AcrB et al.: Obstinate contaminants in a picogram scale.

One more bottleneck in the membrane protein structure pipeline.

Georgios Psakis\textsuperscript{1,2}, Julia Polaczek\textsuperscript{1}, Lars-Oliver Essen\textsuperscript{1}

\textit{Department of Chemistry}\textsuperscript{1}, Philipps-University Marburg, Hans-Meerwein-Strasse, 35033 Marburg, Germany

\textit{Department of Plant Physiology}\textsuperscript{2}, Justus-Liebig University Giessen, Schenkenbergstrasse 3, D-35390 Giessen, Germany

Address correspondence to: Lars-Oliver Essen, Department of Chemistry, Philipps-University Marburg, Hans-Meerwein-Strasse, 35033 Marburg, Germany; e-mail: essen@chemie.uni-marburg.de; fax: +49 6421 2822012.
Abstract

Heterologous expression of integral membrane proteins from *Helicobacter pylori* 26695 in *Escherichia coli* enabled the identification of 17 candidates for purification and subsequent crystallization. 45% of the purified proteins were contaminated with what was later identified as the multidrug efflux pump (AcrB) of *E. coli*, and 17% with the succinate dehydrogenase. While additional purification steps ensured removal of succinate dehydrogenase, they failed to remove AcrB completely, leaving picogram amounts present in fractions intended for 3D-crystallization. Two of these targets, the Na⁺ dependent D-glucose/D-galactose transporter (GluP-HP1174) and the carbon starvation protein A (CstA-HP1168), produced small crystals (<40 µm). Crystals from the GluP preparation diffracted to 4.2 Å resolution and belonged to the rhombohedral space group *H*32. Subsequent molecular replacement proved that these crystals were derived from a contaminant, the efflux transporter AcrB. This unexpected crystallization of AcrB from picogram amounts was observed in six new conditions. The systematic occurrence of AcrB in membrane preparations stems from the upregulation of its transcription in response to the stress induced by the expression of a selected target. This, along with its tendency to crystallize in the picogram scale, poses a serious concern in membrane protein expression using heterologous hosts harbouring AcrB.

*Keywords:* AcrB; succinate dehydrogenase; heterologous expression; picogram crystallization; membrane protein contaminants; structural genomics.
In recent years membrane protein structural genomics has focused on increasing the number of targets going through the pipeline from gene to structure in a rapid and cost-efficient way. Often, however, many of the standard quality controls are bypassed in order to ensure the fast screening of the highest number of targets possible. In previous work (Psakis et al., 2007), we proposed that a medium-throughput genomics approach can be more effective, as it allows for the fine-tuning and control of parameters that can interfere with the efficiency of expression and solubilisation of a selected target and with the quality of its purification. The latter approach was used for the identification of suitable *H. pylori* 26695 integral membrane proteins for structural studies (Psakis et al., 2007). The Na\(^+\)/H\(^+\) D-glucose/D-galactose symporter (GluP-HP1174) and the carbon starvation protein A (CstA-HP1168) were two of those proteins.

GluP is the only porter responsible for Na\(^+\)-driven D-glucose/D-galactose transport in *H. pylori* (Mendz et al., 1995; Tomb et al., 1997). Since no mammalian hexose-transporters show significant alignments with GluP of *H. pylori*, and glucose transport into *H. pylori* is not susceptible to inhibitors that affect mammalian transporters (Mendz et al., 1995), this protein represents an ideal target for structure-based drug design. CstA, on the other hand, belongs to a family of putative integral membrane transporters predicted to be responsible for the entry of amino acid and/or carboxylic-based substrates into the bacterium (Tomb et al. 1997). The latter substrates provide the main and preferred carbon energy source in *H. pylori*. Despite extensive research on the carbon-starvation metabolism of bacterial hosts, however, little is known about the structure-function relationship of these transporters making them thus to attractive targets for future research.

Here, we report the co-purification of selected targets with the multi-drug efflux pump (AcrB) and/or succinate dehydrogenase. Both contaminating proteins display similar features: 1) the levels of their gene transcription can be induced under standard experimental conditions, 2) they show affinity to Ni\(^{2+}\) immobilised matrices, and 3) they crystallize in the
same space group. Although additional purification steps can result in the complete removal of succinate dehydrogenase, only partial removal of AcrB is possible and the presence of even picogram amounts of the latter is enough to drive formation of its rhombohedral crystal form.

GluP(His)$_6$ was purified at 4°C by high-flow Ni$^{2+}$-NTA affinity chromatography (Qiagen) compatible with the Äkta Prime purifier (Amersham Biosciences). The protein was extracted from the inner membranes of BL21(DE3) E. coli cells in 50 mM Tris-HCl (pH 8.5), 300 mM NaCl, 0.2 mM MgCl$_2$, 10% glycerol, 20 mM β-mercaptoethanol, 10 mM imidazole and 0.1% dodecyl-β-D-maltopyranoside (DDM-Applichem) – buffer A- by gentle mixing at 4°C for 16 hours. Following centrifugation at 109,000 g for 1 hour at 4°C, the supernatant containing the solubilised fractions was filter-sterilized (0.22 μm filter-Millipore) and was loaded to a Ni$^{2+}$-NTA column previously equilibrated with 10 column volumes of buffer A. The bound protein was further washed with 60-80 column volumes of buffer A, and eluted with a step gradient created by mixing buffers A and B (buffer A with 500 mM imidazole). Fractions containing the tagged protein were collected and desalted in a PD-10 column (Amersham Biosciences). The protein was eluted in 25 mM sodium-citrate buffer (pH 5.0), 5% glycerol, 20 mM β-mercaptoethanol and 0.05% DDM and concentrated to 6-8 mg.ml$^{-1}$ by using a centricon Amicon with a 50 kDa cutoff (Millipore). For the final polishing step, the purified protein was subjected to ion-exchange chromatography (Sartorius) to ensure the removal of excess detergent and other protein contaminants. CstA(His)$_6$ was solubilised in 2% decyl-β-D-maltoside (DM; Applichem), and was concentrated to 4-6mg.ml$^{-1}$ in a buffer containing 50 mM Hepes (pH 7.5), 5% glycerol, 2 mM β-mercaptoethanol, and 0.2% DM.

The purification efficiency for both proteins was assessed by 15% SDS-PAGE electrophoresis (Fig. 1). The identity of both the desired and contaminating proteins was confirmed by Western blotting and MALDI-TOF analysis of their trypsin-generated products.
Crystallization screenings were carried out with a Cartesian Mycrosys 4004 robot (Genomic Solutions), using the sitting drop method in a 96-well plate set-up. Protein solution (4-8 mg.ml\(^{-1}\)) was mixed in a 1:1 ratio with mother liquor (0.05 µl protein plus 0.05 µl liquor). More than 864 conditions from screens from Qiagen and Sigma were tested per batch of purified protein. Samples were incubated at 289 K. In the case of GluP(His)\(_6\) preparation, crystals appeared after 4 months and kept growing for further 2 months (Fig. 2). Crystals were reproducible with different protein batches, and optimisation of the original conditions (lowering the PEG content by 2%; Pos and Diederichs, 2002) resulted in mature crystals (30-50 µm) within two weeks. In the case of the CstA(His)\(_6\) preparation, crystals appeared after 2 weeks and reached a maximum size of 40 µm. Successful crystallisation conditions are listed in Table 1. Several pyramidally shaped crystals were screened at beam-lines ID23 (ESRF, Grenoble) and PX06SA (SLS, Villigen) and found to diffract to 3.8-4.6 Å resolution. One complete dataset was recorded from a single putative GluP crystal at beam-line PX06SA at the SLS, Villigen (Table 2). After indexing (Powell, 1999) scaling (Bolotovsky et al. 1998) and taking radiation damage, during data collection, into account (300 images; 0.5\(^{\circ}\) oscillation, 1 s exposure) the crystals were assigned to the rhombohedral space group \(R32\) (hexagonal setting \(H32\), \(a=b=146.5\) Å, \(c=515\) Å) and diffracted up to at least 4.2 Å resolution (Fig. 2B). Overall, the crystals comprise three or four monomers (MW 43951 Da) per asymmetric unit with a predicted solvent content of 69% or 59%, respectively.

Initial attempts to solve the GluP(His)\(_6\) structure by molecular replacement (MOLRE - Vagin and Teplyakov, 1997; ccp4i - Lesslie, 1992; Potterton et al., 2003) using the \(E. \ coli\) lactose permease (1PV6) or the \(E. \ coli\) glycerol-3-phosphate transporter (1PW4) models failed. This was not surprising given the low sequence identity (≤12%) to the corresponding proteins and the limited data quality. At this stage, however, and in order to exclude the possibility of the crystallization of a contaminating protein, we searched for membrane proteins crystallizing in the \(R32\) space group or its hexagonal setting \(H32\). The \(E. \ coli\)
acriflavine resistance protein B (AcrB, PDB code 1IWG, a=b=144.5 Å, c=519.2 Å) and *E. coli* respiratory complex II (succinate dehydrogenase, PDB code 1NEK, a=b=138.8 Å, c=521.9 Å) were two of the proteins we identified. Parallel to this approach we isolated three barely visible bands of the SDS-PAGE gel (Fig. 1A) and subjected their trypsin-generated fragments to MALDI-TOF mass-spectrometric analysis. The top band was identified as AcrB (Suppl. S1), whilst the other two failed to produce a signal. Molecular replacement by PHASER using the AcrB structure as model accordingly produced an unambiguous solution (Z-scores for rotation / translation function: 10.3 / 39.4), confirming the scenario that the crystals were not of GluP(His)$_6$ but of AcrB. The abundance of AcrB in the fractions was <0.1% as judged by densitometry of scanned SDS-PAGE lanes. The total amount of AcrB present in the crystallization drop was hence in the range of 100-500 pg, providing as yet the first example of a membrane protein crystallizing from picogram amounts. Additionally, six new crystallization conditions for AcrB were identified (Table 1). Crystallization of AcrB, so far, only occurred in our screens in the presence of maltosides and/or glucosides.

AcrB poses a serious problem, when heterologous proteins are expressed in *E. coli* or other hosts harboring multi-drug resistance proteins, because its expression is induced by stress. Consequently, upon expression of a system that is energy-consuming or membrane destabilizing, transcription of *acrAB* will be up-regulated (Eaves *et al.*, 2004), thus enriching the *E. coli* membranes with AcrB. In the scope of our medium-throughput genomics project (Psakis *et al.*, 2007), AcrB comprised the major contaminant in 45% of the purified targets. Major facilitator superfamily (MFS) porters, ion-channels, proteins of the Sec machinery and the hypothetical carbon starvation protein A (CstA) from *H. pylori* were the targets regularly contaminated with AcrB. Of course, co-purification of AcrB with the selected His$_6$-tagged target was facilitated by its ability to bind to the immobilized Ni$^{2+}$, but AcrB is neither the only histidine-containing protein nor the only protein that binds to immobilized nickel.
Extensive washing with higher concentrations of imidazole, prior to protein elution, achieved the reduction in the concentration of the contaminant (Fig. 1B) but not its complete removal.

From the detergent screens performed on GluP and CstA we identified the following effects on AcrB’s solubility: n-octyl-β-D-glucopyranoside (OG) > Triton X-100 = polyoxyethylene-9-laurylether (C12E9) > octanoyl-n-methylglucamide (MEGA8) > decanoyl-n-methylglucamide (MEGA10) = n-undecyl-β-D-maltoside (NUM) (Suppl. S2A-S2B). OG solubilised AcrB completely but not selectively, thus resulting in 50% loss of GluP in the same fractions. Triton X-100, C12E9, MEGA8 and MEGA10 solubilized AcrB more selectively. The use of the latter detergents, however, often caused the irreversible aggregation of either GluP or CstA, significantly lowering, thus, the yield of their second solubilization in a more stabilizing detergent. What appeared to be more effective in removing and/or reducing the concentration of AcrB in the membrane fractions was their extensive washing with 0.2 M (NH4)2SO4 (Suppl. S2D) or 0.2 M MgSO4 prior to detergent-solubilization of the selected targets. Unfortunately, the presence of ≤100 pg of AcrB in the crystallization drops was enough to yield crystals reproducibly. Furthermore, the huge excess of another membrane protein like GluP has no significant effect on the AcrB structure and packing (Fig. 2C, 2D).

CstA fractions were also contaminated with the succinate dehydrogenase complex. The succinate dehydrogenase expression levels were shown to increase when E. coli cells grew under aerobic conditions and at temperatures ranging from 30-37 °C (Cheng et al. 2006). The latter conditions resulted in the successful expression of several H. pylori genes in E. coli (Psakis et al. 2007), but at a cost of generating E. coli membrane fractions rich in the respiratory complex II. Extensive washing of the inner membrane fractions with 0.2-0.4 M NaBr (Bondos and Bicknell, 2002), prior to their solubilisation, succeeded in removing all traces of succinate dehydrogenase, but also caused a partial loss of CstA. About 17% of our purified targets (MFS porters) were contaminated with succinate dehydrogenase.
The flavoprotein subunit of the complex (SdhA) has 19 histidines, several of which are freely exposed and could interact with the Ni\(^{2+}\)-NTA matrix. SdhA and CstA co-migrate in the SDS-PAGE gel (Fig. 1B, lanes I & II), but they can be distinguished due to the orange-yellow colour the quinone confers to the SdhA solution. Unlike AcrB, complete removal of succinate dehydrogenase from CstA could be achieved by extensive imidazole washing (40-50 mM), prior to its elution, and ion exchange chromatography. After the successful removal of the contaminating complex II, CstA was subjected to the same crystallization screening as GluP(His)_6. Traces of AcrB, however, were still present (Figure 1B, lane III), allowing the contaminant to crystallize in an additional condition (Table 1). Although these crystals did not diffract, their rhomboidal shape (Fig. 2A) and the presence of AcrB (Figure 1B, lane III) suggested once again that the contaminating protein had crystallized. Crystallization of either CstA or AcrB alone under the same conditions did not yield any crystals, suggesting that CstA may have induced nucleation of AcrB.

The real reason that AcrB is so problematic for the structural analysis of recombinant membrane proteins, and especially for high-throughput membrane protein genomics projects, is that: 1) its expression increases in response to global stress and 2) it is a membrane-protein that tends to crystallize in picogram quantities. These properties of AcrB should be taken into account, not only for heterologous expression of membrane proteins in *E. coli*, but also for heterologous expression in other bacterial hosts where multi-drug efflux homologues or orthologues are present.

Induction of gene transcription for AcrB and/or succinate dehydrogenase was not only observed in B-type (e.g. BL21 or BLR) but also in K-type *E. coli* strains [e.g. JM110 or RE707 (Essenberg et al., 1997)] (Suppl. S3). Whereas induction of gene transcription for succinate dehydrogenase was independent of the media used, when employing K-type *E. coli* strains, induction of gene transcription for AcrB was mostly observed in rich media (Suppl. S3).
The fact that 1) AcrB has a significant affinity for immobilized Ni\(^{2+}\) matrices, and 2) interacts non-specifically with anti-His antibodies, contributes to the seriousness of the problem, but AcrB is not the only protein with such tendencies. In this report we showed that the respiratory complex II of *E. coli* also displayed that behaviour. Extensive washings with increased imidazole concentrations plus additional purification steps ensured complete removal of succinate dehydrogenase from preparations but not of AcrB.

Whereas selective solubilization of AcrB from the membrane fractions by use of certain detergents was not particularly successful, washing of the latter with 0.2 M (NH\(_4\))\(_2\)SO\(_4\) or 0.2 M MgSO\(_4\) prior to solubilization, resulted in significantly cleaner preparations, where AcrB could not be detected. So long as even a picogram amount of AcrB, however, was present in a drop containing the appropriate mother liquor, crystals could be produced. Whether the selected target extrinsically induced the crystallization of AcrB under the newly observed conditions (as was the case with CstA) is a matter open to further investigation. In our experience, both silver-staining and Coomassie Blue staining SDS-PAGE gels had to be overloaded with the purified sample for AcrB to be visible. Veesler et al. (2008) failed to detect AcrB in Coomassie Blue staining SDS-PAGE gels. In our preparations, however, and according to our data the concentration of AcrB did not exceed 0.01mg.ml\(^{-1}\) (<0.1% of the total fractions analyzed on SDS-PAGE gel), which much more clearly exemplifies the severity of the problem.

Although, combinations of different detergents and/or sulphate salts, and the use of the appropriate growth media, severely reduced the AcrB contamination in the purified fractions, to the extent that it was no longer detectable by silver staining, in several cases AcrB crystals grew from these preparations within 8-10 months. In our view, the only foolproof solution to the reported problem would be either the use of different tags, or the use of *acrB* and or *acrAB* knockouts of standard laboratory strains. Knockouts have been successfully produced.
in *E. coli* (Miller et al., 2004) and in *Haemophilus influenza* (Dean et al., 2005) and are currently in use.

The unwanted crystallization of protein contaminants in structural genomics projects has always been an important concern. There are several examples for low-level contaminants (<5%) in the preparations of recombinant soluble proteins, which gave diffraction-quality crystals like bacterioferritin and the Gab protein from *E. coli* (van Eerde et al., 2006, Lohkamp & Dobritzsch, 2008). Whereas the outstanding tendency of crystallization by highly symmetric protein complexes like bacterioferritin can be easily rationalized, one might wonder about the peculiarities of the formation of the rhombohedral AcrB crystal form. From its crystal packing (Fig. 2D) the layer-like arrangement of AcrB trimers is apparently resulting in large lamellar sheets that comprise the membrane region and associated detergent/lipids from the crystallization condition. Accordingly, the crystallization of AcrB from picogram amounts might be more dictated by phase-transitions of the detergent/lipid content within the crystallization drops than from simple oversaturation of solubilized AcrB trimers. This might explain, why crystallization of AcrB contaminants appears to be highly affected by the presence of particular recombinant membrane proteins as not all AcrB-contaminant comprising targets yielded AcrB crystals under identical conditions.

**Acknowledgments**

This work was supported by the Bundesministerium für Bildung und Forschung (BMBF) as part of the Proteome Analysis of Membrane Proteins Initiative (Pro-AMP) and the European Membrane Protein Consortium (e-MeP), on whose 2007 meeting in Marseille these results were presented. We would like to thank Essenberg, R. C. and Henderson, P. J. F. for the *E. coli* RE707 cells. We also thank Uwe Linne for MALDI-TOF and ESI mass-spectrometric analyses and Klaas Pos for valuable discussion about the issue of AcrB contaminations.
References


Figure legends

**Fig. 1.** Purification profiles of GluP(His)$_6$ (A) and CstA(His)$_6$ (B) on 12% Coomassie Blue stained SDS-PAGE gels. (A) Lane M, molecular weight markers (Fermentas, 2 µg); Lane 1, *E. coli* BL21(DE3) membranes containing GluP(His)$_6$ (30 µg); Lane 2, insoluble material after detergent extraction (30 µg); Lane 3, detergent extract of *E. coli* BL21(DE3)/GluP(His)$_6$ (30 µg); Lane 4, Ni$^{2+}$-NTA chromatography (10 mM imidazole) flowthrough fraction (30 µg); Lane 5, Ni$^{2+}$-NTA chromatography (40 mM imidazole) washing fraction (30 µg); Lane 6, purified and concentrated GluP(His)$_6$ (30 µg). (B) I, Ni$^{2+}$-NTA co-purification of CstA(His)$_6$, succinate dehydrogenase complex (subunits SdhA, B, C, and D indicated) and AcrB (30 µg); II, purified complex, contaminated with AcrB after cation-exchange of the previous fractions (20 µg); III, purified and concentrated CstA contaminated with AcrB (30 µg); IV, purified native *E. coli* AcrB (0.05 µg) shown as control; V, detection of native *E. coli* AcrB (0.05 µg) by anti-His$_5$ antibody (bottom band-dye front).

**Fig. 2.** (A) Crystals of AcrB in crystallization screens of CstA(His)$_6$ (A1) and GluP(His)$_6$ (A2-A4). Scale; 50 µm (B) Diffraction pattern of crystal A2 using synchrotron radiation on beam-line PX06SA at the SLS. Arrows indicate the minimum and maximum resolution in Å. (C) Model of AcrB refined against 4.2 Å data from crystal A2 ($R_{work}/R_{free}$: 0.318/0.363, r.m.s.d. for bond angles: 0.020 Å). Given the available resolution the AcrB model (red ribbon) is almost identical to the reference structure (gray ribbon, PDB code 1T9U, r.m.s.d. 0.84 Å for 1013 C$_\alpha$-positions). The other monomers of the AcrB trimer as generated by crystal symmetry are indicated in yellow and orange. (D) Molecular packing of AcrB trimers within the $R32$ crystal form. The arrangement of AcrB trimers within these crystals results in extensive...
lamellar regions occupied by the transmembrane domains of AcrB and associated detergents & lipids (boxed in ochre).