

## Purification of recombinant BtpA and Ycf3, proteins involved in membrane protein biogenesis in *Synechocystis* PCC 6803

Tatjana M.E. Schwabe\*, Kirsten Gloddek<sup>1</sup>, Daniela Schluesener, Jochen Kruij<sup>2</sup>

*Plant Biochemistry, Faculty of Biology, Ruhr-University Bochum, Universitaetsstr. 150, ND 3/130, 44801 Bochum, Germany*

### Abstract

The gene products Ycf3 (hypothetical chloroplast open reading frame) and BtpA (biogenesis of thylakoid protein) are thought to be involved in the biogenesis of the membrane protein complex photosystem I (PSI) from *Synechocystis* PCC 6803. PSI consists of 12 different subunits and binds more than 100 cofactors, making it a model protein to study different aspects of membrane protein biogenesis. For a detailed biophysical characterization of Ycf3 and BtpA pure proteins must be available in sufficient quantities. Therefore we cloned the corresponding genes into expression vectors. To facilitate purification we created His-tagged versions of Ycf3 and BtpA in addition to the unmodified forms. Immobilized metal affinity chromatography (IMAC) yielded His-tagged proteins which were used for the production of antibodies. Purification strategies for non-tagged proteins could also be established: Ycf3 could be purified in soluble form using a two-step purification in which ammonium sulfate precipitation was combined with anion-exchange chromatography (IEC). BtpA had to be purified from inclusion bodies by two-consecutive IEC steps under denaturing conditions. An optimized refolding protocol was established that yielded pure BtpA. In all cases, MALDI-TOF peptide mass fingerprinting (PMF) was used to confirm protein identity. Initially, size exclusion chromatography and CD-spectroscopy were used for biophysical characterization of the proteins. Both Ycf3 and BtpA show homo-oligomerization in vitro. In summary, purification protocols for Ycf3 and BtpA have been designed that yield pure proteins which can be used to probe the molecular function of these proteins for membrane protein biogenesis.

© 2002 Elsevier Science B.V. All rights reserved.

**Keywords:** Biogenesis; *Synechocystis*; Purification; Membrane proteins; Photosystem I

### 1. Introduction

Membrane protein complexes perform many crucial biological reactions in the cell. While the molec-

ular function and structure of these complexes is becoming more and more clear, almost nothing is known about the molecular steps in membrane protein biogenesis. One factor for this is certainly the experimental difficulties associated with handling hydrophobic membrane proteins.

Cyanobacterial photosystem I (PS I) serves as a model system for the investigation of the assembly of multi-subunit membrane protein complexes. It is localized in thylakoid membranes of cyanobacteria and chloroplasts, where it functions as one of two light-driven electron pumps, essential for the photo-

\*Corresponding author.

E-mail address: [tatjana.schwabe@ruhr-uni-bochum.de](mailto:tatjana.schwabe@ruhr-uni-bochum.de) (T.M.E. Schwabe).

<sup>1</sup>Present address: RWTH Aachen, Inorganic Chemistry and Electrochemistry, 52056 Aachen, Germany.

<sup>2</sup>Present address: Aventis Pharma Germany, DI&A Functional Genomics, 65926 Frankfurt, Germany.

synthetic electron transport chain [1]. It is a major component of these membranes and its function can easily be monitored by spectroscopic techniques in vitro or in vivo. These, among other reasons, make PS I an ideal model system. The two main subunits PsaA and PsaB form a central heterodimer which is the functional heart of PSI. Ten additional subunits are assembled around this core. Collectively, the 12 subunits bind well over 100 cofactors, specifically around 96 chlorophylls, 22 carotenoids, three Fe–S clusters, two phylloquinones, four lipids and one calcium ion [2].

This presents a major challenge for PSI biogenesis: all cofactors must be integrated in a specific molecular environment presumably at a specific point in time to yield functional PSI. In general, the PSI life cycle can be divided into four distinct steps: (i) PS I subunit synthesis from gene sequence, (ii) PSI complex formation from individual subunits, (iii) dynamic interaction of the mature complex with other membrane proteins and (iv) degradation of the PS I complex. In our group each step is investigated, but in this paper we concentrate on the second step, PSI complex assembly.

In recent years, few gene products which may play important roles in the assembly process have been discovered (for reviews, see Refs. [3,4]). Two were found when Rochaix and coworkers knocked out hypothetical chloroplast open reading frames (*yfc*) in *Chlamydomonas reinhardtii*, a green alga containing chloroplasts. The phenotype of both *yfc3* and *yfc4* deletion mutants showed that no PS I accumulated in these eukaryotic cells [5].

Bartsevich and Pakrasi [6] reported another protein, which appears to be important for PS I biogenesis: a point mutation in the so-called *btpA* (biogenesis of thylakoid proteins A) gene leads to reduced PS I levels in *Synechocystis* PCC 6803.

We constructed deletion mutants of these three genes in *Synechocystis* PCC 6803, a unicellular prokaryotic cyanobacterium. Here, loss of *Ycf3* led to the complete absence of assembled PS I, whereas the deletion of *yfc4* still allowed low amounts of PS I to be made (to be published). The absence of *BtpA* gave similar results, which are significantly different from those reported by Pakrasi and co-workers (to be published elsewhere).

Our final goal is to understand the functions of

these proteins on a molecular basis. Therefore, we decided to overexpress and purify the proteins *Ycf3*, *Ycf4* and *BtpA*. In this paper, purification strategies and preliminary characterization of His-tagged and native *Ycf3* and *BtpA* are presented.

## 2. Experimental

### 2.1. Construction of expression plasmids

Genes *yfc3* (slr0823) and *btpA* (sll0634) were amplified from *Synechocystis* PCC 6803 DNA by PCR using specific primer pairs CGCGGATCCATATGCCCCGCACCCAACGTAAC/CGCGGATCCAAGCTTTTAGAAAAACACATCCATTTTCAG (*yfc3*) and CGCGGATCCAAGCTTTCAACA-AACCACTGATTTTGTTC/CGCGGATCCCATATGGATTTATTTCAAACCTTTTCAGACCC (*btpA*) which introduced additional restriction sites. Using these sites (BamHI and NdeI), the amplified inserts were cloned into similarly restricted pRSET5a (Invitrogen, Karlsruhe, Germany) for expression of non-tagged and into pET16b (Novagen, Madison, WI, USA) for expression of N-terminally His-tagged protein. Protein expression was thus controlled via a T7 promoter system.

### 2.2. Protein expression

The *Escherichia coli* BL21 (DE3) strain was used for heterologous overexpression. Small scale cultures were inoculated with freshly transformed cells and were used to optimize protein yields. Expression efficiency was monitored by SDS-PAGE [7] for which the cell pellet was boiled for 5 min in SDS-sample buffer. Optimization was carried out with respect to medium, temperature and expression time (see Results section). In each case, induction of T7 promoter system with isopropylthiogalactoside (IPTG; 1 mM) was done after the cultures reached an OD<sub>600</sub> of 0.6.

### 2.3. Purification of proteins

*E. coli* cultures over-expressing the desired protein were harvested by centrifugation and cells lysed by lysozyme/freeze treatment. After DNase I digestion,

a detergent solution was added (1% Na-deoxycholate in 20 mM Tris–HCl pH 8.0, 200 mM NaCl) and the inclusion bodies were pelleted by low speed centrifugation (15 000 *g*). The resulting pellet was washed with 20 mM Tris–HCl pH 8.0, 1 mM EDTA and 1% (w/v) Triton X-100. The supernatant was centrifuged at high speed (60 000 *g*) to separate soluble proteins from membranes. To ascertain in which fraction the desired proteins accumulated, we used SDS–PAGE and Western blots. For this, polyclonal antibodies against purified His-tagged protein (which had been further purified by preparative SDS–PAGE) were raised in rabbit.

#### 2.4. Immobilized metal affinity chromatography (IMAC)

For purification of His-tagged proteins, chromatography was done on a GradiFrac system, comprising GradiFrac, Pump P-1, Monitor UV-1, Recorder REC 102, Valve IV-7 and Switch Valve PSV-50 (Amersham Biosciences, Uppsala, Sweden) or on ÄKTAprime (Amersham Biosciences). Columns were pre-packed chelating sepharose (HiTrap 1 ml, Amersham Biosciences) for analytical and a self-packed 12 ml column of the same material for preparative purifications. Before each run, the columns had to be loaded with metal: either 0.1 *M* NiSO<sub>4</sub> or 0.1 *M* CuSO<sub>4</sub> in 10% acetic acid. After equilibration with running buffer (typically 6 *M* urea, 20 mM Tris–HCl pH 8.0, 20 mM NaCl, 20 mM imidazole), the sample was applied. An increase in imidazole concentration—either by a linear gradient or by discrete steps—led to elution of bound proteins. After each run, bound metal was completely removed through an EDTA-washing step (50 mM).

Refolding on column was carried out by a slow exchange of binding buffer (6 *M* urea, 20 mM Tris–HCl pH 8.0, 0.5 *M* NaCl, 5 mM imidazole) with refolding buffer (20 mM Tris–HCl pH 8.0, 0.5 *M* NaCl, 20 mM imidazole, within 12 column volumes, flow-rate was 0.5 ml/min). Protein was eluted in a two-step gradient from 5 to 150 mM, to 0.5 *M* imidazole (1 ml/min, elution buffer composition as refolding buffer). All buffers for this procedure contained 1% β-mercaptoethanol as reducing agent.

#### 2.5. Ion exchange (IEC) and size exclusion chromatography (SEC)

IEC was performed on a Waters system (two pumps, model 510 fitted with preparative pump heads, Rheodyne injector, Model 9125) coupled to a diode array detector (PDA 996, Waters, Milford, MA, USA) and a conductivity monitor (Amersham Biosciences). Columns were kept in a column oven (Jasco, Gross-Umstadt, Germany) at constant temperature (10 °C). Columns and run conditions are described in the Results section. For size exclusion chromatography (SEC), a Waters 550 pump was used, allowing precise flow-rate control at low flow-rates. SEC was calibrated using Bio-Rad gel filtration standard, diluted 1:5 in running buffer (typically 20 mM Tris–HCl pH 8.0, 150 mM NaCl).

#### 2.6. Refolding of purified protein

Refolding of solubilized proteins was done by dialyzing against urea-free buffers. To optimize the refolding process, different pH and ion conditions were tested in small scale experiments. The success of a refolding experiment was judged based on the ratio of protein amount between pellet and supernatant using SDS–PAGE analysis after separation by centrifugation (Beckman Optima TL, 150 000 *g*, 60 min).

#### 2.7. Analysis of purified protein

Circular dichroism measurements were conducted on a J715 spectropolarimeter (1-mm cell length, Jasco, Gross-Umstadt, Germany). Five measurements were averaged for final results. Jasco J-700 software was used to calculate the amount of secondary structure elements from the CD-spectra [8]. These data were compared to predicted secondary structure data from several algorithms (GORIV [9], SOPMA [10] and HNN [11]), all available through the ExPasy server (<http://www.expasy.org>).

#### 2.8. MALDI-TOF peptide mass fingerprinting (PMF)

Proteins separated on analytical SDS–PAGE were stained with colloidal Coomassie [12] and spots cut

from the gel. The gel pieces were dried after destaining (50% acetonitrile, 2.5 mM NaHCO<sub>3</sub>, three times at 37 °C) in a speedvac and rehydrated with 2.5 mM NaHCO<sub>3</sub> buffer containing 12.5 ng/μl porcine trypsin (Promega, Madison, WI, USA). Peptide fragments were eluted by adding 50% acetonitrile, 0.5% trifluoroacetic acid after digestion overnight at 37 °C. For MALDI-TOF analysis on a Voyager-DE PRO (Applied Biosystems, Foster City, CA, USA), the sample was mixed with matrix (10 mg/ml α-hydroxycinnamic acid in 60% acetonitrile, 1% trifluoroacetic acid) directly on the steel sample plate.

### 3. Results

#### 3.1. Production of recombinant proteins

To exclude PCR errors or cloning artifacts (see Experimental section), the expression plasmids were

verified by sequencing and comparison with published sequence information [13]. To optimize protein yields, several different growth media for *E. coli* were tested. Only minor differences in protein yields could be found, with rich media having a positive effect (data not shown). Based on these experiments, LB medium was used for overexpression of HBtpA, whereas 2×YT medium was used for BtpA, HYcf3 and Ycf3. Although the T7 promoter system is IPTG-inducible, expression analysis on SDS–PAGE showed that induction is not necessary for the expression of our proteins. Indeed, it is rather detrimental. Induction leads to reduced growth rates, which in turn are not compensated for by sufficient increase in overexpression (Fig. 1). Thus IPTG induction was omitted in routine protein preparation, still leading to accumulation of the desired *Synechocystis* proteins to over 50% of total *E. coli* cell protein.

In order to ascertain where the proteins accumulated in the cell, *E. coli* cells were harvested and

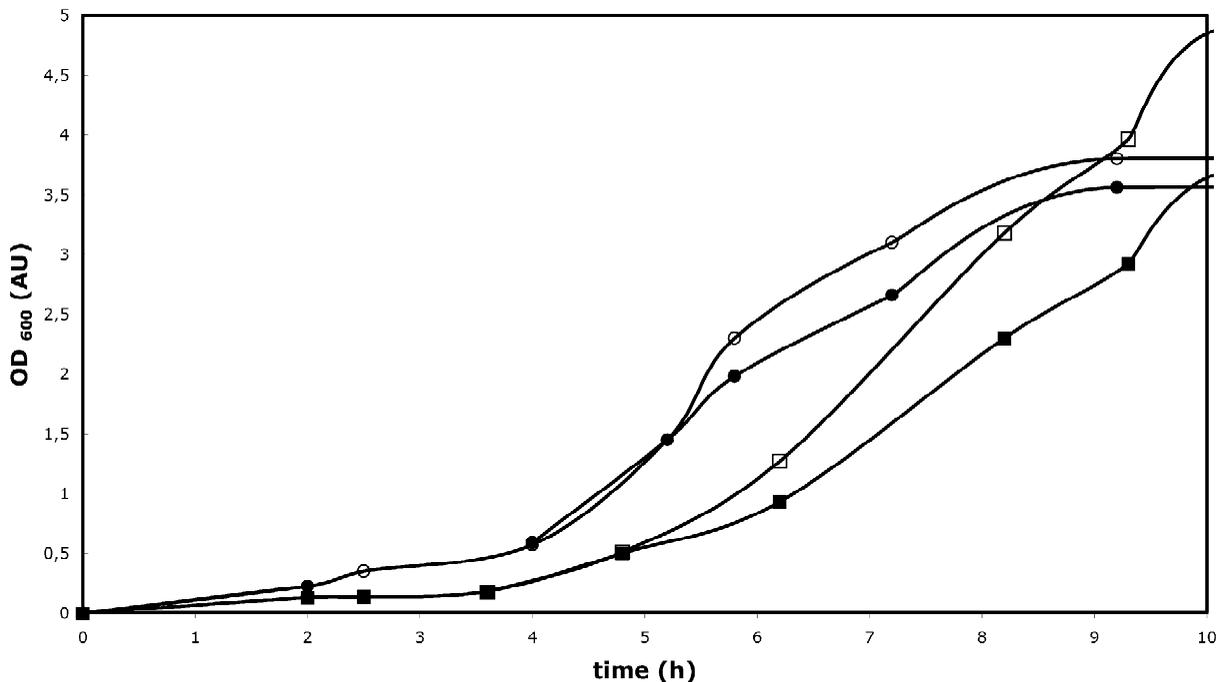


Fig. 1. Influence of IPTG-induction on growth rates of *E. coli* overexpressing HBtpA and BtpA, respectively. Circles: *E. coli* expressing HBtpA (open circles: no IPTG added). Squares: *E. coli* carrying BtpA expression plasmid (open squares: no IPTG added). All cultures were grown to an OD<sub>600</sub> of 0.6. Expression was induced (closed squares and circles) by adding IPTG to a final concentration of 1 mM. Data represent typical growth curves.

fractionated. A membrane fraction, a fraction rich in inclusion bodies and a soluble fraction were separated and analyzed by SDS–PAGE for their protein content. Although a certain percentage of BtpA is associated with *E. coli* membranes, the vast majority of BtpA was found in inclusion bodies making them an ideal source for further purification (Fig. 2). All efforts to avoid its accumulation in inclusion bodies were unsuccessful. For example, lowering the temperature during expression to 24 °C did not yield significantly more soluble BtpA (Fig. 2). Also co-expression with GroEL and GroES from *Synechocystis* did not increase the amount of soluble protein (data not shown). The same is true for HBtpA—it accumulates almost exclusively in inclusion bodies (data not shown).

In contrast, Ycf3 accumulated partially in soluble form and could be purified from soluble cell extract, if *E. coli* cells were grown at 30 °C. HYcf3 could only be isolated from inclusion bodies. Attempts to obtain HYcf3 in a soluble form failed (data not shown).

### 3.2. Purification of His-tagged proteins

First, inclusion bodies were completely solubilized in a chaotropic buffer (6 M urea, 20 mM Tris–HCl pH 8.0, 20 mM NaCl) and subjected to IMAC on Chelating Sepharose Fast Flow (Amersham Biosciences). We tested different metals for their ability to bind specifically HBtpA or HYcf3. Whereas copper was most efficient in purifying HBtpA, nickel proved to be a better alternative in the case of HYcf3.

As can be seen on the SDS–PAGE (Fig. 3), HBtpA could be purified to 95% purity in a two-step elution with imidazole (first step: 20–66 mM, main HBtpA peak elutes at 250 mM imidazole, elution profile not shown). Additional purification steps, like ion-exchange chromatography or hydrophobic interaction chromatography, were unsuccessful (data not shown). However, by IMAC we could purify ~10 mg of HBtpA starting with 600 ml of *E. coli* culture.

Similarly HYcf3 was purified under denaturing conditions by a two-step gradient (5 mM imidazole to 50 mM, to 250 mM imidazole) yielding almost

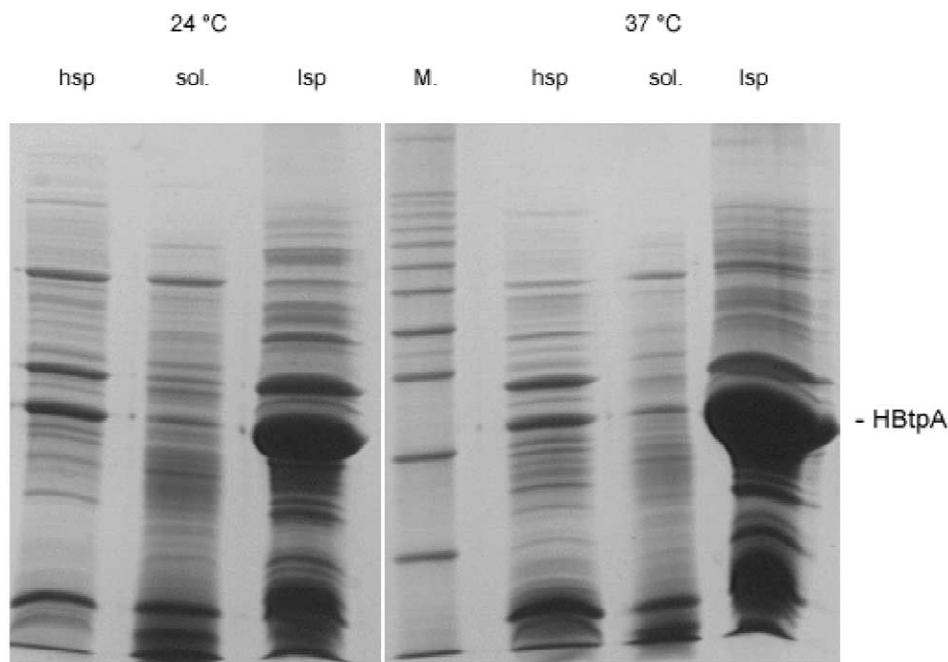


Fig. 2. Localization of overexpressed BtpA in *E. coli*. Cells were fractionated (see Experimental section), solubilized in SDS-buffer and applied to SDS–PAGE. The gel was stained with Coomassie brilliant blue. Expression was done in 2×YT medium at the temperatures indicated. hsp, high speed pellet; lsp, low speed pellet; M:  $10 \cdot 10^3$  molecular mass ladder; sol, soluble fraction.

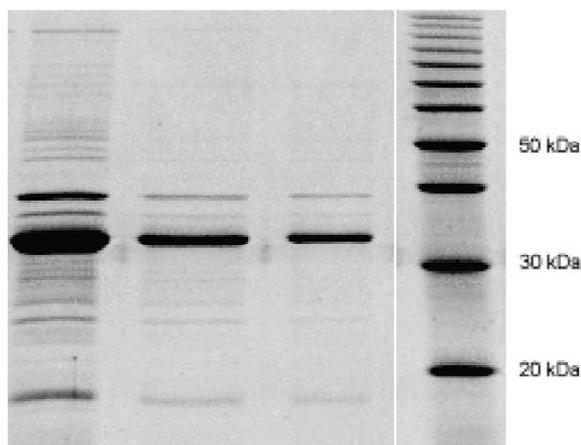


Fig. 3. SDS–PAGE of IMAC-fractions in HBtpA-purification. Collected fractions were TCA-precipitated and resuspended in SDS-buffer. The gel was stained with Coomassie. In lane 4 a molecular mass marker was applied ( $10 \cdot 10^3$  ladder).

pure protein (data not shown). Attempts to purify the small amounts of soluble HYcf3 by IMAC under native conditions were not successful (data not shown).

### 3.3. Purification of Ycf3

Soluble protein extract from *E. coli* was subjected to ammonium sulfate precipitation as a first purification step. After optimization, it was found that most Ycf3 precipitated at 50% saturation (ammonium sulfate) while most *E. coli* proteins remained soluble. Precipitated Ycf3 was washed once in 50%  $(\text{NH}_4)_2\text{SO}_4$ , 20 mM Tris–HCl pH 8.0, 20 mM NaCl and centrifuged again. The pelleted Ycf3 was solubilized by dialysis against 20 mM Tris–HCl pH 8.0, 20 mM NaCl; this step reduced the ion concentration and prepared the sample for further purification by HPLC. Several chromatography materials were tested for purification of Ycf3. The protein bound neither to weak nor strong cation-exchange material (Macro-Prep CM and High S, Bio-Rad, Hercules, CA, USA), and nor to hydrophobic interaction material (Poros HP2, Applied Biosystems). Variations in the buffer composition and pH did not change these results (data not shown). However, anion-exchange chromatography on Poros HQ50

(Applied Biosystems) proved to be efficient for Ycf3 purification. Starting with a continuous sodium chloride gradient (20 mM to 2 M NaCl within 7 column volumes) we could improve purification by applying a two-step gradient (from 20 mM to 1 M NaCl in 5 column volumes, then in 2 column volumes to 2 M NaCl). Ycf3 eluted in a sharp peak at  $\sim 0.8$  M NaCl (Fig. 4). Analysis by SDS–PAGE confirms that our purification protocol yields soluble Ycf3 with only minor impurities (Fig. 5).

### 3.4. Purification of BtpA

A different purification strategy had to be developed for non-tagged BtpA: inclusion bodies were solubilized in urea buffer (6 M urea, 20 mM Tris–HCl pH 8.0, 20 mM NaCl) and purified in a denatured state by multiple chromatography steps. Ammonium sulfate precipitation of solubilized inclusion bodies was not suitable in this strategy, because BtpA precipitated at low concentrations, together with most *E. coli* proteins (data not shown).

Therefore, solubilized BtpA was directly applied to a strong ion-exchange column (Poros HQ50, Applied Biosystems). High flow-rates permitted rapid enrichment of BtpA from inclusion bodies. BtpA-containing fractions, which elute late in the linear sodium chloride gradient (20 mM to 1.5 M NaCl; Fig. 6), were concentrated and subjected to either hydrophobic interaction chromatography (Poros HP2, Applied Biosystems) or a second ion-exchange chromatography using a different material (Uno-Q column, Bio-Rad).

For HIC, the sample had to be first adjusted to a concentration of 3 M  $(\text{NH}_4)_2\text{SO}_4$ . The running conditions used were 6 M urea, 20 mM Tris–HCl pH 8.0, 20 mM NaCl and 2 M  $(\text{NH}_4)_2\text{SO}_4$ ; elution was done by decreasing the  $(\text{NH}_4)_2\text{SO}_4$  concentration from 2 to 0 M within 3 column volumes. This step provided no further purification of the sample, as SDS–PAGE of elution fraction was similar to the starting material (data not shown).

For the second IEC column, we changed the buffer from Tris–HCl pH 8.0 to a phosphate buffer pH 7.0. In this solution, the sample was loaded onto an UNO-Q column (Bio-Rad) for final purification (6 M urea, 20 mM sodium phosphate buffer pH 7.0, 20 mM NaCl). Elution was achieved with a linear

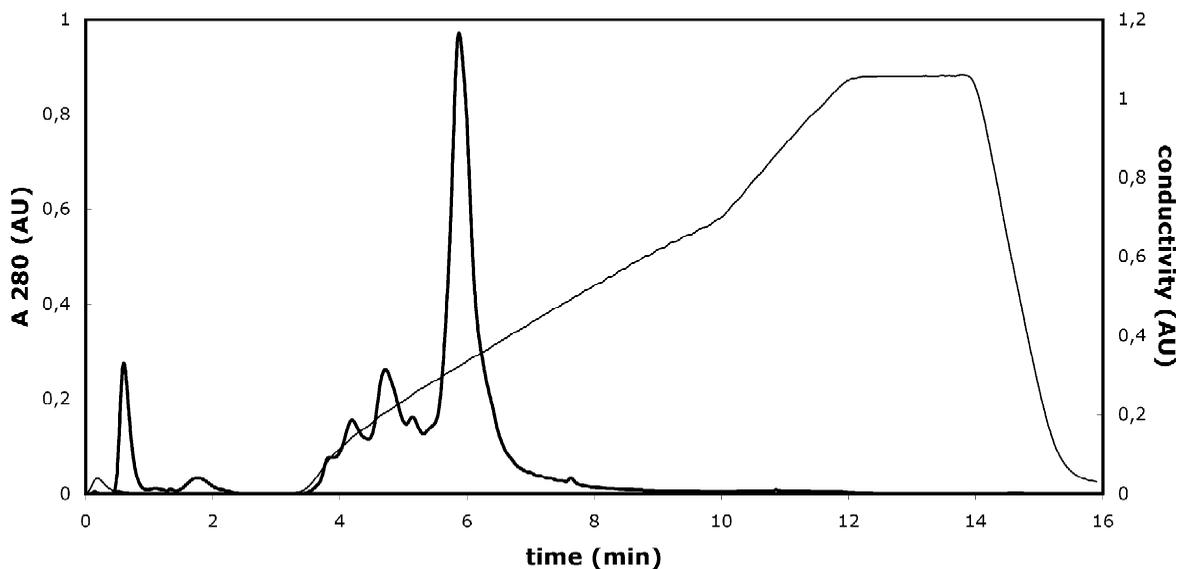


Fig. 4. Elution profile of soluble Ycf3 in anion-exchange chromatography. Ycf3 sample after ammonium sulfate precipitation was subjected to IEC on Poros HQ50 material (column volume: 6.5 ml). Binding buffer was 20 mM Tris-HCl pH 8.0, 20 mM NaCl, flow-rate 4.5 ml/min. Ycf3 eluted in a salt gradient (thin line) at  $\sim 0.8$  M NaCl (gradient from 20 mM NaCl to final concentration of 2 M NaCl). Bold line: absorbance at 280 nm (AU).

NaCl-gradient (from 20 mM to 0.47 M NaCl in 3 column volumes, then to 1.5 M NaCl in 0.7 column volumes) yielding almost pure BtpA with no further contaminations visible on silver stained gels (Fig. 7).

### 3.5. Refolding of purified protein

All proteins (HYcf3, HBtpA, BtpA) which were purified under denaturing conditions had to be refolded. This could be achieved in all cases by dialysis against buffers without chaotropic reagents.

To maximize the yield of soluble or correctly folded proteins, various parameters like pH or buffer composition had to be optimized.

A Tris buffer (20 mM Tris-HCl pH 8.0, 20 mM NaCl) proved to be most suitable for HYcf3. In contrast, an alkaline buffer (20 mM CAPS pH 10, 20 mM NaCl) yielded a slightly less favorable ratio of soluble protein to aggregates and an even more acidic buffer (20 mM MES pH 6.5, 20 mM NaCl) lead to almost quantitative precipitation. Similarly, the presence of any divalent cation induced precipitation of HYcf3 (data not shown). Therefore a Tris-

HCl buffer of pH 8.0 containing 20 mM NaCl was used for all further refolding experiments.

The best results for either HBtpA or BtpA were also obtained by dialyzing against a Tris buffer (pH 8.0) containing NaCl (Fig. 8). As we could see already for Ycf3, the addition of divalent ions led to decrease in soluble protein obtained after centrifugation. But even under optimized conditions,  $\sim 30\%$  of the protein was lost due to irreversible aggregation.

### 3.6. Refolding on IMAC-column

As an alternative to purification in the presence of urea and subsequent refolding of the protein by dialysis, we tested the refolding of HYcf3 directly on the column, before elution with imidazole. The quality of the preparation was as pure as that of the one-step IMAC (data not shown). Size exclusion chromatography showed, however, that we could recover more protein in oligomeric form than by dialysis-refolding (see inset in Fig. 9). Using this method, we can purify 8 mg pure HYcf3 from 500 ml *E. coli* culture. Fast refolding, by changing from binding to refolding buffer within 1 column volume

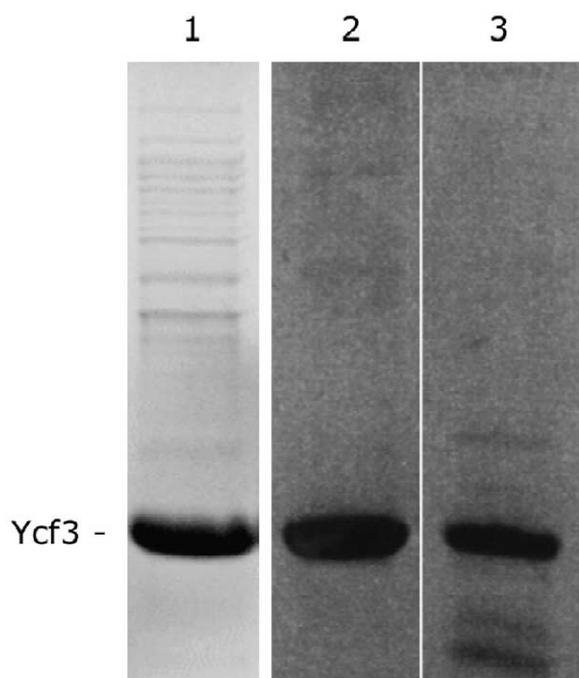


Fig. 5. SDS-PAGE of different fractions during Ycf3-purification. Lane 1: Ycf3 sample after IEC on Poros HQ50, Coomassie stain. Lanes 2 and 3: pure Ycf3 after size exclusion chromatography (lane 2: oligomer; lane 3: monomer, silver stain).

did not lead to successfully refolded protein (data not shown).

### 3.7. Characterization of purified proteins: size exclusion chromatography

By size exclusion chromatography (SEC) of the purified proteins we could show that BtpA as well as Ycf3 can form oligomeric complexes, suggesting a successful refolding process.

Most Ycf3 elutes at a time corresponding to an oligomer of  $260 \cdot 10^3$ ; only minor amounts of Ycf3 can be found in a monomeric form ( $20 \cdot 10^3$ ; Fig. 9). Refolded His-tagged Ycf3, however, consists mainly of monomeric protein (data not shown). SEC of refolded HYcf3 shows another not well defined peak eluting partly in the void volume which corresponds to irreversibly aggregated HYcf3. Different ions ( $\text{CaCl}_2$  and  $\text{MgCl}_2$ ) had a small effect on the elution time of Ycf3 (data not shown). This effect is most probably due to changes in elution characteristics

however, and does not represent a true change in protein conformation.

Refolded BtpA and HBtpA are present in a monomeric form of  $\sim 30 \cdot 10^3$ ; in HBtpA this peak is just a shoulder of a peak most likely consisting of degraded protein. However, large amounts of BtpA as well as HBtpA occur in the form of a high molecular complex (between  $600$  and  $700 \cdot 10^3$ ) which is well below the upper separation limit of the column. This peak is very broad, indicating that the complex is made up of different numbers of subunits (Fig. 10).

### 3.8. CD-spectroscopy

The CD-spectra for Ycf3 and HYcf3 are very similar, as can be seen in Fig. 11. Calculation of secondary structure content according to Yang et al. [8] shows a ratio of 45–65%  $\alpha$ -helix to 0–12%  $\beta$ -sheet and 20–22% turns. These results concur with theoretical predictions for secondary structure of Ycf3 by a variety of different algorithms (Table 1). The additional tag of HYcf3 has little influence on these values.

The CD-spectrum for BtpA, shown in Fig. 12, indicates 41%  $\beta$ -sheets and 33% random coils. This disagrees somewhat with theoretical data, which predict a larger percentage of  $\alpha$ -helical and random structures (Table 1).

### 3.9. Antigen elution

In order to purify native BtpA and Ycf3 directly from *Synechocystis* PCC 6803, antigen specific affinity columns will be employed. To test which conditions would be suitable for elution of antigen from antibody, a dot blot technique was employed [14]. In brief, the purified antigen is blotted onto nitrocellulose and incubated with the antibody. The blots can then be washed using a variety of solutions or conditions. After this washing step a secondary, enzyme-coupled antibody is incubated with the blot and the amount of bound antibody is visualized using a colorimetric reaction. In the case of Ycf3 and BtpA elution with either very acidic (100 mM glycine pH 2.5) or very alkaline (100 mM triethylamine pH 11.5) conditions was most successful (Fig. 13).

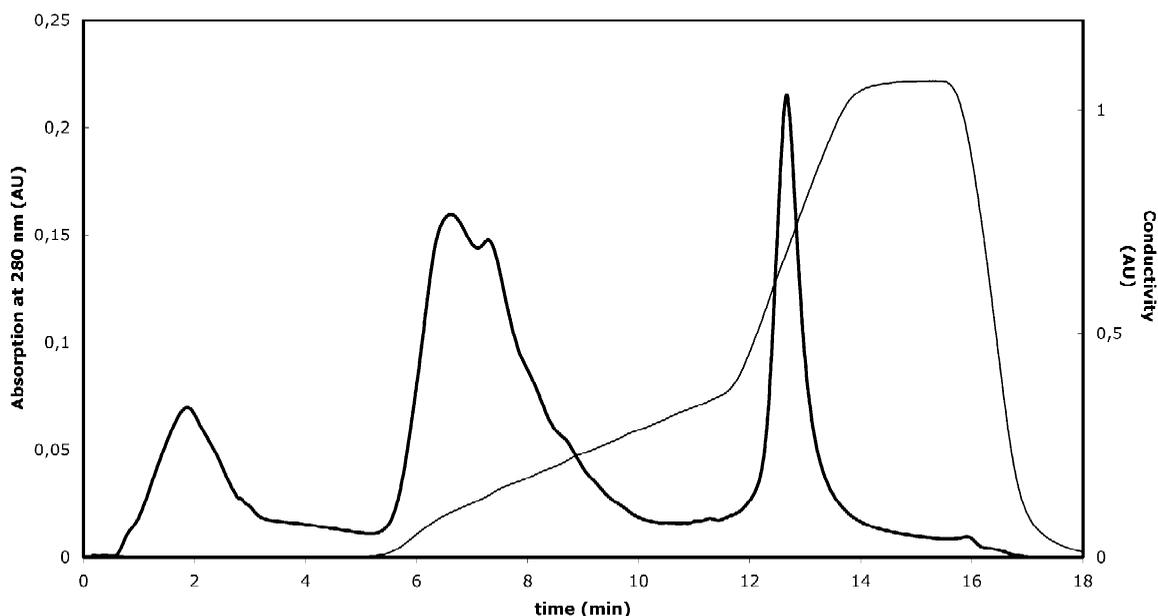


Fig. 6. Elution profile of solubilized BtpA-inclusion bodies on an anion-exchange column. Inclusion bodies were solubilized in 6 M urea, 20 mM Tris-HCl pH 8.0, with running buffer of 6 M urea, 20 mM Tris-HCl pH 8.0, 20 mM NaCl and loaded onto the column (Poros HQ50, column volume 6.5 ml) under denaturing conditions. Elution was achieved by raising the salt concentration in a two-step gradient to 1 M NaCl at a flow-rate of 4.5 ml/min. Bold line: absorption at 280 nm (AU); thin line: salt gradient (in % of maximal conductivity).

### 3.10. MALDI-TOF PMF identification

SDS-PAGE of BtpA preparations always showed faint bands of impurities or co-purified proteins. In order to identify these, we performed tryptic digests

and peptide mass fingerprinting (PMF). In Western blots, our HBtpA antibody had reacted with bands below  $30 \cdot 10^3$ . Now MALDI-TOF PMF showed that one prominent band at  $15 \cdot 10^3$  corresponds to an N-terminal BtpA-fragment of  $\sim 144$  amino acids.

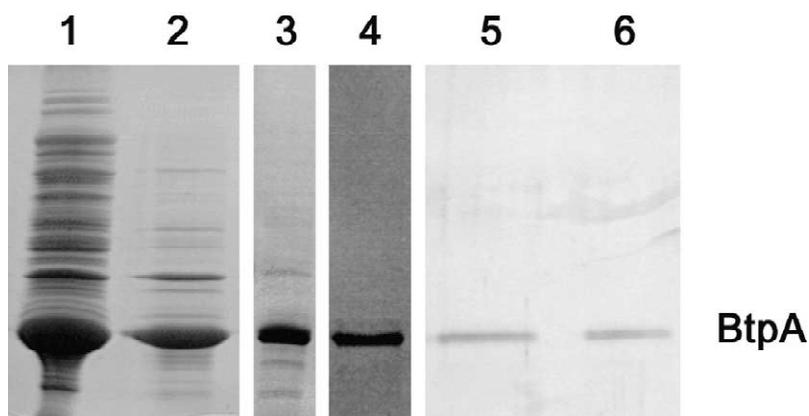


Fig. 7. SDS-PAGE analysis of different steps in BtpA-purification. Lane 1: total *E. coli* cell extract after overexpression in  $2 \times$ YT medium, growth for 16 h. Lane 2: purified inclusion bodies from whole cells shown in lane 1. Lane 3: after IEC on Poros HQ50. Lane 4: purified BtpA after Uno-Q column. Lane 5: pellet after refolding of denatured BtpA. Lane 6: final sample: refolded soluble BtpA.

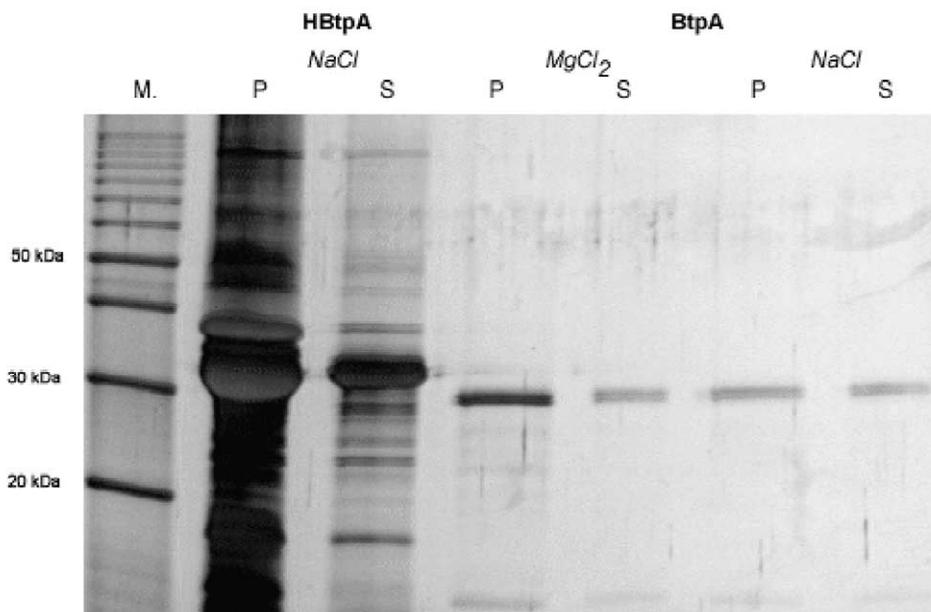


Fig. 8. SDS-PAGE of (H)BtpA after dialysis. Refolding by dialysis against 20 mM Tris-HCl pH 8.0 with either 20 mM NaCl or 20 mM MgCl<sub>2</sub>. HBtpA was renatured in the presence of 20 mM Tris-HCl pH 8.0, 20 mM NaCl. Lanes show either pellet (P) or supernatant (S) after centrifugation of renatured sample (150 000 g, 60 min). The gel was silver stained.

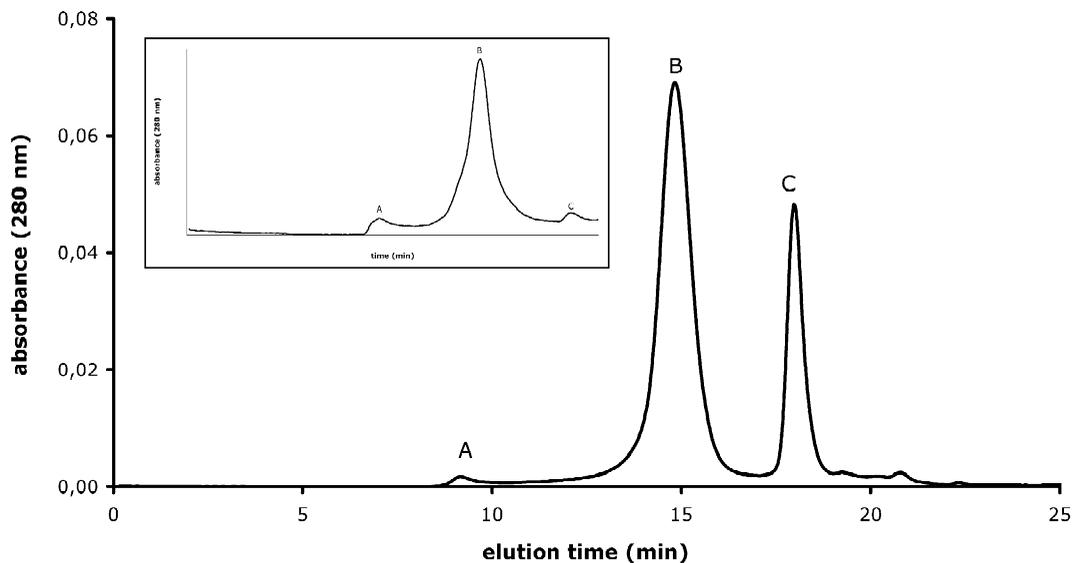


Fig. 9. Size exclusion chromatography profile of Ycf3. Protein was applied to Bio-Silect 400 column (column volume 14 ml, Bio-Rad) in running buffer 20 mM Tris-HCl pH 8.0, 150 mM NaCl at a flow-rate of 0.7 ml/min. Peaks correspond to molecular masses of A: void volume, B:  $260 \cdot 10^3$  (oligomer), C:  $20 \cdot 10^3$  (monomer). Calibration with Bio-Rad gel filtration standard. Inset: SEC of HYcf3 (refolded on the column) under the same conditions, peak labeling as for Ycf3.

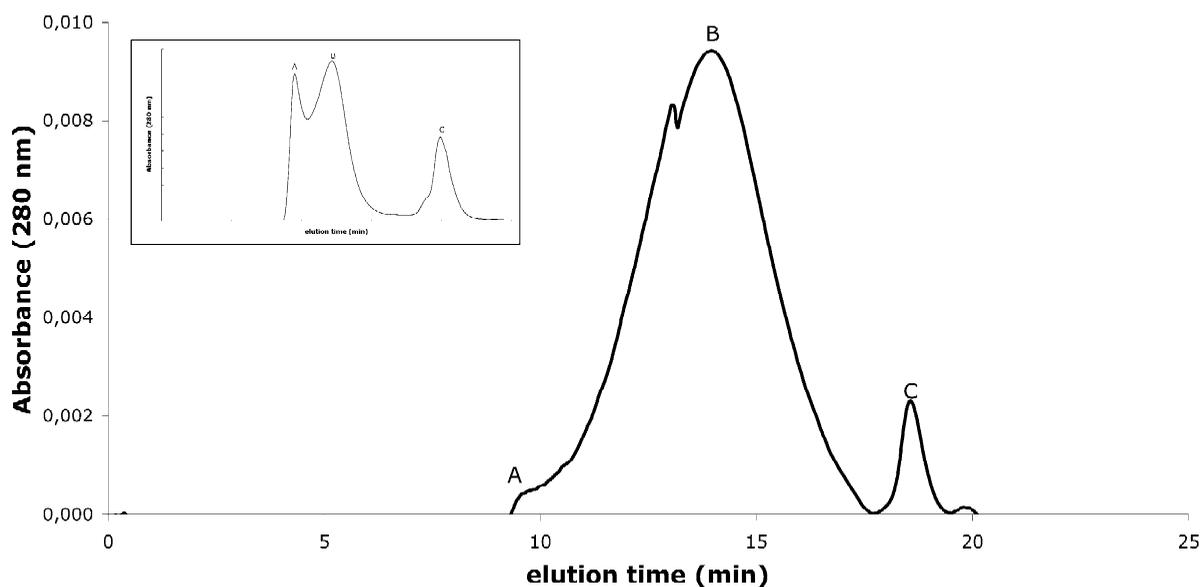


Fig. 10. Size exclusion chromatography of purified and refolded BtpA. Refolded protein was applied in running buffer (20 mM Tris–HCl pH 8.0, 150 mM NaCl at 0.7 ml/min) to a Bio-Silect 400 column (Bio-Rad, column volume 14 ml). Peaks correspond to molecular masses of A: void volume, B:  $650 \cdot 10^3$  (oligomer), C:  $30 \cdot 10^3$  (monomer). Calibration with Bio-Rad gel permeation standard. Inset: SEC of HBtpA under the same conditions, peak labeling as for BtpA.

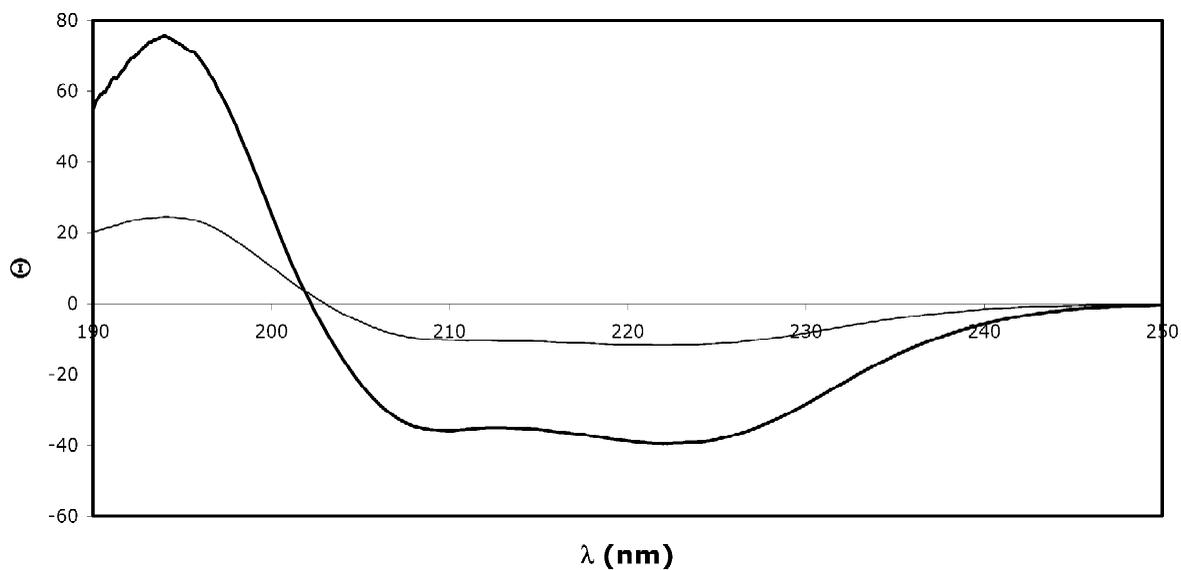


Fig. 11. CD-spectroscopy of HYcf3 and Ycf3. Thin line: Ycf3 in 4 mM Tris–HCl pH 8.0, 4 mM NaCl. Bold line: refolded HYcf3 in 4 mM Tris–HCl pH 8.0, 4 mM NaCl. Spectra shown have been arrived at by subtracting buffer spectra from those of the measured samples.

Table 1  
Secondary structure predictions for Ycf3 and BtpA

(H)Ycf3	Predictions for Ycf3			CD-spectroscopy	
	GOR IV	SOPMA	HNN	HYcf3	Ycf3
$\alpha$ -Helix	67.63	50.29	51.45	46	66
$\beta$ -Sheet	3.47	16.18	9.25	12	0.0
Turn	0	8.09	0	22	19
Random	28.90	25.43	39.31		15

(H)BtpA	Predictions for BtpA			CD-spectroscopy	
	GOR IV	SOPMA	HNN	HBtpA	BtpA
$\alpha$ -Helix	40	42	44	22	18
$\beta$ -Sheet	16	24	21	33	41
Turn				10	8
Random	43	34	36	35	33

The algorithms of different programs predict the procentual shares of several secondary structure classes. For computational analysis, values for non-tagged proteins only are shown. The CD-spectra were used to predict these structures as well, according to Ref. [8].

Further bands at  $74 \cdot 10^3$  and  $43 \cdot 10^3$  could be identified as the chaperone Hsp70 (DnaK) and the elongation factor EF-Tu A or B, respectively.

MALDI-TOF PMF was also used to confirm the identity of purified HYcf3 and Ycf3 (data not shown).

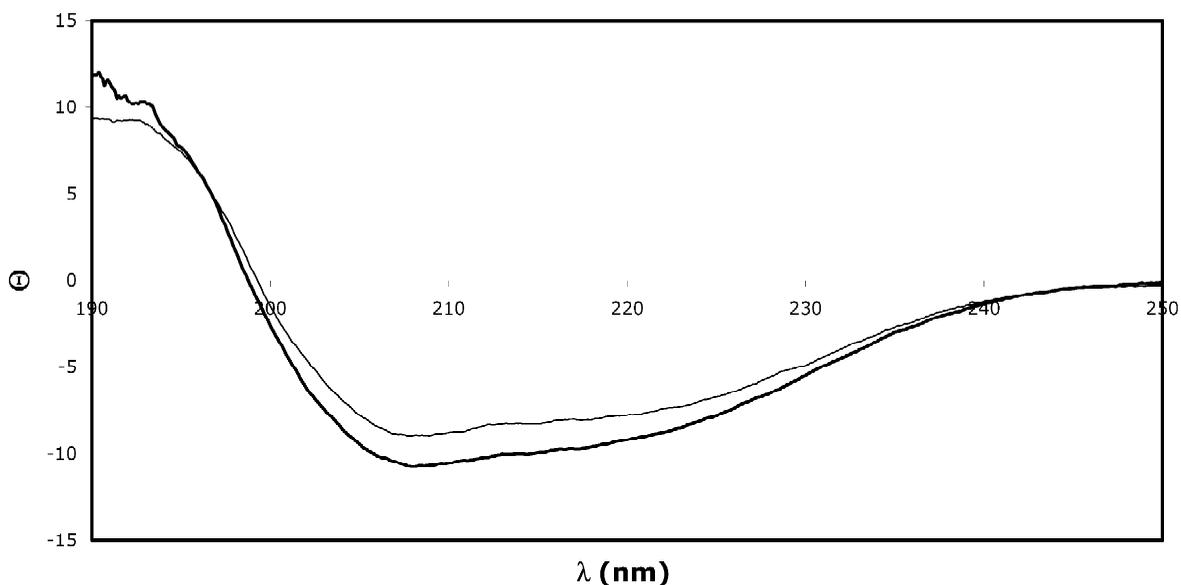


Fig. 12. CD-spectroscopy of HBtpA and BtpA. Bold line: CD-spectrum of refolded BtpA in 15 mM Tris–HCl pH 8.0, 7.5 mM NaCl; thin line: refolded HBtpA in same buffer. Data represent typical CD-spectra for the purified proteins, five measurements were accumulated for each spectrum.

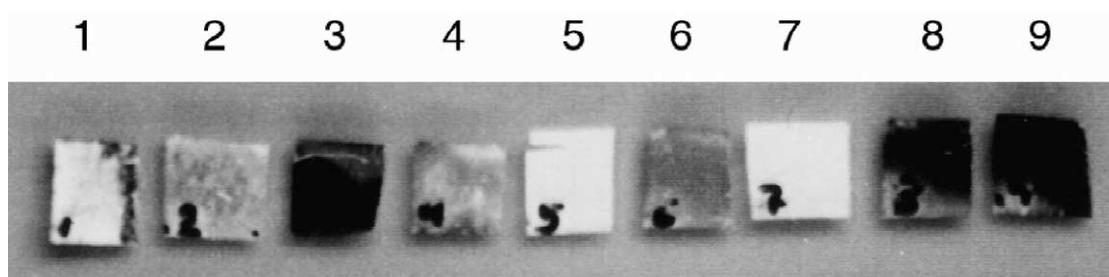


Fig. 13. Optimization of antibody elution from HBtpA. Different washing buffers have been tested on antibody bound to HBtpA-nitrocellulose as follows: (1) 5 M LiCl, 10 mM Tris-HCl pH 7.0; (2) 2 M guanidiniumhydrochloride, 10 mM Tris-HCl pH 8.0; (3) 2 M urea, 10 mM Tris-HCl pH 8.0; (4) 2 M MgCl<sub>2</sub>, 20 mM MES-NaOH pH 6.5; (5) 100 mM glycine pH 2.5; (6) 50% (v/v) ethyleneglycol, 10% (v/v) glycerol, 5 mM Tris-HCl pH 8.0; (7) 100 mM triethylamine pH 11.5; (8) A. dest.; (9) no elution.

## 4. Discussion

### 4.1. Purification of recombinant proteins

IMAC-purification strategies were very efficient for the His-tagged proteins, yielding almost pure protein very fast and on a large scale. We are able to purify ~10 mg HBtpA from 5 g fresh weight *E. coli* (amount after SEC). The yield of HYcf3 was 16 mg protein from 50 ml *E. coli* culture.

Purification of Ycf3 was accomplished by a combination of ammonium sulfate precipitation with IEC. This strategy yields 40 mg pure protein starting from 500 ml *E. coli* culture.

Purification of BtpA proved to be extremely difficult and could be done only by using two different anion-exchange matrices (Poros HQ and UNO-Q). The strategy yielded ~32 mg protein from 500 ml *E. coli* culture. This amount is reduced after refolding and SEC, so that only ~1 mg pure, refolded protein can be obtained.

### 4.2. Refolding

The refolding of HYcf3 by dialysis is very efficient and yields mainly monomeric protein with a CD-spectrum almost indistinguishable from that of soluble Ycf3. In addition, its derived secondary structural data correspond well with data from prediction programs, showing that  $\alpha$ -helices dominate. If HYcf3 is slowly refolded while bound to the nickel-loaded column, oligomerization is increased

compared to dialysis-refolding. Apparently, if the His-tag is bound to the column, it interferes less with refolding and subsequent homo-oligomerization of HYcf3 than in solution.

Sequence analysis shows that Ycf3 consists of three so-called TPR motifs [15]. This degenerate sequence motif occurs in a wide variety of proteins and functions as a module for protein-protein interactions—either via TPR-TPR or via TPR-non-TPR interaction. Members of this family have already been crystallized and their structure solved at high resolution [16]. It was shown that most of the proteins containing TPR-motifs consist to large extent of  $\alpha$ -helices. This strongly suggests that the refolding of HYcf3 was indeed successful. A striking observation was the fact that incorporation of divalent cations inevitably leads to precipitation of most of the protein. This might be explained by the metal-chelating properties of the His-tag: we used a tag comprised of ten His residues which should allow a more specific IMAC purification than with the commonly used shorter tag of six residues. The downside of this strategy is apparently a tendency to bind divalent cations which could lead, as in this case, to protein precipitation. Therefore, if choosing a His-tag the number of His residues should be kept to a minimum to reduce solubility problems.

Keeping (H)BtpA in solution after removing urea from the buffer proved to be difficult. Even at low protein concentrations the protein precipitates readily, so that a loss of ~30% of protein has to be accepted when renaturing BtpA. Divalent cations

increased the amount of precipitated protein significantly and might be explained for HBtpA as outlined for HYcf3. Nevertheless, the resulting solution contains specifically folded BtpA, whose measured secondary structure disagrees with the predicted one. This might be explained by the fact that the measured solution contains large amounts of BtpA-oligomers, and that the secondary structure prediction algorithms fail to take this into account. Thus, the prediction might be correct for the monomer but not for the oligomer of BtpA. But unless we manage to purify soluble BtpA directly from *Synechocystis*, we will have no reference data to judge that the refolded state is indeed the native one. BtpA is much more hydrophobic than Ycf3 and problems concerning solubility could be expected. Inclusion of detergents into the refolding solution could be an alternative which we will pursue further.

#### 4.3. Characterization of purified proteins

SEC shows that both proteins exist in a monomeric as well as in an oligomeric state. Whereas Ycf3 forms well-defined oligomers of  $\sim 260 \cdot 10^3$ , BtpA exhibits a rather larger oligomer of more than 20 subunits. This peak is very broad, indicating a heterogenic oligomer. In both cases, the presence of an N-terminal His-tag inhibits oligomerization to some extent. Only one fourth of BtpA is present in monomeric form in contrast to half of tagged BtpA. This effect is even more pronounced in the case of Ycf3 and HYcf3: non-tagged Ycf3 can almost exclusively be detected as oligomer. The well-defined peak of  $\sim 260 \cdot 10^3$  would correspond to an oligomer of  $\sim 13$  subunits. Ycf3 shows three so-called TPR-motifs [15,17], which form short double-helices. Three of these motifs together would form one turn of a super-helix, which might be involved in protein–protein interactions, as could be shown in the case of protein phosphatase 5 [16]. These modules for protein–protein interactions could be either involved in binding other proteins or in forming homo-oligomers. The latter cases is obvious for Ycf3.

To verify that both Ycf3 and BtpA form homo-oligomers in vivo we conducted yeast-2-hybrid experiments and could corroborate these findings: Ycf3 and BtpA show homo-oligomerization using this system (data to be published elsewhere).

For Ycf3, we could see that the addition of ions like  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$  to the running buffer either slightly influences the conformation of the protein or modifies a slight interaction with the SEC matrix. However, this slight effect might not be significant or important in vivo.

#### 4.4. MALDI-TOF PMF identification

The identification of additional bands in some of the BtpA samples presents interesting observations. Detection of an N-terminal fragment of BtpA supports our finding that BtpA degrades quite rapidly, even in vivo. We could also detect degradation bands in Western blots from *E. coli* extract. Since the preparation for Western blot assay was very fast and allowed little time for in vitro degradation, the degradation must have taken place already in the cells. This could also be observed in Western blots of *Synechocystis* extracts (data not shown). Thus, the use of protease inhibitors would probably not improve the stability of BtpA.

The chaperone Hsp70 co-purifies with BtpA. If this is a specific interaction, it would be very interesting. *E. coli* has a gene encoding a BtpA homologue (*sgcQ*), which might be the reason why the overexpressed protein is processed and targeted to the membrane of the host cells. We can however not exclude the possibility that this co-purification is entirely unspecific. The same holds true for the potential association of BtpA with the elongation factors EF-Tu from *E. coli*. Since these are among the most abundant proteins in the bacterium, their presence might be just due to the fact that they are hard to get rid of during any purification of a protein from *E. coli*. We expect affinity chromatography, which we want to undertake using either the antibody columns presented in this paper or columns for which His-tagged protein has been coupled to chelating sepharose, to answer these questions and to expose more interaction partners of our proteins.

In summary, we have developed purification methods for Ycf3 and BtpA with and without a His-tag. This is a first step in the elucidation of the molecular role of these proteins for membrane protein assembly. Initial structural characterization of the proteins has already shown that both can form homo-oligomers which might be important for their function, as

these results were confirmed by a yeast-2-hybrid system.

## 5. Nomenclature

IEC	ion-exchange chromatography
IMAC	immobilized metal affinity chromatography
IPTG	isopropylthiogalactoside
MALDI–TOF-MS	matrix assisted laser desorption/ionisation time-of-flight mass spectrometry
PMF	peptide mass fingerprinting
PS I	photosystem I
SEC	size exclusion chromatography

## Acknowledgements

We thank Professor Dr M. Roegner for his continuous support, encouragement and valuable discussions. Financial support by the Deutsche Forschungsgemeinschaft (SFB 480, T.M.E.S. and J.K.) and the Fonds der Chemischen Industrie (J.K.) is gratefully acknowledged.

## References

- [1] P.R. Chitnis, *Plant Physiol.* 111 (3) (1996) 661.
- [2] P. Jordan et al., *Nature* 411 (6840) (2001) 909.

- [3] T.M.E. Schwabe, J. Kruip, *Indian J. Biochem. Biophys.* 37 (6) (2000) 351.
- [4] P.R. Chitnis, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52 (2001) 593.
- [5] E. Boudreau et al., *EMBO J.* 16 (20) (1997) 6095.
- [6] V.V. Bartsevich, H.B. Pakrasi, *J. Biol. Chem.* 272 (10) (1997) 6382.
- [7] U.K. Laemmli, *Nature* 227 (259) (1970) 680.
- [8] J.T. Yang, C.S. Wu, H.M. Martinez, *Methods Enzymol.* 130 (1986) 208.
- [9] J. Garnier, J.F. Gibrat, B. Robson, *Methods Enzymol.* 266 (1996) 540.
- [10] C. Geourjon, G. Deleage, *Comput. Appl. Biosci.* 11 (6) (1995) 681.
- [11] Y. Guermeur, Ph.D. Thesis, University Pierre et Marie Curie, Paris, 1996.
- [12] V. Neuhoff, R. Stamm, H. Eibl, *Electrophoresis* 6 (1985) 427.
- [13] T. Kaneko, S. Tabata, *Plant Cell Physiol.* 38 (11) (1997) 1171.
- [14] S.M. Rutkowska, P.M. Skowron, *Biotechniques* 27 (5) (1999) 930.
- [15] M. Goebl, M. Yanagida, *Trends Biochem. Sci.* 16 (5) (1991) 173.
- [16] A.K. Das, P.W. Cohen, D. Barford, *EMBO J.* 17 (5) (1998) 1192.
- [17] J.R. Lamb, S. Tugendreich, P. Hieter, *Trends Biochem. Sci.* 20 (7) (1995) 257.