Chapter 4

Structural divergence of paralogous S-components from ECF-type ABC transporters


4.1 Abstract

Energy Coupling Factor (ECF) proteins are ATP Binding Cassette (ABC) transporters involved in the import of micronutrients in prokaryotes. They consist of an ECF module, which contains the two nucleotide-binding subunits and the integral membrane subunit EcfT, and a second integral membrane subunit that captures the substrate (the S-component). Different S-components, unrelated in sequence and specific for different ligands, may interact with the same ECF module. Here, we present a high-resolution crystal structure at 2.1 Å of the biotin-specific S-component BioY from Lactococcus lactis. BioY shares only 16% sequence identity
with the thiamin-specific S-component ThiT from the same organism, of which we recently solved a crystal structure (Erkens et al., 2011). Consistent with the lack of sequence similarity, BioY and ThiT display large structural differences (RMSD = 5.1 Å), but the divergence is not equally distributed over the molecules: the S-components contain a structurally conserved N-terminal domain that is involved in the interaction with the ECF module, and a highly diverged C-terminal domain that binds the substrate. The two-faced nature of these proteins explains how the S-components with large overall structural divergence can interact with the same ECF module, while at the same time specifically bind very different substrates with subnanomolar affinity. Solitary BioY (in the absence of the ECF module) is monomeric in detergent solution, binds D-biotin with a high affinity, but does not transport the substrate across the membrane.

4.2 Introduction

Energy Coupling Factor (ECF) proteins are an abundant class of ATP Binding Cassette (ABC) transporters, involved in the import of vitamins and transition metal ions in prokaryotes (chapter 1, 2 and (D. A. Rodionov et al., 2006; Hebbeln et al., 2007; D. Rodionov et al., 2009)). Like all ABC transporters, ECF transporters consist of two cytosolic nucleotide-binding domains (NBDs), which are associated with integral membrane subunits that form the translocation pore. In ECF transporters the two NBDs (EcfA and EcfA' which may be identical or homologous) and a single membrane subunit (EcfT) form a so-called energizing or ECF module. A second integral membrane protein (the S-component) binds the substrate and forms a complex with the ECF module to create a functional transporter. This organization is typical for ECF transporters (chapter 1, 2 and 3), since other ABC importers utilize a soluble substrate-binding protein to capture ligands (chapter 1 and (Higgins, 1992; Berntsson et al., 2010)). In many ECF transporters multiple S-components (specific for different substrates) can interact with the same energizing module (chapter 2, 3 and (D. Rodionov et al., 2009)). Strikingly, S-components from a single organism, which interact with the same ECF module, are generally not homologous at the sequence level. To gain insight in the characteristic modularity of ECF transporters, one needs to compare crystal structures of different S-components that interact with the same ECF module, i.e. S-components from a single organism. Crystal structures of the S-components ThiT from Lactococcus lactis (thiamin-specific) and RibU from
Structural divergence of paralogous S-components

*Staphylococcus aureus* (riboflavin-specific) have recently been determined (Erkens et al., 2011; Zhang et al., 2010). We now present the crystal structure at 2.1 Å of the S-component BioY from *Lactococcus lactis*, which we show to bind biotin. BioY and ThiT form complexes with the same ECF module (see chapter 3) and share only 16 % sequence identity.

### 4.3 Materials and methods

#### 4.3.1 Protein expression.

Selenomethionine-substituted BioY containing a N-terminal decahistidine tag was expressed in *Lactococcus lactis* strain NZ9000 (Ruyter et al., 1996), as previously described (Berntsson et al., 2009). Briefly, the cells were grown semi-anaerobically in chemically defined medium (CDM) to an OD$_{600nm}$ of 1.5. At this point, the cells were spun down and resuspended in CDM with selenomethionine instead of methionine. After 20 minutes, *bioY* expression was induced by the addition of 0.1% (v/v) of culture supernatant from the nisin A-producing strain NZ9700 (Ruyter et al., 1996). The cells were grown to an OD$_{600nm}$ of 4, and then harvested by centrifugation and subsequently resuspended in buffer A (50 mM Na-Hepes pH 7.5, 300 mM NaCl and 10% (v/v) glycerol). Cell lysis was performed by passing the cells twice through a cell disruptor (Constant Systems Ltd) at a pressure of 39 kPsi, 4 °C. Prior to the disruption, MgSO$_4$ (5 mM) and DNase (100 µg/mL) were added. Unbroken cells were removed by centrifugation at 6000 xg, 15 min, 4 °C. Membrane vesicles were collected by a subsequent centrifugation at 267,000 xg for 80 min at 4 °C, and resuspended and homogenized in buffer A to a protein concentration of 40 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C.

#### 4.3.2 Protein purification.

Membrane vesicles (100 mg total protein) were thawed, and diluted in buffer A to approximately 5 mg/mL total protein. Solubilization was done by the addition of 1% (w/v) of dodecyl-maltoside and incubation at 4 °C for 1h (the mixture was mixed by gentle rotation). Unsolubilized material was spun down at 267,000 xg and 4 °C for 20 min. 0.5 mL Ni-Sepharose plus 15 mM imidazole pH 7.8 were added to the supernatant, and the mixture was incubated at 4 °C for 1h.
(under gentle rotation). The suspension was poured into a 10 mL disposable column (Bio-Rad), and the flow-through was discarded. The column was washed with 20 column volumes of buffer B (50 mM Na-Hepes pH 7.5, 300 mM NaCl, 50 mM imidazole pH 7.8 plus 0.35% (w/v) n-nonyl-β-D-glucopyranoside (NG, Anatrace)). The protein was eluted from the column in 2 fractions (0.35 and 0.75 mL, respectively) with buffer B supplemented with 500 mM imidazole pH 7.8. The second elution fraction was loaded onto a Superdex 200 10/300 GL gel filtration column (GE Healthcare), equilibrated with buffer C (20 mM Na-Hepes pH 7.5, 150 mM NaCl plus 0.35% (w/v) NG). Peak fractions were concentrated to 7 mg/mL, using a Vivaspin 30 kDa molecular weight cutoff concentrator (VVR International), and immediately used for crystallization trials or other biochemical assays. For biochemical characterization of BioY, the protein was purified from cells grown in CDM without biotin (for the isolation of substrate-free protein). The purification protocol was slightly modified: solubilization was done in a buffer containing 50 mM potassium-phosphate, 300 mM NaCl, 10% glycerol plus 1% maltose-neopentyl glycol 3 (MNG-3) (Chae et al., 2010), pH 7.5. The nickel-sepharose column was washed with 20 column volumes of 50 mM potassium phosphate, pH 7.5, 300 mM NaCl, 10% glycerol, 50 mM imidazol plus 0.03% MNG-3, and eluted with the same buffer supplemented with 500 mM imidazol. Size-exclusion chromatography was done in 50 mM potassium phosphate, pH 7.5, 150 mM NaCl plus 0.03% MNG-3.

4.3.3 Crystallization.

Initial crystal hits of BioY were found in several conditions, all containing high concentrations of PEG and pH values between 7 and 9. Optimization of the conditions yielded diffraction-quality crystals with a size of ca 100x50x50 µm. The best crystals were grown at 5 °C with the reservoir solution containing 0.1 M Tris pH 8.0, 0.05-0.2 mM CaCl₂ plus 45-50% PEG400. Due to the high PEG400 concentration, no further cryo-protectant was needed, and the crystals were directly fished from the drop and flash frozen in liquid nitrogen.

4.3.4 Structure determination.

Diffraction data was collected at the PX1 beam-line at the Swiss Light Source. Single-wavelength Anomalous Dispersion (SAD) data on SeMet-BioY was collected to 2.1 Å, at 100 K with a wavelength of 12.657 keV. Data processing and reduction
Structural divergence of paralogous S-components

were carried out, using XDS (Kabsch, 1993) and programs from the CCP4 suite (Collaborative Computational Project, 1994). Relevant statistics can be found in Table 4.1. Initial phase information was found using autoShap (Vonrhein et al., 2007), and an initial model containing 95% of the residues could be built using ARP/warp (Langer et al., 2008). 15 selenium sites were found within the asymmetric unit, corresponding to all 5 of the possible sites per BioY molecule. A few cycles of refinement using Refmac5 (Murshudov et al., 1997), and non-crystallographic symmetry with loose restraints, interspersed with manual model building in Coot (Emsley & Cowtan, 2004), were necessary to finish the model. Water molecules were automatically placed in F_O-F_C Fourier difference maps at 3σ cutoff levels and validated to ensure correct positioning, using Coot. The final protein model contains residues 1-188 for all 3 molecules in the asymmetric unit. Electron density that could correspond to acyl chains (without visible density for headgroups) was not modeled. R_work and R_free of the final model after refinement were 18.6% and 20.6%, respectively. All structure figures were prepared using PyMOL (Schrödinger, LLC).

4.3.5 Fluorescence titration.

Tryptophan fluorescence was measured in a stirred quartz cuvette on a SPEX Fluorolog 322 fluorescence spectrophotometer (Jobin Yvon) at 25 °C. Purified biotin-free BioY was diluted in size-exclusion chromatography buffer to the indicated concentration (final volume 1000 µL). D-biotin was added in 0.5 µL steps. The excitation and emission wavelengths were 280 nm and 360 nm, respectively. The data was analyzed as described (Erkens & Slotboom, 2010) (see also chapter 6). Because of the high-affinity binding by BioY, the protein was diluted to ~10 nM for titrations with biotin. For BioY from R. capsulatus, the fluorescence measurements were done in the same way, except that the emission wavelength was 349 nm and that the protein concentration was ~50 nM.

4.3.6 Light scattering.

The oligomeric state of BioY was determined via size-exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) as described before (see chapter 3). We used BioY isolated from cells grown on rich medium (with biotin) that was purified in the same way as the biotin-free protein.
4.3.7 Uptakes by proteoliposomes containing BioY.

Substrate-free BioY (in a buffer of 50 mM potassium phosphate, pH 7.5, 150 mM NaCl plus 0.03% MNG-3) was reconstituted into proteoliposomes at a protein:lipid ratio of 1:250, essentially as described in chapter 3. Proteoliposomes were subjected to three cycles of freeze thawing using liquid nitrogen, extruded through a 400 nm pore size polycarbonate filter (Avestin), and centrifuged (267,000 rcf, 4 °C, 20, Beckman TLA 100.4 rotor) in a 50 mM potassium phosphate buffer, pH 7.5. For transport assays, 2 µL of proteoliposomes (125 µg/µL lipid concentration) were diluted into 200 µL of buffers containing 25 nM [3H]biotin and 1 µM non-labeled biotin. The buffer compositions were varied in order to create different membrane gradients: (i) 57 mM sodium phosphate pH 6.5 containing 1 µM valinomycin (diluted from a 3 mM stock in ethanol) to obtain gradients for protons and sodium ions in combination with membrane potential; (ii) 57 mM potassium phosphate pH 6.5 (proton gradient); (iii) 57 mM sodium phosphate pH 7.5 (sodium ion gradient); (iv) 50 mM sodium phosphate pH 7.5 containing 1 µM valinomycin (gradient of sodium ions in combination with membrane potential); (v) 50 mM potassium phosphate pH 7.5 (no gradients). Buffers were pre-warmed to 25 °C and the suspension was briefly vortexed after addition of proteoliposomes. After 15, 60 or 180 seconds 2 mL of ice-cold 50 mM potassium phosphate pH 7.5 was added followed by rapid filtration over 0.45 µm pore-size cellulose nitrate filter (Whatman Maidstone UK). The filters were washed once with 2 mL ice-cold 50 mM potassium phosphate pH 7.5. An extra timepoint was taken (240 s) after the radiolabeled biotin had been chased with 1 mM of unlabeled biotin. Background signal was determined by using liposomes without BioY and with all the gradients present. Radioactivity trapped on the filters was measured by addition of 2 mL of emulsifier plus scintillation liquid, and subsequent counting in a Perkin Elmer 1600CA scintillation counter.

4.3.8 Uptakes by proteoliposomes containing EcfAA'T-BioY.

EcfAA'T-BioY complexes were expressed in E. coli and purified with Nickel-Sepharose and size exclusion chromatography as described in chapter 3. Purified complexes were reconstituted into proteoliposomes and transport assays were done as described for EcfAA'T-NiaX in chapter 3, with [3H]biotin (American Radiolabeled Chemicals) at a concentration of 62.5 nM for uptake and 600 nM for efflux experiments, no unlabeled biotin was added.
4.4 Results

Selenomethionine (SeMet)-substituted BioY was produced in the expression strain *L. lactis* NZ9000 (Kunji et al., 2003) and purified using the detergent n-nonyl-β-D-glucopyranoside, which was also used for the crystallization of RibU and ThiT (Zhang et al., 2010; Erkens et al., 2011). SeMet-BioY crystals of space group C2 diffracted to 2.1 Å and were used to solve the structure using multi-wavelength anomalous dispersion (MAD) phasing (Table 4.1).

The electron density was of high quality and allowed for modeling of the entire amino acid sequence of BioY, except for the N-terminal tag (MHHHHHHHHHA), which was used for metal-affinity purification. After refinement well-defined residual density was observed inside the protein, which could be assigned unambiguously to a D-biotin molecule. In addition, five complete detergent molecules were modeled into densities around the protein. The asymmetric unit contained three copies of BioY that were virtually identical (r.m.s. deviation <0.2 Å), with molecule A rotated ~90° with respect to B, and molecule C rotated ~160° with respect to B (figure 4.1a).

Table 4.1. Data collection and refinement statistics.
Values in parenthesis are for highest-resolution shell.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>SeMet-BioY1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>C2</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>89.8, 57.4, 166.9</td>
</tr>
<tr>
<td>a, β, γ (°)</td>
<td>90.0, 91.1, 90.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>48.3 - 2.1</td>
</tr>
<tr>
<td>R_sym (%)</td>
<td>6.5 (50.4)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>15.3 (2.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.4 (90.7)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>10.8</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>48.3 - 2.1</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>47255</td>
</tr>
<tr>
<td>R_work/R_free</td>
<td>18.5/20.5</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>4335</td>
</tr>
<tr>
<td>Biotin</td>
<td>48</td>
</tr>
<tr>
<td>Water</td>
<td>123</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>33</td>
</tr>
<tr>
<td>Biotin</td>
<td>38</td>
</tr>
<tr>
<td>Water</td>
<td>52</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.015</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.51</td>
</tr>
</tbody>
</table>
Figure 4.1. **Crystal structure of BioY from Lactococcus lactis**  
A Secondary structure cartoon showing the relative orientation of the three molecules of BioY in the asymmetric unit of the BioY crystals. The three molecules are colored differently (orange, dark grey and light grey) and the approximate membrane boundaries are indicated with dotted lines for the orange and dark grey molecules. The trimeric arrangement is incompatible with a membrane-embedded oligomer.  
B A monomer of BioY in secondary structure cartoon representation colored from blue (N-terminus) to red (C-terminus). The bound biotin molecule is shown in stick representation with carbon atoms in orange. The left and right views are from the plain of the membrane and along the membrane normal (from the outside), respectively.  
C Binding site of biotin. The biotin molecule is shown in orange and the interacting residues from BioY in gray. Electron density for biotin ($2F_o - F_c$ map contoured at 1.5$\sigma$) in blue mesh.  
D Sliced surface representation of BioY showing the binding cavity, with the bound biotin shown in orange. Coloring and viewpoints as in panel B.

The relative orientation of the three proteins in the asymmetric unit is incompatible with a membrane environment, and the observed crystallographic trimer is very likely only due to crystal contacts. Indeed, light-scattering experiments (SEC-MALLS) confirmed that BioY is monomeric in detergent solution (figure 4.2).
BioY has six membrane-spanning $\alpha$-helices (figure 4.1a, b). The fold of BioY resembles the folds of RibU and ThiT, but superimposition of all backbone atoms of BioY and either ThiT or RibU revealed large structural differences with root mean square deviations of 5.1 and 4.4 Å, respectively. The structural divergence is not equally distributed over the length of the proteins. BioY, ThiT and RibU have structurally very similar N-terminal domains (Helices 1-3) and highly variable C-terminal domains (Helices 4-6) (figure 4.3).

The long membrane-embedded loop between helices 1 and 2 (indicated in black in figure 4.3) bridges the two domains and likely has an important functional role (see below). The N-terminal domain of the S-components contains a conserved motif (AxxxA, with x mostly hydrophobic amino acids) in helix 1. For ThiT we have shown that the motif is essential for the interaction with the ECF module (Erkens et al., 2011). The structural conservation of the N-terminal halves of S-components likely forms the basis of their shared use of the ECF module. The variable domain contains the substrate-binding site. BioY was crystallized with a biotin molecule bound to a site near the extracellular face of BioY (figure 4.1b, d). The ligand is
Figure 4.3. Superimposed structures of BioY (orange), ThiT (gray) and RibU (yellow), viewed from the plane of the membrane (left panel) and a top-view from the outside of the cell (right panel). The structures have been superimposed on helices 1 & 3 in order to highlight the structural similarities of helices 1-3 and the differences of helices 4-6. Loops 1 are indicated in black. The N- and C-terminus are marked with N and C, respectively (left panel). Helices 1-6 are marked with H1-H6 (right panel). See also figure 1.13.

mostly occluded, except for the carboxylate tail that has access to the solvent via a narrow tunnel (figure 4.1d). The tunnel is too small to allow passage of the biotin molecule. Biotin interacts with helices 4, 5 and 6, and the loop between helix 3 and 4 in the variable domain (figure 4.1c). In addition, the loop between helices 1 and 2 from the N-terminal domain directly binds the ligand. Most of the residues involved in biotin binding are conserved among BioY homologues (figure 4.4).

The side-chains of Asp163 and Lys166, as well as the backbone carbonyl of Pro37 and the backbone NH of Ile39 interact with the imidazole ring. The sidechains of Phe159 and Tyr92 stack with the imidazole and thiophene rings of biotin, respectively, and Arg93 Interacts with the carboxylate of the pentanoic acid group (figure 4.1c).

Biotin was bound to BioY in the crystals, even though the substrate had not been present during the purification or crystallization. Apparently, biotin originating from the growth medium remained associated with the protein, indicating slow off-rates and high-affinity binding. To produce biotin-free BioY, the expressing cells were cultivated in defined growth medium in the absence of biotin. Biotin binding to the purified apo-protein was measured using the intrinsic protein fluorescence titrations. These measurements revealed a protein:biotin binding stoichiometry of 1:1 and a dissociation constant $K_d$ of 0.3 nM (figure 4.5).
We reconstituted purified apo-BioY into proteoliposomes to determine if BioY could mediate transport of biotin in addition to binding. The reconstituted protein mediated rapid binding of radiolabeled biotin to the proteoliposomes, but in none of the conditions that we tested (presence/absence of membrane gradients of protons, sodium ions and a membrane potential of -120 mV) substrate accumulated inside the liposomes (figure 4.6). We conclude that BioY, just like ThiT and RibU, is a substrate-binding protein, which alone cannot facilitate transport of biotin across the membrane.

When BioY was co-purified with the ECF module as described in chapter 3 and the purified EcfAA’T-BioY complexes were consecutively reconstituted in proteoliposomes, MgATP-dependent uptake was observed (figure 4.7a).
Figure 4.5. Biotin binding to BioY. A. Titration of 10 nM BioY with D-biotin. The intrinsic protein fluorescence was measured (excitation wavelength 280 nm, emission wavelength 360 nm). Inset: fluorescence spectrum of 300 nM BioY in the absence of biotin (black line) and in the presence of a saturating amount of biotin (1 mM, gray line) B. Biotin binding to BioY\textsubscript{Rc}. Titration of 50 nM BioY from \textit{Rhodobacter capsulatus} with D-biotin. The intrinsic protein fluorescence was measured (excitation wavelength 280 nm, emission wavelength 349 nm). Inset: fluorescence spectrum of BioY in the absence of biotin (black line) and in the presence of a saturating amount of biotin (100 nM, gray line).

Figure 4.6. Biotin binding to liposomes containing purified and reconstituted BioY. Biotin association with the liposomes was measured in the presence of a proton gradient (10-fold; white circles), a sodium ion gradient (100-fold; white triangles), a sodium ion gradient plus a membrane potential (~120 mV; white squares), combined gradients of protons and sodium ions plus a membrane potential (black circles), and in the absence of any ion gradient (black triangles). Black squares indicate empty liposomes in the presence of all gradients. After 210 seconds an excess of unlabeled biotin was added.
Figure 4.7. Biotin transport in liposomes with purified and reconstituted EcfAA'T-BioY. A Biotin uptake via EcfAA'T-BioY reconstituted in proteoliposomes loaded with 10 mM MgATP (black circles) or with 10 mM MgADP (white circles). B Biotin efflux with proteoliposomes containing EcfAA'T-BioY and loaded with biotin. At 4 minutes the sample was split and either buffer was added (white circles) or MgATP (black circles).

We calculated that biotin is accumulated 29-times in the lumen of the liposomes containing MgATP (based on an internal volume of the liposomes of \(\sim 0.73 \text{ mL/mg lipid}\) (Geertsma et al., 2008)). However, the calculated number of transported biotin molecules is lower than the number of EcfAA'T-BioY complexes (assuming that all purified protein is reconstituted quantitatively, which is probably incorrect). To exclude the possibility that the observed accumulation of biotin is not due to transport but rather to binding in a MgATP dependent way, proteoliposomes were loaded with radiolabeled biotin and the efflux was measured (figure 4.7b). Although there is leakage of biotin from the liposomes upon dilution into buffer without ATP, the presence of MgATP increased the rate of efflux, showing that a significant fraction of biotin was transported across the membrane.

4.5 Discussion

The ECF module from *L. lactis* can interact with eight different S-components, six of which share less than 20% sequence identity with any of the other S-components in the organism (chapter 3). For two of the S-components from *L. lactis*, BioY
and ThiT, we now have determined crystal structures at a high-resolution. There is large structural variation between BioY and ThiT, with the two proteins displaying an RMSD of 5.1 Å. The structural differences are in line with the lack of sequence conservation (16% identity between BioY and ThiT). Most variation is in the C-terminal domain (helix 4-6), which is involved in substrate binding. The substrates thiamin and biotin are chemically very different explaining the large structural variation in the binding domains. Similarly, the C-terminal domain of the riboflavin-binding S-component RibU from *S. aureus*, for which a crystal structure of moderate resolution is available, is structurally divergent. The N-terminal domains (consisting of helices 1-3) are more similar in all structures and contain the AxxxA motif that was found to be essential for thiamin transport by the ECF module ThiT complex (Erkens et al., 2011). This domain interacts with the ECF module. Although BioY, ThiT from *L. lactis* and RibU from *S. aureus* are structurally very different, there are also similarities. In all three proteins, not only residues from the non-conserved C-terminal domain interact with the substrate, but also residues from the loop between helices 1 and 2 in the N-terminal domain (residues 36-39 in BioY). Loop 1-2 forms a bridge between the substrate-binding domain and the N-terminal domain that interacts with the ECF module. The loop therefore may mediate coupling between conformational changes in the ECF module induced by ATP binding/hydrolysis, and substrate binding/release in the S-component (figure 4.3). In response to ATP binding/hydrolysis in the ATPase domains, the loop may rearrange and thereby perturb the substrate-binding site while at the same time opening up the pathway for substrate translocation. Because loop 1-2 is located on the external face of the S-components, whereas the nucleotide-binding domains are cytoplasmic, the ATP-dependent conformational changes will have to be communicated/transduced via a transmembrane protein, e.g. the EcfT subunit.

BioY from *L. lactis* is homologous with BioY from *Rhodobacter capsulatus* (BioY<sub>Rc</sub>, 35% sequence identity). The latter protein is the archetypical member of the S-component superfamily (Hebbeln et al., 2007), and based on characterization of BioY<sub>Rc</sub> it has been suggested that the BioY-family of S-components may be different from other S-components both in terms of oligomeric structure, and in terms of transport mechanism (Hebbeln et al., 2007; Finkenwirth et al., 2010). In contrast, our biochemical and structural characterization of BioY suggest a unifying structure and mechanism.

First, the stable structural unit of BioY in detergent solution is a monomer. Analysis of the crystal packing showed that adjacent BioY molecules in the crystals
do not represent physiological oligomers, because their relative orientations are incompatible with embedding in a membrane environment. In addition, static light scattering experiments (SEC-MALLS) confirmed that BioY is monomeric, just like other S-components such as ThiT (Erkens et al., 2011) and RibU (Zhang et al., 2010). Similarly, we found that BioY<sub>Rc</sub> from <i>Rhodobacter capsulatus</i> is monomeric in detergent solution (ref (Finkenwirth et al., 2010), and unpublished data). Therefore, from a structural point of view the BioY-family is not different from other S-components. In contrast, <i>in vivo</i> fluorescence lifetime measurements have indicated that BioY<sub>Rc</sub> forms oligomers when heterologously overproduced in <i>E. coli</i>. We cannot exclude that the oligomeric state of S-components is different in membranes, but none of the available structures (BioY, ThiT and RibU) show possibilities for extensive interfaces, which would be required for stable oligomer formation. Therefore the structure in membranes must be different, were oligomerization to occur.

Second, D-biotin bound to BioY with a high affinity (K<sub>d</sub> 0.3 nM). The affinity measured <i>in vitro</i> compares well to the K<sub>d</sub> of 0.15 nM that was found in the 1970s for biotin binding to whole cells of <i>Lactobacillus casei</i> (Henderson et al., 1985), which was likely mediated by a BioY orthologue. Fluorescence titrations showed the presence of a single binding site per BioY molecule, which is consistent with the crystal structure. Subnanomolar K<sub>d</sub> values and 1:1 binding stoichiometry are conserved features of S-components (Duurkens et al., 2007; Erkens & Slotboom, 2010; Eudes et al., 2008; Henderson & Zevely, 1978; Henderson et al., 1985), and these properties are also conserved in BioY<sub>Rc</sub> from <i>R. capsulatus</i> (K<sub>d</sub> of ~0.7 nM, figure 4.5b).

Third, for many S-components it has been shown that they cannot transport their substrates without the ECF module (chapter 1, 2, 5, 7 and (D. Rodionov et al., 2009; Erkens et al., 2011; Zhang et al., 2010)). The exception is BioY<sub>Rc</sub> from <i>R. capsulatus</i>, which was reported to be a low affinity, high capacity transporter in the absence of the ECF module when overexpressed heterologously in <i>E. coli</i> cells (Hebbeln et al., 2007). Although such transport kinetics is not incompatible with the high affinity binding that we find for BioY, it would require input of energy. We could not detect transport activity of BioY from <i>L. lactis</i> reconstituted in proteoliposomes. The reconstituted protein supported binding of radiolabeled biotin only. Therefore, we conclude that BioY from <i>L. lactis</i> is not different from the other S-components, in that it acts as a high-affinity substrate-binding protein for the ECF-transporters, and that it cannot transport biotin on its own. The lack of transport activity is
consistent with the absence of an obvious translocation path through BioY. Most of the conserved residues in BioY (figure 4.4) are involved in either ligand binding (near the extracellular face of the protein), or tight packing of the helices (many glycines), or interaction with the ECF module (AxxxA motif). Although a riboflavin pathway through RibU was proposed based on the presence of conserved residues (Zhang et al., 2010), we do not consider the proposed residues to be well-conserved (see also figure 7.2 and the discussion in chapter 7). In addition, the same authors showed that RibU does not allow uptake of riboflavin in cells (Zhang et al., 2010). Similarly, ThiT does not support transport of thiamin, and the structure does not show a translocation pathway (Erkens et al., 2011). Since no translocation pathway can be observed through BioY, ThiT or RibU, the pathway is hypothesized to be located at the interface of the S-component and the transmembrane protein of the ECF module, as is the case in other ABC transporters (see chapter 1 and 7 and figure 7.2 therein).

For reconstituted BioY in the absence of the ECF module 1 biotin molecule binds per 5.7 BioY molecules. The biotin binding stoichiometry is not 1:1 as it is in detergent solution. There are two probable explanations for the deviation First, BioY may insert in both the right-side-out and inside-out orientation in the liposomes. Second, the protein concentration is based on the assumption that there is quantitative recovery of active protein after reconstitution, which is probably not the case. Reconstituted EcfAAT-BioY complexes transported biotin in a MgATP-dependent manner. However, the number of transported biotin molecules was lower than expected since less than one biotin molecule is transported per complex (1 biotin per 43 EcfAAT-BioY complexes). Again, it is likely that not all EcfAAT-BioY complexes are inserted properly into the proteoliposomes during the reconstitution procedure: a part may be inserted in an inside out orientation, and a part may be lost due to aggregation / complex dissociation. The latter is very likely, because the complexes are not very stable in detergent solution (chapter 6).

In the efflux assays with EcfAAT-BioY (figure 4.7b), efflux of biotin is observed when the liposomes are diluted in buffer without biotin. Since no MgATP was present, transport via the EcfAAT-BioY is not expected to take place. The decrease is probably due to unspecific leakage of the proteoliposomes. Importantly, biotin was released more rapidly from the proteoliposomes containing EcfAAT-BioY when external MgATP was present, indicating that ATP-dependent transport does occur. Overall, our data indicate that BioY is a high-affinity biotin binding protein but depends on the ECF module for transport, just like the other well-characterized S-components (see chapter 1, 2, 3, 5 and 7).