
| PrrB            | Rhodobacter sphaeroides | Redox potential |
| SpdB            | Streptomyces coelicolor | DNA? |
Secondary active transport

Symport

$\alpha$-KG

$H^+$ Lactose

$H^+$

Antiport

GlycerolP

$P_i$

Phosphotransferase

Sensor kinase

Respiration

ATP synthase

Primary active transport

Nutrient, K$^+$

Toxin

SugarP

Sugar
Figure 2
©2005 Biochemical Society
Figure 3A
Figure 3B

Western blotting