

Structures of Carbohydrate, Cysteine and Cystine Binding Receptors of ATP-Binding Cassette (ABC) Transporters

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy

of Freie Universität Berlin

by **Haydar Bulut**

from Mazgirt, Turkey

Berlin, 2012

1. Reviewer: Prof. Dr. Wolfram Saenger

Institut für Chemie-Biochemie/Kristallographie, Freien Universität Berlin

2. Reviewer: Prof. Dr. Volker Haucke

Institut für Chemie und Biochemie, AG Membran-Biochemie & Molekulare Zellbiologie, Freie Universität Berlin

Date of the oral examination: 25/01/2012

Table of Contents

1. Abstract	1
2. Introduction	3
2.1 ATP-Binding Cassette (ABC) Transporters	3
2.2. Structural organization of ABC transporters	5
2.2.1. Transmembrane Subunits	5
2.2.2 ABC Subunits	7
2.5. A Translocation Mechanism of ABC Transporters	12
2.6. The Energy-coupling Factor (ECF) Transporters	13
2.7. X-ray Structure Determination of Membrane Proteins	14
2.8. Detergents for Crystallization of Membrane Proteins	16
2.8 Aims of this Work	17
3. Chemicals	18
3.1 Buffer, Solutions and Media	18
4. General protocol for protein purification	21
4.1. Designing and Cloning of Constructs for Protein Expression	21
4.2. Expression Systems	22
4.3. Membrane Preparations and Solubilization	23
4.5. Metal Affinity Chromatography	24
4.6. Size-exclusion Chromatography (SEC)	24
5. Protein characterization	26
5.2. Dynamic Light Scattering (DLS)	26
5.2. Mass Spectrometric Analysis	26
5.3. Fluorescence Spectroscopy	27
5.4. CD Spectroscopy	28

6. X-ray Crystallography methods	30
6.1. Crystallization	30
6.2. Structure Determination	31
6.2.1. Data Collection	31
6.2.2. Cryoprotection	32
6.2.3. Radiation Damage	32
6.2.4. The Phase Problem in Crystallography	33
6.2.5. Molecular Replacement (MR)	33
6.2.6. Isomorphous Replacement (IR)	33
6.2.7. Anomalous Dispersion (AD)	34
6.2.8. Model Building and Refinement	35
7. Purification of ABC importers	36
7.1. Overexpression and Purification of Art M2P2/J	36
7.2. Overexpression and Purification of HisP2QM	37
7.3. Overexpression and Purification of Pro(WV) ₂ V ₂	38
7.4. Overexpression and Purification of the NGO 0373-74 Complex	39
8. Structural Analyses of Binding Receptors	41
9. Chapter I	42
10. Chapter II	70
11. Chapter III	98
12. Discussion	126
13. Literature	130
14. List of Abbreviations	138
15. Acknowledgement	139
16. Curriculum Vitae	140



Abstract

1. Abstract

ATP-Binding Cassette (ABC) transporters are integral membrane proteins carrying a variety of substrates across the cell membrane using energy provided by ATP hydrolysis. The minimal core complexes of ABC transporters consist of two transmembrane subunits (TMSs) that form a specific ligand transport pore and two cytosolic ATP-binding subunits (ABSs) that bind and hydrolyse ATP to provide the energy for the translocation of substrate across the cell membrane. In addition to these core subunits, canonical prokaryotic ABC-importers require a periplasmic binding protein (PBP) or receptor, which traps the substrate in the periplasm and delivers it to the external face of the transport complex.

In this study, general purification methods were established for obtaining functional prokaryotic ABC importers ArtM₂P₂ (arginine, lysine and histidine), HisQMP₂ (arginine, lysine, ornithine and histidine), Ngo0373/0374 (L-cystine) and Pro(WX)₂V₂ (glycine betaine, proline betaine). These importers were biochemically characterized and used for crystallization trials that were, however, not suitable for X-ray analyses.

In parallel, several conjugate periplasmic substrate receptors were co-crystallized with their ligands and further structurally analyzed concerning their substrate specificity. The structural analyses included the receptor proteins of: **(i)** The acarbose receptor GacH from *Streptomyces glaucescens* that was co-crystallized with specific acarbose homologs and acarbose, and the maltose receptor MalE from *Salmonella typhimurium* was cocrystallized with acarbose and the structure elucidated. These structures provide insight into the acarbose carbophor cycle of *Streptomyces glaucescens* and suggest that the uptake of acarbose by competitor microorganisms is coupled with ABC transporter. **(ii)** The structure of the galactose receptor AcbH from *Actinoplanes sp* was obtained with β-D-galactopyranose in the binding pocket, and the receptor was found to exhibit a very specific substrate preference for this sugar. This opposes the previous finding which regarded AcbH as an acarbose binding protein. **(iii)** Lastly two solute receptors for L-cystine and L-cysteine, respectively, from *N. Gonorrhoeae* were solved in complex with their ligands, and their binding pockets were compared with respect to their substrate specificity.

These X-ray structures of receptors in complexes with their ligands provide detailed insight into the arrangements of binding pockets for substrate specificity and information regarding their extracellular role in the translocation of these substrates.

Abstract

Zusammenfassung

ATP-Binding Cassette (ABC)-Transporter sind integrale Membranproteine, die verschiedene Substrate durch die Zellmembran unter ATP-Verbrauch transportieren. Die minimalen Kern-Komplexe von ABC-Transportern bestehen aus zwei Transmembran-untereinheiten (TMU), die eine spezifische Pore für Liganden bilden sowie aus zwei ATP-bindenden Untereinheiten (ABU), die die Energie für den Transport durch Hydrolyse von ATP gewinnen. Zusätzlich zu diesen Kern-Untereinheiten benötigen prokaryotische kanonische ABC-Importer ein periplasmatisches Bindeprotein (PBP) oder einen Rezeptor, der das Substrat im Periplasma bindet und an die Außenseite des Transport-Komplexes transportiert.

In dieser Arbeit wurde die Expression und Aufreinigung der prokaryotischen ABC Importer ArtM₂P₂ (Arginin, Lysin und Histidin), HisQMP₂ (Arginin, Lysin, Ornithin und Histidin), Ngo0373/0374 (L-Cystin) and Pro(WX)₂V₂ (Glycin Betain, Prolin Betain) etabliert. Die Importer wurden biochemisch charakterisiert und für Kristallisationsversuche verwendet, jedoch wurden keine für die Röntgenstrukturanalyse geeigneten Kristallen erhalten.

Parallel dazu wurden mehrere Komplexe von periplasmatischen Substrat-Rezeptoren mit ihren Liganden co-kristallisiert und ihre Strukturen hinsichtlich Substratspezifität analysiert. Es wurden folgende Strukturen erhalten: **(i)** Acarbose-Rezeptor GacH aus *Streptomyces glaucescens* mit verschiedenen Acarbose-Homologen und Acarbose, und der Maltose-Rezeptor von *Salmonella typhimurium* wurde mit Acarbose co-kristallisiert und die Strukturen aufgeklärt. Diese Strukturen geben Einsichten in den Acarbose Carbophor Zyklus von *Streptomyces glaucescens* und lassen vermuten, dass die Aufnahme von Acarbose durch konkurrierende Mikroorganismen an ABC-Transporter gekoppelt ist. **(ii)** Die Struktur des Galactose-Rezeptors AcbH aus *Actinoplanes sp* wurde mit β -D-galactopyranose in der Bindungstasche erhalten. Es konnte festgestellt werden, dass dieser Rezeptor hohe Spezifität für diesen Zucker zeigt. Dies widerspricht der vorherigen Ansicht, dass AcbH ein Acarbose bindendes Protein ist. **(iii)** Des weiteren wurden die Strukturen zweier Substratrezeptoren für L-Cystin bzw. L-Cystein aus *N. gonorrhoeae* im Komplex mit ihren Liganden aufgelöst und ihre Bindungstaschen wurden in Bezug auf ihre Substratspezifität verglichen.

Die gewonnenen Daten dieser Arbeit geben einen neuen und detaillierten Einblick in den Aufbau verschiedene Bindungstaschen und liefern Informationen über deren Substratspezifität sowie die extrazellulären Funktionen in der Translokation dieser Substrate.

2. Introduction

2.1 ATP-Binding Cassette (ABC) Transporters

All living cells regularly import a wide variety of nutrients from their environment to feed their metabolic activity, and they eliminate waste products and toxic elements. The translocation of solutes across the cell membrane is generally mediated by specialized membrane proteins. The group of integral membrane proteins responsible for transport were initially divided into channels and carriers (Saier, 2000). ATP binding cassette (ABC) transporters belong to carriers that are ubiquitously expressed in all kingdoms of life, from prokaryotes to humans (Higgins *et al.*, 1992). Active transporters use the energy from binding and hydrolysis of ATP to drive the translocation of diverse substrates across the cell membrane. ABC transporters work as bidirectional machineries and maintain either export out of or import into the cells. ABC exporters pump many toxic molecules and metabolite wastes out of the cell and ABC importers take up useful molecules such as amino acids, trace elements and vitamins. The diversity of molecules that are transported demands different transport systems that are specific for their substrates and therefore a large part of genomes is required for encoding transport proteins. For instance, in *E. coli* 80 different ABC transporters have been characterized that participate in transport processes (Blattner *et al.*, 1997), which corresponds to ~5% of the total genome (Linton *et al.*, 1998).

The human genome contains 49 genes that are responsible for encoding ABC transporters (Vasiliou *et al.*, 2009). Based on sequence homology and domain organization, they have been categorized into seven distinct subfamilies (A to G). Although several members have dedicated roles in host detoxification and protection of the body against dietary toxins, they also decrease the bioavailability of orally administered drugs and reduce drug disposition to physiological compartments. Increasing drug efflux due to overexpression of ABC transporters from cancer cells leads to failure of chemotherapy. P-glycoprotein (P-gp; ABC1) is the best characterized ABC transporter that causes multi-drug resistance (MDR) in cancer cells. P-gp actively extrudes many types of drugs from cancer cells, keeping their intracellular levels below a cell-killing threshold. Beside anticancer drugs, P-gp also extrudes HIV protease specific inhibitors, antibiotics, antidepressants, antiepileptics and analgesics. P-gb is highly expressed at the apical membranes of the intestinal epithelial cells, where it can prevent the absorption orally ingested amphipathic drugs from the gastrointestinal lumen to the blood (Szakacs *et al.*, 2006).

Introduction

Other members of ABC drug transporters that play a major role in multi-drug resistance are multi-drug resistant protein 1 (MRP1 or ABCC1), breast cancer resistance protein (ABCG2), and mitoxantrone resistance protein (ABC-P). MRP1 transports less toxic but more water soluble drugs compared to xenotoxins which are extruded by P-glycoprotein. ABCG2 protects specific body compartments such as the placenta and the blood–brain barrier (BBB) by extruding xenobiotics (Ozvegy *et al.*, 2001).

A lipid flippase (MsbA), which shows a high sequence homolog to human P-gp (Tomii *et al.*, 1998) is responsible for the translocation of essential lipid-A component from inner to the outer layer of *E. coli* membrane. Dysfunction of the transport mechanism results in a lethal accumulation of lipid A in the cytoplasmic leaflet. The transport complex consists of two homodimers of membrane spanning region fused with a NBD. The structure of the MsbA from *E. coli* solved by Xray crystallography that reveals an open conformation of a lipid binding chamber located at the inner surface of the transporter. The size and shape of the chamber could accommodate and transport a wide variety of amphipathic molecules, which explains the affinity of Eco-MsbA and its homolog MDR1(human) transporters towards a wide range of substrates (Ward., 2007).

Since the living environment of microorganisms contains numerous toxins and antibiotics, microorganisms have developed active extrusion mechanisms that are mediated by ABC transporters to remove drugs from the cytoplasm. In the last decade the number of drug-resistant pathogens increased dramatically, including *Neisseria gonorrhoeae*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Pseudomonas aeruginosa*, and *Vibrio cholera* (Putman *et al.*, 2000). Ongoing research in this field focuses on the development of more effective inhibitors targeting the action of multi-drug transporters with higher affinity. These inhibitors are applied in three different cell levels. First, inhibition targets directly the pumping mechanism of ABC transporter at the cell membrane. The second approach is based on blocking the signaling pathway that controls the expression of ABC transporters. The third approach is focused on the inhibition of specific transcription factors to block gene expression (Shukla *et al.*, 2008, Szakacs *et al.*, 2006). However, cytotoxic effects and adverse pharmacokinetics have limited their use.

Dysfunction of ABC transporters may cause inherited human diseases such as cystic fibrosis, hypercholesterolemia or diabetes. Cystinuria is an autosomal recessive disease caused by dysfunction of human ABC transporters that are responsible for uptaking cystine, arginine, lysine and ornithine. High concentration of L-cystine in urine may lead to formation of cystine stones in the kidneys, ureter and bladder (Trinchieri *et al.*, 2004). Besides medical treatment, drug targeting

Introduction

aims to lower the amount of free cystine in the urine by forming chemical bonds with the sulphhydryl groups formed by reduction of the cystine molecule (Fjellstedt et al., 2003).

2.2. Structural organization of ABC transporters

General structural features of all ABC transporters comprise a core complex that is minimally required for transport. The core complex consists of two transmembrane subunits (TMSs) and two ATP binding subunits (NBSs). Two domains are encoded either as a single polypeptide chain (e.g., the cystic fibrosis transmembrane regulator, CFTR) or NBDs and TMDs are encoded as separate subunits (e.g., the vitamin B12 permease of *E. coli*). In canonical prokaryotic ABC importers, substrate translocation is dependent on an additional protein component. This protein functions as receptor that binds specifically to its substrate on the cell surface or in the periplasm and delivers it to the appropriate ABC transporter (Parcej et al., 2007).

2.2.1. Transmembrane Subunits

The transmembrane subunits or domains consist exclusively of α -helices that are embedded in a phospholipid bilayer. Substrate specificity channels that open to both sides of the plasma membrane are formed in the core of α -helices. Depending on membrane spanning segments per subunit, ABC importers are divided into two subgroups (Locher 2009). The majority of transporters are classified in type I (Parcej et al., 2007) that contain six helices per subunit (Fig. 1). Moldydate, maltose and methionine transporters are among the type I transporters that are structurally characterized. The methionine transporter (MetI) with five transmembrane helices per subunit has the simplest TMS architecture (Kadaba *et al.*, 2008). The transmembrane subunits of moldydate transporter (ModB) have six helices per TMS, with high structural symmetry (Gerber *et al.*, 2008). The transmembrane subunits (MalF and MalG) from the maltose transporter contain six and eight transmembrane helices, respectively, in a heterodimer association (Oldham et al., 2007).

Type II transporters consist of larger TMS with 10 to 12 α -helices per subunits. Examples of Type II importer are the vitamin B₁₂ and metal-chelate transporter, each containing 10 TM helices per subunit with a total of 20 TM segments in the assembled transporters (Locher *et al.*, 2002, Pinkett *et al.*, 2007).

The crystal structure of Sav1866 was solved in the ATP-bound configuration and structure of MsbA solved both in nucleotide bound and apo configurations. Recently structure of P-gp obtained in Apo and drug-bound P-gp state. Common features observed in all exporters are formation of each TM segment from 6 helices that extend beyond the membrane boundary and

Introduction

protrude far into the cytoplasm (Dawson and Locher, 2006). Because of these differences that presence in TM segments ABC exporters classified as Type III transporters (Parcej and Tampé, 2010).

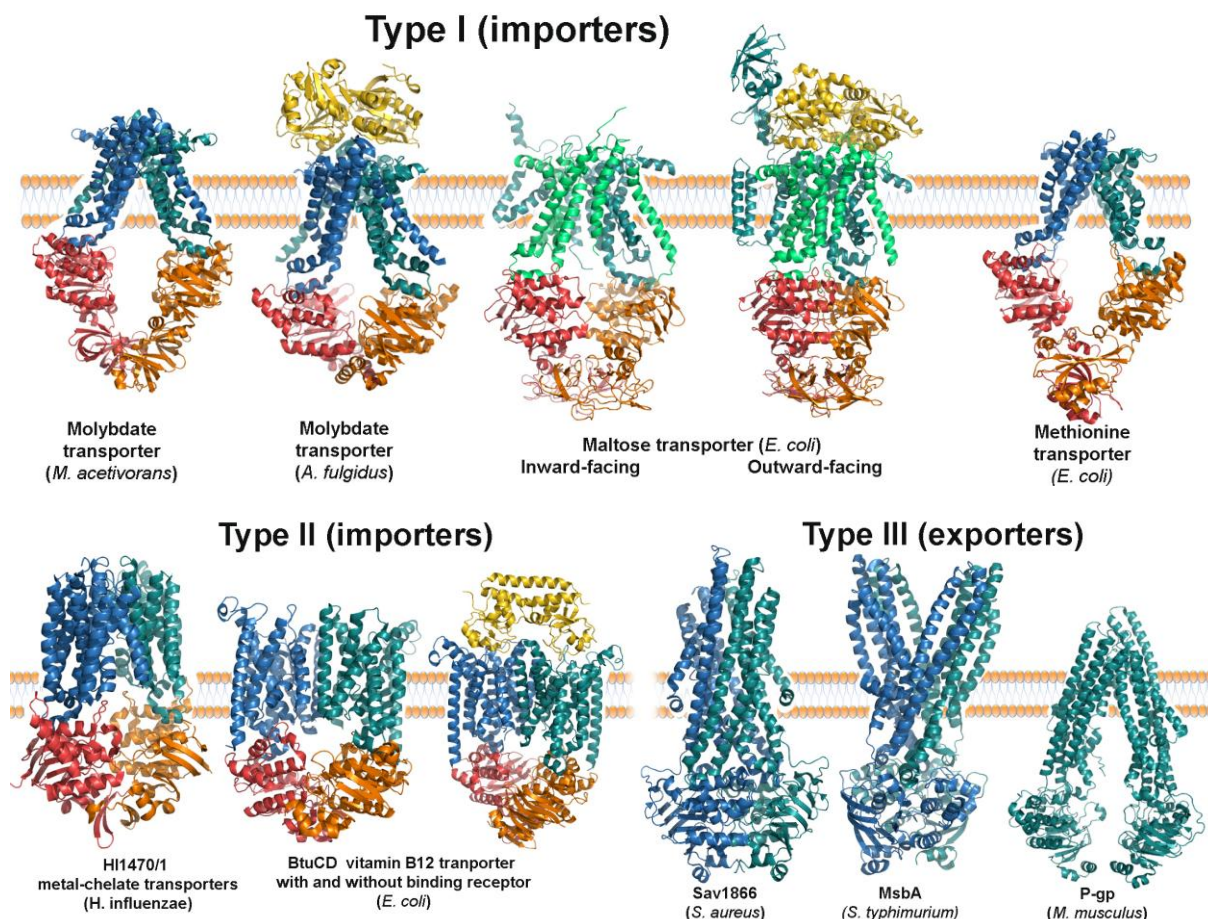


Fig 1: Ribbon representation of the currently available crystal structures of ABC transporters, with the source organism indicated below each transporter. The subunits of ABC importers are shown with different colors, red and orange are NBSs, blue and light blue are TMDs and yellow are SBPs. In two exporters (Sav1866 and MsbA) TMD is fused to NBD in homodimer association and represented as one color. P-gp however is encoded as a single polypeptide chain. The type I importers contain 10-14 α -helices. Type II transporters contains up to 20 helices, and the large cytoplasmic extension of transmembrane helices observed in these exporters are classified as Type III transporters. The Pdb codes of transport complex used for the illustration of Type I were ModBC (3D31), ModABC (2ONK), MalFG (3FH6), MalFGK (2R6G), MetNI (3DHW); Type II were HI1470/1 (2NQ2), BtuCD (1L7V), BtuCDF (1Q19); Type III Sav1866 (2HYD), MsbA (3B60) and P-gp (3G60). (Picture was reproduced from Parcej and Tampé, 2010)

Introduction

2.2.2 ABC Subunits

The ATP-binding subunits are located on the cytoplasmic side of the transporter and consist of a highly conserved RecA-like subdomain and an α -helical subdomain (Fig. 2). This arrangement creates a binding groove for the hydrolysis of two ATP molecules between the Walker A motif of the first subunits and the LSGGQ motif, the so called “ABC signature motif” of the second subunit (Hollenstein, *et al.*, 2007). The Q-loop’ is located at the interface of two subunits and and play role in interactions with TMS (Hollenstein, *et al.*, 2007). Hydrolysis of ATP is mediated by two residues are located on the second subunits around the active site (Fig. 2cd). While histidine residue forms a hydrogen bond with the γ -phosphate of ATP (Linton and Higgins, 2007), the acidic residue (a glutamate in most ABC transporters) at the end of the Walker B motif, activates a proton of a hydrolytic water that attacks and activates the γ -phosphate of ATP (Davidson *et al.*, 2008, Schneider and Hunke, 1998).

In the absence of ATP, water molecules are able to access the nucleotide-binding sites throughout the dimer interface (Chen *et al.* 2003). Each transport cycle requires the positive cooperativity of two ATP molecules as driving force (Senior & Bhagat 1998). Most of the isolated NBDs exist in monomeric form, and addition of ATP analogues drive the dimerization. The situation is different in full transporter, where in the absence of ATP, NBDs remain in contact through coupling helices of the transmembrane segments (Hollenstein, *et al.*, 2007). In some transporters, such as molybdate, maltose and methionine, the NBD is fused to the regulatory domain (Kadaba *et al.*, 2008, Gerber *et al.*, 2008, Davidson *et al.*, 2008). In the structures of full ABC transporters (ModBC or MetNI), molybdate or methionine, respectively, are found at the solute-binding sites located at the C-terminal parts of the NBDs. The solute binding in this region prevents the dimerization of the NBDs and causes an inhibition of the ATPase activity at high concentrations of the substrate (Parcej *et al.*, 2009).

Introduction

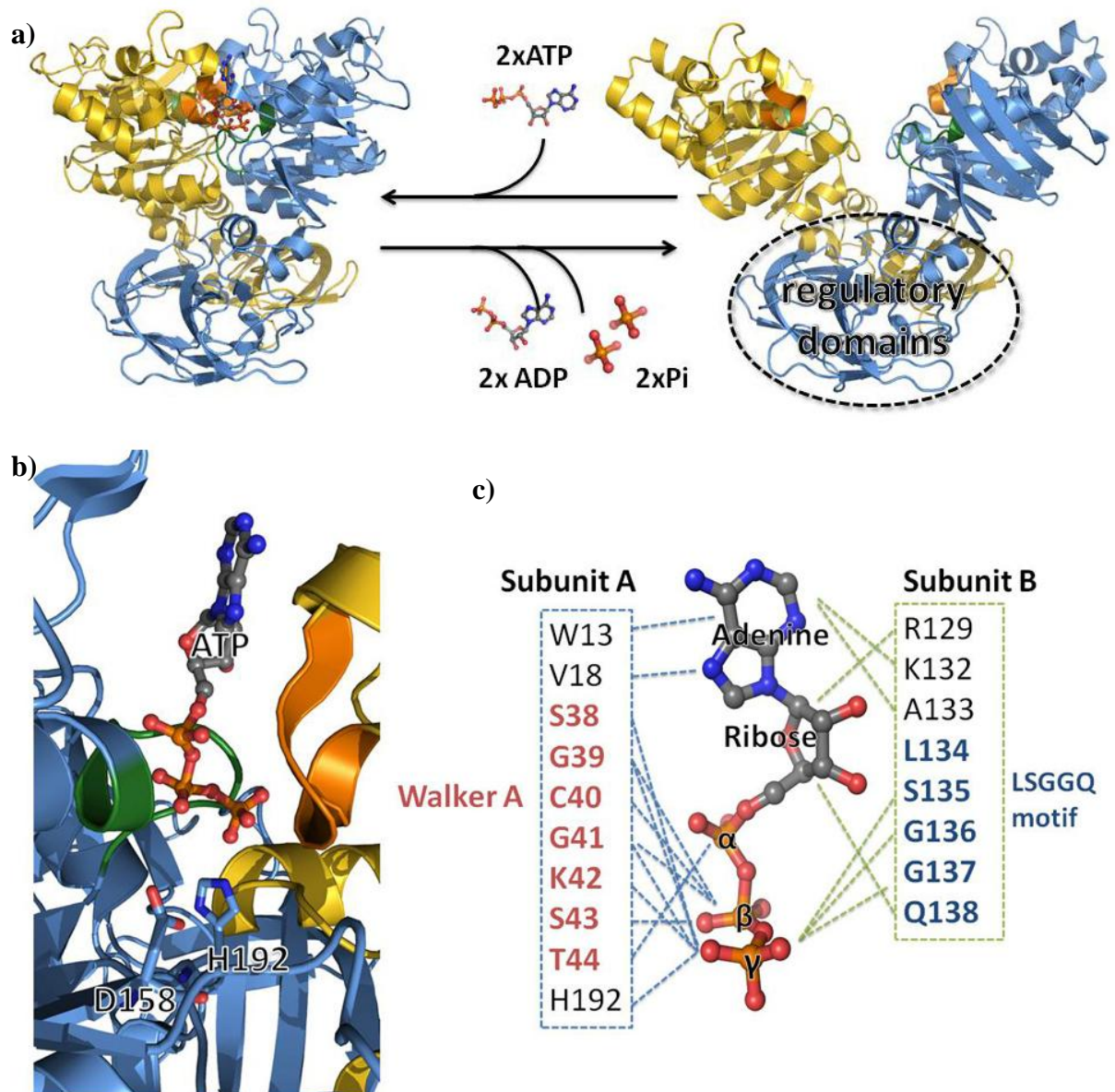


Fig. 2: **a)** Side view of the NBS of maltose transporter in two different states. On the left two ATP molecules are bound to dimer interface of MalK, and the Walker A and LSGGQ motifs are shown in green and orange, respectively (Pdb:1Q1E). When ATP is bound, the interface closes and the ATPs are sandwiched between the NBDs. Hydrolysis of the ATPs and disassociation of the hydrolysis products result in opening up of the dimer interface (Pdb:1Q1E) (Chen *et al.*, 2003). **b)** Top view of an ATP molecule inside the binding cleft that is formed by two subunits, His192 and Asp165 residues are indicated. These residues are responsible for the ATP hydrolysis. **c)** An interactive diagram of ATP with neighboring residues is shown. ATP is sandwiched between the Walker A and the LSGGQ motifs (Chen *et al.*, 2003 and Linton *et al.*, 2007)

Introduction

2.2.3. Periplasmic solute binding proteins (SBPs)

Solute binding proteins (SBPs) are extracellular components of ABC transporters and responsible for capturing and delivering ligand to the transport complex (Eitinger *et al.*, 2011, Berntsson *et al.*, 2010). Their ligands range from small inorganic compounds such as Zn^{2+} or phosphate to much larger substrates such as carbohydrates or polypeptide chains (Wilkinson *et al.*, 2003, Felder *et al.*, 1999).

Although most of procaryotic ABC importers use one solute binding protein (receptor) per complex, members of two transporter families of the ATP-binding cassette (ABC) superfamily use two or even four extracytoplasmic substrate-binding domains for uptake (Biemans-Oldehinkel *et al.*, 2003).

In gram-negative bacteria, SBPs diffuse freely in the periplasm, while in gram-positive bacteria, lacking an outer membrane, solute binding proteins are anchored to lipids of outer surface of the cell membrane via acyl chains at the N-terminal cysteine residue or, in the case of archaea, they are fused to the transmembrane subunits (van der Heide T., and Poolman B., 2002, Sutcliffe *et al.*, 1995) (Fig. 3) Uptake of nutrients across the outer membrane of gram-negative bacteria is mediated by porins, which share a β -barrel motif, that appears to be a hallmark structure for outer membrane transporters (Delcour *et al.*, 2002). Although small ligands such as amino acids, sugars, and short peptides can diffuse through porin channels, some essential nutrients and vitamins larger than ~600 Da require high-affinity transporter complexes that are driven by an electrochemical gradient across the cytoplasmic membrane (Postle, K. and Kadner, R. J. 2003, Shultis *et al.*, 2006).

A considerable number of three-dimensional structures of SBPs isolated from bacteria in the ligand-free and ligand-bound states have been obtained using X-ray crystallography. The secondary structure elements of SBPs exhibit common features (Berntsson *et al.*, 2010) that consist of two lobes containing a central β -sheet flanked by α -helices. The two lobes are interconnected by a hinge region with the substrate binding pocket located in a cleft at the interface between the lobes. Upon ligand binding, the SBPs undergo a hinge-bending motion, consisting of a rotation and twist of the lobes, resulting in the transition from an open to a closed conformation (Quiocho *et al.*, 1996, Shilton *et al.*, 1996). The hinge-bending motions of SBPs make them ideal target for constructing biosensors, and even with reagentless electrochemical sensors it has been possible to obtain a significant signal to detect their specific ligand (Dwyer and Hellings, 2004, Tolosa *et al.*, 2003).

Introduction

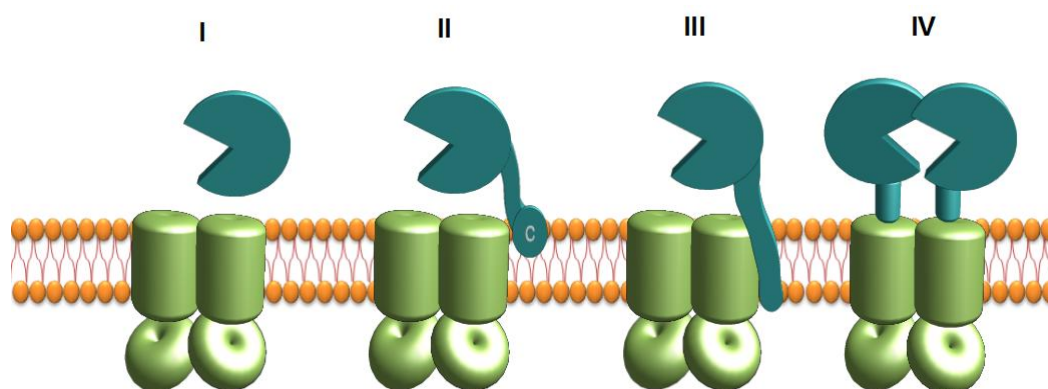


Fig. 3. Schematic representation of transport complexes with differently organized SBPs. Core transport complexes are shown in green color, TMS are shown as cylindrical and NBDs are shown as spheres. SBPs are drawn in blue. In gram-negative bacteria, SBPs reside in the periplasm (I) while in gram-positive bacteria and archaea, they are anchored to the extracellular side of the cytoplasmic membrane via an N-terminal lipid moiety (II). In some archaea, the N-terminal peptide of SBP is associated with the membrane (III). In some bacteria and archaea, including *Lactococcus lactis*, *Helicobacter pylori* and *Sulfolobus solfataricus*, SBP are fused to the TMS of transporters

Although the secondary structure elements and type of movement upon ligand binding are conserved in the SBPs, the versatile sequences in the binding pocket enable them to bind different ligands. In the binding pockets hydrogen bond networks are found determining substrate selectivity e. g. with respect to carbohydrates and their epimers like D-glucose and D-galactose (Licht et al., 2010), or with respect to small ligands such as phosphate and sulfate (Luecke et al., 1990). While hydrogen bond interactions play the main role when trapping small ligands in the binding pocket, hydrophobic interactions also contribute to binding the non-polar portions of larger ligands such as carbohydrates (Vahedi-Faridi et al., 2010).

2.3. Classification of SBPs

An initial structure based classification of SBP divided them into two groups according to the formation of the sheet topology within the lobes. The classification proposed an evolutionary origin of the SBPs from dimerization of a CheY-like α/β featured protein. While Type I SBP contains seven structures, type II SBPs contained 8 structures of SBP (Fukami-Kobayashi *et al.*, 1999). In class I and II the lobes differ in the number and order of β -strands. In addition, Type III SBP appeared first in the structure of a Zinc-binding protein, in which two domains are interconnected by a single long α -helix (Lee et al., 1999).

Introduction

In the following decade many more of SBP structures appeared in the protein database, which were collected and analyzed by Berntsson and his colleagues (Fig. 4). Here Type III SBPs were classified in Cluster A, this group of SBPs binds different metal ions, but also Vitamin B₁₂. X-ray Structure of this receptor exhibits a very small rotation (4Å) up on ligand binding (Wilkinson *et al.*, 2003). Cluster B binds mainly carbohydrates and and hydrophobic amino acids such as Leu, Ile and Val. SBPs classified in Cluster E are observed essentially in Tripartite ATP-independent periplasmic transporters (TRAP transporters)TRAP-transporters, this type of transporters require an electrochemical ion gradient for the uptake of substrate. Most of the amino acid binding proteins exhibit Cluster F features and depending on their substrate specificity they are further subclassified in four different groups (Berntsson *et al.*, 2010).

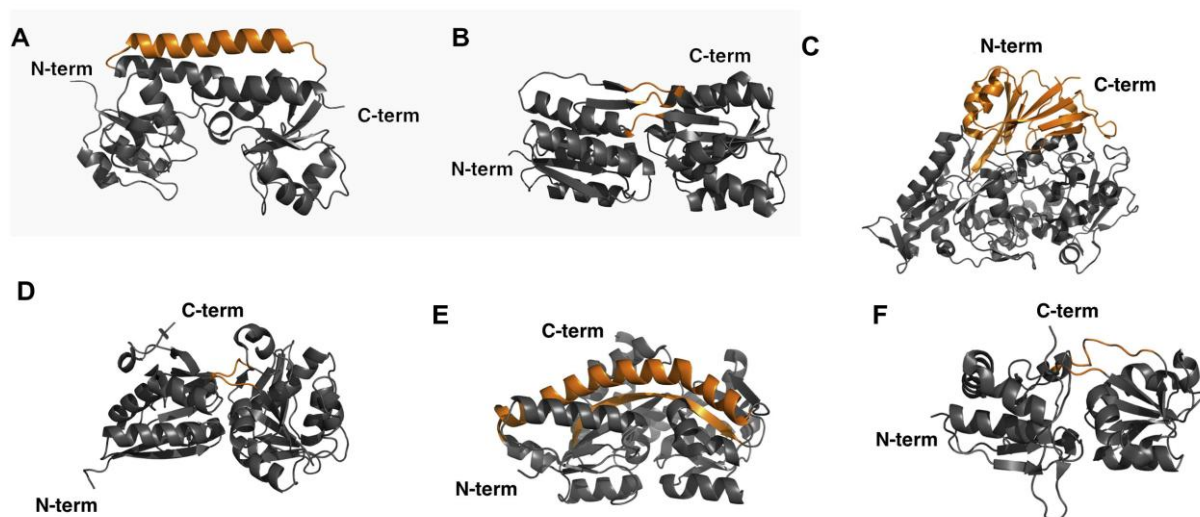


Fig. 4. The SBPs were classified in 6 different Clusters (A-F) based on structural features and formation of hinge regions that are shown in orange color. In Cluster A two lobes are interconnected by a rigid helix, which allows only small movement during substrate trapping. Cluster B however contains three short loops interconnecting two domains. The characteristic feature of Cluster C is the presence of an extra domain that is fused to the N-terminal lobe. Hinges of Cluster D are formed by two short loops. SBPs of Cluster E contain a large helix functioning as hinge region. Cluster E also contains two hinges, but compared to hinges of Cluster D, the loops are larger and more flexible. The PDB coordinates for the clusters A–F are as follows: BtuF (PDB code: 1N2Z), RBP (PDB code: 1DRJ), OppA (PDB code: 3DRF), ModA (PDB code: 1ONR), UehA (PDB code: 3FXB) and HisJ (PDB code: 1HSL). (Picture modified from Berntsson *et al.*,

Introduction

2.4. Receptor-Transporter interaction

The lobes of an SBP are in an “open” configuration without bound ligand and change configuration to a “closed” form when they trap the specific ligand. These receptors are known as “Venus flytrap” due to similarities in the trapping mechanism. The high affinity ligand binding by the two lobes of SBPs elicits a dramatic conformational change. The resulting complex possesses a protein binding surface not present in the open conformation, which can be recognized by the extracellular face of TMS (Hollenstein *et al.*, 2007). Upon interaction of the SBP with the outward facing TMSs of the transporter, the substrate enters the transmembrane pore for the subsequent translocation process.

So far, three of seven ABC transporters were co-crystallized with their receptors (Schneider *et al.*, 2011). The Vitamin B₁₂-importer from *E. coli* crystallized in complex with and without its substrate binding receptor (Locher *et al.*, 2002; Hvorup *et al.*, 2007). At the interface main interactions take place between glutamate residues of the substrate binding protein (BtuF) that are potential hydrogen bond acceptors of arginine residues that are located on the surface of the TM segments (BtuC). Both glutamates and all three arginines are found to be conserved in iron siderophore/cobalamin transport systems from various organisms (Borths *et al.*, 2002). Similar charged interaction also observed molybdate transporter, here an aspartate residue on the outer surface of transmembrane subunits (ModB) interact with the arginine of one lobe and lysine residues of the second lobe of binding receptor (ModA) (Hollenstein *et al.*, 2007). In case of maltose transporter however one arginine residue is located on binding receptor MalE forms a multiple hydrogen bonds with transmembrane subunits (MalFG) (Oldham *et al.*, 2007).

2.5. A Translocation Mechanism of ABC Transporters

The presently known structures of substrate importers exhibit two different configurations (Fig. 5). In the ATP binding state they adopt an outward-facing configuration to take up substrate and in the resting state they adopt an inward-facing configuration to deliver substrates into the cytoplasm (Locher *et al.*, 2002, Dawson *et al.*, 2007) (Fig. 5). An opposite mechanism appears in exporters, which receive their substrates in the inward-facing state and release them in an outward facing state (Oldham *et al.*, 2008).

Alignment of the TMSs sequences indicates a few matches in the cytoplasmic helices of almost all the TMSs of ABC import systems. The so called “coupling helices” features an EAA motif, which is regarded as a contact region between TMS and NBS, in which the signal

Introduction

transmission between the transport complex subunits takes place (Hollenstein *et al.*, 2007, Locher *et al.*, 2002, Roger *et al.*, 2007).

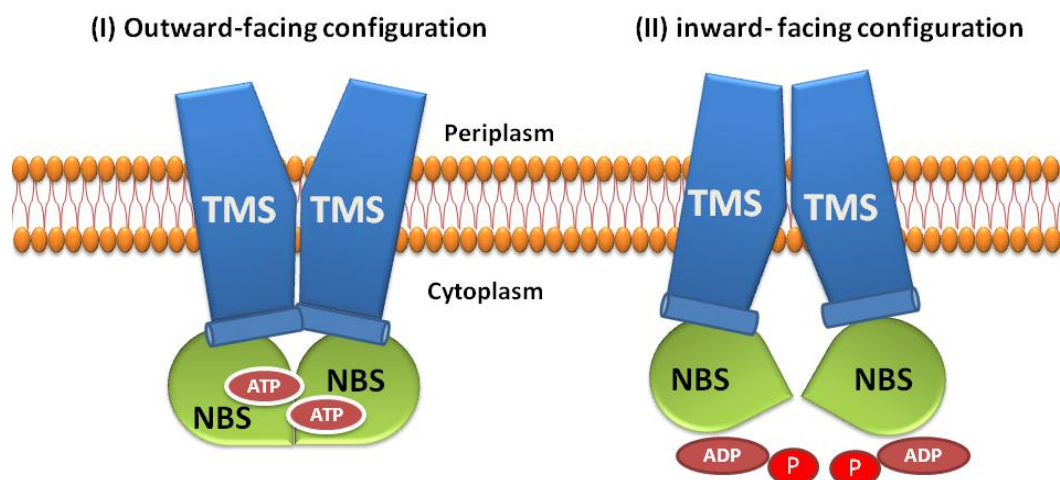


Fig. 5. Schematic coupling mechanism of ABC transporters is shown on the top. The coupling helices are shown as cylindrical shapes that attach the TMSs to NBSs (I) In this model, binding of two ATP molecules in the interface between the two NBSs triggers a tight dimer formation of NBSs, at the same time coupling helices at the cytoplasmic part of TMSs approach each other, which drives a configurational change in TMSs in a so called “outward-facing configuration” (II) Hydrolysis of ATP and release of the hydrolysis products reverts the TMSs to the opposite direction, the so called “inward-facing configuration” (Dawson *et al.*, 2007).

2.6. The Energy-coupling Factor (ECF) Transporters

Although ECF type of transporters have a similar organization as the ABC transporters, they lack an extra binding protein in the periplasm and instead one of membrane subunits functions as substrate binding protein (known as S component). Recently, the two crystal structures of S-component were published. The X-ray structure of the riboflavin-binding protein RibU from *Staphylococcus aureus* and thiamine-specific S-component ThiT from *Lactococcus lactis* were found to be unrelated to the membrane subunits of other canonical ABC transporters. The S-component tightly binds to its ligand, and the translocation of the substrate is subsequently driven by forming complex with an S-specific A-T module that consists of one transmembrane and two nucleotide binding subunits (Zhang *et al.*, 2010, Erkens *et al.*, 2011).

The ECF transporters are divided into two groups, in group I ECF transporters S and an A-T module are encoded by adjacent gene clusters while in group II transporters, the A-T modules are shared by many S components (Rodionov *et al.*, 2009) (Fig. 6). The substrates for ECF transporters contain in particular dozens of water soluble vitamins, cofactors, transition-metal ions such as Ni^{+2}

Introduction

and Co^{+2} , also hydrophobic amino acid tryptophan, queuosine and its precursors queuosine, aminomethyl-7-deazaguanine (preQ₁) (Eitinger *et al.*, 2011).

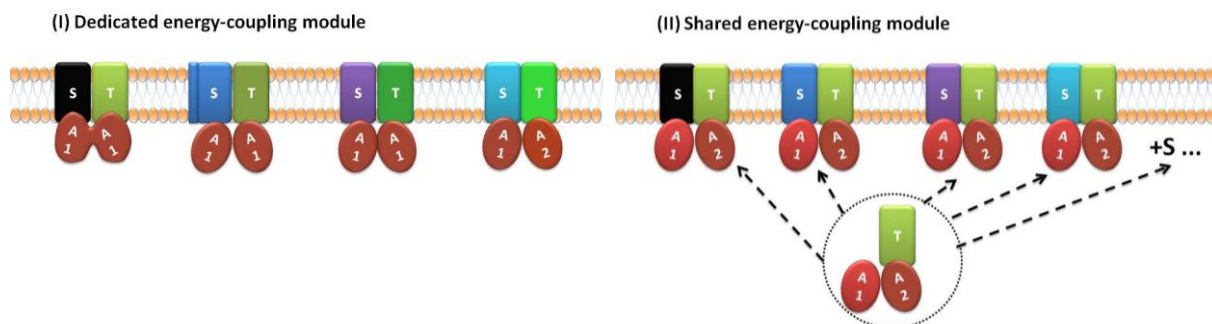


Fig. 6. Structural organization of the ECF type importers. Substrate is directly recognized by the S component of the ECF transporters, AT modules contain a transmembrane subunit (T) associated with either two different (A1, A2) or copies of the same ATPase protein (A1, A1). Group I transporters have a substrate-specific S component and a dedicated AT module encoded by linked genes. Group II transporters have individual S components and shared AT modules that are not linked to S components (Picture was modified from Eitinger *et al.*, 2011).

2.7. X-ray Structure Determination of Membrane Proteins

Around thirty percent of *Homo sapiens* and *Escherichia coli* genomes are dedicated for encoding integral membrane proteins (Wallin and von Heijne, 1998). Membrane proteins comprise channels, transporters and receptors, that are playing fundamental roles in cell physiology such as transport, signaling, adhesion, recognition and trafficking. Those crucial functions make integral membrane proteins ideal targets for drug design and biotechnological applications. Despite their abundance in the proteome, only a limited number of three-dimensional structures of membrane proteins were solved so far (260 unique entries by 2010 can be found in database under <http://blanco.biomol.uci.edu/mpstruc/listAll/list>).

The transmembrane segments of membrane protein structures are featured by either α -helical (75%) or β -barrel fold (25%). Porins are typical examples of β barrel proteins, which function as diffusion pores at outer membranes of Gram-negative bacteria. This fold reflects remarkable stability, and protein expression easily works out in bacterial hosts (Delcour *et al.*, 2002).

Several difficulties are encountered during purification and structural characterization of membrane proteins. First of all, most of membrane proteins are naturally not abundant except for

Introduction

photosynthetic membrane proteins, which are present at sufficient level and therefore structurally well-characterized. Despite rare achievement in membrane protein overexpression, recombinant techniques are still an alternative method for protein production (Griesshammer and Tate, 1995). In addition the expressed protein could occur in an inactive conformation in detergent solution after isolation from its natural lipid environment. The crystallization process is also hardened by detergent molecules, which form a disordered “belt” like surface around membrane proteins (Ostermeier and Michel, 1997).

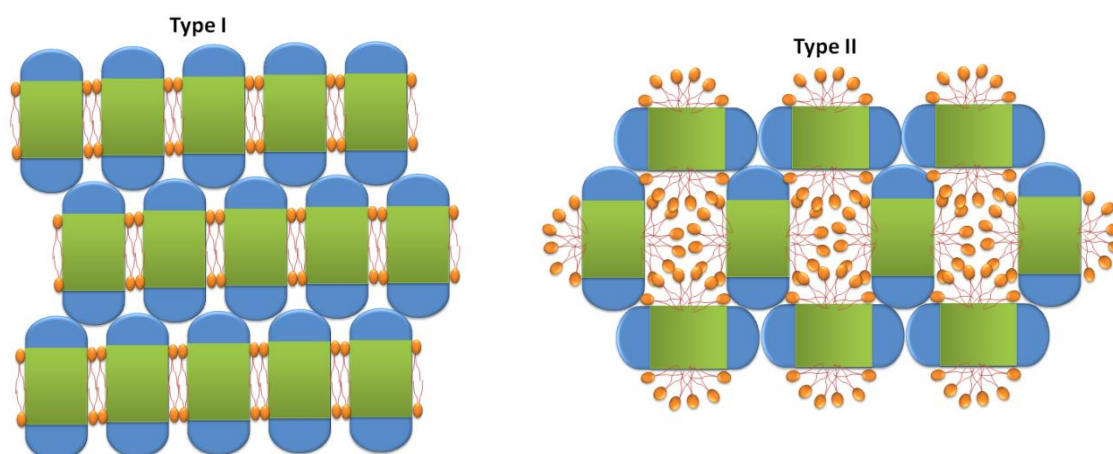


Fig. 7. The two basic types of crystal packing observed in membrane proteins. Hydrophobic transmembrane segments of membrane protein are shown in green color and detergent molecules are attached to them. In Type I crystals packing the hydrophilic portions (blue) are in contact with each other forming 2D crystalline array, which then builds up the third dimension. In Type II crystal packing, detergents bound to hydrophobic portion form a detergent pool, and crystal packing is driven essentially by hydrophilic surface interactions (Reproduced from Michel, 1983).

Difficulties still exist during the data collection step, most of the membrane protein crystals are not well-ordered and do not permit collection of high quality data sets for structural analyses (Ubarretxena-Belandia and Stokes, 2010). Time consuming optimization attempts include generally detergent exchange or engineering of protein for well-ordered crystals. Since the hydrophobic part of membrane proteins form main crystal contacts (Fig. 7), increasing the hydrophilic portion could be eventually helpful for crystal packing (Iwata, 2003). For this purpose, generating antibody fragments to create extra hydrophilic surface attached to membrane proteins was successfully applied on several membrane proteins (Hunte and Michel, 2002).

Recently a major breakthrough was obtained in membrane protein crystallization using the lipidic cubic phase (Caffrey et al., 2008). Compared to detergent based crystallization method, such

Introduction

as vapor diffusion or microbatch approaches, the lipidic cubic phase (LCP) based crystallization methodology was found to be more robust to impurities in various membrane protein crystallization (Kors et al., 2009). Using the LCP-based method, it was possible to produce crystals of photosynthetic reaction center (RC) of *Rhodobacter sphaeroides* with protein solutions containing up to 50% of lipid material and membrane fragments (Kors et al., 2009)

2.8. Detergents for Crystallization of Membrane Proteins

In order to purify membrane proteins from their surrounding native membrane they need to be solubilised by so-called “mild detergents” without causing any denaturation. The detergent mimics the lipid bilayer covering the hydrophobic surface of the protein by forming protein-detergent micelles. It is essential to use a suitable detergent that disrupts the membrane but not the tertiary structure of the protein under study. The protein solubility is a key parameter in crystallization, which proceeds by the removal of proteins from solution into a crystalline array. The effect of detergents on the protein solubility was examined for ABC transport complexes by measuring the amount of protein remaining in solution after high-speed centrifugation. The test consistently indicated that highest overall protein solubility was achieved by longer chain detergents (10-12 carbon chain). Many membrane proteins including maltose transporter were successfully solubilized and crystallized using β -dodecylmaltoside (DDM) (Davidson et al., 1991). Indeed, MalFGK2 was inactivated by OG (octylglucoside) when used in purification; activity was observed when the protein complex was isolated and purified in β -DDM, and reconstitution was mediated by OG (Davidson et al., 1991). For many membrane proteins, comparison of the maltoside containing detergents was found to be milder and more effective than members of glucoside containing detergents (Fig. 8) (Ames et al., 2001, Knol et al., 1996). Therefore β -DDM was chosen mainly for solubility and crystallization purpose.

Introduction

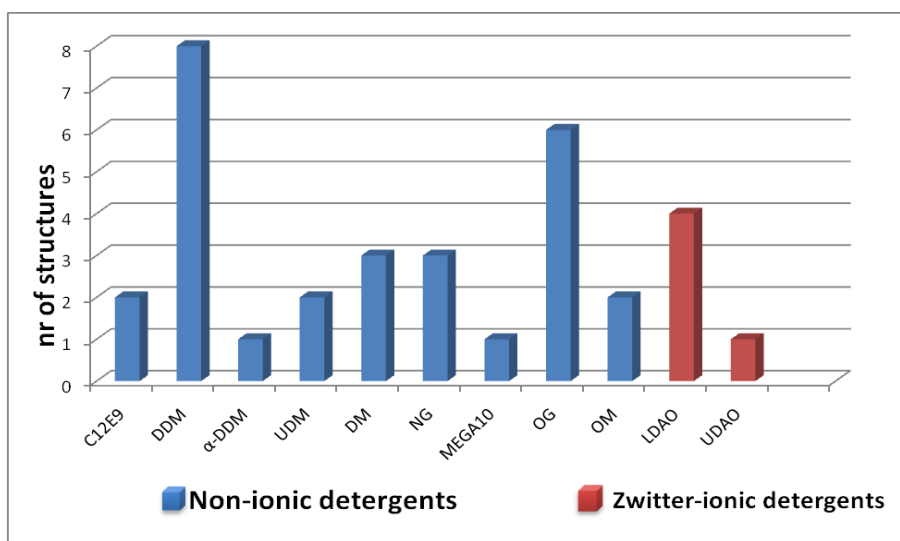


Fig. 8. Most common detergents used for membrane protein crystallization (Iwata S, 2003). DDM: dodecylmaltoside, OG: octylglucoside, DM: decylmaltoside, UDM: undecylmaltoside, LDAO: *N,N*-dimethyldodecylamine-*N*-oxide, UDO: *N,N*-dimethylundecylamine-*N*-oxide, NG:nonylglucoside, OM: octylmaltoside.

2.8 Aims of this Work

The here described prokaryotic ABC transporters consist of three components, an extracellular receptor, membrane-intrinsic proteins forming a pore, and two cytosolic ATPases. The arginine transporter ArtM₂P₂ of *Geobacillus stearothermophilus*, histidine transporter HisQMP₂ of *Salmonella Typhimurium*, the Proline-betaine transporter Pro(WX)₂V₂ of *Helicobacter pylori*, and putative cystine transporter *ngo0372-74* of *Neisseria gonorrhoeae* were chosen as model systems. Attempts to crystallize these transporters with various substrates or ATP analogs to capture the transport complex either in- or outward facing conformation to archive conformational homogeneity.

Crystallographic studies should be performed including purification and characterization of complex transporters with the aim to obtain I) intact structure of ABC importers to shed light on their translocation mechanism, and II) structural analyses of several conjugate water soluble receptors with focus on their substrate specificity.

Chemicals

3. Chemicals

Chemicals were purchased from Merck, Sigma-Aldrich, Invitrogen or Carl Roth if not stated otherwise.

3.1 Buffer, Solutions and Media

Buffers were prepared using MilliQ-water unless noted otherwise. The pH was adjusted using either NaOH or HCl.

LB medium

1 % (w/v) tryptone

0.5 % (w/v) yeast extract

0.5 % (w/v) NaCl

Terrific Broth

1.2% Tryptone,

2.4% yeast extract,

1.0% glycerol,

72mM K₂HPO₄, 17mM KH₂PO₄

LB agar

1 % (w/v) tryptone

0.5 % (w/v) yeast extract

0.5 % (w/v) NaCl

1.5 % (w/v) agar

2xYT medium

1.6 % (w/v) tryptone

1 % (w/v) yeast extract

0.5 % (w/v) NaCl

Coomassie stain

1 g/l Coomassie

10 % Acetic acid

Chemicals

25 % Methanol

4x SDS separating gel buffer 1.5 M Tris

0.4 % (w/v) SDS

pH 8.8

4x SDS stacking gel buffer

0.5 M Tris

0.4 % (w/v) SDS

pH 6.8

10x SDS running buffer

246 mM Tris

1.92 M Glycine

10 % (w/v) SDS

Cell lysis buffer

50 mM Tris pH 7.5

150 mM NaCl

5% (v/v) glycerol

1 mM PMSF

0.3 % protease inhibitor cocktail

and DNase (0.05 mg/L)

Antibiotic Stocks

Ampicillin stock 50 mg/ml sterile filtered

Chloramphenicol stock 34 mg/ml in EtOH sterile filtered

Kanamycin stock 50 mg/ml sterile filtered

Streptomycin stock 50 mg/ml sterile filtered

Tetracyclin stock 5 mg/ml in EtOH sterile filtered

Chemicals

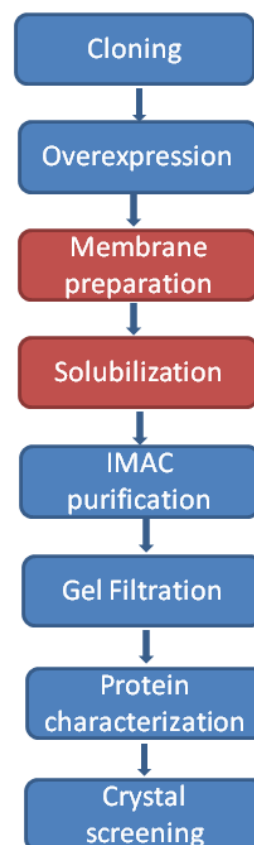
Transporter Complex	Membrane Solubilization buffer	Elution buffer	Crystallization buffer
Art M₂P₂	50 mM Tris/HCl, pH 7,5, 5% (v/v) glycerol, 0.1 M NaCl 0.1 mM PMSF 1.0 % DDM	20 mM Tris/HCl, pH 7,5 5% (v/v) glycerol, 0.2 M NaCl, 0.1 mM DDM, 0,2 mM L-Arg, 250 M imidazol	10 mM Tris/HCl, pH 7,5 5% (v/v) glycerol, 0.2 M NaCl 0.1 mM DDM 0,2 mM L-Arg 5 mM AMP-PCP 0,5 mM EDTA
HisQMP₂	50 mM Tris, pH 7.6, 10 % glycerol, 0.1 mM PMSF, 1,2%, decanoilsucrose	50 mM Tris, pH 7,6, 5 % glycerol, 10 mM ATP, 0,1 mM PMSF, (4 mM mercaptoethanol), 0.2 % decanoilsucrose	50 mM Tris, pH 7.6, 5 % Glycerol, 0.2 % Decanoilsucrose
NGO0373/4	50 mM Tris, pH 7.6, 20 % glycerol, 0.1 mM PMSF, 1.0 % DDM	50 mM Tris, pH 7.6, 10 % glycerol, 300 mM NaCl, 0.1 mM PMSF, (4 mM mercaptoethanol), 200 mm imidazol, 0.02 % DDM	20 mM Tris, pH 7,5, 10 % glycerol, 300 mM NaCl, 0,1 mM PMSF, 0,02 % DDM
ProWXV₂	50 mM Tris, pH 7.6, 20 % glycerol, 0.1 mM PMSF, 1.0 % DDM	50 mM Tris, pH 7.6, 10 % glycerol, 500 mM NaCl, 0.1 mM PMSF, (4 mM mercaptoethanol), 200 mm imidazol 0.02 % DDM	20 mM Tris, pH 7,5, 10 % glycerol, 300 mM NaCl, 0,1 mM PMSF, 0,02 % DDM

General protocol for protein purification

4. General protocol for protein purification

Purity and homogeneity of protein samples are among the most important parameters to obtain protein crystals.

General workflow of purification for the crystallization of transport complexes and their substrate binding protein is shown on the right side. Blue highlighted purification steps were taken as common purification both for core complex and substrate binding receptors. Since the transport complex contains two transmembrane subunits embedded in lipid bilayer, purification protocol required additional two extra steps of purification that are shown in red color. These steps were needed to extract protein from the lipid membrane using a mild detergent.



4.1. Designing and Cloning of Constructs for Protein Expression

ABC transporter core domains were cloned with a cleavable poly-histidine tag fused to their ATP-binding subunits. For expression of conjugate binding receptor, plasmids were designed by removing predicted signal sequence and cloned with N-terminal cleavable poly-histidine tag in pET vectors. The removal of the polyhistidine tag was required for protein crystal growth.

For expression, genes of transport core complexes or associated receptors were inserted in pET plasmids transformed into bacterial strains. Those strains typically contain a single copy of T7 polymerase on the chromosome in a lambda lysogen (DE3), which is under the control of the Lac-UV5 promoter. In the absence of lactose in growth media, the lac repressor (lacI) binds to the lac operator and blocks the transcription of lac promoter. When lactose is the sole carbon source, or

General protocol for protein purification

when the lactose analog IPTG is added to the medium, lactose (or IPTG) binds to the repressor and induces its dissociation from the operator, thereby permitting transcription from the promoter.

Although protein expression was mainly induced by IPTG, in case of core complexes higher IPTG concentration caused cell death or unproductive cell growth, therefore the concentration of IPTG, reduced to 0.1mM. Transcription of all other receptors were induced with 0.5mM concentration of IPTG, only exception was an AcbH, there IPTG induction cause an insoluble inclusion bodies. Therefore this protein was alternatively produced in auto-induction media. This protein expression does not need inducers such as IPTG during mid-log phase of the cultures. The method is based upon a buffered medium enriched with trace elements that contain a mixture of carbon sources, including lactose. The bacteria initially use glucose; when the glucose is exhausted, they start to import lactose that induces expression of the T7 polymerase from the DE3 lambda lysogen (Studier et al., 1990).

4.2. Expression Systems

While solute binding were proteins successfully expressed in almost all *E. coli* strains, the quantity of the expressed transport complexes heavily depended on expression strains. Therefore after cloning of transporter complexes, available *E. coli* strains were tested to find out suitable expression strain.

The following list of *E. coli* strains with strain supplier characteristics were used for cloning and expression of proteins in this study.

- TOP10 Invitrogen *E. coli* strain is ideal for high-efficiency cloning and plasmid propagation. It allows stable replication of high-copy number plasmids.
- BL21 Star™(DE3)pLysS strain contains a mutation in the gene encoding RNaseE (*rne131*), which is one of the major sources of mRNA degradation. Therefore, BL21 Star™ cells significantly improved the stability of mRNA transcripts and increased protein expression yield from T7 promoter-based vectors.
- Rosetta 2 Novagen protein expression strains were engineered to contain extra copies of genes that encode tRNAs that most frequently limit translation of heterologous proteins in *E. coli*.

General protocol for protein purification

- JM109 is a K strain bacterium that carries *recA1* and *endA1* mutations. The *recA1* genotype provides minimized recombination and aids in plasmid stability while *endA1* provides for high quality plasmid DNA preparation.

Many laboratories generally use BL21(DE3) for high-level protein production purposes. It has the advantage of being deficient in both *lon* and *ompT* proteases and is compatible with the T7 *lacO* promoter system (Studier et al., 1990). However, for expression of transport complex, it was often important to use BL21 (DE3) derivatives carrying additional tRNAs to overcome the effects of codon bias. In some cases, the yield of recombinant protein obtained from the RNase E-defective BL21 Star strains was 2- to 10-fold greater than that from the conventional BL21 strain. These strains are encoding an extra resistance gene for chloramphenicol that is used for selection in expression media. In addition, reduction of temperature and enriching the growth media increased the quantity of protein being expressed. Some successful expression conditions for transport complexes are listed in below.

Recombinant Protein	Expression Vector	Strain (constructed plasmid) (reference)	media	temperature	induction
HisQMP ₂	pET22b	Rosetta pLysS(pVE26) (V. Eckey)	Terrific broth	22°C	0,1 mM IPTG
ArtM ₂ P ₂	pQE60	JM109(pRF2) (R. Fleischer)	Terrific Broth	37°C	0,5 mM IPTG
Ngo0373/0374	pET22b	BL21starpLysS(pCJ58) (C. Jacobi)	Terrific Broth	22°C	0,1 mM IPTG
ProWXV ₂	pET22b	BL21starpLysS (pFSA111) (F. Scheffel)	Terrific Broth	22°C	0,1 mM IPTG

4.3. Membrane Preparations and Solubilization

After expression of transport complexes, the cells were harvested and resuspended in homogenization buffer. The resuspended cells were passed thorough a French Press at maximum 1250 bar. Cell debris were pelleted by centrifugation at 10000 g for 10 min. The upper fraction containing *E. coli* cell membrane was isolated from soluble fraction by ultracentrifugation at 200 000 g for 1 h. The soluble fraction was discarded and isolated membrane was collected and gently solubilized in membrane buffer containing mild detergent (1% β -dedocylmaltoside, DDM). Membrane fraction was stirred in membrane buffer for 1h until entirely solubilised.

General protocol for protein purification

Since the isolation of membranes enriched in membrane proteins, less detergent is required for solubilisation. In addition this process eliminates proteolytic degradation often caused by proteases in soluble fractions. At the end of solubilization a second ultracentrifugation was used for 30 min at 200 000 g to pellet unsolubilized membrane parts, and the supernatant was subjected to Co^{2+} -affinity chromatography (QIAGEN).

4.5. Metal Affinity Chromatography

Metals ions immobilized on a matrix form a complex with poly histidine tag fused to the protein of interest. Although the mechanisms are the same for all metal ions, the precise geometry of coordination sites and the binding strength differ for different metals.

Ni^{2+} is often used for proteins with a 4-8 histidines long tag. However, unspecific proteins containing histidines are able to bind the resin unspecifically as well. Experimental observation indicates that TALON resin binds with higher selectivity over nickel-based resins (e.g., Ni-NTA), in addition cobalt ions bind to the resin at four sites, resulting in lower metal ion leakage. Therefore immobilized cobalt loaded TALON matrix was used as first step purification of transport-complex and periplasmic substrate binding proteins.

A column loaded with TALON resin was equilibrated with a membrane buffer. The concentration of solubilized membrane was adjusted to ~10 mg/ml in membrane buffer and applied to Co^{2+} affinity chromatography. Non-specific proteins were removed from the column by washing with several column volumes of membrane buffer. Bound transport complex was then eluted with elution buffer.

4.6. Size-exclusion Chromatography (SEC)

SEC is an effective method to separate proteins by their molecular mass. The principle of size exclusion chromatography (SEC) is based on a fine pored resin material with well defined pore size. While small molecules and solutes pass through the pores and utilize basically the entire bed volume, large molecules do not enter the small pores and are thereby effectively excluded from parts of the column bed volume. As a result, large molecules elute first, whereas ions and other small solutes come off last. SEC may therefore be used to combine buffer or detergent exchange platform besides improvement in purification. Since membrane proteins are in complex with detergent molecules they usually run at molecular weights higher than predicted.

Size-exclusion chromatography was applied to all transport complexes or their solute binding receptors for further purification. Elution buffer was exchanged to desired composition for

General protocol for protein purification

crystallization trials by removing small compounds such as ATP or imidazole from protein. The concentration of DDM was reduced to the desired critical micelle concentration (CMC) for crystallization. Although the purity of solute binding receptors was satisfying after affinity chromatography, SEC was applied essentially to remove the affinity tag after cleavage. Removing the affinity tag from the proteins was an essential step to obtain crystals. Indeed crystallization trials carried out using protein fused with histidine tag, apparently flexible poly histidine tag was obstacle that prevented protein crystallization by blocking possible crystal contacts.

The purified transport complexes essentially were applied on gel-filtration columns containing Sephacryl S200 column and characterised by using SDS-PAGE. This was followed by concentrating the protein solution to approximately 10 mg/ml for crystallization using centricon devices (Millipore, 100,000 MW cutoff).

5. Protein characterization

5.2. Dynamic Light Scattering (DLS)

Dynamic Light Scattering (DLS) is a widely used technique in the field of protein crystallography. Using DLS it is possible to measure the size of protein, polydispersities and consequently aggregation of protein samples. Polydispersity of the sample and the presence of protein aggregates are obstacles for the crystallisation of proteins. The DLS measurement is based on laser light that is scattered from dissolved macromolecules or suspended particles. Due to the Brownian motion of the molecules and particles in solution, fluctuations of the scattering intensity can be observed. To determine monodispersity of membrane proteins that are further used for crystallization trials, required ~0,5-2 mg/ml protein solution in microcuvette (cell volume 20-30 μ l).

Conformational homogeneity of core complex NGO 0373-74 of Cystine transporter was analysed with DLS and is shown in Fig. 9. The results indicated no aggregation and monodisperse protein.

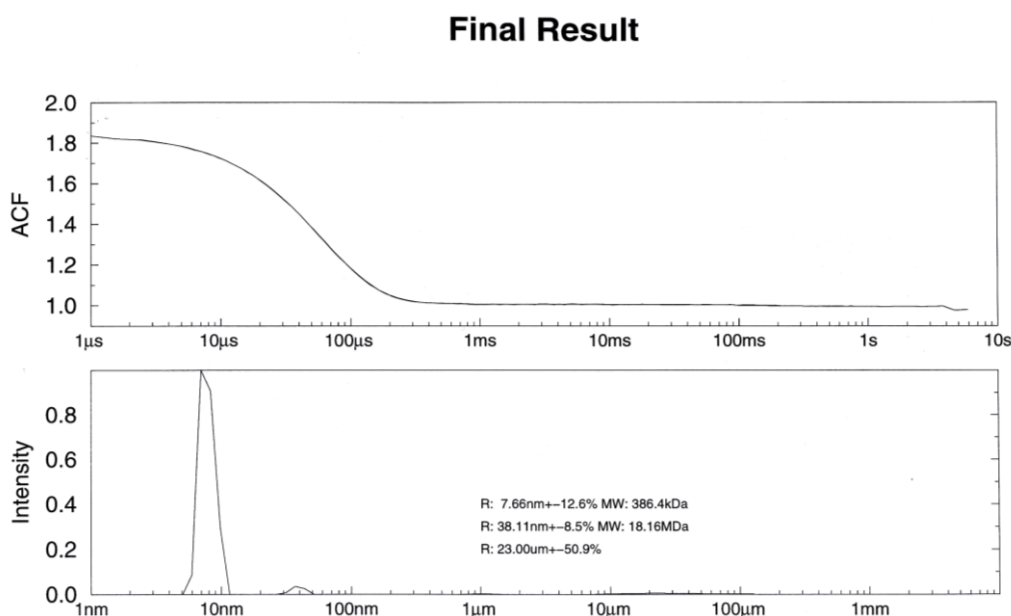


Fig. 9. Auto correlation function (average from 10 measurements) (top) and calculated hydrodynamic radius for the NGO 0373-74 transport complex (below).

5.2. Mass Spectrometric Analysis

Matrix-assisted laser desorption/ionization (time-of-flight) MALDI(TOF) used to identify purified protein. In prensiple, peptide co-crystalized with matrix molecules are inserted into the

Protein characterization

vacuum chamber of the mass spectrometer and pulsed by laser beam. Transfer of a high energy from laser beam into the matrix molecules promotes the transition of matrix molecules and peptides from the solid state into the gas state. The power of the laser beam is usually adjusted in a way that it has enough energy to ionize the biomolecules and matrix molecules but does not split the large analyte molecule. Ionized molecules become accelerated in the electric field of the mass spectrometer and fly towards an ion detector where their arrival is detected as an electric signal. Their mass is proportional to their time of flight (TOF) in the drift tube and can be calculated accordingly.

The protein of interest was extracted from SDS-PAGE and digested into several fragments using proteolytic enzyme trypsin. A typical sample: protease ratio is 50:1. The proteolysis is carried out overnight and the resulting peptides are extracted with acetonitrile and dried under vacuum. A small amount of the peptide (usually few microliters) was delivered to MALDI-TOF facility at Free University Berlin. Further experiments were conducted by Dr. Christoph Weise.

5.3. Fluorescence Spectroscopy

In principle, the sample is excited with UV light (295 nm), and the longer wavelength emitted signal, called fluorescence, is measured (310-450 nm). The magnitude of the fluorescence signal is strongly dependent on tryptophan residues and on their environment. Thus, a change in their environment is translated into a decrease of the fluorescence signal emitted by the protein (quenching).

Since the structure Cys-CysRec was obtained in the “open” conformation in the absence of L-cystine, fluorescence spectroscopy experiments were conducted to identify the ligand of this receptor. The addition of L-cystine (but not any other amino acids) to the protein solution induced a significant quenching of the fluorescence signal emitted by the protein (Fig. 10).

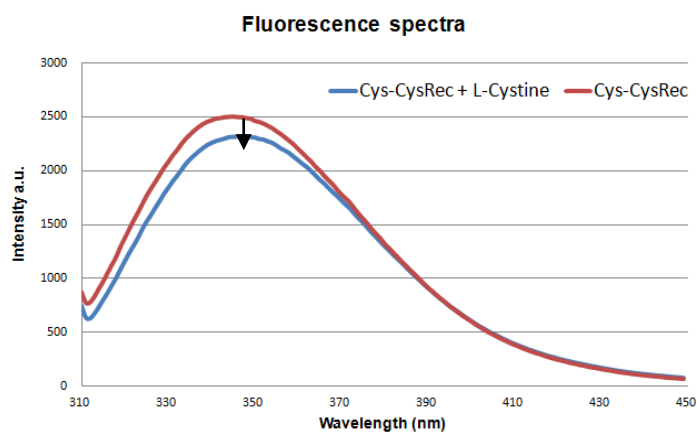


Fig. 10. Fluorescence emission spectra of Cys-CysRec, in absence and presence of 1 μ M L-cystine at 25⁰C. A fluorescence signal was emitted from Trp-97, Tyr-59, and Tyr-176 that are within binding pocket of Cys-CysRec.

Protein characterization

5.4. CD Spectroscopy

Asymmetric substances have different absorption coefficients for of right- and left-circularly polarized light (Sharon *et al*, 2005). As a result of this physical phenomenon, circular dichroism (CD) spectroscopy is a widely used technique to determine secondary structure composition of proteins (and DNA) in solution. The peptide backbone (with minor influences from side chains) produces difference in absorption spectra for left- and right handed circularly polarized light. Therefore different secondary structural types of proteins, notably α -helix, parallel and anti-parallel β -sheet, turn, and random coil can be analyzed by CD spectra between 250 and 180 nm wavelength (Fig. 11). Thus, CD spectroscopy is a very powerful and rapid method for the analysis of structural compositions in proteins as well as structural changes upon different types of perturbation (Greenfield NJ. 2006).

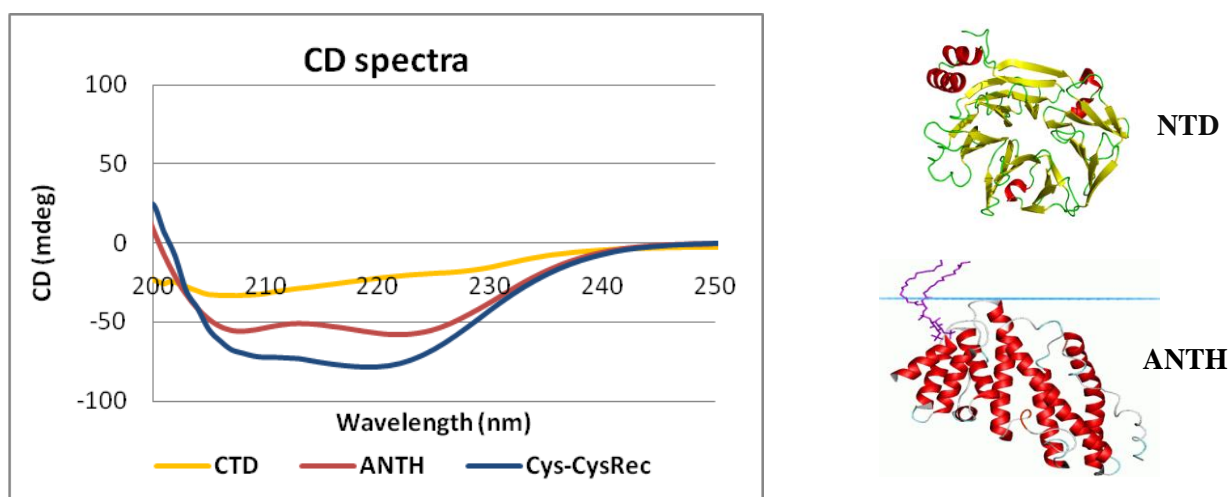


Fig. 11. CD spectra of two proteins with representative secondary structure elements. In this example, the α -helical ANTH domain of protein AP180 (at pH 7.5) have negative bands at 222 nm and 208 nm and a positive band at 193 nm . A clathrin terminal domain(CTD) with well-defined antiparallel β -sheet conformation (at pH 7.5) shown a negative band at 218 nm and positive bands at 195 nm. Cys-CysRec contains typical α/β fold is shown with blue line.

5.4.1. Melting curves

When a folded protein in solution is heated, the equilibrium between the native (N) and unfolded structure (U) is shifted towards the unfolded form. This transition can be followed by monitoring the change in CD of an e.g. α -helical protein (at 222nm) that becomes a random coil polypeptide chain during denaturation.

Protein characterization

In this study spectropolarimeter J-710 (Jasco, Tokyo, Japan) was used. The instrument was first calibrated with a standard solution of (+)-10-camphorsulfonic acid. Homogeneous samples of Cys-CysRec at concentrations of 0.5 mg/mL in 20 mM Tris pH 7.5, 100 mM NaCl were prepared in the absence of L-cystine, and, alternatively, with a 5-fold molar surplus of L-cystine. Thermal stabilities of the two samples were monitored at 222 nm (Fig. 12).

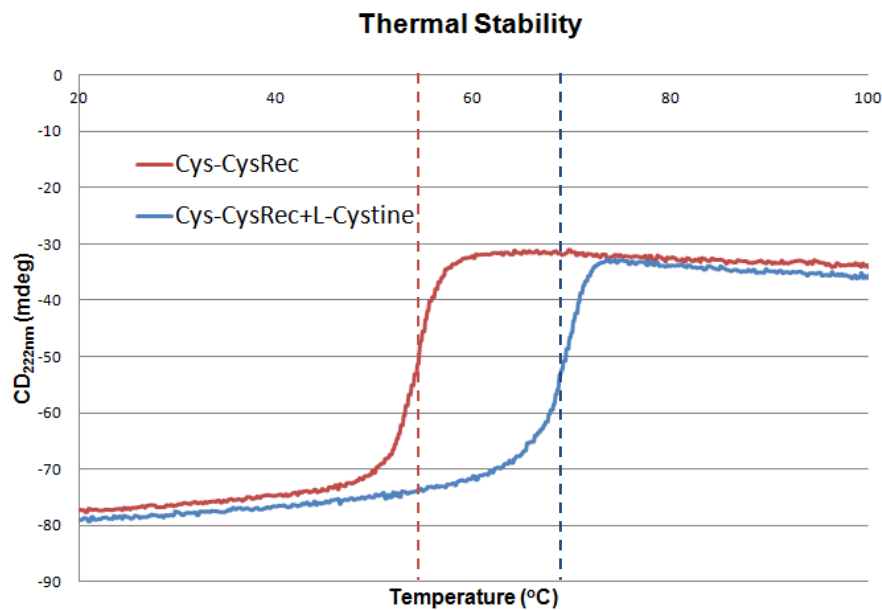


Fig. 12. Circular dichroism spectrum of Cys-CysRec (red) and with its ligand cystine (green) monitored in range of temperatures of 20-100 °C.

6. X-ray Crystallography methods

6.1. Crystallization

An X-ray crystallographic study requires diffraction quality crystals. In order to find ‘hits’ or ‘leads’ that point to conditions where crystallization occurs conditions that start to optimize, the target protein solution is exposed to a variety of precipitating agents. Finding crystallization conditions for a new protein, particularly membrane proteins, requires many trials which are usually conducted in trial-and-error experiments to find right conditions for crystallization. Small crystals, crystalline precipitate and phase separation are considered leads that are worth pursuing (Chayen *et al.*, 2008).

There are several common techniques to obtain suitable crystals for X-ray diffraction measurements. The most widely used crystallization technique is vapor diffusion. These techniques are based on equilibrating a protein solution against a reservoir containing crystallizing agents at higher concentration than in the drop. As the equilibration proceeds conditions vary continuously (Fig. 13). Trials are usually conducted in 96-well plates. Although classic vapor diffusion screens are most commonly used for membrane proteins, several special kits based on crystallization conditions mined from the Protein Data Bank are available, such as MemGold from molecular dimensions. These kits contain 96 conditions covering a range of pH, PEGs and salt additives. Once a lead is identified, optimization is performed. Fine-screening of the crystallization conditions was performed by varying the concentration of protein, type and concentration of precipitant, pH, temperature or additive concentration.

Alternative methods to vapor diffusion are microbatch method, microfluidics and lipidic cubic phase or lipidic mesophase. Starting point of free interface diffusion (FID) and dialysis contain either undersaturated protein solution or protein mixed with a low concentration of precipitating agent. The solubility is defined as the concentration of protein in the solute that is in equilibrium with crystals. The supersolubility curve is defined as the line separating conditions where spontaneous nucleation (or phase separation, precipitation) occurs from conditions where the crystallisation solution is undersaturated and remains clear (Chayen *et al.*, 2004).

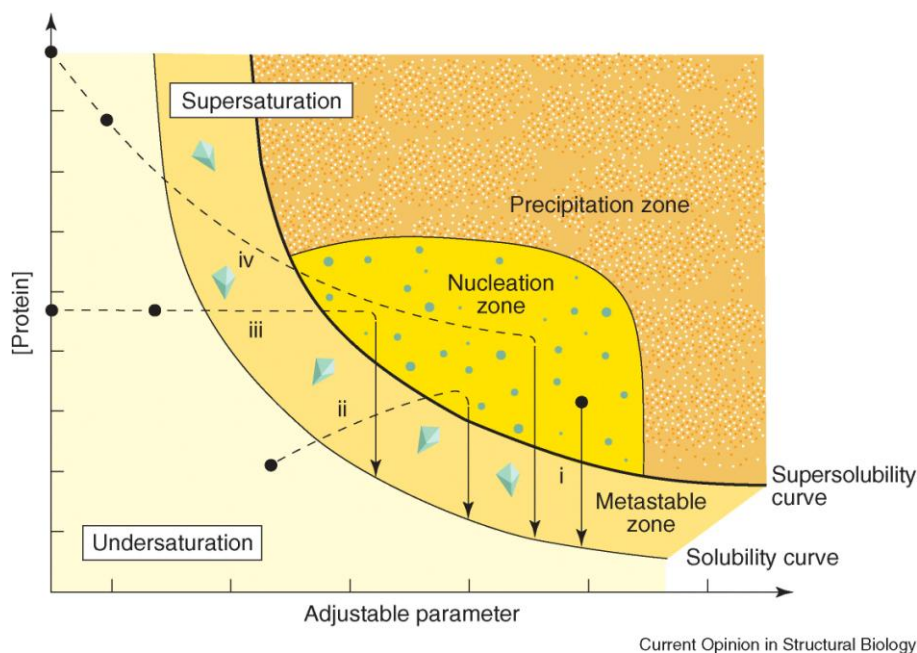


Fig. 13. Schematic illustration of the phase diagram of protein crystallisation. The four major crystallisation methods are represented (i) Batch, (ii) Vapour diffusion, (iii) Dialysis (iv) interface diffusion. Their different routes to nucleation and metastable zones are, assuming that the adjustable parameter is precipitant concentration. (Picture was taken from Chayen *et al.*, 2004).

6.2. Structure Determination

6.2.1. Data Collection

During an X-ray diffraction experiment the electromagnetic X-radiation excites the electrons of the atoms in a crystal structure. The excited electrons scatter radiation with the same wavelength as the used X-radiation in all directions and return to their ground state. Since in crystals the atoms are arranged in regular three-dimensional arrays, the scattered X-rays interfere with each other either constructively or destructively. The relationship between the diffraction pattern and the geometrical crystal parameters is described by Bragg's law. This law explains the conditions to obtain constructive interference between scattered electrons depending on the wavelength (λ) of the incident X-ray beam, the glancing angle (2Θ) formed between the incident and scattered beam in dependence of a set of planes (separated by the interplanar distance d_{hkl}), and n is an integer

$$n\lambda = 2d_{hkl}\sin\Theta$$

X-ray Crystallography methods

The scattered X-rays are usually recorded on an electronic area detector (CCD, charge coupled device). To collect all possible diffraction peaks, or reflexions, the crystal has to be rotated (0.1-2.0 degree per image) in the X-ray beam during data collection to vary the parameter (Θ), thereby fulfilling Bragg's law for all sets of planes d_{hkl} . The rotation range required for complete data collection depends on the internal symmetry of the respective crystal that is defined by the space group of the crystal.

6.2.2. Cryoprotection

High energy X-ray beamlines cause radiation damage on protein crystals, therefore data collection is conducted at cryogenic temperatures. X-ray data collection at low temperature of crystals reduces thermal vibrations in crystals and improves data quality and completeness. However another problem that is usually faced during data collection is losing crystal integrity or data quality due to ice crystal formation that is prevented by cryoprotectants in the mother liquor during cooling to around 100K. Most commonly used cryoprotectants include glycerol, ethylene glycol, MPD, PEGs, sucrose, erythritol and xylitol (Garman and Mitchell, 1996). It is usually necessary to screen several cryoprotectants, or to increase cryoprotectants concentration gradually in several steps to avoid crystal damage. In some cases significant amount of cryoprotectants already exist in the mother liquor. In the presence of an efficient cryoprotectant, protein crystals are transferred to liquid nitrogen for flash cooling.

6.2.3. Radiation Damage

Radiation damage is one of the major limitations in protein crystallography, even for cryocooled crystals. Radiation damage causes increasing mosaicity, R_{sym} , Wilson B factors that reduce resolution of diffraction pattern during data collection. Structural damage can be observed in protein crystals, such as breaking disulfide bonds or losing carboxylate groups from acidic residues (Ravelli et al., 2006).

Photoelectrons are the predominant source of radiation damage at synchrotron sources. Primary radiation damage occurs when a photon directly interacts with solvent or protein electrons. This creates free radical species which then create secondary free radicals known also as secondary radiation damage (Murray and Garman, 2002).

X-ray Crystallography methods

6.2.4. The Phase Problem in Crystallography

In an X-ray diffraction experiment we obtain intensities of waves scattered from planes (hkl) in the crystal, which are proportional to the number of electrons in the crystal. The electron density at a certain point $\rho(xyz)$ in the crystal unit cell can be calculated from the summation over all the (hkl) reflexions once all the phase angles α_{hkl} associated with each reflexion (hkl) are known. The intensities provide structure amplitudes (that are proportional to the square roots of the intensities measured) but not the phase angles which are essential for the calculation of the electron density within a crystal (Taylor GL., 2010).

There are several phasing techniques to solve this problem, molecular replacement, isomorphous replacement and anomalous dispersion.

6.2.5. Molecular Replacement (MR)

Molecular replacement (MR) is the most commonly used method to obtain phase information for structure determination (Evans *et al.*, 2008). MR uses a previously known molecular structure to solve the crystallographic phase problem of a related new structure. With the rapid growth in the number of protein structures available as search models, the use of this method will keep increasing in the future. Molecular replacement requires a high level of structural homology (typically >25% sequence identity). The homologous search model can be processed prior to its use in molecular replacement to enhance the chance of success (Lebedev *et al.*, 2008). Most common options include (i) a poly-alanine model, (ii) a model with modified atomic B-factors that increase proportionally with the atom accessibility by solvent, and (iii) a model with the corrected amino acid sequence according to the alignment of the search and target protein sequences. The rigid part of the model should first be tried in case conformational changes have been predicted. The MR algorithm tries to fit the Patterson distribution derived from the model structure with that derived from the diffraction data using first a rotation and then a translation function (Vagin *et al.*, 2010)

6.2.6. Isomorphous Replacement (IR)

The isomorphous replacement (IR) method is based on introducing heavy-atoms that function as markers to determine the phase angles into a protein crystal by soaking the crystals in mother liquor supplemented with a compound containing a heavy-atom. Unit cell parameters and packing of the molecules in the crystal unit cell must not be disrupted by the introduction of these

X-ray Crystallography methods

heavy-atoms, i.e. isomorphous conditions have to be fulfilled. The differences between the reflection's amplitudes are measured using the native crystal (F_P) and the derivative crystal (F_{PH}) that provide the contribution of the heavy-atom (F_H) by $|F_H| = |F_{PH}| - |F_P|$. Once the heavy-atoms are located in the crystal using Patterson methods, the phase angles of the reflexions (hkl) can be calculated (Taylor. 2010).

Some examples of traditionally soaked heavy-atoms into protein crystals:

- Mercurial compounds tend to bind to thiol groups of cysteines or nitrogen of histidine,
- Platinum compounds bind to cysteine, histidine and methionine residues,
- Uranyl salts bind to carboxyl groups of aspartic and glutamic acids.

6.2.7. Anomalous Dispersion (AD)

Heavy-atoms have the capacity to absorb of X-rays and re-emit them at specific wavelengths. Due to resonance, the absorption dramatically drops just below the *absorption edge* at a wavelength that is characteristic of each heavy-atom. As a result of this phenomenon, a difference exists between the intensity of symmetry-related reflections hkl and $-h-k-l$, which is called *anomalous dispersion*. Since the other atoms such as carbon, nitrogen, oxygen, and hydrogen in a protein do not contribute significantly to anomalous dispersion at the wavelengths used in X-ray crystallography ($\sim 1.0 \text{ \AA}$), the wavelength of the X-radiation can be varied to maximize the anomalous dispersion of a particular heavy-atom naturally present in the crystal or introduced by crystal soaking.

A typical Multi-wavelength Anomalous Dispersion (MAD) experiment requires data collection at different X-ray wavelengths (edge, peak and remote) around the absorption edge of the respective heavy-atom that can be finely tuned using synchrotron sources. The most preferred heavy-atom is selenium that is not soaked into protein crystals, but instead the target protein is expressed using a Met(-) organism, and the medium is supplemented with selenomethionine. During expression methionines are replaced by selenomethionines having a theoretical absorption peak at $\sim 0.979 \text{ \AA}$ wavelength (Rupp 2010).

Thanks to the progress achieved in data collection procedures and in the development of powerful algorithms the Single-wavelength Anomalous Diffraction (SAD) strategy can nowadays be applied with reasonable chances of success. The advantage of this method is that data are collected at a single wavelength at the absorption peak of the employed heavy atom.

X-ray Crystallography methods

6.2.8. Model Building and Refinement

After initial phase angles of all the measured intensities (hkl) were obtained, an electron density map of the protein can be calculated. When the map is interpretable, the secondary structure elements (α -helices and β -sheets) are first to be traced and modeled. The structure amplitudes calculated from this model can be compared with the measured structure amplitudes. The agreement between the calculated structure amplitudes (F_{calc}) from the model and the observed structure amplitudes (F_{obs}) from experimental data is represented by the factor $R = (\Sigma(F_{obs} - F_{calc}) / (\Sigma F_{obs}))$.

In each round of model building, chemical (bond lengths, bond angles, torsion angles, etc) and physical parameters (atomic overlaps, energetically favored interatomic interactions, etc) are taken into account. During structure refinement the improvement of the model can be monitored by R-factor values, and as better agreement between calculated and observed structure factors is obtained, the R-factor decreases.

The reliability of the finally obtained structural models has to be evaluated using the MolProbity server. It analyses the potential geometrical errors in the model and gives the most reliable view of hydrogen bonding, steric interactions, and different rotamers of main chain and amino acid side chains. The inclusion of hydrogen atoms in this process is needed to determine the orientation of Asn/Gln/His side-chain flips on the basis of potential hydrogen bond formation. Sterically disallowed conformations in the protein backbone are detected by a Ramachandran plot. This diagram monitors torsion angles ϕ (the angle of rotation around $C\alpha$ -N bonds) and ψ (the rotation angle around $C\alpha$ -C bonds) and indicates outliers that have to be corrected based on the electron density map (Chen *et al.*, 2010).

Purification of ABC importers

7. Purification of ABC importers

7.1. Overexpression and Purification of Art M₂P₂/J

ABC transporters mediating the uptake of amino acids have been classified into three separate subfamilies designated as PAAT (Polar Amino Acid Transporters), HAAT (Hydrophobic Amino Acid Transporters), and MUT (Methionine Uptake Transporters) (Saier, 2000; Hosie and Poole, 2001; Zhang *et al.*, 2003). The prokaryotic ABC transport system ArtMP-J from *G. stearothermophilus* belongs to PAAT subfamily, which imports arginine and histidine. The importer complex consists of an extracellular substrate-binding receptor, ArtJ, which is required for the transporter function and the transporter pore which is composed of two heterodimers; the membrane-integral subunit, ArtM, and the ATP-hydrolysing subunit, ArtP. The crystal structure of the extracellular substrate-binding receptor, ArtJ, has been recently solved (Vahedi-Faridi *et al.*, 2008). Interaction between substrate-binding receptor (ArtJ) and the membrane intrinsic transport pore (ArtM) was analyzed by mutating several charged residues at the putative binding surface. Mutational analyses indicated the importance of Glu-163 residue on the surface of receptor (Vahedi-Faridi *et al.*, 2008) and much higher transport activity observed when Lys-159 on the surface of transmembrane subunits (ArtM) was exchanged to Glu-159 or Ise-159 (Schneider *et al.*, 2011). Purification and screening experiments aimed to solve the crystal structure of the transport pore ArtM₂P₂ and the complex with binding receptor ArtJ were performed (Fig 14).

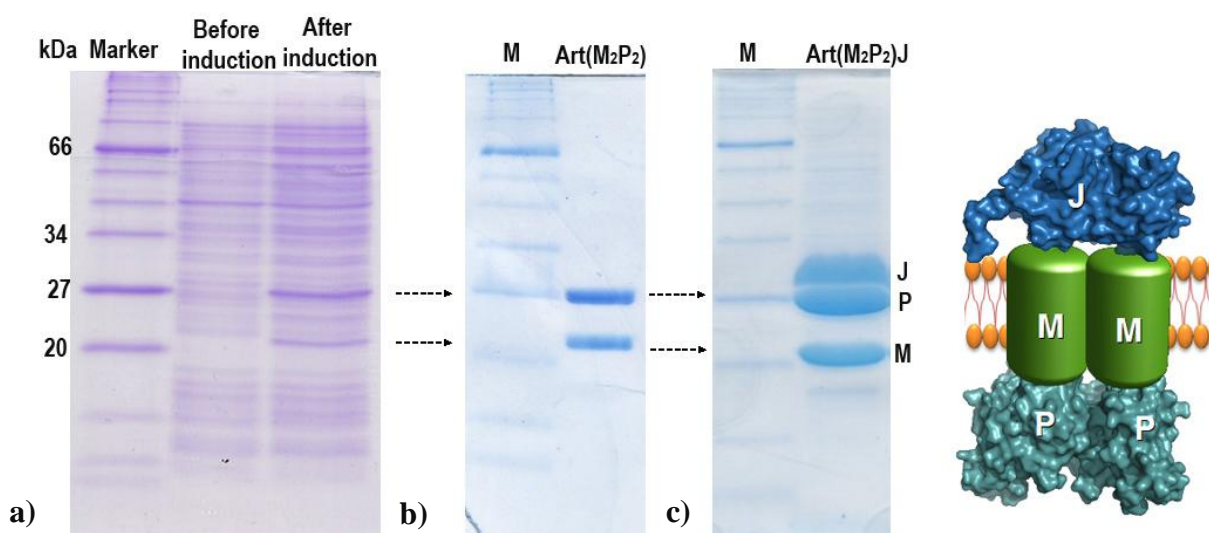


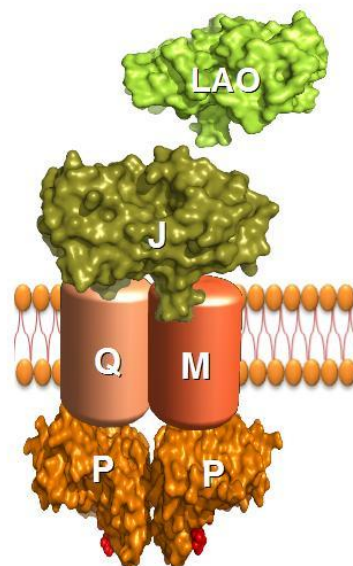
Fig. 14. Summary of overexpression and purification of arginine transporter; the core complex contains His₆ fused to ArtP, 27kDa and ArtM, 24kDa purified using Co⁺² metal affinity chromatography and run on gel-filtration S-200. **a)** High yield of Art(MP)₂ is visible on SDS gel after 4 hours of overexpression. **b)** Purity of protein samples that were used for crystallization trials judged on an SDS gel. Crystallization trials were set up using either the complex ArtM₂P₂ or **c)** in complex with arginine binding protein ArtJ.

Purification of ABC importers

7.2. Overexpression and Purification of HisP₂QM

Histidine-transporter (HisP₂QM) of *S. Typhimurium* has been extensively studied from the biochemical, physiological, and genetic points of view. It consists of a core complex, HisQMP₂, which comprises different integral membrane subunits, HisQ and HisM, each spanning the membrane five times, and two copies of HisP, the ATP-binding subunit (Fig. 15). The overall crystal structure of the HisP monomer consist of two thick arms (arm I and arm II) that form an 'L' like shape (Hung *et al* 1998). Interestingly, the transport complex has two periplasmic binding proteins, HisJ for histidine and LAO for lysine-, arginine-, ornithine (Oh *et al* 1994). The sequences of these binding receptors share 70% identity and therefore might have been evolved by gene duplication. The structures of the two receptors are quite similar, being composed of two lobes held together by two peptide loops forming a hinge. (Oh *et al* 1993).

Fig. 15. Domain organization of histidine transporter comprises two conjugate binding receptors, HisJ and LAO, two transmembrane subunits, HisQ and HisM, and a homodimer of the ATPase subunit, HisP. The crystal structure of HisP shows two monomers contacting each other through hydrophobic interactions on the outer side of arm I.; the ATP-binding pocket is near the end of arm I



A vector was designed by fusing a C-terminal his-tag to the nucleotide-binding subunit HisP_(His6). An *E. coli* strain containing plasmid pVE26 (Eckey *et al.*, 2010) (freshly introduced by transformation) was grown at 30 °C until the culture reached an A_{650 nm} of 0.6. IPTG was added to a final concentration of 0,1 mM and growth continued at 22 °C overnight. The slower growth at this temperature allowed a higher percentage of transport complex to remain soluble. The core complex (HisQMP₂) was purified by metal affinity chromatography. All buffers were supplemented with 10% glycerol and protease inhibitor cocktail that was critical for the stability of transmembrane subunit HisQ. The protein was solubilized first in decanoylsucrose and during gel filtration run it was exchanged to β -DDM for crystallization trials (Fig. 16). Protein was

Purification of ABC importers

concentrated to around 15mg/ml using centricon devices (Millipore, 100,000 MW cutoff). Initial screening of the crystallization conditions gave rise to primary crystals of the HisP₂QM that are further tried for optimization to get crystals suitable for X-ray diffraction data collection.

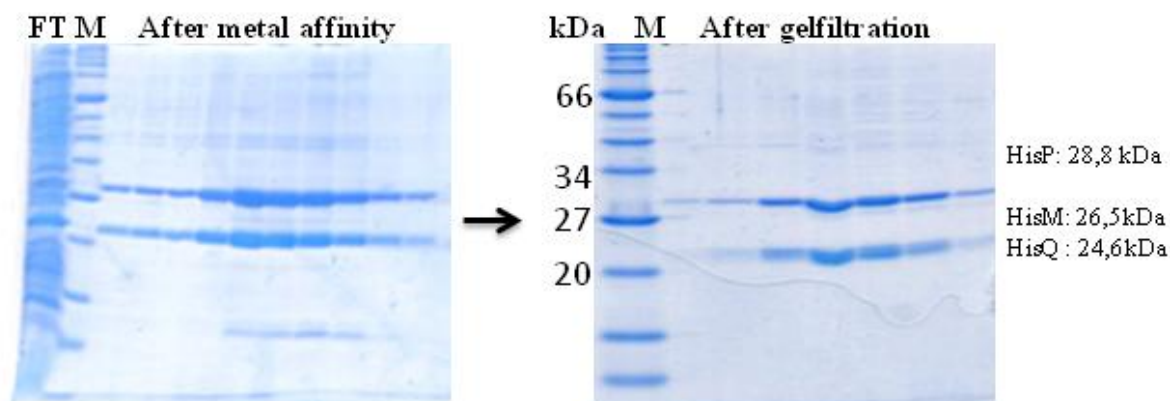


Fig. 16. SDS–polyacrylamide gel of the purified HisQMP₂ complex, after affinity and SEC chromatography, respectively. Small bands of impurity were eliminated by SEC chromatography.

7.3. Overexpression and Purification of Pro(WX)₂V₂

Cytoplasmic water content is essential for cell survival and division. Micro-organisms living in high saline environment have developed an ATP-binding cassette uptake system that imports osmoprotectants such as glycine betaine and proline betaine to avoid unfavourable water efflux (Kempf and Bremer, 1998). In addition these solutes act as a protein stabilizer (Caldas, *et al.*, 1999).

The Pro(WX)₂V₂ gene cluster was predicted to be responsible for glycine betaine and proline betaine uptake in *H. pylorii*. In contrast to other model transporters in this study, transmembrane segments covalently fused with the substrate binding receptors (ProX) that are associated with motor domain (ProV). A structural homolog of binding receptor ProX was solved in complex with both of its natural ligands glycine betaine and proline betaine, as well as in complex with the artificial ligand trimethylammonium (Schiefner *et al.*, 2004). The structure exhibits the typical “venus fly trap” fold found in many importers, however, the interaction of the receptor (ProX) with the transmembrane domains was not investigated.

The transmembrane domain (ProW) was predicted to have 6 helices using the TMHMM server, thus Pro(WX)₂V₂ can be classified as type I transporter. Two intact structures (maltose and molybdate) from this type transporter co-crystallized with their binding receptors, indicate only one binding receptor able to bind to the surface of TMs (Gerber *et al.*, 2008, and Oldham *et al.*, 2007). This leaves questions about their binding mechanism presumably switching one to other binding

Purification of ABC importers

receptor in actual time. Transporter complexes were expressed and purified in β -DDM up to 15 mg/ml (Fig. 17) as described previously and drystal trials were conducted using sitting drop vapor diffusion.

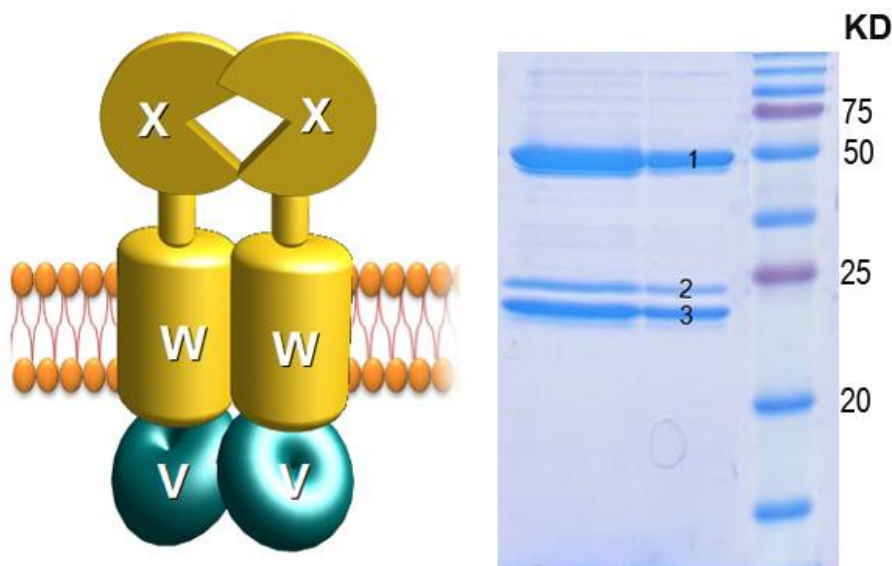


Fig. 17. On the Right, the scheme display of the domain organization of Pro(WV)₂V₂. The transport complex contains two ATPase subunits (ProV) and two homodimer membrane (ProW) subunits attached to substrate binding receptors (ProX). The purified transport complex was analysed by MALDI-TOF. Band 1 comprises the ProWX and bands 2 and 3 ProV subunits.

7.4. Overexpression and Purification of the NGO 0373-74 Complex

Based on sequence identity the NGO 0372-74 gene clusters from *Neiseria gonorrhoeae* were identified as amino-acid ABC transporters. The ligand of receptor component (NGO 0372) was structurally and biochemically characterized as being L-cystine. Since the substrate selectivity of the receptor component determines the substrate of the transport complex, it appears that the NGO 0372-74 transport complex is responsible for L-cystine uptake in *N. Gonorrhoea*. Core components of the transport complex contain a homodimer of membrane subunits (NGO 0373) and a homodimer of ATPase subunits+6xhis (NGO 0374) purified by Co-TALON chromatography and applied to gelfiltration column for further purification (Fig. 18).

Purification of ABC importers

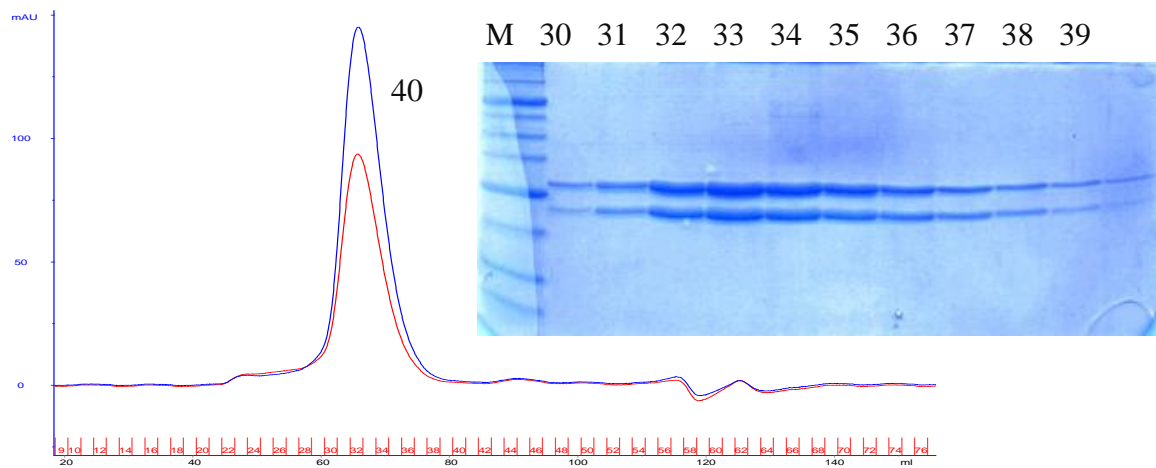


Fig. 18. Size-exclusion chromatography of the NGO 0373-74 complex that was purified by Co-TALON chromatography as described and concentrated to 20 mg/ml on Amicon Ultra (100,000 MWCO), and 40 mg of protein was loaded onto Sephadex S200 in 20 mM Tris.HCl, pH 8.0, 100 mM NaCl, 5% (v/v) glycerol, and 0.15 % (w/v) DDM. On top right a coomassie brilliant blue-stained SDS-polyacrylamide gel of the individual protein fractions is shown.

8. Structural Analyses of Binding Receptors

Crystal of GacH, MalE co-crystalized with acarbose and acarbose analogs. Structure of AcbH obtained with monosaccharide D-Galactose. The periplasmic binding proteins of the *N. gonorrhoeae* were overexpressed independently in *E. coli* and their x-ray structures were solved in complex with their ligands. Results are subdivided as following 3 different parts; chapters I and II were published in the Journal of Molecular Biology and chapter 3 was accepted to the Journal of Molecular Biology.

9. Chapter I

Crystal Structures of the Solute Receptor GacH of *Streptomyces glaucescens* in Complex with Acarbose and an Acarbose Homolog: Comparison with the Acarbose-Loaded Maltose-Binding Protein of *Salmonella typhimurium*

Ardeschir Vahedi-Faridi^{1§}, Anke Licht^{2§}, Haydar Bulut¹, Frank Scheffel², Sandro Keller³, Udo F. Wehmeier⁴, Wolfram Saenger^{1*}, and Erwin Schneider^{2*}

¹ Institut für Chemie und Biochemie/Kristallographie, Freie Universität Berlin, Takustr. 6, D-14195 Berlin, Germany

² Institut für Biologie, AG Bakterienphysiologie, Humboldt Universität zu Berlin, Chausseestr. 117, D-10115 Berlin, Germany

³ Leibnitz-Institut für Molekulare Pharmakologie (FMP), Robert-Roessle-Strasse 10, D-13125 Berlin, Germany

⁴ Institut für Chemische Mikrobiologie, Bergische Universität Wuppertal, D-42097 Wuppertal, Germany

* Corresponding authors:

saenger@chemie.fu-berlin.de

erwin.schneider@rz.hu-berlin.de; Phone: +49(0)30-2093-8121; Fax: +49(0)30-2093-8126

§ These authors contributed equally to this work.

Abbreviations: ABC, ATP-binding cassette; ecMBP, *Escherichia coli* maltose/maltodextrin binding protein; tIMBP, *Thermococcus litoralis* trehalose/maltose-binding protein; ttMtBP, *Thermus thermophilus* maltotriose-binding protein; ttGBP, *Thermus thermophilus* glucose-binding protein; PDB, Protein Data Bank; BSA, buried surface area; CV, column volume

Chapter I

Abstract

GacH is the solute binding protein (receptor) of the putative oligosaccharide ATP-binding cassette transporter GacFG, encoded in the acarbose biosynthetic gene cluster (*gac*) from *Streptomyces glaucescens* GLA.O. In the context of the proposed function of acarbose (acarviosyl-1,4-maltose) as a ‘carbophor,’ the transporter, in complex with a yet to be identified ATPase subunit, is supposed to mediate the uptake of longer acarbose homologs and acarbose for recycling purposes. Binding assays using isothermal titration calorimetry identified GacH as a maltose/maltodextrin-binding protein with a low affinity for acarbose but with considerable binding activity for its homolog, component 5C (acarviosyl-1,4-maltose-1,4-glucose-1,1-glucose). In contrast, the maltose-binding protein of *Salmonella typhimurium* (MalE) displays high-affinity acarbose binding. We determined the crystal structures of GacH in complex with acarbose, component 5C, and maltotetraose, as well as in unliganded form. As found for other solute receptors, the polypeptide chain of GacH is folded into two distinct domains (lobes) connected by a hinge, with the interface between the lobes forming the substrate-binding pocket. GacH does not specifically bind the acarviosyl group, but displays specificity for binding of the maltose moiety in the inner part of its binding pocket. The crystal structure of acarbose-loaded MalE showed that two glucose units of acarbose are bound at the same region and position as maltose. A comparative analysis revealed that in GacH, acarbose is buried deeper into the binding pocket than in MalE by exactly one glucose ring shift, resulting in a total of 18 hydrogen-bond interactions *versus* 21 hydrogen-bond interactions for MalE_{acarbose}. Since the substrate specificity of ATP-binding cassette import systems is determined by the cognate binding protein, our results provide the first biochemical and structural evidence for the proposed role of GacHFG in acarbose metabolism.

Keywords: acarbose; ABC transporter; acarbose-binding protein; protein crystallography; *Streptomyces glaucescens*

Chapter I

Introduction

The pseudomaltotetraose acarbose (acarviosyl-1,4-maltose), produced by strains of the genera *Actinoplanes* and *Streptomyces*, is an inhibitor of various α -glucosidases, including those in the human intestine.^{1,2} Therefore, acarbose, which is industrially produced by optimized strains of *Actinoplanes* sp. SE50/110, is clinically used in the treatment of patients suffering from type II diabetes because it slows the release of glucose from starch-containing diets.³

The characteristic core structure of acarbose-like secondary metabolites, the pseudodisaccharide acarviosine, consists of an unsaturated C7 cyclitol bound via an imino bridge to 4-amino-4,6-dideoxyglucose (Fig. 1a and b) and is primarily responsible for the inhibitory effect.

The acarbose structure is intracellularly completed by a maltose residue and phosphorylated to acarbose-7-phosphate in order to prevent inhibition of cytosolic enzymes. Presumably due to the action of the ATP-binding cassette (ABC) exporter AcbWXY, acarbose-7-phosphate is subsequently secreted into the natural environment, where it becomes dephosphorylated.^{5,6} Here, acarbose not only impairs the activity of starch-degrading enzymes from competing microorganisms but also, due to its structural similarity to maltotetraose, is recognized as a substrate by their maltose/maltodextrin transport systems.^{7,8} However, as demonstrated in the case of *Escherichia coli*, acarbose is not metabolized since it is not a substrate of amyloamylase, an essential enzyme for maltose/maltodextrin utilization.⁷ As a consequence, acarbose effectively inhibits the growth of competing organisms on maltose/maltodextrins as sole sources of carbon and energy.

Besides these inhibitory effects, which might mediate a growth advantage for the slowly growing acarbose producers, an additional role as a ‘carbophor’ (carbon storage molecule) for the uptake of carbon sources was recently proposed.⁶ The carbophor cycle includes the following steps: (i) enzymatic transfer of glucose or oligosaccharides derived from starch to (extracellular) acarbose by the secreted acarviosyltransferase AcbD; (ii) uptake of the longer derivatives of acarbose into the cytoplasm by the ABC import system AcbHFG-MsiK2; (iii) liberation and subsequent glycolytic degradation of added glucose units; and (iv) reexport of acarbose-7-phosphate. In consequence, this would result in a net gain of carbon and energy. Indeed, strains of *Actinoplanes* produce various acarbose homologs, which, depending on the carbon source, differ in the number and chemical nature of sugars linked to the acarviosine core unit.⁹

Recently, a second acarbose production gene cluster (*gac*) with a high level of similarity to the *Actinoplanes* *acb* gene cluster was identified in *Streptomyces glaucescens* GLA.O.⁴ Biochemical

Chapter I

characterization of several gene products has subsequently led to the proposal of a functionally analogous acarbose biosynthetic pathway, which differs only in details but leads to the same endproduct. Moreover, this pathway shares all functions for the predicted ‘carbophor’ cycle, including genes for two ABC transport systems (Fig. 1c). Whereas the gene cluster encoding the putative ABC exporter GacWXY shows the highest level of sequence similarity to the postulated export system AcbWXY from *Actinoplanes* sp., the putative ABC import system GacHFG is more related to the proposed maltose uptake systems from *Streptomyces* than to AcbHFG. In a previous study, it was demonstrated that the solute receptor AcbH, when purified from inclusion bodies, binds acarbose and longer homologs, but no maltooligosaccharides.¹⁰ The substrate specificity of GacH is unknown.

To elucidate the role of the putative GacHFG transporter, we focused on GacH, since solute receptors of ABC importers are the major determinants of substrate specificity.¹¹ Here we present the first structural evidence for the postulated ‘carbophor’ function in *S. glaucescens*. Moreover, comparison with the maltose-binding protein (MalE) of *Salmonella typhimurium*, in complex with acarbose, revealed significant differences in acarbose binding to both receptors.

Results and Discussion

Substrate-binding activities of GacH

Homology searches indicated that GacH is a component of an ABC import system belonging to the CUT 1 subfamily,¹² the members of which predominantly transport disaccharides and oligosaccharides. Thus, we first investigated the binding of maltose, maltotetraose, acarbose, and component 5C (acarviosyl-1,4-maltose-1,4-glucose-1,1-glucose) to purified and dialyzed tagless GacH (Fig. 2g) by isothermal titration calorimetry. GacH showed considerable differences in its affinities for maltose ($K_d=2.9\pm 0.3\ \mu\text{M}$; Fig. 2a), maltotetraose ($K_d=1.7\pm 0.2\ \mu\text{M}$; Fig. 2b), and component 5C ($K_d=10.0\pm 0.6\ \mu\text{M}$; Fig. 2c), in comparison to its affinity for acarbose ($K_d=248\pm 9\ \mu\text{M}$; Fig. 2d). Furthermore, we assayed the binding of maltotetraose and acarbose to *E. coli* MalE, yielding similar affinities for both oligosaccharides (K_d values of $2.5\pm 0.1\ \mu\text{M}$ and $2.9\pm 0.3\ \mu\text{M}$, respectively) (Fig. 2e and f). Although the latter value should be taken as an estimate due to the small differences in enthalpy change, it is noteworthy that the dissociation constant of MalE for acarbose is in the range of the inhibition constant ($1.1\ \mu\text{M}$) for maltose uptake in *E. coli*.⁷

Chapter I

These results suggest that the substrate spectrum of GacH resembles that of maltose/maltodextrin-binding proteins from other prokaryotes.^{8,13,14} However, and in striking contrast to MalE proteins of *E. coli* and *Alicyclobacillus acidocaldarius*,⁸ GacH displays only a low affinity for acarbose. This property is shared with AcbH of *Actinoplanes sp.*, which, other than GacH, does not bind maltodextrins.¹⁰ On the other hand, GacH binds component 5C considerably better than acarbose. This finding is in line with the notion that longer acarbose homologs are the preferred substrates of the GacHFG transporter, while uptake of acarbose is for recycling purposes only.⁴

Overall structure of GacH

A model of GacH was built based on X-ray diffraction data from selenomethionine-substituted crystals and further refined using native data sets of GacH cocrystallized with maltotetraose, acarbose, and the acarbose homologs component 5C and component mixture 6AB (acarviosyl-1,4-maltose-1,4-glucose-1,4-fructose and acarviosyl-1,4-maltose-1,4-glucose-1,4-glucose) at 1.55 Å, 1.40 Å, 1.55 Å, and 1.35 Å resolutions, respectively. Interestingly, cocrystallization of GacH and acarbose homolog mixture 6AB resulted in an unliganded open conformation of GacH. Thus far, the crystal structures of several carbohydrate-binding proteins displaying specificities for different glucose-containing substrates, represented by *E. coli* maltose/maltodextrin-binding protein (ecMBP),^{15–17} *Thermococcus litoralis* trehalose/maltose-binding protein (tlMBP),¹⁸ *Thermus thermophilus* maltotriose-binding protein (ttMtBP),¹⁹ and *T. thermophilus* glucose-binding protein (ttGBP),²⁰ have been determined. All share a common structural motif that consists of two domains joined by a hinge region flanking the central sugar-binding pocket. The overall structure of GacH shows a similar fold that is composed of two lobes/domains: domain I (residues 17–132 and 289–354), which forms an α -helix bundle, and domain II (residues 137–273 and 364–404), which contains a central β -sheet flanked by two layers of α -helices (Fig. 3). The domains are separated by a tripartite hinge region (residues 133–136, 274–288, and 355–363). Similar to the MalE receptor, GacH undergoes a remarkable conformational change upon ligand binding (Fig. 3).

The crystal structure of GacH, in complex with acarbose, shows clear electron density for the four sugar residues (A–D; see Table 1) in the binding pocket. Acarbose is bound in a U-shaped fashion within a hydrophobic keyhole formed by the aromatic residues Phe59 of domain I, Phe368, Tyr182, and Trp254 (Fig. 4), with Phe59 located in the center of the binding pocket. Phe368,

Chapter I

Tyr182, and Trp254 belong to domain II of GacH and form the outer portion of the U-shaped keyhole. As with other carbohydrate receptors, hydrogen bonds and van der Waals contacts play key roles in oligosaccharide recognition and binding. The initial interaction of acarbose presumably occurs with the open conformation of GacH. In GacHacarbose, only Asp180 and Asn365 of domain I bind with two hydrogen bonds to glucose units A and C (Fig. 4), whereas all the other hydrogen-bond interactions are contributed by domain II (Table 1).

The electron density clearly shows the acarviosine moiety of acarbose near the surface of the binding pocket, with the unsaturated cyclitol moiety partially protruding from the GacH surface, thereby limiting the ability of the protein to specifically bind acarbose. In fact, a comparison of the binding of acarbose to GacH versus the binding of maltotetraose to GacH shows that the maltose moiety of acarbose coincides perfectly with the second maltose group of maltotetraose (Fig. 5a). In contrast, an overlay of the first maltose moiety in maltotetraose with the acarviosyl group shows a 180° rotation of the cyclitol moiety, thereby inducing a small shift at the 4-amino-4,6-dideoxyhexose residue of acarbose. These differences and the absence of one hydroxyl group at 4-amino-4,6-dideoxyhexose result in the deletion of five hydrogen bonds for the binding of acarbose (Table 1). Similarly, our structural data obtained for GacH_{component 5C} clearly show that the acarbose analog also binds to GacH, with its acarviosyl group located at the outer part of the binding pocket. Surprisingly, the electron density shows only five of the six sugar rings present in component 5C. The electron density identifies both ends of the acarbose analog and, as stated above, the acarviosyl group is at the entrance of the binding pocket, while the maltose moiety and 1,4-glucose are located in the binding pocket (Fig. 5b). The absence of the sixth α -1,1-glucose and the absence of the α -1,1-glycosyl bond suggest that the residue was hydrolyzed prior to or during crystallization because component 5C, as used in the crystallization experiment, showed the expected molecular mass of 969.4 g/mol by electrospray ionization tandem time-of-flight mass spectrometry (data not shown). Moreover, HPLC analysis by the manufacturer provided no clue for impurity. Likewise, SDS-PAGE analysis does not indicate the presence of a protein band contaminating the GacH preparation that might have cleaved the compound (Fig. 2g). Unfortunately, the amount of component 5C that was made available to us by the manufacturer was insufficient to explore this phenomenon further (e.g., by thin-layer chromatography analysis).

Interestingly, the resulting ligand (component 5C) bound in the pocket of GacH is, in fact, one of the acarbose homologs detected in the culture supernatant of *S. glaucescens* GLA.O by Rockser and Wehmeier.⁴ These findings indicate that GacH specifically binds maltose with α -1,4-

Chapter I

glycosyl bond at the inner part of the binding pocket and, therefore, cannot bind the disaccharide trehalose with α -1,1-glycosyl bond at the end of component 5C. This notion is consistent with the failure of trehalose to compete with [U-¹⁴C]maltose for binding to GacH in a filtration assay (data not shown) and is further reinforced by the observation that cocrystallization of GacH with component mixture 6AB (Fig. 1) did not lead to a complex with GacH, but resulted in the open conformation of the protein. This finding seems to contradict the result of competition experiments clearly showing an inhibitory, albeit weak, effect of component mixture 6AB on the [U-¹⁴C]maltose binding activity of GacH (data not shown). Additional experiment(s) will be required to explain the unexpected discrepancy.

The binding mode of component 5C by GacH features 21 hydrogen bonds, while maltotetraose is bound by 23 hydrogen bonds (Table 1). The two interior saccharide rings C and D (Fig. 4) show nearly identical bonding interactions, whereas ring B of component 5C (4-amino-4,6-dideoxyhexose) provides no interactions with Asn365 because Asn365 is twisted away from the ligand when compared to GacHacarbose. Ring A of component 5C (cyclitol moiety) protrudes from the binding pocket and shows no direct interactions with any amino acid of GacH (Fig. 5b). However, when compared to the interactions with acarbose, component 5C exhibits three additional hydrogen-bond interactions (Table 1) involving Asp180 (two interactions with O6B) and Arg358 (one additional interaction with O3B). These findings are consistent with results from the binding assays described above. The large increase in the K_d value of GacH for acarbose could be due to the reduction in the number of hydrogen bonds, especially the missing interactions with Asp180.

In conclusion, we identified GacH as a maltose/maltodextrin receptor with additional affinity for maltodextrin-like acarbose homologs, with possible reduction in affinity due to the lower number of hydrogen bonds. GacH does not seem to be specific for the binding of the acarviosyl group; rather, it shows specificity for the binding of the maltose moiety at the interior part of its binding pocket.¹

* Please note that ITC measurements with component mixture 6AB were not feasible due to insufficient amounts of material.

Chapter I

Comparison of acarbose binding in GacH and MalE

The previous finding that acarbose is transported by the *E. coli* maltose/maltodextrin ABC transporter (MalE-FGK2),⁷ together with the observed high affinity of MalE for acarbose (Fig. 2), prompted us to solve the crystal structure of MalE in complex with acarbose. Similar to GacH, crystallization in the presence of acarbose led to a MalE structure in which acarbose is bound in the binding pocket via the maltose moiety. This is very similar to the complex formed between MalE and maltotetraose [Protein Data Bank (PDB) ID 4MBP], while the acarviosine moiety protrudes from the binding pocket; in contrast to GacH, there is no 180° rotation of the cyclitol moiety. The lack of acarviosine rotation is also the reason that the hydrogen bonds do not differ significantly in MalE_{acarbose} and MalE_{maltotetraose} complexes (21 and 22, respectively; Table 2), which may account for the similarity in the respective K_d values. However, an overlay of the GacH_{acarbose} and MalE_{acarbose} binding pockets (Fig. 6) shows significant differences, although the overall structure of GacH is similar to that of MalE (Fig. 3), with an overall RMSD of only 1.886 Å. Most strikingly, acarbose is buried deeper in the binding pocket of GacH by exactly one glucose ring shift compared to MalE. The differences in substrate binding mode and those resulting from the low amino acid sequence identity (Tables 1 and 2) result in a completely different bonding environment for acarbose within these two binding pockets. Surprisingly and in spite of the different binding geometry, the number of hydrogen bonds for MalE_{acarbose} is larger than that for GacH_{acarbose} (21 versus 18; Tables 1 and 2). A comparison of the three-dimensional GacH structure with those of tlMBP, ttMtBP, and ttGBP revealed that the specificity depends on the presence or the absence of four protein subsites (A–D) that bind individual glucose units and are located in the same region of the binding pocket.¹⁹ The most deeply buried subsite A is absent in ecMBP but is present in the other three proteins, whereas the most exposed subsite D is present only in the ecMBP. This could be the reason that the bound maltotriose in ecMBP is shifted exactly by one glucose ring out of the pocket compared with the bound substrates of the other complexes, as shown in the overlay with GacH (Fig. 6). This is further reinforced by the identification of specific subsite A residues in the GacH sequence (Thr27, Glu32, Phe59, Arg81, Trp254, and Trp290), while no residue homologous to ecMBP subsite D was found in GacH (Fig. 7).

Figure 8 shows that only four amino acids exhibit a major movement in the orientation of their side chains due to the binding of acarbose to MalE when compared to the MalE_{maltose} structure (PDB ID 1ANF): Tyr341, Arg344, Lys42, and Arg66. Of these, only Tyr341 moves towards the substrate and forms a hydrogen bond with the hydroxyl group (O3B) of acarbose sugar B. Arg344

Chapter I

and Arg66 move “away” from the substrate to avoid clashes with sugar moiety B. Lys42, which is located at the entrance of the binding pocket, shifts away to avoid clashes with sugar A located near the entrance. When compared to the binding of maltose, acarbose contributes five additional hydrogen bonds: four through sugar B and one through sugar A. Furthermore, five water-mediated hydrogen bonds contribute to the binding of sugars A and B to MalE. Sugars C and D of acarbose coincide with both sugar rings of maltose, with no significant differences in positions and interactions with the residues of the binding pocket. The binding of acarbose also increases the buried surface area (BSA) to 667.8 Å² when compared to that of maltose (BSA: 432.78 Å²). However, in comparison to maltose where almost the entire molecule is buried in the binding pocket (BSA/accessible surface area: 432.8 Å²/465.0 Å²=0.93), sugar ring A of acarbose lies at the MalE surface (BSA/accessible surface area: 667.8 Å²/796.1 Å²=0.84). Interestingly, the additional interactions of the receptor with sugar moieties A and B of acarbose and the flexibility of MalE to avoid clashes with the longer substrate via movement of the relevant amino acid side chains do not result in a noticeable increase in binding affinity when compared to the binding of maltose.

Conclusions

Together, our findings provide the first experimental evidence for the proposed role of acarbose as a carbon storage molecule for the producing bacteria. At low (micromolar) concentrations, longer acarbose homologs are recognized by GacH and likely internalized by the cognate Gac transporter. In contrast, acarbose itself is only a substrate at high (micromolar) concentrations (i.e., when no longer needed as an inhibitor of the maltose/maltodextrin metabolism of competitors and thus accumulating in the environment). Under these conditions, acarbose is imported for recycling purposes, enabling the bacteria to feed on its own metabolite, as demonstrated for *Actinoplanes* sp.²² The operon structure of the gacHFG gene implies that GacH delivers its substrates to an ABC transporter composed of the membrane proteins GacF and GacG and an as yet unidentified ATPase subunit for which, however, MsiK (a multiple sugar import protein) is a good candidate.^{23–25} Purification of the transport complex GacFG–MsiK2 and reconstitution of its functions in proteoliposomes would provide the ultimate proof for its role in acarbose metabolism. Experiments to achieve this goal are underway in our laboratories but are hampered by the low expression of the genes in *E. coli* host strains

Chapter I

Materials and Methods

Chemicals

Maltose, maltotriose, and maltotetraose were obtained from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany) and were of 98 % to 96 % purity. D-(+)-Trehalose was purchased from Carl Roth GmbH (Karlsruhe, Germany). [U-¹⁴C] maltose (22.2 GBq/mmol) was obtained from Hartmann Analytik (Braunschweig, Germany). Acarbose (acarviosyl-1,4-maltose), component 5C (acarviosyl-1,4-maltose-1,4-glucose-1,1-glucose), component mixture 6AB (acarviosyl-1,4-maltose-1,4-maltose-1,4-fructose and acarviosyl-1,4-maltose-1,4-maltose-1,4-glucose), all derived from *Actinoplanes* sp., were kind gifts from H. Wehlmann (Bayer AG, Wuppertal, Germany). According to the manufacturer, acarbose was of 97.8 % purity and contained trace amounts of acarbose derivatives (0.9 %) but was devoid of glucose, maltose or maltodextrins.

Purification, crystallization and structure determination of GacH

For overproduction purposes, the *gacH* gene (codons 36–425) lacking the signal sequence and the putative linker region were amplified by PCR from genomic DNA of *S. glaucescens* GLA.O using the oligonucleotide primers *algf1* (ATAATACATATGACGGAGCTGTCGGGGACCGTCACCTTC) and *algr1* (ATAATAGGATCCCTACTTGTAGTCCTTGAGCAGCTTGCG) and cloned into the expression vector pET15b providing an Nterminal His6-tag (Novagen, Germany). The described modifications allowed the recovery of soluble GacH from the cytoplasm of the *E. coli* host strain. The resulting plasmid pAL09 was transferred into *E. coli* strain Rosetta (plysS) (Stratagene) containing the *ptacSL1* plasmid,²⁶ thereby allowing coexpression of *gacH* and the genes of the GroEL/ES chaperonin complex. Cells were grown in Luria–Bertani medium supplemented with ampicillin, chloramphenicol, and kanamycin at 37 °C. At OD₆₅₀=1, expression was induced by adding 0.5 mM IPTG, and growth continued for 4 h. Subsequently, cells were harvested, resuspended in buffer 1 [50 mM 4-morpholinepropanesulfonic acid–KOH (pH 7.0), 300 mM NaCl, 5% (vol/vol) glycerol, and 0.1 mM PMSF], and disintegrated by passage through a French press. After ultracentrifugation, the supernatant containing the truncated GacH (Δ Cys1–Gly15) fused to an N-terminal His6- tag was subjected to metal-affinity chromatography using a 5-ml TALON-HiTrap resin and the ÄKTA™ purifier FPLC system (GE Healthcare, Freiburg, Germany). The purification protocol included equilibration of the matrix with 5 column volumes (CV) of buffer 1, loading of the cytosolic extract

Chapter I

of 1 L of culture containing His6-GacH, washing of the column with 10 CV of buffer A, and elution of retained proteins with 10 CV of buffer 1 containing 100 mM imidazole. All buffers were filtered through 0.45- μm filters before use, and chromatography was performed at a flow rate of 3 ml/min at 4 °C. Subsequently, GacH-containing fractions were pooled, and the His6-tag was cleaved by treatment with thrombin (Thrombin CleanCleave™ Kit; Sigma-Aldrich Chemie GmbH) in accordance with the manufacturer's instructions. To remove endogenously bound sugar, we dialyzed the purified GacH protein against 10 L of buffer 2 [20 mM 4-morpholinepropanesulfonic acid–KOH (pH 7.0), 100 mM NaCl, and 1% (vol/vol) glycerol], with several changes. Routinely, 60 mg of protein per liter of cells was obtained. For crystallization purposes, the GacH dialysate was concentrated up to 60 mg/ml and supplemented with acarbose (5 mM), maltotetraose (5 mM), or the acarbose homologs component 5C and component mixture 6AB (2 mM each). Crystals were grown at 18 °C using the sitting-drop vapor-diffusion method. The reservoir solution contained 2.5 M $(\text{NH}_4)_2\text{SO}_4$ and 100 mM citric acid (pH 4.0), and drops were prepared by mixing 3 μl of reservoir solution and 3 μl of protein/sugar solution. Large tetragonal-shaped crystals grew approximately in 2 weeks to a maximum size of 1.0 mm \times 0.5 mm \times 0.3 mm. Before data collection, crystals were briefly soaked in a cryoprotection medium [reservoir solution with 15% (vol/vol) glycerol], mounted on a nylon loop, and then flashcooled in liquid N₂. X-ray data were collected at beamline BL1 at BESSY-II (Berlin, Germany). The space group was determined as P212121 with a solvent content of 52%, resulting in an occupancy of two molecules per asymmetric unit. The unit cell parameters were a=71.53 Å, b=73.37 Å, and c=163.315 Å. The crystals diffracted to the 1.58-Å-resolution shell. The data set was 99% complete up to 1.58 Å resolution, with an overall Rmerge of 5.6%. Multiwavelength anomalous diffraction data were collected at two wavelengths (1: 0.97973 Å; 2: 0.9799 Å) at beamline BL1 at BESSY-II. The individual data sets were processed with the programs DENZO and SCALEPACK²⁷ and merged/scaled by the CCP4 program suite.²⁸ The positions of the anomalous scatterers were located with the HySS (Hybrid Substructure Search) software of the PHENIX project.²⁹ Heavy-atom positional refinement,

Purification, crystallization and structure determination of MalE

For crystallization, the maltose-binding protein (MalE) of *S. typhimurium* was purified from the osmotic shock fluid³² of the overproducing strain MM134 (pES35)³³ using an amylose resin and complexed with acarbose by a final elution step with buffer containing 10 mM Na-citrate (pH 6.2) and 10 mM acarbose. The protein yield was routinely 25 mg/L cells. Crystals of acarbose-bound

Chapter I

MalE were obtained as described previously.¹⁵ Single crystals were soaked briefly in cryoprotectant solution [reservoir solution containing 15% (vol/vol) glycerol], mounted on a nylon loop, and then flash-cooled in liquid N₂. X-ray data were collected at beamline BL2 at BESSY-II and processed using HKL2000²⁷ and SCALEPACK. The space group was determined as C2, with unit cell parameters a=76.92 Å, b=89.23 Å, and c=64.16 Å. The crystals diffracted to 1.75 Å resolution, and the data set was 93% complete, with an overall Rmerge of 4.3%. The phase problem was solved by molecular replacement with the CCP428 program MOLREP using the 1.67-Å structure of the maltodextrin-binding protein with bound maltose (PDB ID 1ANF) as model.¹⁵ All water molecules and ligand atoms were omitted from the starting model. Subsequent cycles of isotropic B-value and positional refinement to 1.83 Å resolution were performed using REFMAC5.³⁴ The polypeptide chain and oligosaccharides were built manually using the model building program Coot.³⁰ Orientation of the amino acid side chains and bound water molecules was modeled based on an electron density of N3σ. The final R-factor for the resolution range 40–1.75 Å was 16.7% (R_{free}=20.3%). The statistics of the resulting structure are reported in Table 3. For binding assays, His6-MalE (*E. coli*) was purified by Ni-NTA chromatography from the cytosolic fraction of *E. coli* strain JM109 harboring plasmid pCB06,³⁵ followed by a dialysis step against 10 L of buffer 2 at 4 °C. Comparable with MalE of *S. typhimurium*, about 26 mg of protein per liter of cells was obtained. The MalE proteins of *E. coli* and *S. typhimurium* are 93% identical and functionally fully comparable.

Binding assays

The binding affinities of GacH and MalE for oligosaccharides were determined by high-sensitivity microcalorimetry³⁶ using a VP-ITC device (MicroCal Software, Northampton, MA, USA). To avoid air bubbles, we degassed solutions under vacuum prior to use. Freshly dialyzed GacH or MalE was filled in the reaction cell at a concentration of 50 μM in buffer 2 and titrated in 30 steps of 10 μl each against stock solutions of maltotetraose (0.4 mM) or acarbose (5 mM in case of GacH and 0.6 mM in case of MalE) at 5-min intervals. Power peaks were integrated, and the resulting reaction heats were plotted against the molar substrate/protein ratio and fitted using the ‘one-set-of-sites’ model with origin 5.0 (MicroCal Software), yielding the dissociation constant K_d. The first injection was always excluded from evaluation because it usually suffers from sample loss during the mounting of the syringe and the equilibration preceding the actual titration.

Chapter I

Acknowledgements

We thank Claudia Alings (AG Kristallographie, Freie Universität Berlin) for the initial preparation of crystals, Gerdi Hoelzl (AG Biophysik von Membranproteinen, Leibnitz-Institut für Molekulare Pharmakologie) for assistance with isothermal titration calorimetry measurements, Anne Pohlmann (AG Mikrobiologie, Humboldt Universität zu Berlin) for electrospray ionization tandem time-of-flight analysis of component 5C, and Oliver Daumke (Max-Delbrück-Zentrum für Molekulare Medizin, Berlin) for helpful discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SCHN274/12-1, WE2936/1-1, and SFB 449) and by a fellowship from the German National Academic Foundation (to A.L.).

Accession numbers

The atomic coordinates and structure factors (accession codes [3JZJ](#) (GacH_{acarbose}), [3K00](#) (GacH_{tetraose}), [3K02](#) (GacH_{component 5C}), [3K01](#) (GacH_{open conformation}), and [3JYR](#) (MalE_{acarbose})) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (www.rcsb.org).

Chapter I

References

1. Mueller, L. (1989). Chemistry, biochemistry and therapeutic potential of microbial α -glucosidase inhibitors. In *Novel Microbial Products for Medicine and Agriculture* (Demain, A. L., Somkuti, G. A., Hunter-Creva, J. C. & Rossmore, H. W., eds), pp. 109–116, Elsevier Science, Amsterdam, Netherlands.
2. Truscheit, E., Frommer, W., Junge, B., Müller, L., Schmidt, D. & Wingeder, W. (1981). Chemistry and biochemistry of bacterial alpha-glucosidase inhibitors. *Angew. Chem. Int. Ed.* 20, 744–761.
3. Bischoff, H., Ahr, H. J., Schmidt, D. & Stoltefuss, J. (1994). Acarbose—ein neues Wirkprinzip in der Diabetestherapie. *Nachr. Chem. Tech. Lab.* 42, 1119–1128.
4. Rockser, Y. & Wehmeier, U. F. (2009). The *gac*-gene cluster for the production of acarbose from *Streptomyces glaucescens* GLA.O: identification, isolation and characterization. *J. Biotechnol.* 140, 114–123.
5. Wehmeier, U. F. (2003). The biosynthesis and metabolism of acarbose in *Actinoplanes* SE50/110: a progress report. *Biocatal. Biotransform.* 21, 279–285.
6. Wehmeier, U. F. & Piepersberg, W. (2004). Molecular biology and enzymology of the metabolism of the α -glucosidase inhibitor acarbose. *Appl. Microbiol. Biotechnol.* 63, 613–625.
7. Brunkhorst, C., Andersen, C. & Schneider, E. (1999). Acarbose, a pseudooligosaccharide, is transported but not metabolized by the maltose/maltodextrin system of *Escherichia coli*. *J. Bacteriol.* 181, 2612–2619.
8. Hülsmann, A., Lurz, R., Scheffel, F. & Schneider, E. (2000). Maltose and maltodextrin transport in the thermoacidophilic Gram-positive bacterium *Alicyclobacillus acidocaldarius* is mediated by a high-affinity transport system that includes a maltose-binding protein tolerant to low pH. *J. Bacteriol.* 182, 6292–6301.
9. Schmidt, D. D., Frommer, W., Junge, B., Müller, L., Wingeder, W., Truscheit, E. & Schäfer, D. (1977). Alpha-glucosidase inhibitors. New complex oligosaccharides of microbial origin. *Naturwissenschaften*, 64, 535–536.
10. Brunkhorst, C., Wehmeier, U. F., Piepersberg, W. & Schneider, E. (2005). The *acbH* gene of *Actinoplanes* sp. encodes a solute receptor with binding activities for acarbose and longer homologs. *Res. Microbiol.* 156, 322–327.
11. Boos, W. & Lucht, J. M. (1996). Periplasmic bindingprotein-dependent ABC-transporters. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. & Umberger, H. E., eds), pp. 1175–1209, 2nd edit. American Society for Microbiology, Washington, DC.
12. Schneider, E. (2001). ABC transporters catalyzing carbohydrate uptake. *Res. Microbiol.* 152, 303–310.

Chapter I

13. Duan, X., Hall, J. A., Nikaido, H. & Quioco, F. A. (2001). Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. *J. Mol. Biol.* 306, 1115–1126.
14. Shelburne, S. A., III, Fang, H., Okorafor, N., Sumbly, P., Sitkiewicz, I. & Keith, D. (2007). MalE of group A *Streptococcus* participates in the rapid transport of maltotriose and longer maltodextrins. *J. Bacteriol.* 189, 2610–2617.
15. Spurlino, J. C., Lu, G. Y. & Quioco, F. A. (1991). The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* 266, 5202–5219.
16. Sharff, A. J., Rodseth, L. E., Spurlino, J. C. & Quioco, F. A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry*, 31, 10657–10663.
17. Quioco, F. A., Spurlino, J. C. & Rodseth, L. E. (1997). Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure*, 5, 997–1015.
18. Diez, J., Diederichs, K., Grellner, G., Horlacher, R., Boos, W. & Welte, W. (2001). The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis* at 1.85 Å. *J. Mol. Biol.* 305, 905–915.
19. Cuneo, M. J., Changela, A., Beese, L. S. & Hellinga, H. W. (2009). Structural adaptations that modulate monosaccharide, disaccharide, and trisaccharide specificities in periplasmic maltose-binding proteins. *J. Mol. Biol.* 389, 157–166.
20. Cuneo, M. J., Changela, A., Warren, J. J., Beese, L. S. & Hellinga, H. W. (2006). The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and disaccharide binding sites. *J. Mol. Biol.* 362, 259–270.
21. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
22. Brunkhorst, C. & Schneider, E. (2005). Characterization of maltose and maltotriose transport in the acarbose-producing bacterium *Actinoplanes* sp.. *Res. Microbiol.* 156, 851–857.
23. Schlösser, A., Kampers, T. & Schrepf, H. (1997). The *Streptomyces* ATP-binding component MsiK assists in cellobiose and maltose transport. *J. Bacteriol.* 179, 2092–2095.
24. Hurtubise, Y., Shareck, F., Kluepfel, D. & Morosoli, R. (1995). A cellulase/xylanase-negative mutant of *Streptomyces lividans* 1326 defective in cellobiose and xylobiose uptake is mutated in a gene encoding a protein homologous to ATP-binding proteins. *Mol. Microbiol.* 17, 367–377.

Chapter I

25. Saito, A., Fujii, T., Shinya, T., Shibuya, N., Ando, A. & Miyashita, K. (2008). The *msiK* gene, encoding the ATP-hydrolysing component of N,N'-diacetylchitobiose ABC transporters, is essential for induction of chitinase production in *Streptomyces coelicolor* A3(2). *Microbiology*, 154, 3358–3365.
26. Castanié, M. P., Bergès, H., Oreglia, J., Prère, M. F. & Fayet, O. (1997). A set of pBR322-compatible plasmids allowing the testing of chaperone-assisted folding of proteins overexpressed in *Escherichia coli*. *Anal. Biochem.* 254, 150–152.
27. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
28. Collaborative Computational Project, Number 4 (CCP4). (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D*, 50, 760–763.
29. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. et al. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D*, 58, 1948–1954.
30. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr.* 60, 2126–2132.
31. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. Sect. D*, 57, 122–133.
32. Dean, D. A., Hor, L. I., Shuman, H. A. & Nikaido, H. (1992). Interaction between maltose-binding protein and the membrane-associated maltose transporter complex in *Escherichia coli*. *Mol. Microbiol.* 6, 2033–2040.
33. Hunke, S., Dröse, S. & Schneider, E. (1995). Vanadate and bafilomycin A1 are potent inhibitors of the ATPase activity of the reconstituted bacterial ATPbinding cassette transporter for maltose (MalFGK2). *Biochem. Biophys. Res. Commun.* 216, 589–594.
34. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structure by the maximum-likelihood method. *Acta Crystallogr. Sect. D*, 53, 240–255.
35. Daus, M. L., Berendt, S., Wuttge, S. & Schneider, E. (2007). Maltose binding protein (MalE) interacts with periplasmic loops P2 and P1, respectively, of the MalFG subunits of the maltose ATP binding cassette transporter (MalFGK2) from *Escherichia coli*/*Salmonella* during the transport cycle. *Mol. Microbiol.* 66, 1107–1122.
36. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* 179, 131–137.

Chapter I

Table 1. Summary of hydrogen-bond interactions for GacH

TET/ACR/5C				Protein interface		
Residue	Atom	Residue	Atom	Distance (Å)	Distance (Å)	Distance (Å)
Ring A/ring B	O2A	Trp86	NE1		2.74	2.55
		Asn365	ND2		3.03	
	O3A	Trp86	NE1		3.18	3.24
	O4A	Trp86	NE1	3.21		
	O5A	<i>Asn365</i>	ND2	2.91		
	O6A	<i>Asn365</i>	ND2	2.57		
		<i>Asn365</i>	O	3.52		
Ring B/ring C	O2B	Arg358	NH1	2.84	2.82	3.11
			NH2	2.59	2.83	2.62
		Glu83	OE2	2.38	2.54	2.60
		Glu83	OE1	3.43	3.38	3.63
	O3B	Arg358	NH1	3.19		2.93
		Glu83	OE1	2.79	2.85	2.78
		Glu83	OE2	3.61		3.55
	O6B	<i>Asp180</i>	OD1	2.62		2.63
			OD2	2.97		3.06
Ring C/ring D	O2C	Asp133	OD2	2.63	2.66	2.66
		Trp290	NE1	2.96	2.96	2.95
	O3C	Glu83	OE2	2.66	2.69	2.62
		Gly289	N	3.10	3.21	3.22
	O6C	<i>Asp180</i>	OD2	2.87	2.72	2.76
Ring D/ring E	O1D	Thr27	O	2.74	2.78	2.74
	O2D	Glu32	OE1	2.55	2.57	2.47
			OE2	3.46	3.41	3.43
		Arg81	NH1	2.85	2.83	2.90
	3D	rp290	E1	3.32	3.32	3.36
Total				23	18	21

Chapter I

Table 2. Summary of hydrogen-bond interactions for Male

TET/ACR				Protein interface	
Residue	Atom	Residue	Atom	Distance (Å)	Distance (Å)
Ring A	O2A	Glu44	OE2	2.78	2.72
		Lys42	NZ	3.27	
	O3A	Lys42	NZ	2.50	3.56
	O4A	Glu153	OE2	3.58	
Ring B	O2B	Arg66	NE	3.45	2.85
		Arg66	NH2	3.42	3.66
	O3B	Glu44	OE2	2.50	2.71
		Arg66	NH2	2.82	2.80
		Tyr341	OH		3.22
	O6B	Glu153	OE1	3.28	
Ring C	O2C	Asp65	OD1	2.81	2.67
		Asp65	OD2	3.72	3.53
		Trp62	NE1		3.59
	O3C	Trp62	NE1	3.06	3.00
		Asp65	OD2	2.63	2.66
		Asp65	OD1	3.40	3.43
	O6C	Glu153	OE1	2.49	2.61
		Pro154	N	3.56	3.52
		Tyr155	N	3.14	3.06
Ring D	O1D	Asp14	OD2	2.62	2.70
		Lys15	NZ	3.23	3.09
	O2D	Lys15	NZ	3.02	2.77
		Glu111	OE1	2.57	2.58
			OE2	3.41	3.62
Total				22	21

Chapter I

Table 3. Data collection and refinement statistics

	GacH_{tetraose}	GacH_{acarbose}	GacH_{component 5C}	GacH_{open conformation}	MalE_{acarbose}
<i>Data collection</i>					
Space group	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 4 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>C</i> 2
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	72.17, 72.17, 160.45	72.09, 72.09, 160.37	72.05, 72.05, 160.17	35.52, 92.34, 106.92	76.92, 89.23, 64.16
α , β , γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 112.77, 90.00
Resolution (Å)	50–1.55	50–1.40	50–1.55	50–1.35	40–1.75
<i>R</i> _{sym} or <i>R</i> _{merge}	0.063 (0.401) ^a	0.052 (0.319) ^a	0.046 (0.452) ^a	0.048 (0.141) ^a	0.043 (0.120) ^a
<i>I</i> / σ <i>I</i>	17.18 (2.82) ^a	17.30 (2.15) ^a	23.10 (2.56) ^a	22.55 (8.49) ^a	33.67 (7.99) ^a
Completeness (%)	92.8 (76.9) ^a	95.2 (77.0) ^a	97.0 (83.7) ^a	93.4 (77.4) ^a	93.4 (64.6) ^a
Redundancy	4.5 (3.8) ^a	3.3 (1.6) ^a	4.1 (3.6) ^a	4.0 (2.7) ^a	3.6 (2.3) ^a
<i>Refinement</i>					
Resolution (Å)	65.08–1.55	65.80–1.40	65.80–1.55	69.84–1.35	59.13–1.75
Number of reflections	57,006	77,948	57,005	69,601	35,763
<i>R</i> _{work} / <i>R</i> _{free}	18.8/21.9	20.1/22.7	20.5/24.3	15.5/18.0	16.7/20.3
Number of atoms					
Protein	2933	2964	2933	3007	2881
Ligand	45	44	55	NA	44
Water	573	475	415	629	427
<i>B</i> -factors					
Protein	22.5	21.9	24.8	10.9	13.3
Ligand	23.4	17.7	28.3	NA	14.9
Water	37.0	34.3	34.1	27.7	23.9
RMSD					
Bond lengths (Å)	0.011	0.018	0.012	0.011	0.012
Bond angles (°)	1.306	1.732	1.417	1.425	1.281

Chapter I

Figure legends

Fig. 1. Structures of acarbose and higher homologs used in this study (b), and the atomic numbering of glucose (a). Ac, acarviosyl. (c) Acarbose carbophor cycle in *S. glaucescens* GLA.O. Acarbose is intracellularly synthesized as acarbose-7-phosphate by the glycosyltransferase GacI and exported by the GacWXY ABC export system. In the natural environment, the dephosphorylated acarbose acts as an inhibitor of the α -glucosidases of competitors. Increasing amounts of free unbound acarbose can then be used as acceptor molecules for the uptake of glucose or oligosaccharides resulting from starch degradation, which are possibly transferred to acarbose by the action of the α -amylases GacE2, Z1, and Z2, or by a yet to be identified acarviosyltransferase. The resulting longer homologs are transported into the cytoplasm by the GacHFG importer in complex with a yet to be identified ATPase subunit (MsiK). Glucose units will then be released by the action of GacQ (amylomaltase) and shuttled into metabolism. Acarbose is rephosphorylated by the acarbose-7-kinase GacK and reexported to start a new ‘carbophor cycle’ (modified from Rockser and Wehmeier⁴).

Fig. 2. Oligosaccharide affinities determined by isothermal titration calorimetry. GacH (a–d) and MalE (e and f) were titrated with maltose (diamonds), maltotetraose (circles), component 5C (squares), or acarbose (triangles). Integrated and normalized heats of reaction Q were plotted against the sugar/protein molar ratio R and fitted using a one-site binding model. (g) SDS-PAGE analysis of purified GacH and MalE (*E.c.*). GacH (lane 1) and His₆-MalE (*E.c.*) (lane 3) were purified as described in Materials and Methods. Lane 2, molecular mass marker (kDa).

Fig. 3. Overlay of GacH_{acarbose} with GacH_{open conformation} showing the overall secondary structure features. The structures are overlaid at domain 1 (shown in gray). The large lobe movement upon ligand binding can be observed at domain 2 (shown in yellow for GacH_{acarbose} and light yellow for GacH_{open conformation}). The hinge region is shown in light brown. Acarbose (blue sticks) is shown in the binding pocket. Drawn with Pymol.³⁵

Fig. 4. Binding mode of acarbose to GacH. Oxygen atoms are shaded in red, and nitrogen atoms are shaded in blue. Hydrogen-bond interactions with atomic distances of less than 3.5 Å are shown by broken lines. The aromatic residues (in ball shape) form a hydrophobic keyhole around

Chapter I

acarbose, with Phe59 located in the center of the binding pocket. The other three aromatic amino acids Phe368, Tyr182, and Trp254 belong to domain II of GacH. Acarbose is shown in turquoise. The individual sugar moieties are indicated by boxes (boxes A and B, acarviosyl; boxes C and D, maltose).

Fig. 5. Overlay of GacH_{acarbose} (cyan sticks) with a) GacH_{maltotetraose} (yellow sticks) and b) GacH_{component5C} (salmon sticks). All three oligosaccharides adopt a similar orientation in the binding pocket of GacH. Differences can be observed at sugar ring A (acarviosyl) resulting in a shift in the position of Asn-365 at the base of the binding pocket. An additional hydrogen bond when compared to acarbose binding is shown in red (Asp-180).

Fig. 6. Overlay of GacH_{acarbose} (cyan sticks) and MalE_{acarbose} (yellow sticks) showing that acarbose in GacH is buried deeper into the binding pocket by exactly one glucose ring shift. The indicated amino acids are from GacH.

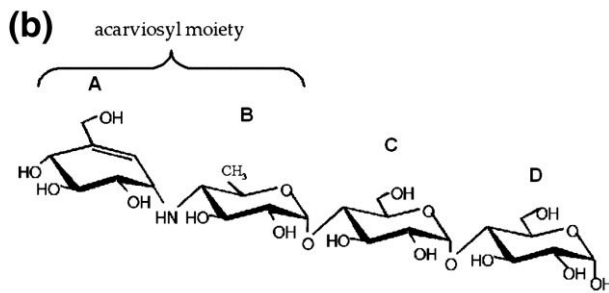
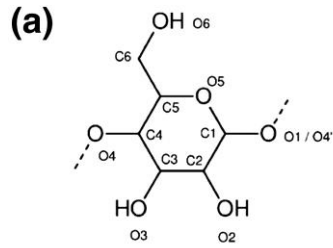
Fig. 7. Amino acid sequence alignment of ecMBP, ttMtBP, tIMBP, ttGBP, and *S. glaucescens* GLA.O acarbose/maltose/maltodextrin-binding protein (sgGacH). The alignment generated from the five carbohydrate-binding proteins—ecMBP_{maltotriose} (3MBP), ttMtBP_{maltotriose} (2GH9), tIMBP_{trehalose} (1EU8), ttGBP_{glucose} (2B3B), and sgGacH (3JZJ)—was performed with CLUSTAL 2.0.10 multiple sequence alignment. Amino acids forming the individual ligand-binding sites are color coded (site A, cyan; site B, red; site C, orange; site D, green). Additionally, amino acids found in two of the sites are boxed using the same color scheme of the one-letter code in one site and a colored box in the other site (modified from Cuneo et al.¹⁹).

Fig. 8. Overlay of MalE_{acarbose} (salmon and white sticks), MalE_{maltotetraose} (cyan sticks), and MalE_{maltose} (yellow sticks) showing substantial movements of the amino acid side chains Tyr341 and Arg344 to accommodate the binding of acarbose and maltotetraose.

Chapter I

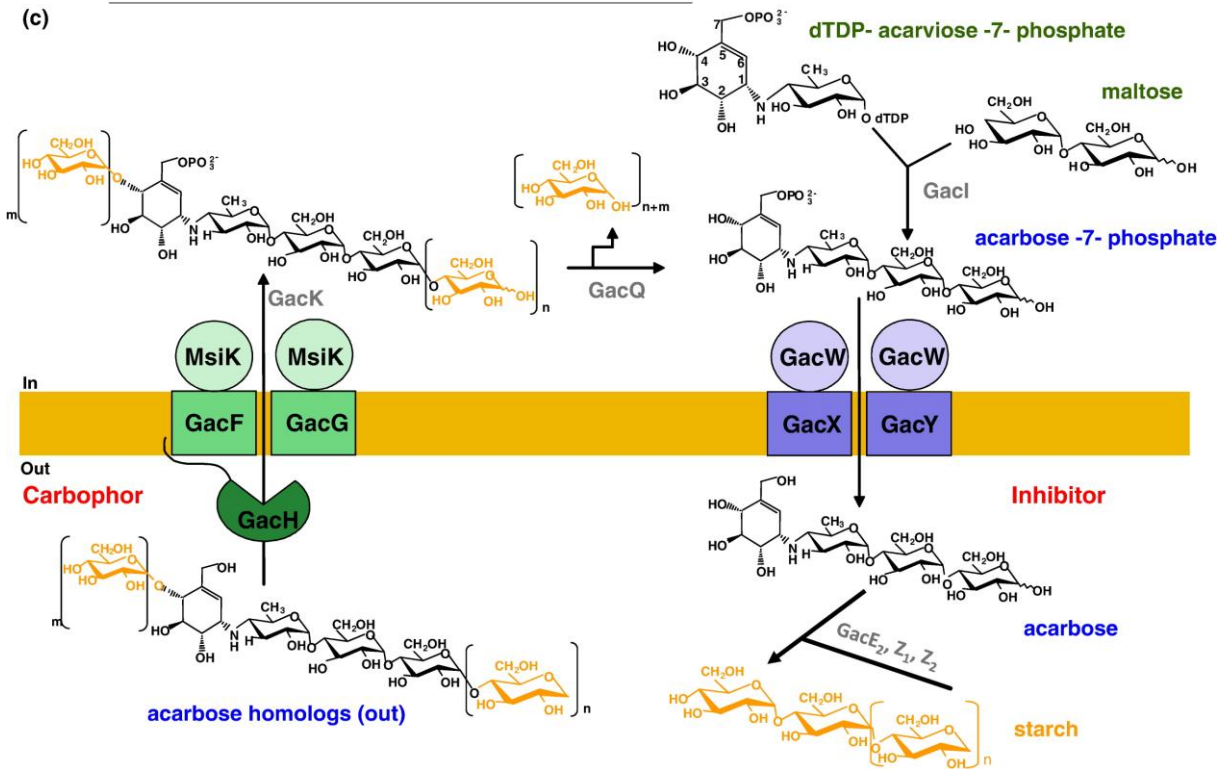
Figures

Fig. 1



Name	Structure
acarbose (component 3)	Ac-1,4-maltose
component 5C	Ac-1,4-maltose-1,4-glucose-1,1-glucose
component 6AB (mixture)	Ac-1,4-maltose-1,4-glucose-1,4-fructose and Ac-1,4-maltose-1,4-glucose-1,4-glucose

(c)



Chapter I

Fig. 2

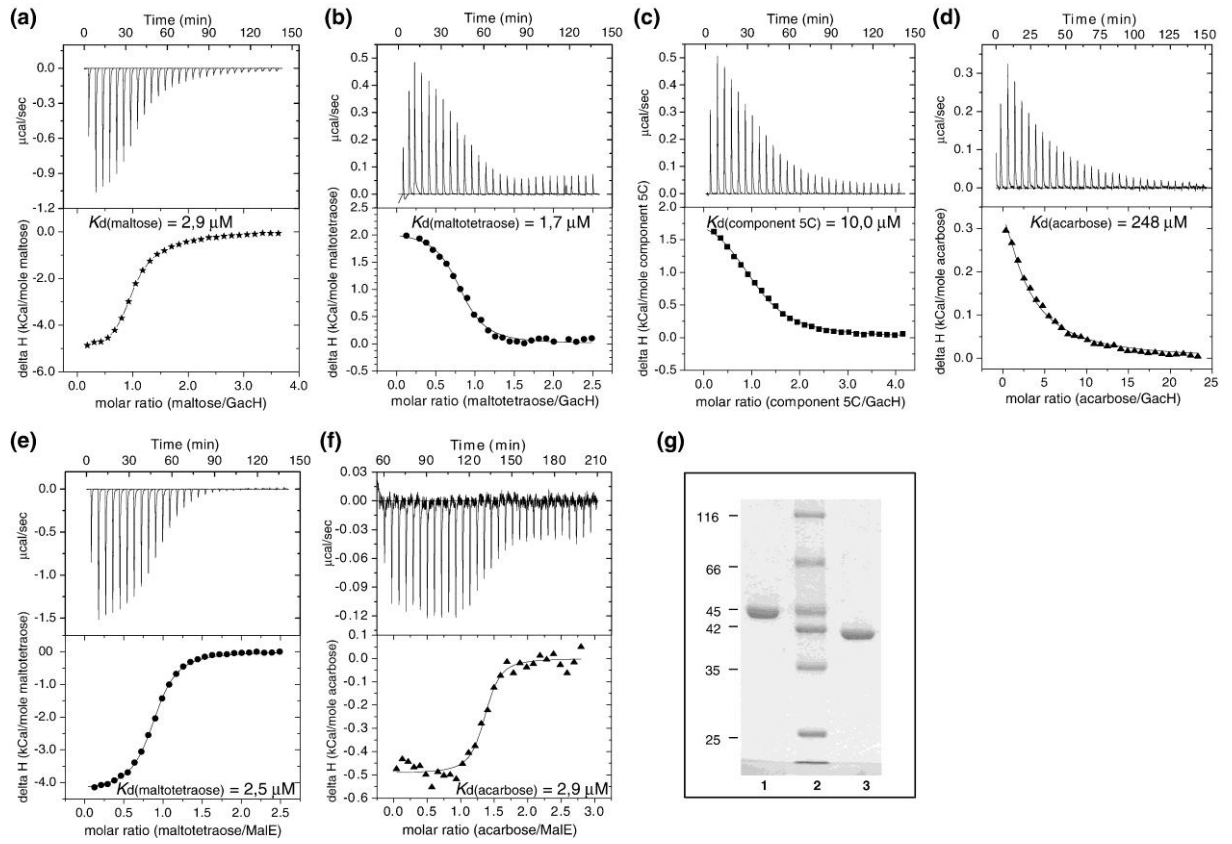


Fig.3

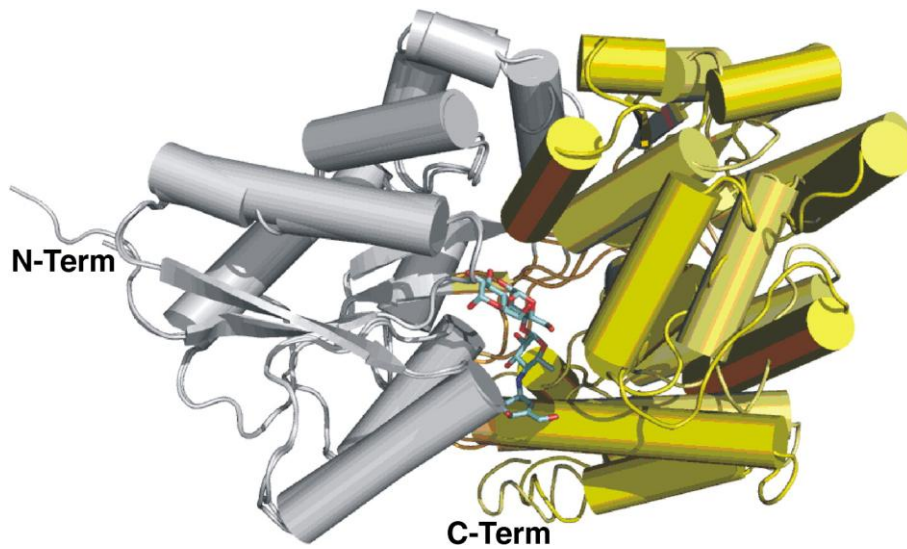


Fig. 4

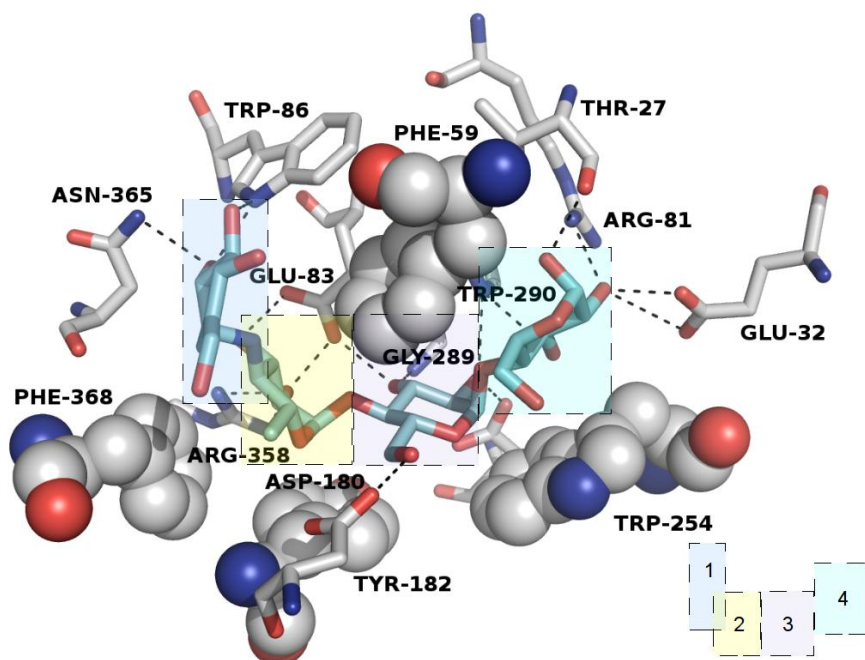


Fig.5

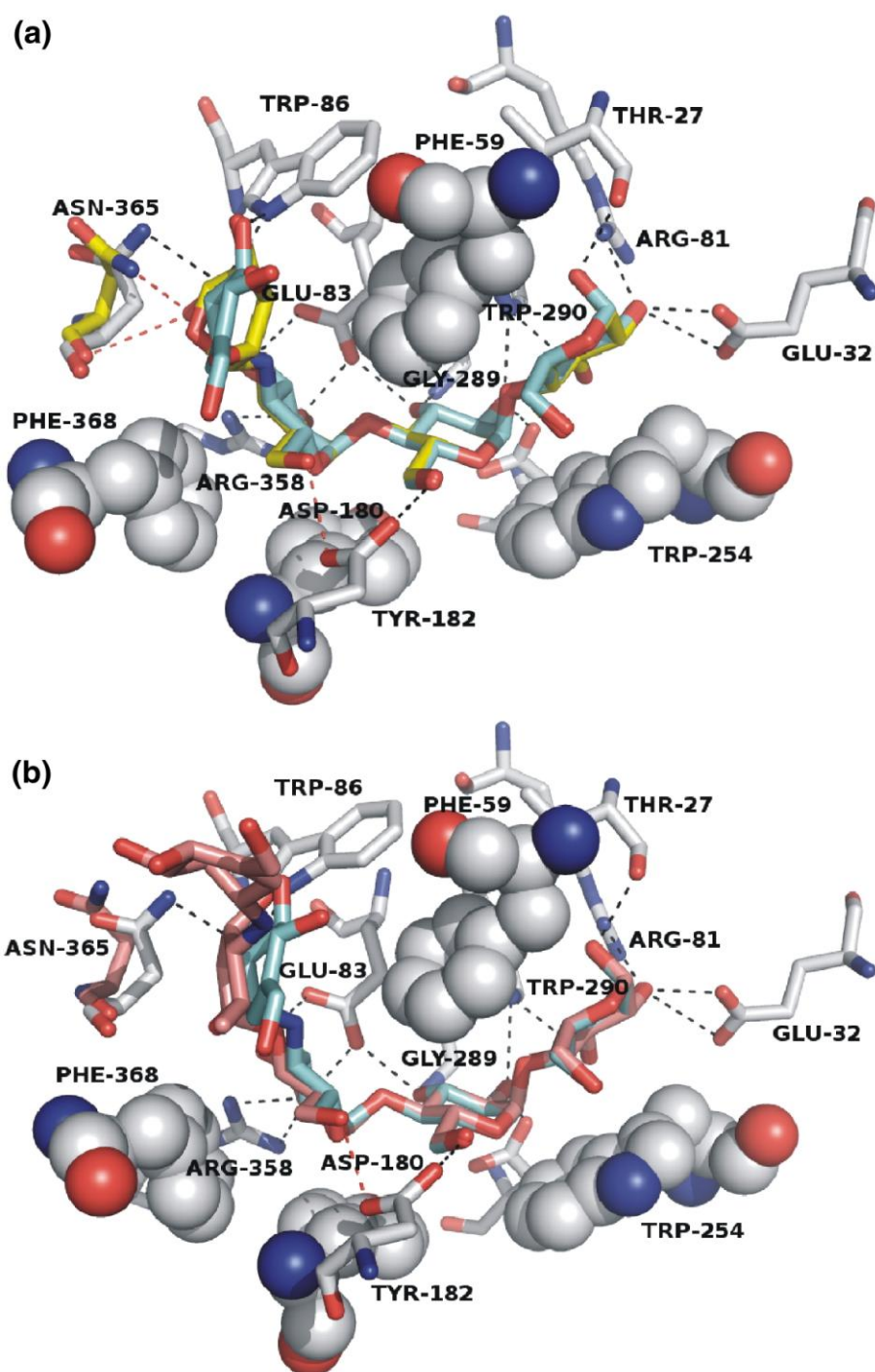
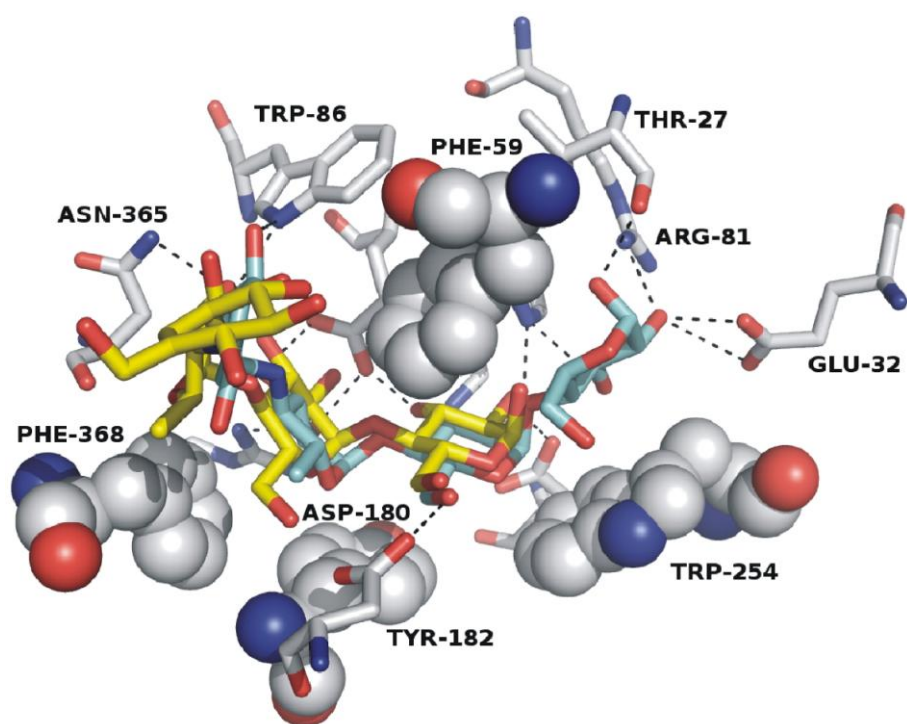


Fig. 6

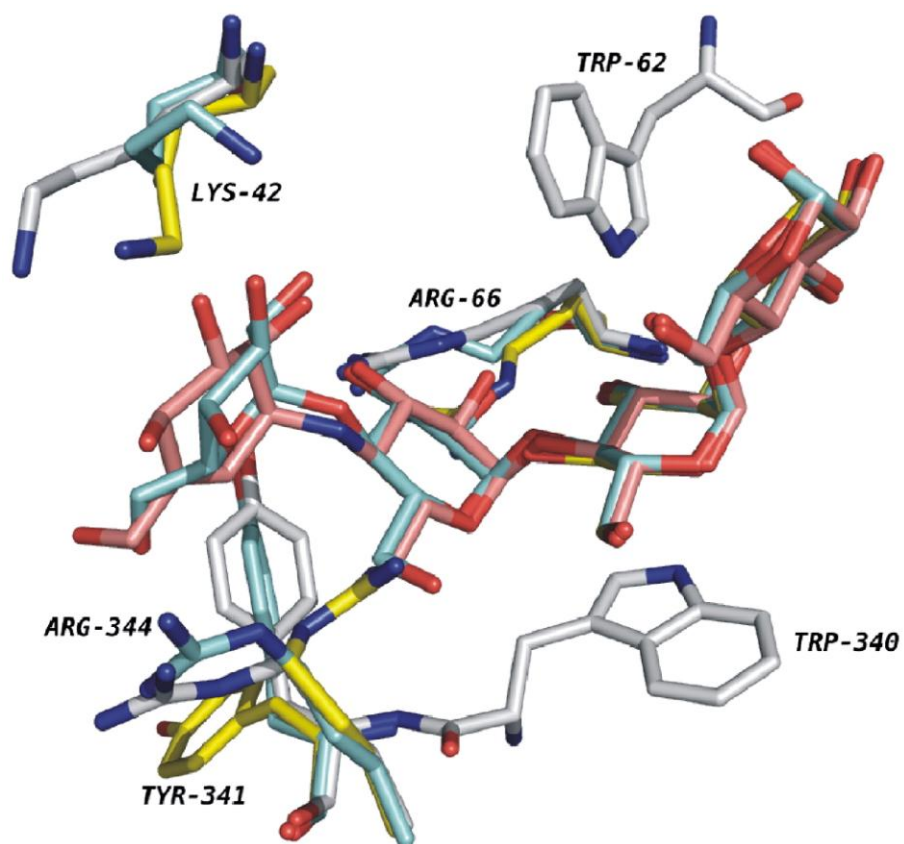


Chapter I

Fig. 7

ecMBP	6	KLVIWINGDK-GYNGLAEVGKKFEKDTGIKVTVEHPD-----KLEEKFPQVAATGDGPD	58
ttMtBP	2	KITVWTHFGGPELEWLKEQARTFERTSGTKVEVVEVP---FAEIKQKFILGAPQQAAD	57
sgGacH	21	TVTTFWDTSNEAEKATYQALAEGFEKEHPKVDVKYVNV--PFGEANAKFKNAAGGNSGAPD	78
tlMBP	6	KIVFAVGGAPNEIEYWKGVIAEFEKKYPGVTVLKRQATDTTEQRRLDLVNALRGKSSDPD	65
ttGBP	2	KLEIFSWWAGDEGPALEALIRLYKQKYPGVEVINATVTGGAGVGNARAVLKTRMLGGDPPD	61
ecMBP		II-FWAHDRFG-GYAQSGLLAEIT--P--DKAFQDKLYPFTWD-AVRYNGKLIAYPIAVE	111
ttMtBP		LVVTVPHDWVG-EMAQAGVLEPVGKYV--TQAYLADLQGVAVE-AFTFGGRLMGLPAFAE	113
sgGacH		VMRTQ-VAWVA-DFASIGYLAPLDGTPALDDGSD--HLPQAAASTRYEGKTYAVPQVID	133
tlMBP		VF-LMDVAWLG-QFIASGWLEPLDDYVQKDNLDL SVFFQSVINLADKQGGKLYALPVYID	123
ttGBP		TFQVHAGMELIGTWVVANRMEDLSALF-RQEGWLQAFPKGLID-LISYKGGIWSVPVNIH	119
ecMBP		AL-SLIYNKDLLP-----NPPKTWEEIPALDKELKA-----KGKSALMFNLQ----EPY	155
ttMtBP		SV-ALIYNKKYVK-----EPPRTWEEFLALAQKLT-----GATFGFLYNIG----DPY	157
sgGacH		TL-ALFYNKELLTKAGV-EVPGSVAELKTAABEITEKT----GATGLYLGRD-----DPY	182
tlMBP		AG-LLYRKDLLEKYGYSKPPETWQELVEMAQKIQSGERETNPNFVGFWQKQYEGGLVC	182
ttGBP		RSNVMWYLPKALKKEWGV-NPPRTWDEFLATCQTLKQKGL--E---APLALGENWT--QQH	171
ecMBP		FTWPLIAADG----G-YAFKYE-NGKYDIKDVGVNDAGAKAGLTFVLVDLIK-N-KHMNADT	208
ttMtBP		FNFGFFKAFG----AENVFAKDAKGNLDPTKLLIGGEVGEKALQFIKDLRFKYNLVPEGV	213
sgGacH		WFLPYLYGEG----GDLVDEKNKT-----VTVDDEAGVRAVRVIKDLVDSKAAITDAS	231
tlMBP		DFVEYVYSNG----GS-LGEFKD----GKWVPTLNKPENVEALQFMVDLIHKYKISPPNT	233
ttGBP		LWESVALAVLGPDDWNNLWN-----GKLFKTPKAVRAWEVFGFVRL---DCANKDA	219
ecMBP		DY----SIAEAAFNKGETAMTINGPWAWSNIDTSKV-----NYGVTVLPTF--KGQPSKP	257
ttMtBP		DY----GVADGAFKDGALAMILNGPQALGDYKKAKV-----DFGIAPFPVPPGAKNPWGP	264
sgGacH		DG--WNNMQNAFKSGKVMAMVNGPWAIEDVKAGARFKDAGNLGVAPVPAGS-AGQGS-P	286
tlMBP		YTEMTEEPVRLMFQQGNAAFERNWPYAWGLHNADDSPVKG-KVGVAPLPHFP--GHKSAA	290
ttGBP		AGLSW-QQAVDRVVQGAAFNVMGDWAAGYMTTTLKLPKGTDFAWAPSPG--TQG--VF	273
ecMBP		FVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYYEELAKD--P	315
ttMtBP		FLGVQGVVNVNAYSKNKTQAVNFAK-TLVTGRNLVAFNQAGG-RIPVSKSAVKQLEKD--P	320
sgGacH		Q-GGWNLSVYAGSKNLDASAYAFVK-YMSSAKVQQQTTEKLSLLPRTSVYEVPSVAD--N	342
tlMBP		TLGGWHIGISKYSDNKALAWEFVK-FVESYSVQKGFAMNLGWNPRVDVYDDPAVVS KSP	349
ttGBP		MMLSDFSFLPKGAKNRQNAINWLR-LVGSKEGQDTFNPLKG-SIAARLSDPSK-YN-AY	329
ecMBP		RIATMENA-QKGEI-----MPNIPQMSAFWYAVRTAVINAASG-RQTVDEALKDAQTRI	368
ttMtBP		VVAGFSKVF-PLGAP-----MPNIPEMGKVGWGPWGNAISLAIQRPDSNVKKIVEDMVAEI	374
sgGacH		EMVKKFKPAVDKAVERPWIAEGNALF--EPIRLQMANVLSGETSPD-EAAANTGDAYRKL	399
tlMBP		HLKELRAVF-ENAVP-----RPIVPYYPQLSEIIQKYVNSALAG-KISPQEALDKAQKEA	402
ttGBP		GQSAMRDWRSNRIVGSLVHGAV---APESFMSQFGTVMEIFLQT--RNPQAAANAQAIA	384
ecMBP		TK----- 396	
ttMtBP		KKAIGR----- 398	Site A
sgGacH		LKDYK----- 424	Site B
tlMBP		EELVKQYS-- 450	Site C
ttGBP		DQVGLGRLGQ 414	Site D

Fig. 8



10. Chapter II

Crystal structures of the bacterial solute receptor AcbH displaying an exclusive substrate preference for β -D-galactopyranose

Anke Licht^{1§}, Haydar Bulut^{2§}, Frank Scheffel¹, Oliver Daumke³, Udo F. Wehmeier⁴,
Wolfram Saenger², Erwin Schneider^{1*} and Ardeschir Vahedi-Faridi^{2*}

¹ Institut für Biologie, AG Bakterienphysiologie, Humboldt Universität zu Berlin,
Chausseestr. 117, D-10115 Berlin, Germany

² Institut für Chemie und Biochemie/Kristallographie, Freie Universität Berlin,
Takustr. 6, D-14195 Berlin, Germany

³ Max-Delbrück-Zentrum für Molekulare Medizin, Kristallographie, Robert-Roessle-Strasse
10, D-13125 Berlin, Germany

⁴ Lehrstuhl für Sportmedizin, Bergische Universität Wuppertal, Pauluskirchstr. 7, D-42285
Wuppertal, Germany

* Corresponding authors

avahedi@odu.edu; Phone: 001-757-683-3409; Fax: 001-757-683-4628

erwin.schneider@rz.hu-berlin.de; Phone: +49(0)30-2093-8121; Fax: +49(0)30-2093-8126

§ These authors contributed equally to this work.

Abbreviations used: ABC, ATP-binding cassette; SBP, solute-binding protein; GBP, glucose/galactose-binding protein; MBP, maltose/maltodextrin-binding protein; GacH, *Streptomyces glaucescens* maltose/maltodextrin-binding protein; ttGBP, *Thermus thermophilus* glucose/galactose binding protein; tlMBP, *Thermococcus litoralis* trehalose/maltose-binding protein; ecMBP, *E. coli* maltose/maltodextrin-binding protein; acarbose, acarviosyl-1,4-maltose; component 5C, acarviosyl-1,4-maltose-1,4-glucose-1,1- glucose; component mixture 6AB, acarviosyl-1,4-maltose-1,4-glucose-1,4-fructose and acarviosyl-1,4-maltose-1,4-glucose-1,4-glucose; PMSF, phenylmethylsulfonyl fluoride; IPTG, isopropyl thio- β -galactoside

Chapter II

Abstract

Solute receptors or binding proteins are indispensable components of canonical ATP binding cassette importers in prokaryotes. Here, we report on the characterization and crystal structures in the closed and open conformations of AcbH, the solute receptor of the putative carbohydrate transporter AcbFG which is encoded in the acarbose (acarviosyl-1,4-maltose) biosynthetic gene cluster from *Actinoplanes* sp. SE50/110. Binding assays identified AcbH as a high affinity monosaccharide-binding protein with a dissociation constant (K_d) for β -D-galactopyranose of 9.8 ± 1.0 nM. Neither galactose-containing di- and trisaccharides, such as lactose and raffinose, nor monosaccharides including D-galacturonic acid, L-arabinose, Dxylose and L-rhamnose competed with ^{14}C -galactose for binding to AcbH. Moreover, AcbH does not bind D-glucose which is a common property of all but one D-galactose-binding proteins characterized to date. Strikingly, determination of the X-ray structure revealed that AcbH is structurally homologous to maltose binding proteins rather than to glucose binding proteins. In the substrate binding pocket, two helices are inserted which reduce the cavity size and allow the exclusive binding of monosaccharides, specifically β -D-galactopyranose, in the ${}^4\text{C}_1$ conformation. Site directed mutagenesis of three residues from the binding pocket (Arg82, Asp361, Arg362) which interact with the axially oriented O4-H hydroxyl of the bound galactopyranose and subsequent functional analyses indicate that these residues are crucial for galactose binding. To our knowledge, this is the first report on the tertiary structure of a solute receptor with exclusive affinity for β -D-galactopyranose. The putative role of a galactose import system in the context of acarbose metabolism in *Actinoplanes* sp. is discussed.

Keywords: ABC transporter; D-galactose-binding protein; AcbH; protein crystallography; *Actinoplanes* sp. SE50/110

Chapter II

Introduction

Extracellular solute binding proteins (SBPs) or receptors are essential constituents of canonical prokaryotic ATP-binding cassette (ABC) importers.^{1,2} SBPs capture a broad range of ligands with high affinities (nM to low μ M range), such as inorganic ions, amino acids, short peptides, carbohydrates or vitamins and deliver them to the cognate transport complex.³ Moreover, intimate interaction of the SBP with extracellular peptide regions of the transporter is a prerequisite for initiating the translocation process.¹

In Gram-negative bacteria, SBPs diffuse freely in the periplasm. Gram-positive bacteria and archaea anchor the binding proteins to the outer surface of the cell membrane via an N-terminal lipid moiety⁴ or, as observed for some archaea, an N-terminal hydrophobic helix. Furthermore, solute-binding domains can also be fused to the transmembrane domains of an ABC transporter, resulting in two or four substrate-binding sites.^{5,6}

Although SBPs exhibit low sequence similarities, their overall structural fold is highly conserved.² They consist of two globular α/β domains or lobes containing a central β -sheet flanked by α -helices. The two lobes are connected by a hinge region with the substrate binding pocket located in a cleft at the interface in between. Upon ligand binding, SBPs undergo a hinge-bending motion, consisting of a rotation and twist of the lobes, resulting in a transition from an open to a closed conformation.^{7,8} Based on the topography of their globular domains and the number of interdomain connections, three structural SBP classes have been proposed.^{3,9} In class I and II the lobes differ in the number and order of β -strands. Additionally, the hinge region is formed by three connecting strands in class I and two connecting strands in class II, whereas in class III SBPs the two domains are connected by a single long helix.

Following the classification of Fukami-Kobayashi and colleagues⁹, class I SBPs comprise mainly monosaccharide binding proteins like the glucose/galactose binding protein (GBP) of *Escherichia coli*^{10,11}, whereas *E. coli* maltose binding protein (MBP), a di-/oligosaccharide binding protein^{12,13}, is the best characterized member in class II. This categorization is consistent with the two defined groups of the carbohydrate uptake transporters (CUT 1 and 2) within the ABC transporter super family.¹⁴ However, an exception is represented by the glucose/galactose binding protein of *Thermus thermophilus* (ttGBP), which is structurally homologous to MBPs rather than to GBPs.¹⁵ The size of the ttGBP binding pocket is partially restricted by structural adaptations, such as the insertion of two loops and one α -helix, resulting in the binding of only monosaccharides.

Chapter II

In a new classification system based on the features of three-dimensional structures of more than 120 SBPs, ttGBP, the maltose/maltodextrin binding proteins (MBPs) of *E. coli* and *Thermoactinomyces vulgaris*, as well as the oligogalacturonide binding protein of *Yersinia enterocolitica* (TogB) are grouped into the same cluster.² They all contain a hinge region of two short strands, 4-5 amino acids long, possess comparable molecular masses (> 40 kDa), and seem to have one small additional subdomain.

Recently, we have solved the crystal structure of GacH, the SBP of the putative oligosaccharide ABC transporter GacFGMsiK2.¹⁶ The transporter is encoded within the acarbose biosynthesis (*gac*)-gene cluster of the soil bacterium *Streptomyces glaucescens* GLA.O.¹⁷ The secondary metabolite acarbose (acarviosyl-1,4-maltose) is a pseudomaltotetraose, consisting of a maltose moiety and the pseudodisaccharide acarviosine.¹⁸ Acarbose is an inhibitor of α -glucosidases¹⁹. It is used in the treatment of patients suffering from diabetes type II.²⁰ In their natural environment, bacteria which secrete acarbose are thought to benefit from the compound's inhibitory action on α -glucosidases from competing organisms. Consistently, the acarbose producer itself is equipped with acarboresistant α -glucosidases.¹⁸ Recently, an additional function of acarbose as 'carbophor' was proposed.¹⁸ Briefly, the carbophor function could be achieved by the enzymatic transfer of glucose or oligosaccharides to acarbose by secreted transferases as well as the re-uptake of the longer derivatives into the cytoplasm by an ABC import system, resulting in a net gain of carbon and energy. Our results on GacH provided first biochemical and structural evidence in support of this hypothesis.¹⁶ We identified GacH as a maltose/maltodextrin binding protein with considerable binding activity for longer acarbose homologs, whereas acarbose is bound with only low affinity, most likely for recycling purposes.

In *Actinoplanes* sp. SE50/110, the bacterium used for industrial production of acarbose, the biosynthesis gene cluster (*acb*) also contains an operon encoding a putative carbohydrate ABC importer, AcbHFG, based on BLAST searches.¹⁸ Despite its rather low identity in primary structure to GacHFG (22-24 %), the transporter is thought to have a similar function. In fact, the solute-binding protein, AcbH, when produced as inclusion bodies in *E. coli* and subsequently refolded was shown to bind acarbose and longer derivatives, albeit with low (millimolar) affinities but neither maltose nor maltodextrins.²¹ These discrepancies prompted us to determine the crystal structure of AcbH. To this end, we established a protocol for overproduction of AcbH in *E. coli* as soluble protein which also allowed re-evaluation of its substrate specificity in a native-like state. Surprisingly, although displaying structural homology to MBP, AcbH was identified as a

Chapter II

monosaccharide binding protein with an exclusive preference for D-galactose. To our knowledge, this is the first report on the tertiary structure of a substrate binding protein with exclusive affinity for β -D-galactopyranose.

Results and Discussion

Substrate specificity of AcbH

Sequence homology searches using the transporter classification database (TCDBBLAST at <http://www.tcdb.org/>)²² indicated that AcbH is the solute binding component of an ABC import system belonging to the CUT1 family, the members of which predominantly transport di- and oligosaccharides.¹⁴ Indeed, the closest homologues of AcbH are α -glucoside disaccharide receptors like the sucrose binding protein AgIE of *Sinorhizobium meliloti*,²³ the trehalose/maltose binding protein (tlMBP) of *Thermococcus litoralis*,²⁴ and the maltodextrin binding protein MdxE of *Pyrococcus furiosus*.²⁵

Strikingly however, AcbH also displays significant similarity to the monosaccharide binding protein of *Thermus thermophilus* for the uptake of glucose and galactose (ttGBP).¹⁵ Previous studies of AcbH refolded from inclusion bodies revealed no affinities for maltose or maltodextrins but very low affinities for acarbose and its homologs 5C and 6AB.²¹

Thus, after having established a protocol to purify soluble ('native-like') AcbH from the cytosolic fraction of the *E. coli* host (see 'Materials and Methods'), we set out to reevaluate the properties of the protein. To avoid a possible prebound ligand, AcbH was purified from minimal medium supplemented with glycerol as carbon source, which resulted in a ligand-free protein, as judged by thin-layer chromatography (data not shown). First, we investigated binding of glucose, galactose, maltose, maltotetraose, acarbose and component 5C (acarviosyl-1,4-maltose-1,4-glucose-1,1-glucose) by isothermal titration calorimetry. Surprisingly, AcbH showed a considerably high affinity for galactose ($K_d = 9.8 \pm 1.0$ nM; **Figure 1a**) but, unlike other bacterial galactose binding proteins²⁶, displayed no binding activity towards glucose (**Figure 1b**). In contrast to previous findings using a refolded protein, native-like AcbH neither bound maltose, maltotetraose nor acarbose and its derivative 5C (data not shown).

Binding constants in the low nanomolar range have been reported previously in case of receptors for amino acids, vitamin B12 or Zn²⁺ ions but not for sugar binding proteins.² Thus, we confirmed our data by a filter binding assay using radiolabelled D-galactose. As shown in **Figure 1c**, a K_d of 10.6 ± 1.2 nM was obtained which is in perfect agreement with the ITC result.

Chapter II

Furthermore, competition studies with different unlabelled mono-, di- and oligosaccharides demonstrated that neither glucose, maltose and maltooligosaccharides nor acarbose and its derivatives 5C and 6AB (acarviosyl-1,4-maltose-1,4-glucose-1,4-fructose and acarviosyl-1,4-maltose-1,4-glucose-1,4-glucose) (at a 100-fold molar excess each) caused a significant decrease in D-[14C]galactose binding (Figure 2). Moreover, the galactosecontaining di- and trisaccharides lactose and raffinose, respectively, or various monosaccharides structurally related to galactose, such as D-galacturonic acid, L-arabinose, D-xylose and L-rhamnose, all known to bind to other monosaccharide binding proteins^{10,27}, failed to compete with D-[14C]galactose binding. Together, these data demonstrate that AcbH is a monosaccharide binding protein with a unique substrate preference for D-galactose only.

To date, the YtfQ binding protein of *E. coli* is the only other known example of a galactose binding protein that lacks affinity for glucose.²⁸ Besides binding of L-arabinose and galactopyranose, YtfQ exhibits a 10-fold higher affinity of approximately 0.13 μ M for galactose in its less abundant furanose form as a distinctive feature. In contrast to YtfQ, AcbH displays very high affinity binding towards commercially available D-galactose, which is predominantly present in the pyranose form in solution and has only a small proportion of around 8 % in the furanose form.²⁹ This suggests that the more abundant galactopyranose is the physiological ligand of AcbH as known from all other bacterial glucose-/galactose binding proteins.

Overall AcbH structure

To elucidate the molecular mechanism by which AcbH, in contrast to other Dgalactose-binding proteins, can discriminate between D-galactose and D-glucose, we solved the X-ray crystal structure of AcbH in complex with D-galactose to 2.15 Å resolution, using a selenomethionine substituted derivative. Thus far, the crystal structures of several carbohydrate-binding proteins displaying specificities for different mono-, di- or oligosaccharides have been determined that are represented by the glucose/galactose-binding proteins from *E. coli* (ecGBP) and *T. thermophilus* (ttGBP), and maltose and/or maltodextrinbinding proteins from *T. litoralis* (tlMBP), *S. glaucescens* (GacH) and *E. coli* (ecMBP). These SBPs share affinities for multiple structurally related substrates and the typical fold consisting of two α/β domains joined by a hinge region flanking the central sugar-binding pocket. The overall AcbH structure shows a similar fold that is characteristic of class II SBPs combined in cluster D-I as defined by Berntsson and co-workers.² The N- and C-terminal domains contain central β -sheets of five and three β -strands each with β -strand order 21347 and 546

Chapter II

(Figures 3a and 3c), that are flanked by α -helices, consistent with the class II SBP topography. The two domains are connected by a tripartite hinge region of two short loops (Hinge I: residues 128-131; Hinge II: residues 287-291) and an additional long loop of 17 amino acids (residues 324-341, Hinge III) (Figure 3). Moreover, the C-terminal lobe is split into a larger C-terminal domain and a small N-terminal domain as a distinct feature of cluster D-I SBPs and described for ecMBP.¹²

Because of the typical class II SBP fold, AcbH has higher structural similarity with diand oligosaccharide-binding proteins such as the raffinose-binding protein from *Streptococcus pneumoniae* (PDB code 2I58), the trehalose/maltose-binding protein (tMBP) from *T.litoralis*²⁴ or the maltose/maltodextrin-binding protein GacH from *S. glaucescens*¹⁶, as identified by a DALI search.³⁰ However, the glucose-binding protein (ttGBP) from *T. thermophilus* exhibits the highest structural identity of 39 % to AcbH, and the root mean square deviation (RMSD) is 2.0 Å. Although ttGBP binds only monosaccharides such as Dglucose and D-galactose, it is structurally homologous to MBPs rather than to GBPs. The superposition of these structures shows that the positions of the secondary-structure elements remain conserved, while the ttGBP substrate binding pocket is spatially reduced compared with that of the homologous tMBP by insertion of two helices (α 2 and α 18) that only allow the binding of monosaccharides.¹⁵

AcbH substrate binding pocket

Although AcbH displays a type II SBP fold and was extensively dialyzed after purification from the cytosolic fraction of the *E. coli* host cells grown in LB medium before crystallization, the electron density shows AcbH with a bound D-galactose. This reflects the extremely high affinity of AcbH to this solute as also determined by ITC and a filtration assay (Figure 1). However, in the selenomethionine-substituted crystals glycerol was found in the substrate binding pocket at positions O4-C4-C5(O5)-C6-O6 of the bound D-galactose in the wild type AcbH structure (Figure 4). Although the cells overproducing selenomethioninesubstituted AcbH were grown with glycerol as source of carbon and energy, glycerol is no substrate of the protein as shown in Figure 2. Thus, we speculate that glycerol has leaked into the crystals from the cryoprotection medium in which they were soaked prior to flash freezing (see ‘Materials and Methods’).

In the native AcbH binding pocket β -D-galactose is bound in the pyranose 4C1 chair conformation, and all the β -D-galactose hydroxyls participate in hydrogen bonds as donors or acceptors (Table 1). Altogether eight amino acids of AcbH and two water molecules form a total of 17 hydrogen bonds with the galactose. His-56, Gln-60, Arg-82, Tyr-129, Asp-361 and Arg-362

Chapter II

form bidentate and tridentate (Gln-60) hydrogen bonds with hydroxyl groups and the ring-O5 of the pyranose while Trp-131 and Gln-255 are engaged in single hydrogen bonds (Table 1, Figure 4). In addition, the indole side chain of Trp-177 stacks against the non-polar portion of the galactose on one side and gives rise to non-polar van der Waals contacts between AcbH and the galactose. Phe-78, Phe-359 and Pro-288 are also involved in van der Waals contacts with the edge of the non-polar portion of galactose. Similar modes of sugar binding are seen in the *E. coli* glucose/galactose binding protein belonging to class I SBPs.¹¹

The β -galactopyranose is bound almost exclusively by the two globular domains, and two residues from the hinges also contribute to the interaction. While Tyr-129 (Hinge I) forms a hydrogen bond with galactose O5, Pro-288 (Hinge II) closes a gap in the binding pocket. The surface presentation of the binding pocket (Figure 5a) reveals a cage-like structure, which is open on the C1 side of the sugar ring, and therefore this part of the binding pocket remains accessible to the solvent. This gate is followed by a small cavity, which is formed between the two domains. Compared with GacH, the AcbH substrate binding cavity is filled by helix 18 (residues 360-363) and partially by helix 2 (residues 56-68), thereby preventing the binding of longer carbohydrates as it was observed for ttGBP (Figure 6).

Substrate specificity of AcbH variants

The unique preference of AcbH for galactose prompted us to investigate the effect of amino acid substitutions in the binding pocket on the substrate specificity of AcbH. To this end, we chose Arg-82, Asp-361, and Arg-362, respectively, for mutagenesis. These residues were found to be implicated in the interaction with the axially oriented O4-H hydroxyl of β -D-galactose (Figure 5b). A similar mode of sugar binding was seen in class II ttGBP. Here, residues Glu-13 and Lys-312 were found to be involved in the interaction with the O4-H hydroxyl of glucose. Since Lys-312 does not form a hydrogen bond with the O4-H hydroxyl in the galactose-bound ttGBP structure, this explains the observed ~100-fold lower affinity for galactose (0.94 μ M).¹⁵

An overlay of D-glucose onto the bound β -D-galactose (Figure 5b) illustrates that the O4-H hydroxyl of D-glucose would not be able to hydrogen bond with Arg-82 and Arg-362 and furthermore would be in close proximity to CD2 of Phe-359 (2.73 Å), resulting in low (or no) specificity of AcbH towards D-glucose.

The respective purified and dialyzed AcbH variants were analyzed for binding of radiolabeled galactose or glucose. As shown in Figure 7, galactose binding is reduced in AcbH(Arg-

Chapter II

82Trp) and AcbH(Arg-82Ala) by 55 % and 65 %, respectively, but is fully abolished in variants carrying mutations of Asp-361 or Arg-362. Moreover, none of the mutants has gained binding affinity for D-glucose. These results indicate that Asp-361, Arg-362 and to a lesser extent Arg-82 are essential in galactose binding as even the replacement by structurally related amino acids inhibited substrate recognition.

Mutations altering the substrate specificity of mono- or oligosaccharide binding proteins have been reported.^{31,32} To elucidate the loss of functionality of AcbH variants, we crystallized Asp-361Lys, Asp-361Glu, Arg-362Lys and Arg-362Ala with identical reservoir solutions consisting of 2.2 M (NH₄)₂SO₄, 100 mM citric acid, pH 4.0. Crystals of the first two variants belonged to space group P212121 and diffracted to 1.9 Å and 1.5 Å resolution, respectively, while variants Arg-362Lys and Arg-362Ala crystallized in space group P21 with two molecules in the asymmetric unit and diffracted to 1.7 Å and 2.0 Å resolution, respectively. In contrast to wild type AcbH, structural analysis of the variant proteins overproduced in the same growth medium (see ‘Materials and Methods’) revealed an open, unliganded conformation of AcbH. This observation underscores the importance of any of the mutated residues for galactose binding. The crystal structure of Asp-361Lys is identical to that of Asp-361Glu with RMSD of 0.12 Å. Similarly, the structure of variant Arg-362Lys is identical to that of Arg-362Ala with RMSD of 0.15 Å (chain A of the crystallographic dimer) and 0.16 Å (chain B of the crystallographic dimer), respectively (Figure 3b). Indeed, comparison of the four variant structures shows only minor differences. Mutations of AcbH are associated with an opening of the cleft between the two domains that provides access for substrates. For example, helix α 2 (residues 56-68) in the substrate binding pocket moves by ~12 Å compared with wild type AcbH. In spite of this movement in the center of the molecule, the N- and C-terminal residues are located in the same positions (Figures 3a and 3b). Domain motions of the variants were analysed using Dydndom software with AcbH in closed conformation as reference. The N-terminal domain of AcbH was chosen as fixed domain and the C-terminal domain as moving domain. In the structures of the two Asp-361 variants significant motions were found at the hinge regions (residues 319-328 and 338-343), and in the two Arg-362 variants significant motions were found at hinge regions, amino acids 319 - 328 and 338 – 347.

Chapter II

Conclusions

The results presented here demonstrate that AcbH is a high-affinity D-galactose-binding protein with high substrate specificity. In contrast to our previous study using refolded AcbH²¹, binding of acarbose or longer homologs of acarbose could not be confirmed with the native-like protein of the present investigation. We assume that AcbH refolded from inclusion bodies, although soluble and displaying secondary structure as judged from CD spectroscopy, had adopted a more labile or slightly different tertiary structure compared with the native-like AcbH used here. As a consequence, longer dextrans might have gained (unspecific) access to the binding cleft at millimolar concentrations. In fact, analysis of the intrinsic tryptophan fluorescence of native-like AcbH prior to and after a denaturation/renaturation treatment did not result in identical spectra, indicating that the original fold was not recovered (not shown).

The X-ray structure reveals that AcbH is homologous to class II MBPs rather than to class I GBPs. The ttGBP is most closely related to AcbH by amino acid sequence and structural homology. In contrast to MBPs, ttGBP and AcbH show structural modifications that accommodate monosaccharides in a class II fold because the insertion of loops and an α -helix reduce the large MBP cavity and allow only the binding of monosaccharides. Similar mechanisms of loop-mediated modifications are seen for interconversion of class II periplasmic ion and sugar binding proteins.^{33,34} However, whereas most glucose/galactose binding proteins exhibit a much higher affinity for glucose, AcbH has a unique and exclusive substrate preference for β -D-galactose in the 4C1 chair conformation. Site-directed mutagenesis of the three residues (Arg-82, Asp-361, Arg-362) interacting with the O4-H hydroxyl of the bound galactose resulted in an open, unliganded AcbH structure with reduced or completely abolished galactose affinity.

Since the solute binding subunits determine the substrate specificity of ABC importers, the physiological role of AcbHFG as a putative galactose transporter within the context of acarbose metabolism remains elusive. One could speculate on D-galactose as a source for glucose-1-phosphate, the precursor of the acarviosine moiety of acarbose¹⁸, via the Leloir pathway. This would include conversion to galactose-1-P by galactokinase and subsequent formation of glucose-1-P by exchange of galactose with the glucose group of UDP-glucose by an epimerase. Regarding the natural habitat of *Actinoplanes* which was isolated from a mud lake, D-galactose might originate from decomposed plants which are rich in galactose-containing oligosaccharides, galactolipids and the polymer galactan.

Chapter II

In this context, the two genes *asp3.1* and *asp3.2*, which are divergently orientated to the *acbHFG* operon (GenBank accession Y18523) come into focus. These genes were previously not considered to belong to the *acb*-cluster because no function for their products was evident for the postulated acarbose metabolism.¹⁸ However, gene *asp3.1* is predicted to encode a galactocerebrosidase (glycosyl hydrolase family 59 at <http://www.cazy.org>) and thus might indicate that galactose-containing lipids may indeed serve as a carbon source in the natural habitat of *Actinoplanes* sp. If so, a high-affinity galactose-specific transport system would be beneficial for this slow growing bacterium. Whether *Actinoplanes* sp. is equipped with additional glucose- or galactose-uptake systems cannot be answered because a genome sequence is not available yet.

In contrast to our results obtained with the GacH protein of *S. glaucescens*¹⁶ which is also encoded within the acarbose gene cluster, the data presented here are not in favour of the proposed role of acarbose as a carbon storage molecule in *Actinoplanes* sp. Here, acarbose obviously seems to be produced only to function as an inhibitor of maltose/maltodextrin transport systems^{35,36} and various α -glucosidases of competitors in the natural environment, thereby mediating a growth advantage for the slowly growing *Actinomycetes*.

Materials and Methods

Chemicals

D-(+)-glucose, D-(+)-galacturonic acid, raffinose, maltose, maltotriose, and maltotetraose were obtained from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany) and were of 96–98% purity. D-(+)-galactose, L-(+)-arabinose, L-(+)-rhamnose and D-(+)-xylose were purchased from Carl Roth GmbH (Karlsruhe, Germany) with > 98 % purity. D-[1-¹⁴C]galactose (56 mCi/mmol) and D-[(U)-¹⁴C]glucose (262 mCi/mmol) were obtained from Hartmann Analytik (Braunschweig, Germany). Acarbose, component 5C and component mixture 6AB, all derived from *Actinoplanes* sp., and starch hydrolysate (mixture of maltose and maltotriose, containing trace amounts of longer dextrans from Cerestar, Krefeld, Germany) were kind gifts of Dr. H. Wehlmann (Bayer Healthcare, Wuppertal, Germany). According to the manufacturer, acarbose was of 97.8 % purity and contained trace amounts (0.9 %) of acarbose derivatives but was devoid of glucose, maltose, and maltodextrins.

Chapter II

Bacterial strains, media and growth conditions

E. coli strain Rosetta 2 (DE3)(pLysS) (Stratagene, Heidelberg, Germany) harboring plasmid pAL8 (expressing *acbH* under control of the pT7lac promoter) and plasmid ptacSL1 (expressing *groEL/ES* under control of the ptac promoter) was grown in ZYP-5052 medium for auto-induction³⁷ supplemented with ampicillin (0.1 mg/ml), chloramphenicol (0.02 mg/ml) and kanamycin (0.1 mg/ml) at 37 °C to an OD₆₅₀ of 0.3 – 0.5. The auto-induction of *acbH* expression occurred during continued growth overnight at 22 °C as the culture approached saturation. Expression of the *acbH* variants from plasmids pAL51 – 56 was carried out as described for wild type.

To prepare selenomethionine-substituted AcbH or to obtain ligand-free AcbH for binding assays, *E. coli* Rosetta2 (plysS, pAL8, ptacSL1) cells or *E. coli* BL21 T1R cells (Sigma Aldrich, Germany) were grown at 39 °C in minimal medium M9³⁸ supplemented with glycerol (4 %), ampicillin (0.1 mg/ml), chloramphenicol (0.02 mg/ml) and kanamycin (0.1 mg/ml) to an OD₆₅₀ of 0.8 – 1.0. At this point, powdered L-lysine, L-phenylalanine, L-threonine (each to 100 mg/ml final), L-isoleucine, L-leucine, L-valine, L-selenomethionine or L-methionine (each to 50 mg/ml final) were added. Subsequently, expression of *acbH* was induced by 0.5 mM IPTG, and growth continued for 4 h.

Cloning of *acbH*

The *acbH* gene (codons 39-433) lacking the signal sequence and the putative linker region was amplified by PCR from genomic DNA of *Actinoplanes* sp. SE50/110 using the oligonucleotide primers alhf2 (5'-ATA ATA CAT ATG ACC AGC GTC AGC GAC GGC AAC GGG CCG-3') and alhr2 (3'-ATA ATA GGA TCC TCA GCG GAA AAT CGT CTT CGC CTG GTC-5'). The primers were designed such that the resulting fragment could be ligated with expression vector pET15b (Novagen, Germany) via NdeI and BamHI recognition sites (underlined), yielding plasmid pAL8. As a consequence, the translated polypeptide has a His₆-tag fused to its N-terminal amino acid.

Purification of soluble AcbHHis₆

Cells of the *E. coli* strain Rosetta2 (plysS, pAL8, ptacSL1) or *E. coli* BL21 T1R (plysS, pAL8, ptacSL1) were harvested, resuspended in buffer containing 50 mM MOPS/KOH, pH 7.0, 200 mM NaCl, 5 % (v/v) glycerol, 0.1 mM PMSF and disrupted by passage through a french press. After a high speed spin, the cytosolic fraction containing His₆-AcbH was subjected to metal affinity

Chapter II

chromatography using a 5 ml TALON-HiTrap resin and the ÄKTATM purifier FPLC system (GE Healthcare, Freiburg, Germany). Following a washing step with 5 column volumes of the same buffer, His6-AcbH was eluted by supplementing the buffer with 0.1 M imidazole. Subsequently, His6-AcbH-containing fractions were pooled and the His6-tag was cleaved by treatment with thrombin (Thrombin CleanCleave™ Kit, Sigma- Aldrich, Taufkirchen, Germany) according to the manufacturer`s instructions. To remove the cleaved His6-tag and imidazol, the sample was concentrated by a centrifugal device (Centrikon YM 10, Millipore), passed through a PD10 desalting column (GE Healthcare), and dialysed against 10 L of 20 mM MOPS-KOH, pH 7.0, 100 mM NaCl, 1 % (v/v) glycerol with several changes.

Purification of selenomethionine-substituted AcbH was performed as described for wild type AcbH except that the AcbH-containing cytosolic fraction was supplemented with 20 mM imidazole and incubated with 1 ml Co²⁺-loaded affinity matrix (TALON, Clontech) for 1 h at 4 °C prior to washing and elution.

Binding assays

The affinities of AcbH for galactose, glucose, maltose, maltotetraose, acarbose, component 5C, lactose and L-arabinose were determined by high-sensitivity microcalorimetry³⁹ using a VP-ITC device (MicroCal, Northampton, USA). To avoid air bubbles, solutions were degassed under vacuum prior to use. Freshly dialyzed AcbH, previously tested for the absence of bound ligand by TLC, was filled in the reaction cell at a concentration of 10 µM in buffer containing 20 mM MOPS-KOH, pH 7.0, 100 mM NaCl, 1 % (v/v) glycerol and titrated in 30 steps of 10 µl each against stock solutions of galactose (80 µM), glucose, maltose, maltotetraose, lactose and L-arabinose (120 µM) or acarbose and component 5C (1 mM) in 5 min intervals at 25 °C. Power peaks were integrated and the resulting reaction heats were plotted against the molar substrate:protein ratio and fitted using the 'one set of sites' model with origin 5.0 (MicroCal Software, Northampton, USA), yielding the dissociation constant K_d. The first injection was always excluded from evaluation because it usually suffers from sample loss during the mounting of the syringe and the equilibration preceding the actual titration.

The affinity of purified AcbH for D-galactose was alternatively determined by a filtration assay as described by Richarme and Kepes.⁴⁰ To this end, samples (100 µl) of 2.5 µM protein in buffer containing 20 mM MOPS-KOH, pH 7.0, 100 mM NaCl, 1 % (v/v) glycerol were pre-incubated at 25 °C for 1 min prior to the addition of increasing concentrations of D-[1-¹⁴C]

Chapter II

galactose ranging from 0.01 to 1000 nM. The reaction was terminated after 1 min by adding 2 ml of ice-cold saturated (NH₄)₂SO₄ solution, and the mixture was immediately passed through a nitrocellulose filter (0.45 μm pore size). After washing with 2 x 5 ml of the (NH₄)₂SO₄ solution followed by 5 ml of distilled water, the radioactivity retained on the filters was determined in a liquid scintillation counter.

The substrate specificity of AcbH variants was determined likewise using 100 μM D- [1-¹⁴C] galactose or D-[U-¹⁴C] glucose, respectively.

For competition studies, samples of purified AcbH were pre-incubated with Dgalactose, D-glucose, D-galacturonic acid, L-arabinose, L-rhamnose or D-xylose, lactose, raffinose, maltose, maltotriose, maltotetraose, acarbose, acarbose homologs 5C and 6AB (500 μM each), respectively, at 25 °C for 1 min prior to the addition of 5 μM D-[1-¹⁴C] galactose.

Crystallization and structure determination

For crystallization purposes, the AcbH dialysate was concentrated up to 25 mg/ml. All of the obtained AcbH crystals (AcbHSEMET, AcbHNATIVE, AcbHD361K, AcbHD361E, AcbHR362K and AcbHR362A) were grown at 18 °C using the sitting drop vapor diffusion method and reservoir solutions containing 2.2 M (NH₄)₂SO₄ and 100 mM citric acid, pH 4.5. Crystallization drops were prepared by mixing 2 μl of reservoir and 2 μl of protein solution to yield rhombus and tetragonal-shaped crystals which grew approximately in two weeks to a maximal size of 1.0 x 0.5 x 0.3 mm. Single crystals were separated from clusters and briefly soaked in a cryoprotection medium (reservoir solution with 15 % (v/v) glycerol), mounted in a nylon loop and then flash-cooled in liquid N₂. X-ray data were collected at beamline BL1 at BESSY-II, Berlin. The space groups were determined as P212121 and P21, resulting in an occupancy of one and two molecules per asymmetric unit, respectively (Table 1). The crystals diffracted to resolution shells ranging from 2.15 to 1.53 Å.

Multiwavelength anomalous diffraction data of AcbHSEMET crystals were collected at three wavelengths (1 = 0.97930 Å, 2 = 0.97935 Å, 3 = 0.98051 Å) at beamline BL1 at BESSY-II, Berlin. The individual data sets were processed with programs DENZO/SCALEPACK⁴¹ and merged/scaled by the CCP4 program suite.⁴² The positions of 16 of the 20 anomalous scatterers (Se) were located with the HySS (Hybrid Substructure Search) software of the PHENIX-project.⁴³ Heavy atom positional refinement, location of additional sites, and phasing was subsequently done with SHARP (Global Phasing, Cambridge, UK). Density modification and model building with

Chapter II

ARP/WARP produced a 94 % complete model. The rest of the model including ligands in the binding pocket and hydration water molecules were built manually in COOT.⁴⁴ Final isotropic individual B-factor refinement using REFMAC5, gave a model with R and Rfree values of 20.5 and 25.9 %, respectively. The structural data of the selenomethionine-substituted AcbH, was subsequently used as model to solve the phase problems of the native and mutant AcbH crystals. Results are shown in Table 2.

Site-directed mutagenesis

Plasmids pAL51 – 56, expressing *acbH*(Arg-R82Trp), *acbH*(Arg-82Ala), *acbH*(Asp-361Lys), *acbH*(Asp-361Glu), *acbH*(Arg-362Lys) and *acbH*(Arg-362Ala), respectively, were constructed by using plasmid pAL8 as template and Stratagene's QuickChange kit according to the manufacturer's instructions or by overlap extension-PCR technique.⁴⁵

Accession numbers

Atomic coordinates and structure factors [accession codes 3006 (AcbHgalactose), 3007 (AcbH(D361K)open), 3008 (AcbH(D361E)open), 3009 (AcbH(R362K)open) and 300A (AcbH(R362A)open)] have been deposited in the PDB, Research Collaboratory for Structural Bioinformatics, Rutgers University (New Brunswick, NJ).

Acknowledgements

We thank Dr. H. Wehlmann (Bayer Healthcare, Wuppertal, Germany) for generous gifts of acarbose, acarbose derivatives 5C and 6AB, and media, and Claudia Alings (AG Kristallographie, FU Berlin) for the initial preparation of the native AcbH crystals. This work was supported by the Deutsche Forschungsgemeinschaft (SCHN274/12-1; WE2936/1-1 and SFB 449) and by a fellowship from the German National Academic Foundation (to A.L.). We acknowledge access to beamline BL14.2 (and/or 14.1 and/or 14.3) of the BESSY II storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt- Universität zu Berlin, the Max-Delbrück Centrum and the Leibniz-Institut für Molekulare Pharmakologie.

Chapter II

References

1. Eitinger T., Rodionov, D.A., Grote, M., and Schneider, E. (2010) Canonical and ECType ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* doi:10.1111/j. 1574-6976.2010.00230.x.
2. Berntsson, R.P., Smits, S.H., Schmitt, L., Slotboom, D.J., and Poolman, B. (2010) A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606-2617.
3. Wilkinson, A.J., and Verschuere, K.H.G.. (2003) Crystal structures of periplasmic solute binding proteins in ABC transport complexes illuminate their function. In: *ABC proteins: from bacteria to man* (Holland, B., Cole, S. P. C., Kuchler, K. and Higgins, C. F., eds.), pp. 187-207, Elsevier, Amsterdam.
4. Sutcliffe, I.C., and Russell, R.R. (1995) Lipoproteins of gram-positive bacteria. *J. Bacteriol.* **177**, 1123-1128.
5. van der Heide T., and Poolman, B. (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep.* **3**, 938-943.
6. Biemans-Oldehinkel, E., Doeven, M.K., and Poolman, B. (2006) ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett.* **580**, 1023-1035.
7. Quijoch, F.A. and Ledvina, P.S. (1996) Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20**, 17-25.
8. Shilton, B.H., Flocco, M.M., Nilsson, M., and Mowbray, S.L. (1996) Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose- and ribose-binding proteins. *J. Mol. Biol.* **264**, 350-363.
9. Fukami-Kobayashi, K., Tateno, Y., and Nishikawa, K. (1999) Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279-290.
10. Vyas, N.K., Vyas, M.N., and Quijoch, F.A. (1991) Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and functional similarity. *J. Biol. Chem.* **266**, 5226-5237.
11. Vyas, M.N., Vyas, N.K., and Quijoch, F.A. (1994) Crystallographic analysis of the epimeric and anomeric specificity of the periplasmic transport/chemosensory protein receptor for D-glucose and D-galactose. *Biochemistry* **33**, 4762-4768.
12. Spurlino, J.C., Lu, G.Y., and Quijoch, F.A. (1991) The 2.3 Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202-5219.
13. Sharff, A.J., Rodseth, L.E., Spurlino, J.C., and Quijoch, F.A. (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **31**, 10657-10663.

Chapter II

14. Schneider, E. (2001). ABC transporters catalyzing carbohydrate uptake. *Res. Microbiol.* **152**, 303-310.
15. Cuneo, M.J., Changela, A., Warren, J.J., Beese, L.S., and Hellinga, H.W. (2006) The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and disaccharide binding sites. *J. Mol. Biol.* **362**, 259-270.
16. Vahedi-Faridi, A., Licht, A., Bulut, H., Scheffel, F., Keller, S., Wehmeier, U.F., Saenger, W., and Schneider, E. (2010) Crystal structures of the solute receptor GacH of *Streptomyces glaucescens* in complex with acarbose and an acarbose homolog. Comparison with acarbose-loaded maltose binding protein of *Salmonella typhimurium*. *J. Mol. Biol.* **397**, 709-723.
17. Rockser, Y., and Wehmeier, U. F. (2009) The *gac*-gene cluster for the production of acarbose from *Streptomyces glaucescens* GLA.O: identification, isolation and characterization. *J. Biotechnol.* **140**, 114-123.
18. Wehmeier, U.F., and Piepersberg, W. (2004) Molecular biology and enzymology of the metabolism of the α -glucosidase inhibitor acarbose. *Appl. Microbiol. Biotechnol.* **63**, 613-625.
19. Schmidt, D.D., Frommer, W., Junge, B., Müller, L., Wingender, W., and Truscheit, E. (1977) α -glucosidase inhibitors: new complex oligosaccharides of microbial origin. *Naturwissenschaften* **64**, 535-536.
20. Bischoff, H., Ahr, H.J., Schmidt, D., and Stoltefuß, J. (1994) Acarbose – ein neues Wirkprinzip in der Diabetestherapie. *Nachr. Chem. Tech. Lab.* **42**, 1119-1128.
21. Brunkhorst, C., Wehmeier, U.F., Piepersberg, W., and Schneider, E. (2005) The *acbH* gene of *Actinoplanes* sp. encodes a solute receptor with binding activities for acarbose and longer homologs. *Res. Microbiol.* **156**, 322-327.
22. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. And Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
23. Willis, L.B., and Walker, G.C. (1999) A novel *Sinorhizobium meliloti* operon encodes an α -glucosidase and a periplasmic-binding-protein-dependent transport system for α -glucosides. *J. Bacteriol.* **181**, 4176-4184.
24. Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W., and Welte, W. (2001) The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis* at 1.85 Å. *J. Mol. Biol.* **305**, 905- 915.
25. Lee, S.J., Moulakakis, C., Koning, S.M., Hausner, W., Thomm, M. and Boos, W. (2005) TrmB, a sugar sensing regulator of ABC transporter genes in *Pyrococcus furiosus* exhibits dual promoter specificity and is controlled by different inducers. *Mol. Microbiol.* **57**, 1797-1807.
26. Henderson, P.J., Giddens, R.A., and Jones-Mortimer, M.C. (1977) Transport of galactose, glucose and their molecular analogues by *Escherichia coli* K12. *Biochem. J.* **162**, 309-320.

Chapter II

27. Tian, Y., Cuneo, M.J., Changela, A., Höcker, B., Beese, L.S., and Hellinga, H.W. (2007) Structure-based design of robust glucose biosensors using a *Thermotoga maritima* periplasmic glucose-binding protein. *Protein Sci.* **16**, 2240-2250.
28. Horler, R.S., Müller, A., Williamson, D.C., Potts, J.R., Wilson, K.S., and Thomas, G.H. (2009) Furanose-specific sugar transport: characterization of a bacterial galactofuranose-binding protein. *J. Biol. Chem.* **284**, 31156-31163.
29. Collins, P., and Ferrier, R. (1995). Monosaccharides: Their chemistry and their roles in natural products, Wiley & Sons, Chichester, UK
30. Holm, L., Kaariaainen, S., Rosenstrom, P., and Schenkel, A. (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780-2781.
31. Vermersch, P.S., Lemon, D.D., Tesmer, J.J., and Quiocho, F.A. (1991) Sugar-binding and crystallographic studies of an arabinose-binding protein mutant (Met108Leu) that exhibits enhanced affinity and altered specificity. *Biochemistry* **30**, 6861-6866.
32. Guntas, G., Mansell, T.J., Kim, J.R., and Osermeier, M. (2005) Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc. Natl. Acad. Sci. USA* **102**, 11224-11229.
33. Heddle J., Scott, D.J., Unzai, S., Park, S.Y., and Tame, J.R. (2003) Crystal structures of the liganded and unliganded nickel-binding protein NikA from *Escherichia coli*. *J. Biol. Chem.* **278**, 50322-50329.
34. Cuneo, M.J., Beese, L.S., and Hellinga, H.W. (2009) Structural analysis of semispecific oligosaccharide recognition by a cellulose-binding protein of *Thermotoga maritima* reveals adaptations for functional diversification of the oligopeptide periplasmic-binding protein fold. *J. Biol. Chem.* **284**, 33217-33223.
35. Brunkhorst, C., Andersen, C., and Schneider, E. (1999) Acarbose, a pseudooligosaccharide, is transported but not metabolized by the maltose/maltodextrin system of *Escherichia coli*. *J. Bacteriol.* **181**, 2612-2619.
36. Hülsmann, A., Lurz, R., Scheffel, F., and Schneider, E. (2000) Maltose and maltodextrin transport in the thermoacidophilic Gram-positive bacterium *Alicyclobacillus acidocaldarius* is mediated by a high-affinity transport system that includes a maltose binding protein tolerant to low pH. *J. Bacteriol.* **182**, 6292-6301.
37. Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* **41**, 207-234.
38. Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Chapter II

39. Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131-137.
40. Richarme, G., and Kepes, A. (1983) Study of binding protein-ligand interaction by ammonium sulfate-assisted adsorption on cellulose esters filters. *Biochem. Biophys. Acta* **742**, 16-24.
41. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
42. CCP4. (1994) Collaborative Computational Project, Number 4; The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **D50**, 760-763.
43. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 1948-1954.
44. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallogr.* **60**, 2126-2132.
45. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.
46. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
47. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucl. Acids Res.* **31**, 3497-3500.

Chapter II

Table 1. Hydrogen bonding interactions between AcbH and galactose

AcbH	Distance (Å)	Galactose
HIS-56[NE2]	3.16	[O1]
HOH-91	3.12	
HOH-636	2.75	
HIS-56[NE2]	2.72	[O2]
GLN-60[OE1]	2.53	
GLN-60[NE2]	3.22	
ARG-82[NH2]	2.48	[O3]
GLN-60[NE2]	3.07	
ASP-361 [OD2]	2.56	
ARG-82[NH2]	3.16	[O4]
ASP-361 [OD1]	2.64	
ARG-362 [NE]	2.74	
ARG-362 [NH2]	2.96	[O5]
TYR-129[OH]	3.29	
GLN-255 [NE2]	3.15	[O6]
TYR-129[OH]	2.86	
TRP-131 [NE1]	2.84	

Chapter II

Table 2. Data collection and refinement statistics

	AcbH_{native}	AcbH_{D361K}	AcbH_{D361E}	AcbH_{R362K}	AcbH_{R362A}
<i>A. Data collection</i>					
Space group	<i>P</i> 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁
Cell parameters					
<i>a</i> (Å)	47.84	39.47	39.51	39.25	39.25
<i>b</i> (Å)	98.81	89.50	88.91	65.76	66.12
<i>c</i> (Å)	95.73	101.58	102.68	177.72	177.71
α	90.00	90.00	90.00	90.00	90.00
β	95.04	90.00	90.00	90.03	89.97
γ (°)	90.00	90.00	90.00	90.00	90.00
Resolution (Å)	50 – 2.15	33.58 – 2.10	33.61 – 1.60	33.71 – 1.76	33.75 – 2.04
<i>R</i> _{sym} (<i>R</i> _{merge})	0.084 (0.234)	0.137 (0.498)	0.063 (0.504)	0.092(0.553)	0.104 (0.538)
<i>I</i> / σ <i>I</i>	11.82 (3.73)	16.92 (5.75)	18.10 (3.55)	8.78 (1.80)	9.48 (2.20)
Completeness (%)	93.1 (83.2)	100.0 (100.0)	99.9 (99.7)	94.4 (88.7)	98.4 (96.3)
Redundancy	2.9 (2.5)	7.2 (7.3)	6.3 (5.8)	2.8 (2.3)	3.3 (2.8)
B. Refinement					
Resolution (Å)	95.35 – 2.15	33.58 – 2.10	33.61 – 1.60	33.71 – 1.76	33.75 – 2.04
No. reflections	44,697	21,736	48,594	84,899	57,306
<i>R</i> _{work} (%)	18.8	16.4	16.7	15.5	16.8
<i>R</i> _{free} (%)	25.1	22.0	21.0	19.5	22.6
No. atoms					
Protein	6110	3045	3060	6179	6156
Ligand	24	n.a.	n.a.	n.a.	n.a.
Water	657	309	445	1004	933
<i>B</i> -factors					
Protein (Å ²)	12.5	16.4	16.4	11.9	15.6
Ligand (Å ²)	10.2	n.a.	n.a.	n.a.	n.a.
Water (Å ²)	20.2	27.2	28.9	26.0	27.1

Chapter II

	AcbH _{native}	AcbH _{D361K}	AcbH _{D361E}	AcbH _{R362K}	AcbH _{R362A}
r.m.s deviation from ideal					
Bond lengths (Å)	0.011	0.009	0.012	0.013	0.010
Bond angles (°)	0.931	1.248	1.488	1.359	1.047

The values in parentheses are for the highest resolution shell.

Figure legends

Fig. 1. Binding affinities of AcbH for D-galactose and D-glucose as determined by isothermal titration calorimetry (**a**, **b**) and a filtration assay (**c**) (see ‘Materials and Methods’ for details). AcbH was titrated with (**a**) D-galactose or (**b**) D-glucose. Integrated and normalized heats of reaction Q were plotted against the sugar / protein molar ratio R and fitted using a one-site binding model. (**c**) Binding of increasing concentrations of [¹⁴C] galactose to AcbH expressed in cpm (*upper part*) and Scatchard plot analysis of the data (*lower part*). Values represent the mean of three independent experiments. Standard deviations are indicated as error bars.

Fig. 2. D-[¹⁴C] galactose binding of AcbH in the presence of various mono-, di-, and oligosaccharides. The indicated carbohydrates were added to the reaction mixture at 500 μM each and D-galactose binding was determined by a precipitation assay as described in ‘Materials and Methods’. Values represent the mean of three independent experiments. Standard deviations are indicated as error bars.

Fig. 3. (**a**) Top view of the overall structure of AcbH in complex with D-galactose (galactose shown in red-yellow). The ordering of β-strands in the N- and C- terminal domains allows AcbH to be categorized as a class II SBP. The N- and C- termini are indicated. Secondary structural elements are numbered. The two domains are interconnected by three hinge regions (H.I, residues 128-131; H.II, residues 287-291; H.III, residues 324-341). (**b**) Structure of AcbH mutant Arg-362Ala. Location of the mutated residue (Arg-362) on helix-18 is highlighted in yellow. Major displacement in the mutant structure takes place at the centre part of the C- terminal domain. Three helices located around the substrate binding pocket in the galactose-bound wild-type AcbH move ~12 Å away from the binding pocket, which gives enough space for the galactose molecule (~ 6 Å

Chapter II

diameter) to access the binding pocket of AcbH. (c) In the schematic topography, arrows represent β -strands, and cylinders represent α - helices. The ordering of the 5- and 3-stranded β -sheets, respectively, in the N- and C-terminal domains is defined by their topography. The strand β 4 links the two domains and comprises hinge II. Molecular graphics were generated using Pymol.⁴⁶

Fig. 4. View of the AcbH binding pocket from the ‘solvent-accessible’ side. Amino acids are shown as sticks. Carbon atoms of galactose are coloured yellow and oxygen atoms are in red. The central cleft between the two domains harbours β -D-galactose in 4C1 conformation. Hydrogen bonding partners that are within 3.4 Å are indicated by dashed lines drawn between β -D-galactose and AcbH amino acid side-chains.

Fig. 5. (a) Surface representation of the galactose binding pocket. (b) Interactions of the axially oriented O4-H hydroxyl of β -D-galactose (gray sticks) with Arg-82, Asp-361 and Arg- 362. A glucose molecule (turquoise sticks) is overlaid onto galactose.

Fig. 6. Amino acid sequence alignment (a) and structural relationship (b, c) of AcbH with maltose/maltodextrin binding protein GacH of *S. glaucescens* (PDB code 3JYR) (Vahedi- Faridi *et al.*, 2010), and glucose/galactose binding protein ttGBP of *T. thermophilus* (PDB code 3B3F). (a) The sequence alignment was generated by CLUSTAL W.⁴⁷ Loops or helices that reduce the oligosaccharide binding site are marked yellow. Residues which participate in galactose binding are shown in red. (b, c) Close-up side-by-side view of the binding pockets of galactose-bound AcbH (red) (b) compared with acarbose-bound GacH (green) (c). Galactose (yellow) and acarbose (blue) are shown in ball and stick mode. Helix α 18 fills the binding pocket in AcbH, thereby decreasing the available binding volume in comparison to the GacH oligosaccharide-binding site.

Fig. 7. D-[14C]galactose (grey bars) and D-[14C]glucose (black bars) binding of AcbH variants carrying amino acid substitutions in the binding pocket. Binding activities of the indicated variants were determined by a precipitation assay as described in ‘Materials and Methods’.

Chapter II

Fig.1

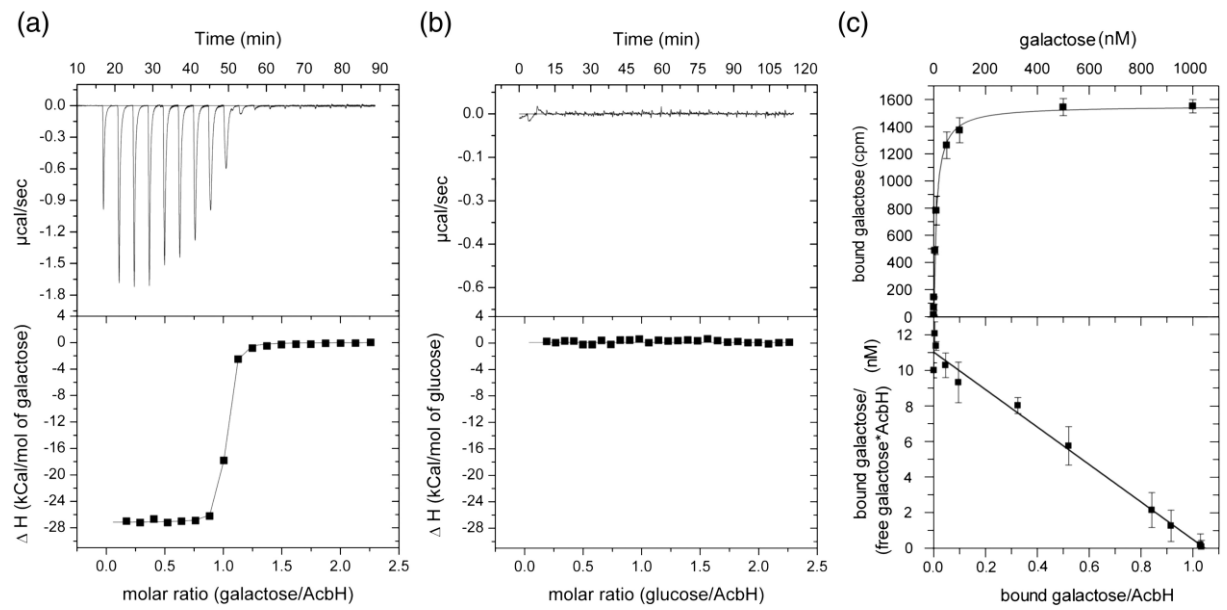
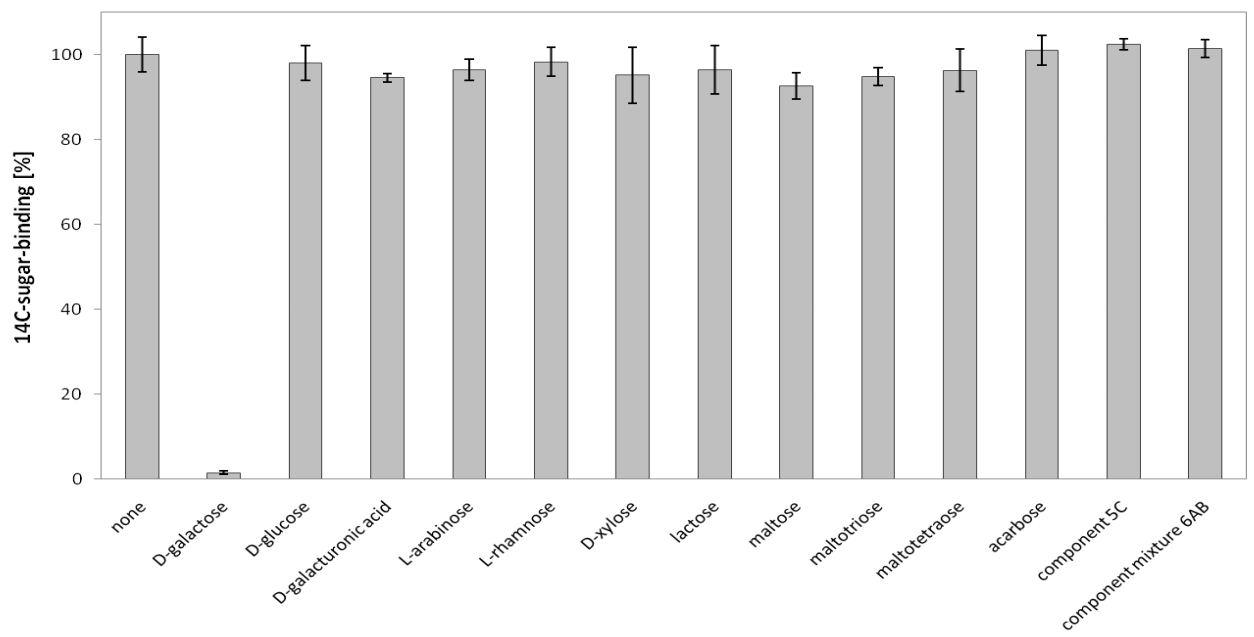


Fig.2



Chapter II

Fig. 3

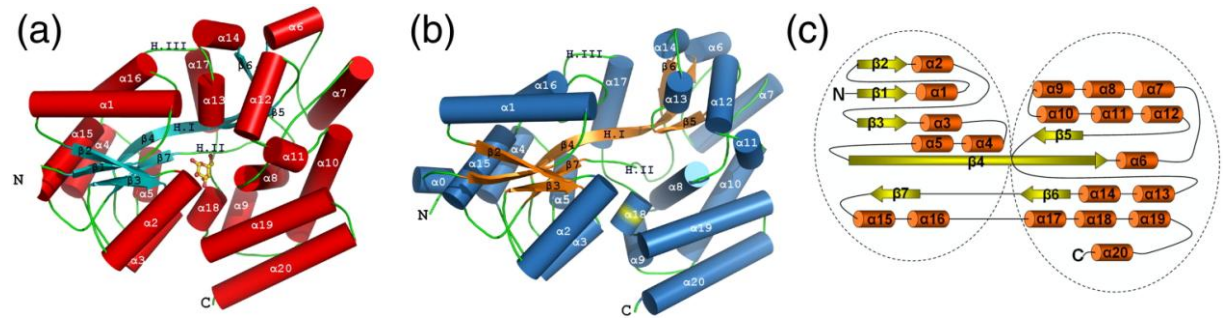


Fig. 4

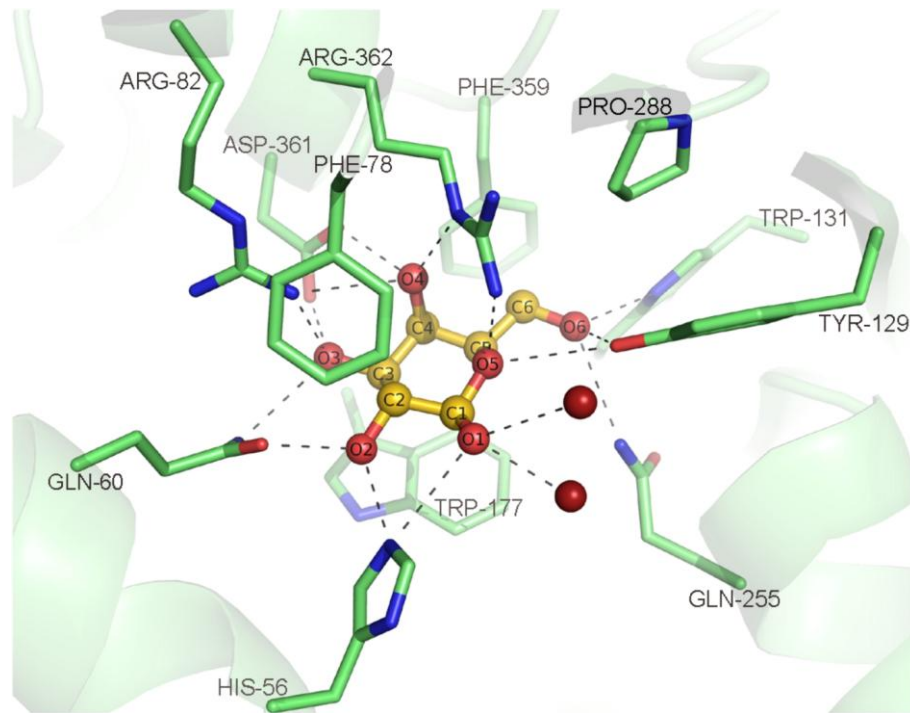
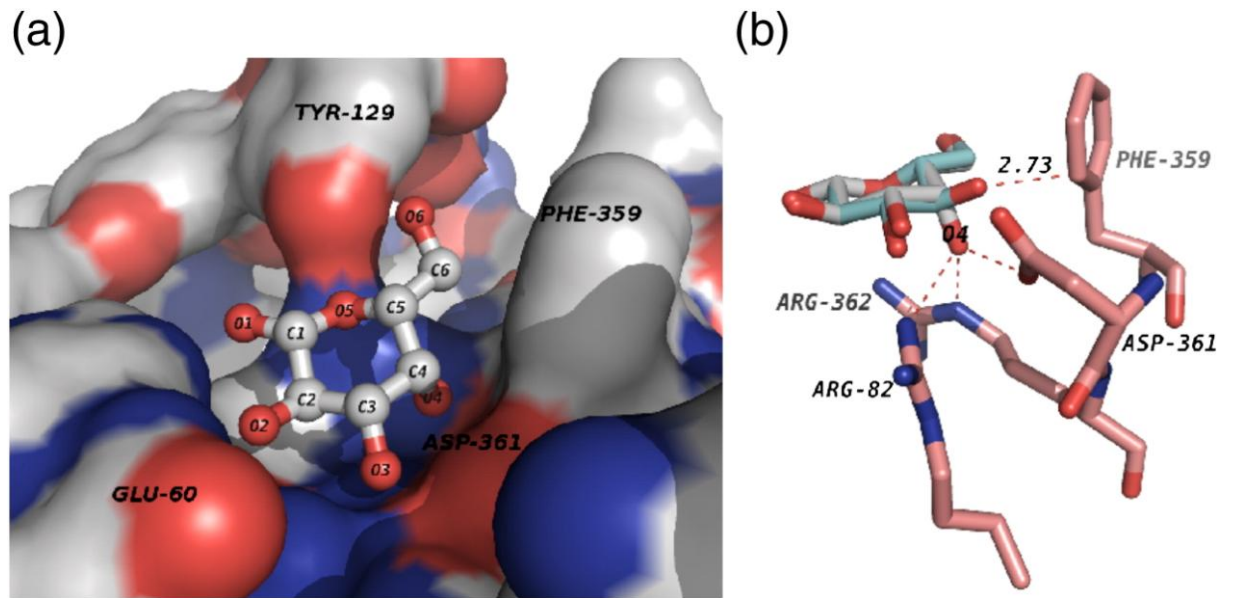


Fig. 5



Chapter II

Fig. 6

(a)

```

Ga cH      C G G D N G G S G T D A G G T E L S G T V T F W D T S N E A E K A T Y Q A L A E G F E K E H P K V D V K Y V N V P F G   60
ttGBP     -----K L E I F S W W A G D ---E G P A L E A L I R L Y K Q K Y P G V E V I N A T V T G G   40
Ac bH     C G S D D K S G T T S V S ---D G N G P I T F G S N Y S ---D E A P K A A F A S L M Q Q A T T S T T V P -V T V N T T   54
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      ---E A N A K F K N -A A G G N S G A P D V M R T E V A W V A D F A S I G Y L A P L D G T P A L D D G S D H L P Q A A   116
ttGBP     A G V N A R A V L K T R M L G G D P D T F Q V H A G M E L I G T W V V A N R M E D L S A L F R Q E G W L Q A F P K G L   100
Ac bH     D H N T F Q N N I S N Y L Q G T --P D S L A T W F A G Y R L Q F F A A Q G L L T P I D D V W D K I G G --T F N D A A   110
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      A S T R -Y E G K T Y A V P Q V I D -T L A L F Y N K E L L T K A G V E V P G S V A E L K T A A A E I T E K T G A T G L   174
ttGBP     I D L I S Y K G G I W S V P V N I H R S N V M W Y L P A K L K E W G V N P P R T W D E F L A T C Q T L K Q K G L E A P L   160
Ac bH     K S L S K G L D G H Y Y L V P L Y N Y P W V V F Y N K S V F Q S K G Y E V P A S W E A F I A L A R K M Q S D G L V P L A   170
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      Y L R G ---D D P Y W F L P Y L Y G E G -G D L V D E K N K T V T V D D E A G V R A Y R V I K D L V D S K A A I T D   229
ttGBP     A L G E N -W T Q Q H L W E S V A L A V L G P D D W N N L W N G K L K F T D P K A V R A W E V F G R V L D C A N K D A A   219
Ac bH     F A D K D G W P A L G T F D I L N L R I N G Y D Y H I K L M K H E V P W T D P G V T K V F D Q W R E L A A Y Q Q K G A N   230
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      A S D G W N N M Q N A F K S G K V A M M V N G P W A I E D V K A G A R F K D A G N L G V A P V P A G S A G Q G S P Q G G   289
ttGBP     G L S -W Q Q A V D R V V Q G K A A F N V M G D W A A G Y M T T T L K L K P G T D F A W A P S P G T Q G V F M M L S D S   278
Ac bH     G R T -W Q D A A K A L E N K Q A G M M F Q G S N Q V A A N Y S A K N L P D L D F F V F P A I N P Q Y G T D Y M D A P T   289
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      W N L S V Y A G S K N L D A S Y A F V K Y M S S -----A K V Q Q Q T T E K L S -L L P T R T S V Y E V P S V A D   341
ttGBP     F G L P --K G A K N R Q N A I N W L R L V G S K E G Q D T F N P L K G S I A A R L D -S D P S K Y N A Y Q S S A M R D   335
Ac bH     D G F I L P K K G K N A A A A K V L Q Y I G T A E A E A A F L K T D H W D V G L A N G L I A P T Y N D I Q K K S V A E   349
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      N E M V K F F K P A V D K A V E R P W I A E G N A L F E P I R L Q M A N V L S G E T S P D E A A A N T G D A Y R K L L K   401
ttGBP     W R S N R I V G S L V H G A V A P E S F M S Q F G T V M E I F L Q T R N P Q A A A N A A Q A I A D Q V G L G R L G Q --   393
Ac bH     I G K C K S V S Q F M D R D T V P D -M A N A M I K L I Q Q F I D Q P T P E T I A T V Q K S A E D Q A K T I F R ----   404
           . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :

Ga cH      D Y K   404
ttGBP     ---
Ac bH     ---
    
```

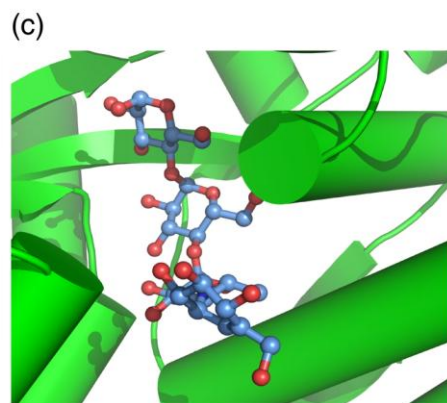
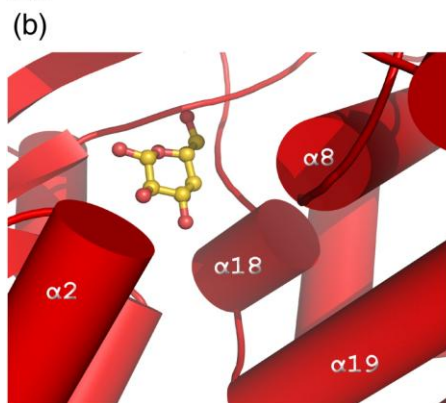
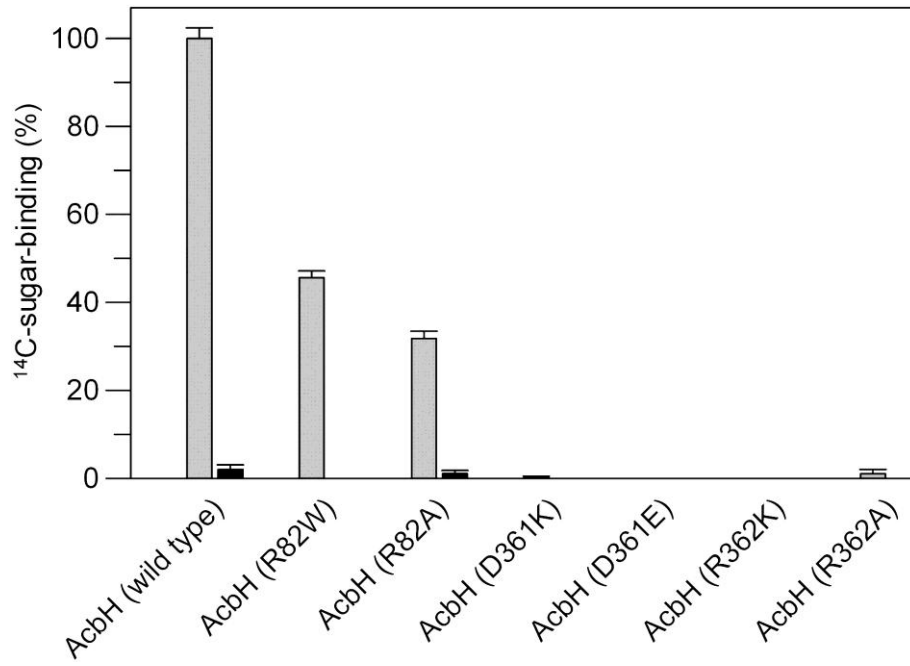


Fig. 7



11. Chapter III

Crystal structures of two solute receptors for L-cystine and L-cysteine, respectively, of the human pathogen *Neisseria gonorrhoeae*

Haydar Bulut^{1*}, Sebastien Moniot^{13*}, Frank Scheffel², Stephanie Gathmann², Anke Licht², Wolfram Saenger^{1#} and Erwin Schneider^{2#}

¹Institut für Chemie und Biochemie/Kristallographie, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany.

²Institut für Biologie/Bakterienphysiologie, Humboldt Universität zu Berlin, Chausseestrasse 117, 10115 Berlin, Germany.

³ Present address : Department of Biochemistry, Universität of Bayreuth, Universitätsstr 30, 95447 Bayreuth, Germany.

Corresponding authors

erwin.schneider@rz.hu-berlin.de; Phone: +49(0)30-2093-8121; Fax: +49(0)30-2093-8126

saenger@chemie.fu-berlin.de; Phone: +49(0)30-838-53412; +49(0)30-838 56702

*These authors contributed equally to this work

Abbreviations used: ABC, ATP-binding cassette; IPTG, isopropyl thio- β -D-galactopyranoside; NBD, nucleotide binding domain; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; SBP, solute binding protein; TMD, transmembrane domain; Tris, tris(hydroxymethyl)-aminomethane

Chapter III

Abstract

ATP-binding cassette (ABC) transporters are integral membrane proteins that carry a variety of substrates across biological membranes at the expense of ATP. The here considered prokaryotic canonical importers consist of three entities: an extracellular solute receptor, two membrane-intrinsic proteins forming a translocation pathway, and two cytoplasmic ATP-binding subunits. The *ngo0372-74* and *ngo2011-14* gene clusters from the human pathogen *Neisseria gonorrhoeae* were predicted by sequence homology as ABC transporters for the uptake of cystine and cysteine, respectively, and chosen for structural characterization. The structure of the receptor component Ngo0372 was obtained in a ligand-free “open” conformation and in a “closed” conformation when co-crystallized with L-cystine. Our data provide the first structural information of an L-cystine ABC transporter. Dissociation constants of 21 and 33 nM for L-cystine and L-selenocystine, respectively, were determined by isothermal titration calorimetry. In contrast, L-cystathionine and L-djenkolic acid are weak binders, while no binding was detectable for S-methyl-L-cysteine. Mutational analysis of two residues from the binding pocket, Trp97 and Tyr59, revealed that the latter is crucial for L-cystine binding. The structure of the Ngo2014 receptor was obtained in closed conformation in complex with co-purified L-cysteine. The protein binds L-cysteine with a K_d of 26 nM. Comparison of the structures of both receptors and analysis of the ligand binding sites shed light on the mode of ligand recognition and provides insight into the tight binding of both substrates. Moreover, since L-cystine limitation leads to reduction in virulence of *N. gonorrhoeae*, Ngo0372 might be suited as target for an antimicrobial vaccine.

Keywords: ABC transporter, L-cystine, L-cysteine, solute binding protein, X-ray crystal structure, *Neisseria gonorrhoeae*

Chapter III

Introduction

The gram-negative bacterium *Neisseria gonorrhoeae* (also known as gonococcus) is an obligate pathogen of man that colonizes primarily the human genitourinary tract, giving rise to the sexually transmitted infection gonorrhoea. Disease caused by this organism poses a threat to human health with an estimated worldwide 62 million new cases per year (http://www.who.int/vaccine_research/diseases/soa_std/en/index2.html).¹ The bacteria mainly inhabit mucosal surfaces of the urethra in males and the cervix in females and as a consequence are exposed to a variety of oxidants. The low-molecular-weight compound glutathione (γ -L-glutamyl-L-cysteinylglycine) (GSH) is considered as one of the first lines of defense against oxidative stress. GSH is a chemical scavenger of radicals and acts as a hydrogen donor to restore oxidized macromolecules. Unusually high concentrations of GSH (> 15 mM) have been reported to be present in *N. gonorrhoeae* which might constitute a powerful antioxidant defense.² Biosynthesis of GSH requires either L-cysteine or L-cystine as precursor which is consistent with the observation that cystine (preferred over cysteine) is an essential metabolite for *N. gonorrhoeae*.³

A connection between cystine limitation due to a defective transport system and sensitivity to hydrogen peroxide, a reagent causing oxidative stress, was previously demonstrated for the aerotolerant bacterium *Lactobacillus fermentum*.⁴ Similarly, blocking L-cystine uptake in animal cells leads to oxidative stress due to depletion of intracellular glutathione.^{5,6}

Nothing is known about putative transport systems mediating the uptake of cystine or cysteine in *N. gonorrhoeae*. However, cystine transporters have been identified in a few other bacteria.^{7,8} In *Bacillus subtilis*, cystine is internalized by two ATP-binding cassette (ABC) transport systems, TcyJKLMN and TcyABC, and a secondary transporter, TcyP.⁹ Likewise, three uptake systems for cystine were identified in *Salmonella enterica* serovar Typhimurium¹⁰, while in *Escherichia coli* two transport systems, including one ABC transporter, seem to be involved in cystine uptake.¹¹ None of the systems have been characterized in greater detail, which also holds for cysteine transporters.⁷

While secondary transporters couple solute translocation across a biological membrane to an electrochemical ion gradient¹², ABC transport systems are powered by the hydrolysis of ATP.¹³ Canonical bacterial ABC importers are composed of two transmembrane domains/subunits (TMDs) forming a translocation pore, two nucleotide binding domains/subunits (NBDs) that bind and hydrolyse ATP, and an extracellular solute receptor (or binding protein) (SBP) which captures the

Chapter III

substrate and delivers it to the translocation pore. SBPs are the main determinants of substrate specificity of their cognate transport system.⁸ The secondary structural elements of SBPs exhibit common features that consist of two large lobes interconnected by a hinge region. The lobes reside in an ‘open’ conformation in the absence of ligand and change conformation to a ‘closed’ form upon trapping their specific substrate (‘venus flytrap’ model).¹⁴ Liganded SBPs bind at the extracellular surface of the translocation pore with high affinity, and concomitant binding of ATP at the NBDs triggers an opening of the lobes of the SBP and release of the bound ligand into the pore.

Subsequent hydrolysis of ATP eventually leads to delivery of the ligand to the cytoplasm and return of the transporter to the resting state.^{8, 15-17}

Here we report on the identification, gene cloning, biochemical and structural characterization of two receptors for L-cystine and L-cysteine, respectively, of *N. gonorrhoeae* that are associated with putative ABC transporters. The crystal structures of both proteins are compared, and molecular details of specificity for cystine observed in the binding pocket were further analyzed by mutagenesis. Our data not only provide the first crystal structures of a cystine receptor but are the first characterization of L-cystine and L-cysteine binding proteins from the same organism.

Results and Discussion

Screening the *N. gonorrhoeae* FA 1090 genome for candidate genes encoding cystine and cysteine receptors, respectively

Substrate specificity of an ABC importer is primarily determined by the extracellular SBP. Thus, in order to identify putative cysteine and cystine transport systems, we searched the genomic sequence of *N. gonorrhoeae* FA 1090 by BLAST[§] using the cysteine binding protein CjaA of *Campylobacter jejuni*¹⁸ (accession number CAA71822.1) and the cystine binding protein BspA of *L. fermentum*¹⁹ (accession number AAC45332.1) as seed sequences. The data revealed Ngo2014 (56% identity to CjaA) and Ngo0372 (42% identity to BspA) as only candidates for cysteine and cystine binding proteins, respectively (Fig. 1). While Ngo2014 is located within a cluster including genes encoding two transmembrane subunits (Ngo2011, Ngo2012) and an ATPase component (Ngo2013) of an ABC transporter, Ngo0372 is part of an operon together with Ngo0373 and Ngo0374, encoding an inner membrane protein and an ATPase subunit, respectively.

In contrast, BLAST search with the secondary cystine transporter, TcyP from *B. subtilis*⁹ (accession number P54596.1), gave no result, indicating that Ngo2011/2012/2013/2014 and

Chapter III

Ngo0372/0373/0374 are the only putative cysteine and cystine transporters, respectively, of *N. gonorrhoeae*. Likewise, no evidence for a homolog of YliB, the receptor of a glutathione ABC transporter of *E. coli*²⁰ (accession number NP_415351.1), was obtained, suggesting that cysteine and cystine are the sole sources of cysteine for the synthesis of glutathione in *N. gonorrhoeae*.

Gene cloning, purification and substrate specificities of Ngo2014 and Ngo0372

The genes encoding the putative cystine and cysteine receptors, respectively, Ngo0372 and Ngo2014, were amplified from chromosomal DNA of *N. gonorrhoeae* FA 1090 by PCR and cloned into expression vector pET19b to yield plasmids pFSA122 and pFSA121 (see Materials and Methods for details) that were transformed into *E. coli* strain BL21 Star™ (DE3) (plysS). Both proteins were expressed and purified from the cytosolic fraction by cobalt affinity chromatography with a yield of about 20 mg per liter of cell culture.

To determine the respective substrate specificities of both proteins, we performed isothermal titration calorimetry (ITC). As shown in Fig. 2a, Ngo0372 exhibited binding of L-cystine with a dissociation constant (K_d) of 21 ± 1 nM. While L-selenocystine was bound with only slightly less affinity (33 ± 2 nM, Fig. 2b), the K_d values for the analogs L-cystathionine (962 ± 22 nM, Fig. 2c) and L-djenkolic acid (56 ± 1 μ M, Fig. 2d) were substantially higher. No binding was observed for methyl-L-cysteine or L-cysteine (data not shown). These data clearly identify Ngo0372 as a cystine binding protein displaying high affinity and high substrate specificity. The K_d value of Ngo0372 for L-cystine is in the same range (10 nM) as that reported for a cystine binding protein purified from the periplasmic fraction of *E. coli* cells¹¹ and about 10 times lower than that found for BspA of *L. fermentum* (0.2 μ M). The latter was calculated from binding of radiolabeled cystine to whole cells expressing *bspA* at their surface,¹⁹ but cystine analogs were not tested. In contrast, no binding affinities for cystine are available for the receptors of the two cystine ABC transporters identified in *B. subtilis* by genetic evidence and transport assays using whole cells.⁹ However, uptake rates for radiolabeled cystine in the presence of competitors revealed the acceptance of cystathionine, djenkolic acid, selenocystine, and methyl-cysteine by the TcyJKLMN transporter, while no such data are available for the TcyABC transporter.⁹

For Ngo2014, which first had to be subjected to a denaturation/renaturation cycle to remove pre-bound ligand (see Materials and Methods for details), high-affinity binding of L-cysteine (26 ± 1 nM, Fig. 2e) was determined, while L-serine was bound with 2000-fold lower affinity (51 ± 1 μ M, Fig. 2f).

Chapter III

These data resemble those reported for the cysteine binding protein CjaA of *C. jejuni*.¹⁸ The latter was shown to bind L-cysteine with a K_d of $0.14 \pm 0.04 \mu\text{M}$ as determined from quenching of the intrinsic tryptophan fluorescence. L-Serine was found to bind with an ~ 200 -fold lower affinity ($32.2 \pm 2.4 \mu\text{M}$).

In a previous study, Friedrich *et al.* showed that loss of protein PilT, which is required for retraction of type IV pili in *N. gonorrhoeae*, results in upregulation of Ngo2011/2012/2013/2014 and Ngo0372/0373/0374.²¹ Interestingly, addition of glutamine reversed this effect, suggesting to the authors that both transporters might be specific for glutamine. Our data, which clearly identified Ngo0372 and Ngo2014 as high-affinity binding proteins with high specificity for cystine and cysteine, respectively, strongly contradict this conclusion. Moreover, during screening for potential substrates of Ngo0372 in the initial stage of this study, glutamine, in contrast to cystine, failed to change the intrinsic protein fluorescence. Thus, the reason for the above finding remains unknown.

Description of overall structures

The crystal structures of the cystine receptor Ngo0372 (termed Cys-CysRec hereafter) and that of its complex with L-cystine were determined at 2.32 Å and 1.12 Å resolutions, respectively (Table 1). While the crystal asymmetric unit of Cys-CysRec contains eight molecules (A to H) with open conformation, that of the complex with L-cystine features only one molecule with closed conformation. The structure of the cysteine receptor Ngo2014 (termed CysRec hereafter) in complex with L-cysteine was solved at 2.26 Å resolution with three molecules with closed conformations in the crystal asymmetric unit. In the three crystal structures, the quality of the electron density maps allowed confident modeling of the complete protein except for the N-terminal segment (residues 1–37/38, 1–35, and 1–37/38 in the structures of Cys-CysRec and the Cys-CysRec/L-cystine and CysRec/L-cysteine complexes, respectively). Since the N-termini of these SBPs are thought to be implied in anchoring of the proteins to the outer leaflet of the cytoplasmic membrane, they are intrinsically flexible and unstructured and often so ill defined in the electron densities that it is not possible to model them.

The overall structures of Cys-CysRec and CysRec adopt the so-called venus flytrap fold typical of receptors associated with ABC transporters¹⁴ (Figs. 3 and 4a and c). The elongated-shaped fold consists of two globular domains (lobes). In both Cys-CysRec and CysRec, the two lobes are built on central β -sheet cores of five strands each and flanked by six and three α -helices. Big lobes (residues 39–135 and 231–275 for Cys-CysRec or residues 39–134 and 232–284 for

Chapter III

CysRec) are about one-third larger than small lobes (136–230 for Cys-CysRec and 135–231 for CysRec) and contain a protruding antiparallel two-stranded β -sheet (β_2 and β_3) closed by a hairpin loop that is proposed to play a role in the recognition and interaction with the transmembrane domains of the cognate ABC transporter.²³ At the interface of the two lobes, a two-stranded antiparallel β -sheet acts as a hinge that undergoes conformational changes upon ligand binding (Fig. 3).

Closure of the ‘Venus fly-trap’ fold of Cys-CysRec

The eight molecules that compose the asymmetric unit of the unliganded Cys-CysRec structure are organized as four homodimers (A/B, C/D, E/F, and G/H), within which two molecules embrace one another to create a globular structure. This quaternary organization of the protein was analyzed on the PISA server, which concluded that the homodimers are a crystallographic artifact rather than representing a biologically relevant association. This notion is further supported by the elution profile of the protein obtained during size-exclusion chromatography purification that clearly indicated Cys-CysRec to be a monomer (data not shown). Each of the pairs X/Y is constituted by molecules that are not conformationally equivalent because rmsd calculations showed that two conformations of type ‘I’ or ‘II’ can be defined (average rmsd on C $^{\alpha}$ atom positions of 0.24 and 0.05 Å between A, C, E, and G and between B, D, F, and H molecules, respectively, whereas this value raises to 0.85 Å when superposing chains of types I and II).

The main difference between the two conformations can be summarized as a rotation of the lobes relative to each other, the ‘type I’ chains being more ‘open’ by a 9° rotation around the hinge region when compared to the ‘type II.’ This rotation corresponds to an ~ 20% closure of the fold as calculated using the webserver DynDom.²⁴ The analysis of the thermal *B*-factor distribution along the protein sequence reveals that part of the fold of type II molecules, corresponding to region 140–200 (large portion of the small lobe), displays higher *B*-factors (average value of 86.8 Å² compared to 25.0 Å² for the rest of the molecule). The extreme flexibility of this region of type II molecules results from molecular packing in the crystal, which leaves more freedom for this region.

The structures of the open conformation (type I) of Cys-CysRec and of the closed conformation of the Cys-CysRec/L-cystine complex were also compared using the DynDom webserver. The movement of closure of the fold upon binding to L-cystine corresponds to a rotation of 60° of the two lobes around the hinge region constituted by residues 135–137 and 225–230 (Fig. 3). It is interesting to note that, except for the interactions mediated by the ligand (see below), the

Chapter III

closure of the fold results in the formation of few extra hydrogen bonds between the two lobes. In the Cys-CysRec/L-cystine complex, five direct hydrogen bonds connect big and small lobes, and five of these bonds cluster near the exit of the binding pocket and might contribute to “lock” the fold in the closed conformation (Table 2). In the same region, an extra set of hydrogen bonds links the two lobes and includes two water molecules that bridge the side chain of Arg123 to the main-chain and side-chain atoms of Thr164 and Glu169. This limited number of interactions between the two domains may explain why the protein is observed in an open conformation when in the absence of L-cystine.

A similar analysis of the interface between the CysRec lobes reveals that, in this protein, the two domains are connected by 11 direct hydrogen bonds and 5 water-molecule-mediated patches. Four of these direct hydrogen bonds cluster at the exit of the binding pocket as observed for Cys-CysRec, and the other interactions are spread all over the interface (Table 3).

Following the recent structural classification of substrate-binding proteins,²⁵ Cys-CysRec and CysRec belong to cluster group F of Class II SBPs. This cluster comprises many amino acid binding proteins including those for positively charged amino acids and glutamine or methionine. Proteins classified in this cluster contain relatively long hinges, associated with more flexibility between the open and closed conformations of these SBPs. The ligand-induced conformational change (60° rotation of the lobes) described here for Cys-CysRec is similar to that of other solute receptors belonging to this class. Indeed, DynDom calculations using open/closed conformations of the structure of lysine-, arginine-, ornithine-binding protein [Protein Data Bank (PDB) code 2LAO/1LST] from *S. Typhi*, glutamine-binding protein from *E. coli* (PDB code 1GGG/1WDN), and compatible SBP from *Archaeoglobus fulgidus* (PDB code 1SW1/1SW5) revealed 51.5°, 55.7°, and 58° rotations of the lobes upon ligand binding, respectively.

Binding mode of L-cystine to Cys-CysRec

Located between the two lobes, the binding site of Cys-CysRec for cystine is almost completely buried and sequesters the co-crystallized L-cystine molecule. The electron density maps revealed that the disulfide bridge connecting the two cysteines had been partially damaged/reduced by X-radiation. L-Cystine is tightly bound to the protein *via* 14 hydrogen bonds that split in two sets of 6 and 8 interactions, respectively, specifically recognizing the amino acid functions of the L-cystine molecule. While one amino acid function forms hydrogen bonds with Gln115, Arg123, and Asn166 and three water molecules, the other one is bound to Glu56, Ser162, Leu184, and Asn201

Chapter III

and to two water molecules (see Fig. 4b and Table 4 for interaction distances). In addition, hydrophobic interactions may occur between the disulfide bridge and the aromatic side chains of Tyr59, Tyr167, and Trp97 and electrostatic interactions between the positively charged amino group of L-cystine and the π electron cloud of the phenyl ring of Tyr59.

Results from ITC experiments reveal lower affinities of the Cys-CysRec toward L-cystine analogs compared to L-cystine itself (Fig. 2). In the cases of shorter molecules (L-cysteine or *S*-methyl-L-cysteine) lacking one of the amino acid functions, no binding could be detected. These molecules might not be able to induce the closure of the venus flytrap fold, which is essential for a perfect binding of the ligand. In the other cases, since both L-cystine and the tested analogs share identical terminal amino acid functions, the differences observed in their binding constants result from the structural variation at the central part of the molecules. Thus, it appears that the accommodation of Se atoms (L-selenocystine) that are somewhat bulkier than sulfur atoms (atomic radius of 100 pm for sulfur *versus* 115 pm for selenium) affects binding only moderately (1.5 times higher K_d). On the other hand, when only one of the sulfurs is substituted by a methylene group (L-cystathionine), the replacement of the hydrophobic and easily polarizable sulfur atom and partial disruption of van der Waals interactions in the central part of the binding pocket has a dramatic effect on the affinity that decreases to the micromolar level. The differences between C–S and S–S bonds (bond length, angle) might also result in an overall shorter conformation of L-cystathionine compared to L-cystine, which could prevent correct positioning of the amino acid functions in their respective sites.

The lower binding affinity observed with djenkolic acid, which contains an extra methylene group between the two cysteines, may result from the increased size of the ligand that the binding pocket cannot accommodate. The fine-tuned hydrogen bonding network recognizing both amino acid functions is disturbed, thus leading to increased binding constants probably through an imperfect closure of the protein fold.

Altogether, these observations suggest that the binding pocket of Cys-CysRec uses in the first place the amino acid functions at both ends of L-cystine to recognize its binding partners, augmented by hydrophobic and van der Waals interactions mediated notably by Tyr59, Trp97, and Tyr167 (see below) at the center of the binding pocket.

Chapter III

Mutational analysis of the cystine binding pocket.

The role of aromatic side chains, in particular Tyr59 and Trp97, in complexing the ligand was further analyzed by site-directed mutagenesis. Variants of Cys-CysRec, carrying mutations of Tyr59 and Trp97, respectively, to phenylalanine, were constructed, purified, and monitored for substrate binding by ITC. The results clearly demonstrate that Trp97 can be replaced by phenylalanine without affecting the binding affinity of the mutant for cystine ($K_d = 23 \pm 4$ nM) and selenocystine ($K_d = 39 \pm 4$ nM) as compared to wild type (Table 5). In contrast, the Tyr59Phe mutation resulted in a decreased affinity for both ligands (60- and 30-fold, respectively). Since the cation- π interaction provided by the phenyl ring is preserved in this mutant, the consequence of the mutation appears to be the loss of the short (2.7 Å) interlobe H-bond between Tyr59 and Asn201 (Table 2). However, since the Trp97Phe mutation, which also disrupts one (longer 3.0 Å) interlobe H-bond, did not affect the affinity, it is likely that the H-bond between Tyr59 and Asn201 is essential to orientate these two side chains and to create the high-affinity binding site rather than to lock the structure in a closed conformation.

Moreover, the affinity of the double mutant (Tyr59Phe, Trp97Phe) for cystine was even stronger affected, most likely by a slightly negative synergistic entropic effect, observed in the ITC measurements. The affinities for the weak binders, L-cystathionine and L-djenkolic acid, were also further decreased, and like wild-type, no binding was detected for *S*-methyl-L-cysteine (Table 5). Together, we conclude that Tyr59 is crucial for high-affinity binding of L-cystine.

Binding mode of L-cysteine to CysRec.

Fourier difference maps of the CysRec/L-cysteine structure revealed clear electron density that was attributed to an L-cysteine molecule. Note that, in contrast to cystine, which is not stable in the (reducing) cytosol of the *E. coli* host cells, cysteine co-purified with CysRec as it was also reported for the cysteine binding protein from *C. jejuni*.¹⁸ The *N. gonorrhoeae* CysRec/L-cysteine complex exhibits a network of 11 hydrogen bonds/salt bridges between L-cysteine and 7 residues of the binding site (i.e., Lys58, Arg102, Asn117, Thr119, Arg124, Thr167, and Asp205) and a water molecule (see Fig. 4d and Table 4 for interaction details). A comparison with the structure of the cysteine binding protein from *C. jejuni* (1XT8) shows that the structures are very similar (rmsd of 0.92 Å), and the same binding mode of the ligand is found (Fig. 4d). The water molecule is hydrogen bonded to the side chains of Asn163 and Asn187 and to the S^γ atom of L-cysteine that also accepts one hydrogen bond from the positively charged side chain of Lys58.

Chapter III

Comparison of the binding-pockets for L-cystine and L-cysteine.

Sequence alignment of Cys-CysRec and CysRec (Fig. 1) reveals 28% identity, and their secondary structure elements are highly conserved with an rmsd value on C ^{α} positions of 1.8 Å. As both Cys-CysRec and CysRec are involved in the recognition and transport of amino acids, some common features are observed in their binding pockets, reflecting the almost identical, very high affinities for their natural ligands. One of the most obvious features concerns the recognition of the amino and carboxylic acid functions that L-cystine and L-cysteine have in common. In both proteins, the carboxylic acid function is primarily recognized *via* the side chain of the essential Arg124 or Arg123, respectively, for CysRec and Cys-CysRec, that establishes a bidentate salt bridge with the carboxylate of L-cystine and L-cysteine, respectively. A second contribution to the carboxylic acid binding is made by the main-chain N–H group of Thr167 or Asn166 in the L-cysteine or L-cystine complex, respectively. The recognition mode of the amino function is also very similar and implies as hydrogen bond acceptor, the main-chain carbonyl oxygens of Asn117 or Gln115, in CysRec and Cys-CysRec, respectively.

The network of hydrogen bonds plays a major role in binding selectivity with respect to L-cysteine and L-cystine (Table 4). While L-cystine forms 9 hydrogen bonds to Cys-CysRec and 4 to water molecules (the two cysteine moieties of cystine form 8 and 6 hydrogen bonds, respectively), for L-cysteine, the number of hydrogen bonds to the residues of the binding pocket is 10 and 1 to a water molecule. It is important to note that, in the case of the L-cysteine complex, the thiol side chain also contributes to the H-bond network. In the case of the L-cystine complex, however, no hydrogen bonds are mediated by the sulfur atoms that are rather implied in van der Waals interactions. These interactions were shown from binding assays to play a crucial role in ligand binding.

However, the interactions of the easily polarizable sulfur of cysteine and cystine come into play as well. As shown by Reid *et al.*, sulfur interacts favorably with Phe, Trp, and Tyr, with sulfur not located directly over the aromatic ring but preferentially shifted toward its periphery and at 4.5–6 Å from the center of the aromatic ring.²⁶ In CysRec, the distance between the cysteine S and Phe55 is 4.59 Å, and there are, in addition, several hydrogen bond interactions with S acting as acceptor and the S–H group as donor. In Cys-CysRec, there are no hydrogen bonds between the protein and the disulfide group, but this is compensated by S–aromatic interactions with Tyr59, Trp97, and Tyr167 at 4.56, 4.80, and 5.6 Å distances, respectively, from one of the two S, suggesting considerable stabilization of the intermolecular interaction.

Chapter III

Conclusions

The here presented data describe for the first time the structural basis and specificity of an L-cystine receptor associated with a putative L-cystine ABC transporter from the human pathogen *N. gonorrhoeae*. Compared to the X-ray structure of an L-cystine receptor from the same organism, also reported in this communication, complex formation with the ligand involves hydrogen bonds and hydrophobic and van der Waals interactions.

Due to their general stability, high binding affinities, and substrate specificities, solute receptors have been proven to be suited for use as biosensors.²⁷ With the structure at hand, the L-cystine receptor might be engineered for such a purpose. In fact, the cystine receptor of *E. coli* has been used in the past to determine intracellular cystine concentrations, for example, in fibroblasts of patients suffering from cystinosis.^{[28] and [29]} Cystinosis is a rare autosomal recessive disorder characterized by an accumulation of intralysosomal cystine due to a defect in cysteine transport across the lysosomal membrane.³⁰ Our results will allow improvement of the sensitivity of the assay by exchanging an amino acid close to the binding cavity of Cys-CysRec for a cysteine residue that permits the attachment of a fluorophore.²⁷ Furthermore, since earlier experiments demonstrated that cystine limitation leads to reduction in virulence of *N. gonorrhoeae*,³ Cys-CysRec might also be used as target for an antimicrobial vaccine.³¹

Material and Methods

Chemicals

L-cystathionine, L-cysteine, L-cystine, L-djenkolic acid, S-methyl-L-cysteine, L-selenocystine, L-serine were purchased from Sigma-Aldrich.

Gene cloning

The *ngo0372* gene was amplified from chromosomal DNA of *N. gonorrhoeae* FA1090 by PCR using a 1:1 mixture of *Taq* polymerase (New England Biolabs) and *Pfu* Reprofast (Genaxxon). The primers used to amplify the gene were 5'-CATATGGCCGCGGCGGTTCGGAAGGCGGCAGCGGAGC-3' (forward) with a NdeI site (in bold) and 5'-GGATCCTTATTGAACACTGATGTCTTTTCCGAAGAATTG-3' (reverse) with a BamHI site (in bold). The resulting DNA fragment translates into a protein lacking the signal sequence (amino acid residues 1-17 as predicted by SignalP 3.0 at

Chapter III

<http://www.cbs.dtu.dk/services/SignalP/>) and carrying a substitution of Cys19→Ala. First, the fragment was ligated with pGEM-T vector DNA (Promega) yielding an intermediate plasmid from which it was excised by treatment with NdeI and BamHI, and subsequently inserted into pET19b (Novagen), previously digested with the same enzymes. The resulting plasmid was named pFSA122. Due to primer design, the translated protein is fused to a His₁₀-tag followed by an enterokinase site at its N-terminus. Gene *ngo2014* was cloned accordingly, resulting in plasmid pFSA121. Primers used for amplification were 5'-CATATGGTTCGGTCTGACCGCCGCCGGGGCGGCTCCGGCG-3' (forward) and 5'-GGATCCTTATTCGGCCAACAATGCTTCCGGTTTGAC-3' (reverse). As a consequence, the translated protein begins with Val18 and carries a Cys23→Ala mutation. Both, Ala19 and Ala23, are far from the substrate binding site and consequently do not significantly influence the binding affinity of the receptors.

Protein purification

For overproduction of Ngo0372 and Ngo2014, cells of *E. coli* strain BL21 Star™ (DE3) (plysS) (Invitrogen) were transformed with plasmids pFSA122 and pFSA121, respectively. Cells were grown in Luria–Bertani medium supplemented with ampicillin (100 µg ml⁻¹), chloramphenicol (20 µg ml⁻¹) and 0.5 % glucose at 30 °C. Gene expression was induced with 0.5 mM IPTG at A₆₅₀ = 0.5, and growth was continued for 3 hours at 30 °C or for 16 hours at 25 °C.

Subsequently, cells were harvested, resuspended in buffer 1 (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 20 % glycerol (v/v), 0.1 mM PMSF) supplemented with protease inhibitor cocktail (complete EDTA-free, Roche) and DNase I, and desintegrated by passage through a French press or by ultrasonication. These and all following purification steps were carried out at 4 °C. Cell debris were removed by ultracentrifugation at 100,000 x g for 90 min, and the resulting supernatant was subjected to metal-affinity chromatography using a 5 ml TALON-HiTrap resin and the ÄKTA™ purifier FPLC system (GE Healthcare, Freiburg, Germany). The matrix was equilibrated with 5 column volumes (CV) of buffer 2 (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20 % glycerol, 0.1 mM PMSF). After sample loading, the matrix was washed with 10 CV of buffer 2 supplemented with 30 mM imidazole, and elution of retained proteins was carried out with 10 CV of buffer 2 containing 300 mM imidazole. All buffers were filtered and degassed before use, and chromatography was performed at a flow rate of 3 ml/min at 4 °C. Subsequently, protein-containing fractions were pooled, concentrated by a centrifugal device (Centrikon YM 10, Millipore), and passed through a

Chapter III

PD10 desalting column (GE Healthcare). The His₁₀-tag was cleaved off by digestion with enterokinase (Enterokinase Cleavage Capture Kit; Novagen®). The resulting tagless protein was further purified by size-exclusion chromatography in 10 mM Tris, pH 7.5, and 150 mM NaCl, using a Superdex S75 column (Amersham Biosciences), and subsequently concentrated as above. The final protein concentration was determined with the bicinchoninic acid protein assay reagent (Pierce) using bovine serum albumin as a standard.

To remove pre-bound ligand, Ngo2014 (0.75 mg/ml) was denatured by exposure to 6 M guanidine HCl in a dialysis bag at 4 °C overnight and subsequently renatured by extensive dialysis against 5 x 2 liters of buffer 3 (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 % glycerol). The renatured binding protein was concentrated by a centrifugal device (Centrikon YM 10, Millipore) followed by ultracentrifugation at 200,000 x g for 30 min.

Site-directed mutagenesis

Mutations were introduced in the *ngo0372* gene (on plasmid pFSA122) using the QuikChange or QuikChange lightning mutagenesis Kit (Stratagene) according to the manufactures protocol.

Crystallization and X-ray diffraction data collection

Crystals of Cys-CysRec (Ngo0372) and CysRec (Ngo2014) were obtained using commercial Hampton research crystallization screens with hanging-drop vapor diffusion at 291 K. The drops were produced by mixing 2 µl of protein solution (in 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl) and 2 µl of the screen solution. Crystals of ligand-free Cys-CysRec and of the complexes Cys-CysRec/L-cystine and CysRec/L-cysteine were grown using protein solutions at 10, 97, and 8 mg/ml, respectively. To obtain the complex Cys-CysRec/L-cystine, we added 1 mM L-cystine to the protein solution (due to endogenously bound ligand, no L-cysteine was added at any time of the purification or crystallization process of CysRec). While the precipitation solution of the ligand-free Cys-CysRec was composed of 20–30% polyethylene glycol 3350 and 100 mM Na-citrate, pH 3.5, the one for the Cys-CysRec/L-cystine complex contained 1.6 M ammonium sulfate and 100 mM Na-citrate, pH 4.0. CysRec/L-cysteine was crystallized using a screen solution containing 16–18% polyethylene glycol 4000, 100 mM Hepes-Na (pH 7.0), and 10 mM ZnCl₂. Further analyses of the CysRec/L-cysteine structure indicated the essential role of Zn²⁺ cations in mediating the crystallographic contacts between Asp, Glu, and His side chains of the different

Chapter III

monomers constituting the crystal asymmetric unit. In all cases, crystals of suitable size grew within 4–7 days and were used directly for X-ray diffraction experiments.

Prior to X-ray diffraction data collection, crystals were cryoprotected using their own mother liquor supplemented with 20% glycerol and flash frozen in liquid nitrogen. Data collections were performed at 100 K at either BL14-1 or BL14-2 beamline at BESSY (Berlin). Data were integrated and scaled using XDS and XSCALE.³² Data collection parameters and data set statistics are presented in Table 1.

Phasing, refinement and analysis of the structures

The structures were solved using molecular replacement³³ and search models Yckb from *B. subtilis* (pdb code 2IEE) and CjaA from *C. jejuni* (pdb code 1XT8) in the case of Cys-CysRec and CysRec, respectively. To take into account the relative flexibility of the two lobes constituting these proteins, all molecular replacement searches were performed considering the lobes independently. Refinement of the atomic coordinates and thermal agitation factors against experimental data were performed with refmac5³⁴ using in each case NCS restraints and TLS parameters obtained from the TLSMD server. Manual rebuilding of the protein models was performed with coot.²⁶ R/R_{free}-factors converged at 21.7/26.7, 10.9/14.2, and 18.5/23.3 %, for Cys-CysRec, Cys-CysRec/ L-cystine, and CysRec/ L-cysteine, respectively. The quality of the models was evaluated using the Molprobity server.³⁵ The refinement statistics of the models are presented in Table 1.

Binding assay

The binding affinities of Ngo0372 and previously de- and renatured Ngo2014 for the indicated amino acids and analogs were determined by high-sensitivity microcalorimetry using a VP-ITC device (MicroCal Software, Northampton, MA, USA). To avoid air bubbles, solutions were degassed under vacuum prior to use. Protein was filled in the reaction cell at a concentration of 50 μM in 50 mM Tris-HCl, pH 7.5 (Ngo0372) or buffer 3 (Ngo2014), and titrated in 30 steps of 10 μl each against stock solutions of L-cysteine, L-cystine, L-selenocystine (0.4 mM each), L-cystathionine (0.6 mM), L-djencolic acid, and L-serine (1 mM each) in 5-min intervals as indicated. With L-serine, Ngo2014 was used at 20 μM. Power peaks were integrated, and the resulting reaction heats were plotted against the molar substrate/protein ratio and fitted using the ‘one-set-of-sites’ model with origin 5.0 (MicroCal Software), yielding the dissociation constant K_d. The first

Chapter III

injection was always excluded from evaluation because it usually suffers from sample loss during the mounting of the syringe and the equilibration preceding the actual titration. Reaction temperature was 37 °C.

Data bank accession numbers

Atomic coordinates and structure factors have been deposited in the PDB, Research Collaboratory for Structural Bioinformatics, Rutgers University (New Brunswick, NJ)^a under codes 3ZSF, 2YLN, and 2YJP for Cys-CysRec, Cys-CysRec/L-cystine, and CysRec/L-cysteine, respectively.

Acknowledgements

We thank Alexandra Friedrich (MPI fuer Infektionsbiologie, Berlin, Germany) for a kind gift of *N. gonorrhoeae* chromosomal DNA, Charlotte Jakobi (Humboldt Universitaet zu Berlin) for contribution in the initial stage of the project, Tobias Werther (Humboldt Universitaet zu Berlin) for assistance with ITC measurements, Oliver Daumke (Max-Delbrück-Zentrum für Molekulare Medizin, Berlin) and Holger Dobbek (Humboldt Universitaet zu Berlin) for making an ITC device available to us and for helpful discussions, and Gabriele Brune (Humboldt Universitaet zu Berlin) for technical assistance. We acknowledge access to beamline BL14.1 and 14.2 of the BESSY II storage ring (Berlin, Germany) *via* the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universitaet zu Berlin, the Max-Delbrueck Centrum and the Leibniz-Institut fuer Molekulare Pharmakologie. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 449, subprojects Z3 and B14).

Chapter III

References

1. Gerbase, A. C., Rowley, J.T., Heymann, D. H., Berkley, S. F. & Piot, P. (1998) Global prevalence and incidence estimates of selected curable STDs. *Sex. Transm. Infect.* 74 (Suppl. 1):S12–S1687.
2. Seib, K.L., Wu, H.-J., Kidd, S.P., Apicella, M.A., Jennings, M.P. & McEwan, A.G. (2006) Defenses against oxidative stress in *Neisseria gonorrhoeae*: a system tailored for a challenging environment. *Microbiol. Mol. Biol. Rev.* 70, 344–361.
3. Keevil, C.W., Major, N.C., Daviess, D.B. & Robinson, A. (1986) Physiology and virulence determinants of *Neisseria gonorrhoea* grown in glucose-, oxygen- or cystine-limited continuous culture. *J. Gen. Microbiol.* 132, 3289-3302.
4. Hung, J., Cooper, D., Turner, M.S., Walsh, T. & Giffard, P.M. (2003). Cystine uptake prevents production of hydrogen peroxide by *Lactobacillus fermentum* BR11. *FEMS Microbiol. Lett.* 227, 93-99.
5. Sagara, J., Miura, K. & Bannai, S. (1993) Cystine uptake and glutathione level in fetal brain cells in primary culture and in suspension. *J. Neurochem.* 61, 1667–1671.
6. Susanto, I., Wright, S.E., Lawson, R.S., Williams, C.E. & Deneke, S.M. (1998) Metallothionein, glutathione, and cystine transport in pulmonary artery endothelial cells and NIH/3T3 cells. *Am. J. Physiol.* 274, L296–L300.
7. Hosie, A.H.F. & Poole, P.S. (2001) Bacterial ABC transporters of amino acids. *Res. Microbiol.* 152, 259–270.
8. Eitinger, T., Rodionov, D. A., Grote, M. & Schneider, E. (2011). Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol. Rev.* 35, 3–67.
9. Burguière, P., Auger, S., Hullo, M.-F., Danchin, A. & Martin-Verstraete, I. (2004) Three different systems participate in L-cystine uptake in *Bacillus subtilis*. *J. Bacteriol.* 186, 4875–4884.
10. Baptist, E.W. & Kredich, N.M. (1977) Regulation of L-cystine transport in *Salmonella typhimurium*. *J. Bacteriol.* 131, 111-118.
11. Berger, E.A. & Heppel, L.A. (1972) A binding protein involved in the transport of cystine and diaminopimelic acid in *Escherichia coli*. *J. Biol. Chem.* 247, 7684-7694.
12. Krämer, R. (1994) Functional principles of solute transport systems: concepts and perspectives. *Biochim. Biophys. Acta* 1185, 1-34.
13. Higgins, C.F. (1992) ABC transporter: from microorganisms to man. *Annu. Rev. Cell Biol.* 8, 67-113.
14. Quioco, F.A. & Ledvina, P.S. (1996) Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* 20, 17–25.

Chapter III

15. Orelle, C., Ayvaz, T., Everly, R.M., Klug, C.S. & Davidson, A.L. (2008) Both maltose-binding protein and ATP are required for nucleotide-binding domain closure in the intact maltose ABC transporter. *Proc. Natl. Acad. Sci. USA* 105, 12837–12842.
16. Davidson, A.L., Dassa, E., Orelle, C. & Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* 72, 317–364.
17. Bordignon, E., Grote, E., and Schneider, E. (2010) The maltose ATP-binding cassette transporter in the 21st century – towards a structural dynamic perspective on its mode of action. *Mol. Microbiol.* 77, 1354-1366.
18. Müller, A., Thomas, G.H., Horler, R., Brannigan, J.A., Blagova, E., Levdikov, V.M., Fogg, M.J., Wilson, K.S. & Wilkinson, A.J. (2005) An ATP-binding cassette-type cysteine transporter in *Campylobacter jejuni* inferred from the structure of an extracytoplasmic solute receptor protein. *Mol. Microbiol.* 57, 143–155.
19. Hung, J., Turner, M.S., Walsh, T. & Giffard, P.M. (2005) BspA (CyuC) in *Lactobacillus fermentum* BR11 is a highly expressed high-affinity L-cystine-binding protein. *Curr. Microbiol.* 50, 33–37.
20. Suzuki, H., Koyanagi, T., Izuka, S., Onishi, A. & Kumagai, H. (2005) The *yliA*, *-B*, *-C*, and *-D* genes of *Escherichia coli* K-12 encode a novel glutathione importer with an ATP-binding cassette. *J. Bacteriol.* 187, 5861–5867.
21. Friedrich, A., Arvidson, C.G., Shafer, W.M., Lee E.-H. & So, M. (2007) Two ABC transporter operons and the antimicrobial resistance gene *mtrF* are *pilT* responsive in *Neisseria gonorrhoeae*. *J. Bacteriol.* 189, 5399–5402.
22. Sutcliffe, I.C. & Russell, R.R.B. (1995) Lipoproteins of gram-positive bacteria. *J. Bacteriol.* 177, 1123–1128.
23. Yao, N., Trakhanov, S., and Quioco, F.A. (1994) Refined 1.89-Å structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport / chemosensory proteins. *Biochemistry* 33, 4769-4779.
24. Qi, G. & Hayward, S. (2009) Database of ligand-induced domain movements in enzymes. *BMC Struct Biol.* 6, 9-13.
25. Berntsson, R.P., Smits, S.H., Schmitt, L., Slotboom, D.J. & Poolman, B. (2010). A structural classification of substrate-binding proteins. *FEBS Lett.* 584, 2606-17.
26. Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallogr.* 60, 2126-2132.
27. De Lorimier, R.M., Smith, J.J., Dwyer, M.A., Looger, L.L., Sali, K.M., Paavola, C.D., Rizk, S.S., Sadigov, S., Conrad, D.W., Loew, L., and Hellinga, H.W. (2002) Construction of a fluorescent biosensor family. *Protein Sci.* 11, 2655–2675.

Chapter III

28. Oshima, R.G., Randall, A., Willis, C., Furlong, C.E., and Schneider, J.A. (1974) Binding assays for amino acids. The utilization of a cystine binding protein from *Escherichia coli* for the determination of acid-soluble cysteine in small physiological samples. *J. Biol. Chem.* 249, 6033-6039.
29. Tietze, F., Rome, L.H., Butler, J.D., Harper, G.S., and Gahl, W.A. (1986) Impaired clearance of free cystine from lysosome-enriched granular fractions of I-cell-disease fibroblasts. *Biochem. J.* 237, 9-15.
30. Thoene, J.G. (1995) Cystinosis. *J. Inher. Metab. Dis.* 18, 380-386.
31. Garmony, H.S., and Titball, R.W. (2004) ATP-binding cassette transporters are targets for the development of antimicrobial vaccines and therapies. *Infect. Immun.* 72: 6757–6763
32. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26, 795-800.
33. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658-674.
34. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Cryst. D* 53, pp. 240–255.
35. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B. 3rd, Snoeyink, J., Richardson, J.S., and Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375-383.
36. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., and Wang, X. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.*, 35, pp. W375–W383
37. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994;22:4673–4680.

Chapter III

Tables

Table 1: Datasets and refinement statistics

Data collection	NgoCys-CysRec 3ZSF	NgoCys-CysRec/L-cystine 2YLN	NgoCysRec/L-cysteine 2YJP
Wavelength (Å)	0.91841	0.91841	0.9395
Space Group	P 2 ₁	P 2 ₁	P 2 ₁ 2 ₁ 2 ₁
Unit cell (Å)	a=74.1; b=99.8; c=128.8; β=91.5°	a=35.7; b=63.4; c=44.6; β=96.9°	a=58.7; b=91.6; c=158.3
Resolution (Å)	48.0 – 2.32	31.0 – 1.12	47.2 – 2.26
Highest shell of resolution (Å)	2.38 – 2.32	1.15 – 1.12	2.38 – 2.26
Unique reflections	81,031 (5,936)	75,069 (5,026)	39,913 (4,974)
Redundancy	3.8 (3.6)	3.8 (2.7)	4.1 (2.7)
Completeness (%)	99.3 (99.7)	99.0 (90.7)	97.8 (86.7)
Rsym (%)	10.6 (58.1)	5.6 (53.9)	10.6 (49.4)
Mean I/σ(I)	9.9 (2.45)	16.3 (2.3)	11.3 (2.25)
Refinement statistics			
R _{work} /R _{free} (%)	21.7/26.7	10.9/14.2	18.5/23.3
Number of protein atoms	14,106	2,250	5,662
Number of water molecules	579	298	358
Average B factor (Å ²)	26.2	8.2	26.1
rms bond length (Å)	0.018	0.023	0.023
rms bond angle (°)	1.564	2.137	1.946

The values in parentheses are for the highest resolution shell.

Chapter III

Table 2: Summary of protein-ligand hydrogen-bond interactions

Ligand atoms	Cys-CysRec	Distance (Å)	CysRec	Distance (Å)
N ¹	Gln115-O	2.8	Asn117-O	2.6
	Asn166-O _{δ1}	2.9	Thr119-O _{γ1}	2.8
	Wat2029	3.0	Asp205-O _{δ2}	2.7
	Wat2173	3.3	-	-
O ¹	Arg123-N _{η1}	2.8	Arg102-N _{η1}	3.0
	Wat2106	2.7	Arg124-N _{η1}	2.8
	-	-	Thr119-N	2.8
OXT ¹	Arg123-N _{η2}	2.9	Arg102-N _{η2}	3.0
	Asn166-N	2.9	Thr167-N	3.0
	-	-	Arg124-N _{η2}	2.9
S _γ ¹	-	-	Lys58-N _ζ	3.1
	-	-	Arg102N _{η1}	-
	-	-	Wat2072	3.2
N ²	Glu56-O _{ε1}	2.7	-	-
	Wat2026	2.8	-	-
O ²	Leu184-N	3.1	-	-
	Asn201-N _{δ2}	2.8	-	-
	Wat2026	3.4	-	-
OXT ²	Ser162-O _γ	2.7	-	-
	Wat2169	2.7	-	-

Table 3: Inter-lobe hydrogen bond interaction of Cys-CysRec

Big Lobe	Distance [Å]	Small Lobe
TYR-59 (O _η)	2.7	ASN-201 (O _{δ1})
TRP-97 (N _{ε1})	3.0	THR-164 (O _{γ1})
ASP-98 (O _{δ2})	2.9	THR-164 (N)
ASP-98 (O _{δ2})	2.7	THR-164 (O _{γ1})
ASP-98 (O _{δ1})	3.2	LEU-163 (N)

Chapter III

Table 4: Inter-lobe hydrogen bond interaction of CysRec

Big Lobe	Distance [Å]	Small Lobe
PRO-270 (O)	3.3	LYS-214 (N _ζ)
VAL-271 (O)	2.9	LYS-214 (N _ζ)
TYR-272 (O _η)	2.6	ALA-207 (O)
LYS-58 (N _ζ)	2.6	ASP-205 (O _{δ2})
ASP-57 (O _{δ2})	3.2	THR-188 (N)
ASP-57 (O _{δ2})	3.3	THR-188 (O _{γ1})
ASP-57 (O _{δ2})	2.9	ASN-187 (N)
ASP-57 (O _{δ1})	3.4	ASN-187 (N _{δ2})
GLU-98 (O _{ε2})	2.6	GLN-186 (N _{ε2})
ARG-102 (N _{η2})	3.3	GLY-165 (O)
ARG-124 (N _{η2})	3.0	GLY-165 (O)

Table 5: Substrate binding affinities of Cys-CysRec wild type and variants determined by isothermal titration calorimetry

Ligand	Cys-CysRec variants			
	Wild type	Trp97Phe	Tyr59Phe	Trp97Phe/Tyr59Phe
	(μM)			
L-cystine	0.021 ± 0.001	0.023 ± 0.004	1.2 ± 0.1	2.6 ± 0.2
L-selenocystine	0.033 ± 0.006	0.039 ± 0.004	1.0 ± 0.1	2.9 ± 0.1
S-methyl-L-cysteine	-	n.a.	n.a.	-
L-cystathionine	0.962 ± 0.022	n.a.	n.a.	91.7 ± 4.1
L-djencolic acid	14.7 ± 1.6	n.a.	n.a.	261 ± 88

n.a.; not analysed

Chapter III

Figure legends

Fig. 1. Sequence alignment of Ngo0372, Ngo2014 and closely related solute receptors.

Sequence alignment of Ngo0372, Ngo2014, and closely related solute receptors. The proteins considered are TcyJ (cystine binding protein of *B. subtilis*, accession number NP_390816), TcyK (cystine binding protein of *B. subtilis*, accession number NP_390815), TcyA (cystine binding protein of *B. subtilis*, accession number P42199), BspA (cystine binding protein of *L. fermentum*, accession number AAC45332.1), and CjaA (cysteine binding protein of *C. jejuni*, accession number CAA71822.1). Conserved and homologous amino acids are indicated by red shading. Residues Tyr59 and Trp97 of Ngo0372, which were mutated in this study, are shown in green boxes. Alignment was performed with CLUSTAL W2, and the figure was drawn using EPSript (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).³⁷

Fig. 2. Substrate binding affinities of Ngo0372 and Ngo2014 determined by isothermal titration calorimetry. Substrate binding affinities of Ngo0372 and Ngo2014 determined by ITC. Ngo0372 (a–d) and Ngo2014 (e and f) were titrated with L-cystine (a), L-selenocystine (b), L-cystathionine (c), L-djencolic acid (d), L-cysteine (e), or L-serine (f). Integrated and normalized heats of reaction Q were plotted against the ligand/protein molar ratio R and fitted using a one-site binding model. The structure of each ligand is indicated in the upper part of the respective graph.

Fig. 3. Overall structure presentation of Cys-CysRec in ligand free, open conformation.

Closure movement of the venus flytrap fold of Cys-CysRec (Ngo0372). The ligand-free open conformation and L-cystine-bound close conformation of Cys-CysRec are represented in cartoon mode in red and blue, respectively. The big (N-terminal) lobes of the two conformations were superposed onto each other to illustrate the large conformational change in the smaller (C-terminal) lobe upon L-cystine binding (indicated by an arrow). L-Cystine (ball-and-stick representation) is shown in the binding pocket. Molecular graphics were generated by using PyMOL (www.pymol.org).

Fig. 4 Overall structure presentation (a, c) and detailed views of the ligand binding sites (b, d). The structures of Cys-CysRec/L-cystine (a) and CysRec/L-cysteine (c) are presented in cartoon mode and colored based on secondary structure elements. Both receptors consist of two

Chapter III

lobes that close around the bound ligand, resembling a venus flytrap. The binding pocket is located in the deep cleft between the two lobes and sequesters the ligand from the exterior. **(b)** Binding pocket of Cys-CysRec loaded with L-cystine (blue and gray, respectively). **(c)** Binding pocket of CysRec loaded with L-cysteine (in orange) on which the cysteine binding pocket of CjaA from *C. jejuni* (PDB code 1XT8) (yellow) was superposed. Superposition of both cysteine binding pockets reveals a perfect match of the amino acid sequence and conformations as well as a conservation of the water molecule position in both structures. Relevant amino acids of the binding pockets are shown in stick mode, and hydrogen bonds (or salt bridges), as broken lines. Distances of the interactions are reported in [Table 4](#).

Chapter III

Figures

Fig. 1.

```

Ngo0372 1  . . . . .MM LKKFVLGGIAALVLA A CGGSEGGSG ASSAPAQSAISGSLIERI NNK GTVT VGT ETE GTYA
TcyA 1  . . . . .M KKALLALFMVVSIA A LAAC . . GA GNDNQSKDNAKDGDLDWASI KKK GVL T VGT ETE GTYE
BspA 1  . . . . .M MKFWKKALLTVAAL TLGTS . . . AG ISSIHAASS . . . AVNSEL KKK GELT IGLE GTYS
TcyJ 1  . . . . .M NKRKG . LVLLSVF A LLG . . . . . G GCSQTNNKT . . . . . DRQA QTVI VGT GTD FFP
TcyK 1  . . . . .M KTKTAFMAILFSLIT VLISA . . . C GAGSQTTGAG . . . . . QKKVQTI VGT GTQ FFP
Ngo2014 1  MKLNAK LKALLASAAIAVGLT A CGGSGSDAQ SSQSSGAAT . . . . . VAAI KEK GVI RIG VFGD KDP
CjaA 1  MK . . . K MLLSIFTTFVAVFLA A CGG . . . . . N NSDSGASNS . . . . . LERI KQD GVR IGV FGD KDP
    
```

```

Ngo0372 61  PFTYHDKD . GKLTGYDVE VTRAVAEK LG . . . V KVEF KET QWDS M MAG LKAGRF DVV ANQVGLT S
TcyA 57  PFTYHDKD TDKLTGYDVE VITEVAKR LG . . . L KVD F KET QWDS M FAG LNSKRF DVV ANQVGT .
BspA 53  PYSYRKNN . . KLTGFVDL GKAVAK MG . . . L KAK F VPTKWDS LVAG LGSG KY DVV MN ITQT .
TcyJ 45  NIAFLNEK . GE L TGYDIE VMK AID K ELPQ . . YTFE F KTM D FSN LLTS LGNK KI DVI AH NMAKN .
TcyK 48  N ICFIDEK . GD L TGYDVELI K ELDKR LPH . . Y KFT F KTM E FSN LLVS LGQH K V D I V AH Q MEKS .
Ngo2014 60  PFGYVDAN . GKNQGF DVE IAKD LAKD LLGSPD KVEF VLTEAANRVEYVRS GKVDLI LANFTQT .
CjaA 51  PFGYVDEK . GVNQGYDIV LAKRIAK EL LGDEN KVEF VLV E AAN RVEF LKSN KVD I I LANFTQT .
    
```

```

Ngo0372 121  P ERQAT FDK . SE P Y SWSGAV LVAHND SN . I IKSIA D I KGV KTAQSLTS NYGEKAK AAG . . . . .
TcyA 117  . DREDK YDF . SDK Y TTSRAV VVTKKD NND I I KSEAD V K GK TSAQSLTS NYNKLATNAG . . . . .
BspA 111  P ERAKQ YNF . ST P Y IKS RFAL IVPAD SK . I I KSLK D I K GK KVVAGTGT NNADI AK KFG . . . . .
TcyJ 105  KEREKR FLYHKV P Y NYSPMY I TVKED NHK I I HTLKD L HGK TVIVGATS NAADYIT KYNKTHGSP I
TcyK 108  KEREKK FLF NKVA Y NHFPLK I TVLQ NND T I I R G I E D L K GK RVITSATS NGALVLK KWNEDNGRPF
Ngo2014 122  P ERAEAVDF . AD P Y MKVALG VVSPKN KP . I I TDMA Q L K D Q TLLVNKGT TADAFFT K SHP . . . . . E
CjaA 113  P ERAEQ VDF . CL P Y MKVALG VAVPQD SN . I I S S I E D L K D K TLLLNKGT TADAYFT K EYP . . . . . D
    
```

```

Ngo0372 176  AQLVPV DGLA Q SLT LIE QKR A DA TLNDELAVL DYL KKN PNAGVKI VWSAPADEKVG SGLIV NKG
TcyA 172  AKVEGV EGMA QALQMI Q QGR VDM TYNDKLAVL NYL KTS GNKNVKIAFET . . GEPQST YFTFR KG
BspA 166  AKLVPNGDFASAVGMIK QGR A DA EINSREAWY AYS KKN STKGLKMIDVSSEQKPAK ISALFN KK
TcyJ 169  HLKYAG QGSND TANQIET GR A DA TIATP . . FAVDF QNK THAFRQKT VGDVLLDTE VYFMFN KG
TcyK 172  EIAYEG QGAN E TANQLKS GR A DA TISTP . . FAVDF QNK TSTIKEKT VGNVLSNAK VYFMFN KN
Ngo2014 179  VKLLKF DQNT E TFDALK DGR GVA LAHDNALLW AWA KEN PNFEVAIGNLG . . PAEF IAPAV QKG
CjaA 170  IKTLKY DQNT E TFAALI DQR G DA LSHDNTLLF AWW KEH PEFKMAIKELG . . NKDV IAPAV KG
    
```

```

Ngo0372 240  NDEAVAKFST A INE LKA DG . . . . . TLKRLG EQF FGD I SVQ . . . . .
TcyA 234  SGEVV DQV NKA LKE MKE DG . . . . . TLSKISKKW FGD VSK . . . . .
BspA 230  DTAIQSSYNKALKE LQQ DG . . . . . TVKRLSEKY FGD I TE . . . . .
TcyJ 230  SQTILA DDT DQA I KK L EK NG . . . . . TLKRLSRKWL GAD YSKSSF EK . . . . .
TcyK 233  EQTLS DDI DKA LQE I ID DG . . . . . TLKRLSLKWL G DYSKEQY . . . . .
Ngo2014 240  NADLLNWVNGEIAAMKK DGR LKAAYE K TLLPV YGEKV KPEALLAE . . . . .
CjaA 231  DKELK E F I D N L I T K L GE EQ F F H K A Y D E T L K S H F G D V K A D D V V I E G G K I
    
```

Fig. 2

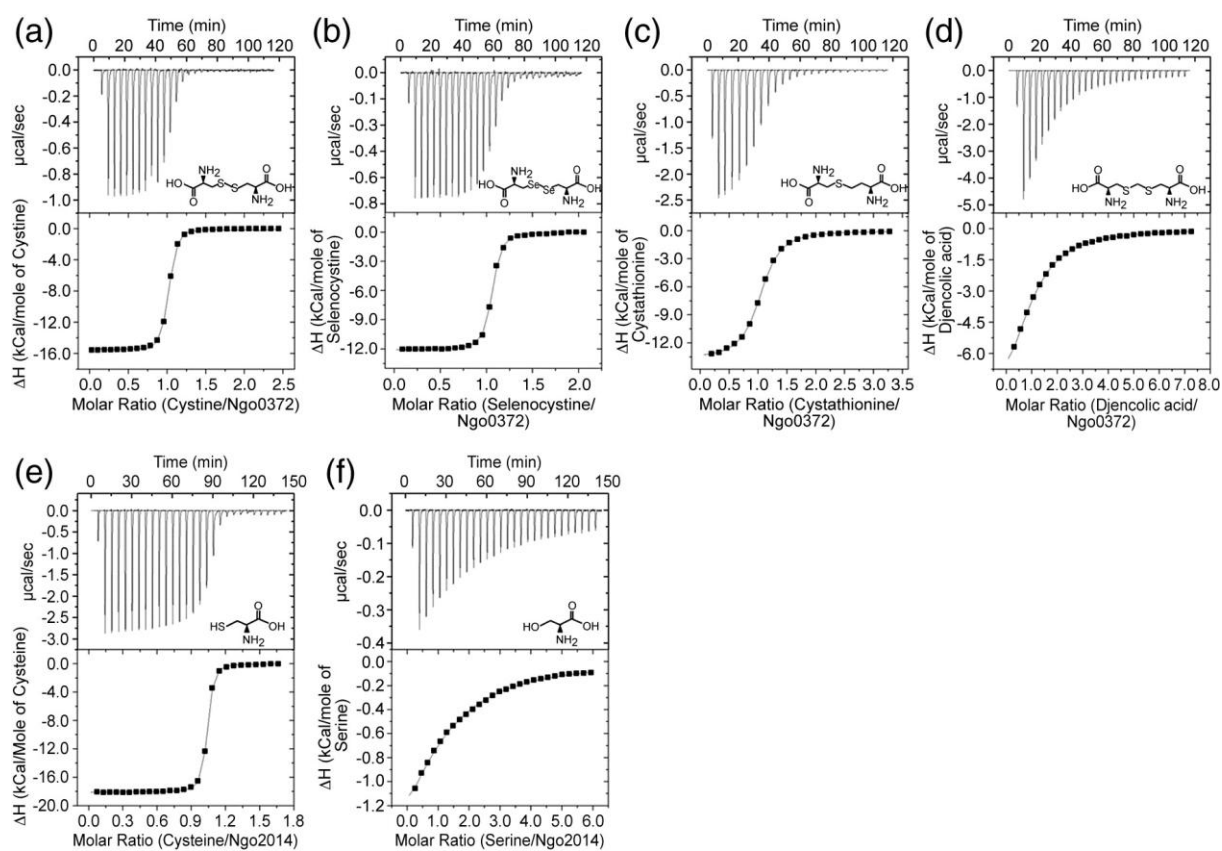
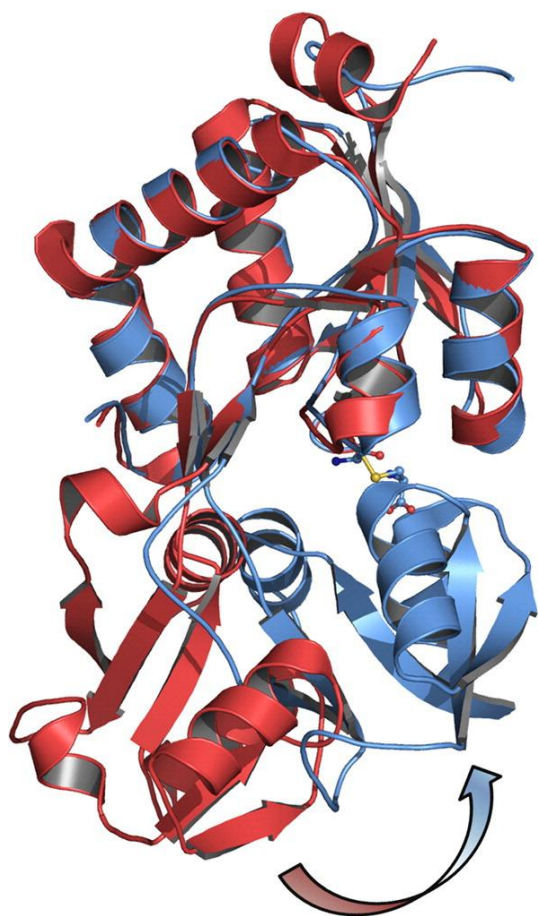
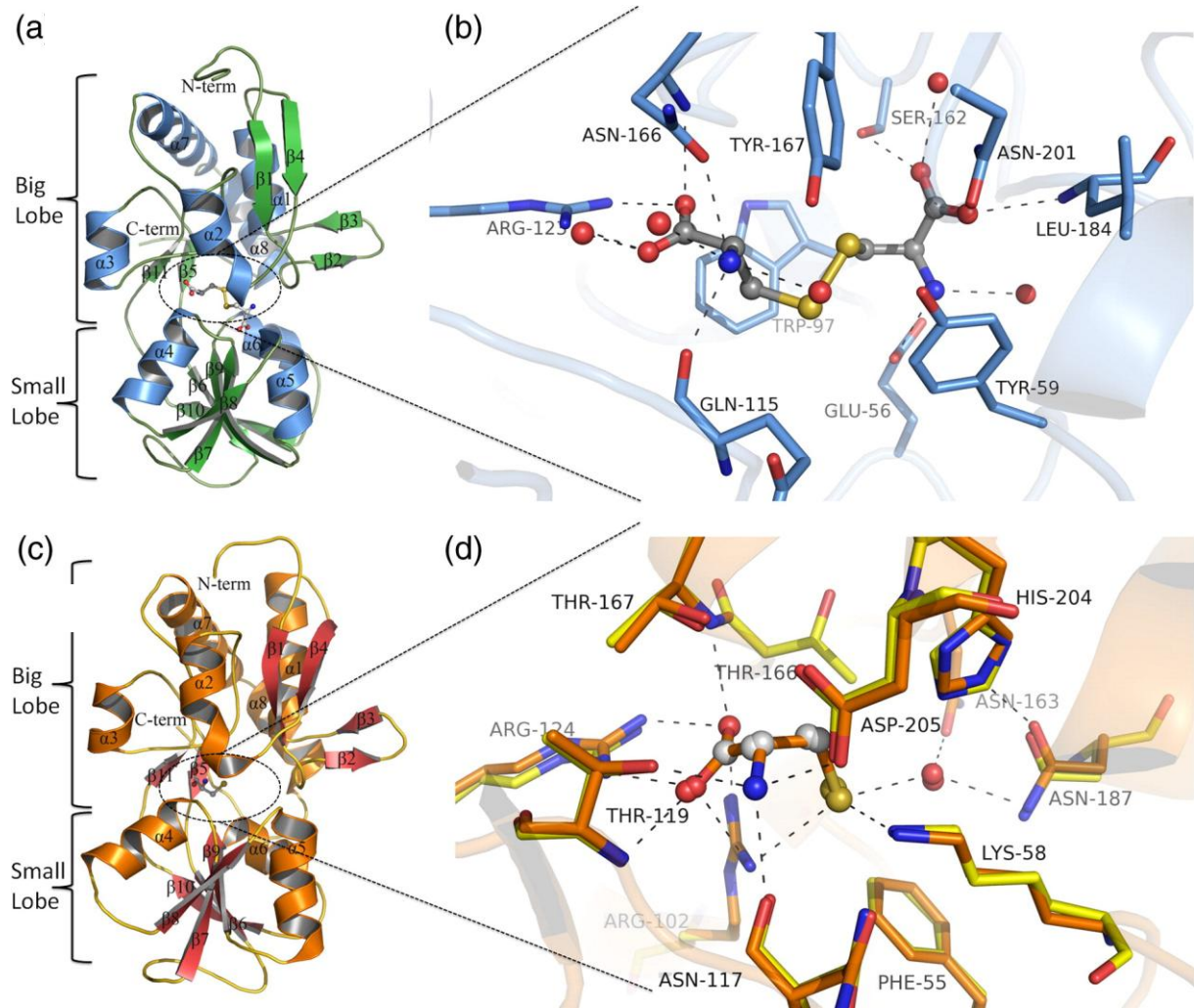


Fig. 3.



Chapter III

Fig. 4.



12. Discussion

Solving X-ray structures of proteins requires successes at different steps, from protein production to purification, from crystal growth to data collection and finally structure determination. Statistics from a structural genomic consortium indicate that only 3.5 % of cloned targets end up with a three-dimensional structure (Chayen *et al.*, 2008), and this portion is even smaller in case of membrane proteins due to their amphipathic nature. Detergents used for isolation and purification of membrane proteins hardly mimic their natural lipid environment (Ostermeier and Michel, 1997). In addition, the here considered ABC transporters exhibit inherent conformational heterogeneity, which could be a further obstacle to obtain well ordered crystals. To ensure conformational homogeneity, non-hydrolysable ATP analogs such as AMP-PNP (Adenylyl-imidodiphosphate) are often used during crystallization trials. However, mutations in the ABC subunits of transporters were found to be effective during crystallization. In case of the maltose transporter, a single variation introduced in the ATP binding subunit locked the whole transporter complex in a unique configuration that provided useful crystals (Oldham *et al.*, 2007). A similar strategy was used during this thesis for the L-Cystine transporter NGO 0373-74. However, although this variant was shown to be inactive and potentially locked in a unique configuration, it was not possible to crystallize this transporter.

Without a crystal structure, the initial steps of structural studies (purification and characterization of a protein) remain mostly unpublished, which leaves many experiences with membrane proteins unknown. During my PhD work, several ABC transporters could be expressed and purified in significant amounts. Despite the lack of well-diffracting crystals, these transporters were structurally characterized using biochemical methods. Especially the arginine transporter ArtM₂P₂ was shown by gel filtration to form a monodisperse homodimer in solution, which was stable for more than one month at 4°C but resisted crystallization. This transporter could be suitable material for the formation of 2-D crystals that would be a subject for electron crystallography based structure determination (Wisedchaisri *et al.*, 2011).

On the other hand, the determination of the X-ray structure of the water-soluble receptors of ABC transporters is a relatively promising target. Since ABC-importers in bacteria require a binding protein or receptor for the uptake of corresponding substrates, detailed structural information of these receptors in complex with their substrates directly

Discussion

reveals the substrate specificity of the associated ABC-transporter. In the first chapter of this thesis, the X-ray structures of GacH are described with acarbose-homologs (high affinity) and acarbose (weak affinity), in addition, the structure of MalE obtained in complex with acarbose. In Chapter II, the molecular basis of sharp discrimination of the AcbH towards the monosaccharide galactose is elucidated and in the last Chapter, the first structural evidence about the import of L-cystine by an ABC-transporter using the receptor, Cys-cysRec, is reported and compared with L-cysteine binding receptor, CysRec, from same organism.

The diversity of substrates in the environment of bacteria requires a large repertoire of diverse substrate importers and the associated substrate-binding receptors. Attempts to classify substrate-binding receptors based on sequence alignment did not give clearcut solution, however, structure based classification revealed a correlation between secondary structure topology of the receptors and substrate groups (Berntsson *et al.*, 2010). Based on this structural classification, carbohydrate-binding receptors and cysteine-/cystine- binding receptors were found in different subgroups of ABC transporters. Cysteine/cystine binding receptors structurally classify with other amino acid binding receptors in cluster F that contain a longer linker between the commonly found two lobes, which confers high flexibility during substrate trapping (refer to section 2.3. for other clusters). Although the secondary structure elements are conserved among the receptors comprising cluster F, the amino acid sequences, which form the unique and specific binding pockets for their substrates, show considerable variation.

When the binding mode of the CysRec for L-cysteine is compared with that of the Cys-CysRec for its dimeric variant L-cystine, significant differences are revealed. While L-cysteine is exclusively stabilized by hydrogen bonds with the CysRec, the centre part of L-cystine formed by the disulfide bridge is surrounded by hydrophobic and aromatic residues of Cys-cysRec. Sulfur atoms are not located directly over the aromatic ring but preferentially shifted towards its periphery with 4.5 to 6 Å distance from the center of the aromatic rings of Tyr59 and Trp97. We performed an amino acid sequence alignment of several L-cystine binding proteins, which indicated that the hydrophobic Tyr59 and Trp97 contribute to the tight binding by forming hydrophobic interactions with the easily polarizable disulfide of L-cystine. Exchanging Tyr59 versus Phe59 resulted in a dramatic decrease in substrate binding affinity. Several L-cystine analogs were also tested by ITC experiments, and differences in the binding affinities pointed to the importance of the disulfide group for ligand recognition.

Discussion

Carbohydrate binding receptors, on the other hand belongs to subcluster (D-I), contain shorter linkers, and only minor structural movements of the two lobes occur upon ligand binding as is shown by the open and closed conformations of galactose receptor AcbH. The analyses of the binding modes of acarbose, galactose and maltose binding proteins from different organisms exhibit high similarities. The hydroxyl groups of pyranose sugars are involved in multiple hydrogen bonds within the binding pocket. In addition, van der Waals interactions are commonly found between tryptophan and/or phenylalanine residues of carbohydrate binding pockets and nonpolar portions of sugar rings.

The roles of hydrogen bond interactions were further investigated in the galactose binding receptor by exchanging essential amino acids that form hydrogen bonds with the O4 hydroxyl of galactose using molecular biology methods. The obtained results indicated an essential role of the hydrogen bond network for substrate recognition and binding, as even minor atomic displacements completely abolished the binding and led to the formation of ligand--free open conformation structures of receptor AcbH.

Further comparison of the ligand binding mode of Cys-cysRec/CysRec with related lysine/arginine/ornithine-binding protein from *Salmonella typhimurium* (pdb code: 1LST), arginine-, lysine-, histidine-binding protein from *Geobacillus stearothermophilus* (pdb code: 2Q2A) and glutamine binding receptor from *Escherichia coli* (pdb code: 1GGG) indicated the essential role of an arginine residue in the substrate binding pockets, which forms a salt bridge with the carboxylic group of the substrates arginine, lysine or glutamine, respectively. The arginine residue, which is found at the surface of the protein in its open conformation is possibly involved in the first contact with the substrate and traps it into the binding pocket.

Although most of substrate-binding receptors show very high selectivity for one substrate, this appears not always to be the rule. Indeed, the receptor GacH was shown to bind several different carbohydrate ligands with similar affinities. This reduced degree of selectivity is essential for an efficient uptake of several acarbose homologs, and at high concentration even acarbose is recognized and bound by this receptor.

Acarbose is a natural agonist produced by *Streptomyces glaucescens* and targets the activity of α -amylase enzymes of competitive bacteria such as *Salmonella typhimurium*. The crystal structure of the maltose-specific receptor MalE of *Salmonella typhimurium* in complex with acarbose provide a structural evidence for uptake of acarbose by this ABC transporter. (The carbophor cycle is described on page 46). Most of the conventionally used antibiotics

Discussion

target the peptidoglycan layer in the cell wall of bacteria. However, acarbose mimics the structure of actual substrate (maltose/maltodextrin), therefore is efficiently imported into the cytoplasm by action of ABC-type importer at the expense of ATP hydrolysis, then acarbose act as inhibitors for α -glucosidases of competitor bacteria.

The here characterized cystine binding receptor could be taken as a target for the design of biosensors for the detection of L-cystine in urine. Cystinuria is characterized by the formation of cystine stones in the kidneys, ureter and bladder due to the high concentration of L-cystine in urine. A biosensor design could be possible by mixing a small amount of urine with a sample of this receptor and then monitor complex formation by a fluorescence signal due to Trp97 in the binding pocket that forms a direct contact with the disulfide bridge of L-cystine and will serve as a sensitive marker. In addition, the here described CD-spectra of the Cys-cysRec indicate that cystine binding protein forms a very stable complex with its specific ligand L-cystine (in section 5.4.1.). The high level of L-cystine in the urine of cystinuria patients should drive the receptor, Cys-cysRec (K_d of 21 nM) towards a more stable liganded conformation, which could be monitored by a shift in thermal stability.

13. Literature

Ames, G F., Nikaido, K., Wang, IX., Liu, PQ., Liu, CE., and Hu, C. (2001) Purification and characterization of the membrane-bound complex of an ABC transporter, the histidine permease. *J. Bioenerg. Biomem.* 33, 79–92.

Berntsson, R.P., Smits, S.H., Schmitt, L., Slotboom, D.J., and Poolman, B. (2010) A structural classification of substrate-binding proteins. *FEBS Lett.* 584,2606-17.

Biemans-Oldehinkel, E., Poolman, B. (2003) On the role of the two extracytoplasmic substrate-binding domains in the ABC transporter OpuA. *EMBO J.* Nov 17;22(22):5983-93.

Blattner, FR., Plunkett, G., Bloch, C.A., Perna, NT., Burland, V., Riley, M., ColladoVides, J., Glasner, J.D., Rode, CK., Mayhew, GF., Gregor, J., Davis, NW., Kirkpatrick, HA., Goeden, MA., Rose, DJ., Mau, B., Shao, Y., (1997). The complete genome sequence of *Escherichia coli* K-12, *Science* 277 1453–1457.

Burguière, P., Auger, S., Hullo, MF., Danchin, A., Martin-Verstraete I. (2004). Three different systems participate in L-cystine uptake in *Bacillus subtilis*. *J Bacteriol.* Aug;186(15):4875-84.

Borths, EL., Locher, KP., Lee, AT., Rees, DC. (2002) The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. *Proc Natl Acad Sci U S A.* Dec 24;99(26):16642-7. Epub Dec 10

Chen, J., Lu, G., Lin, J., Davidson, AL., Quioco, FA. (2003) A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol. Cell.* 12, 651–661.

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. (2010) **MolProbity: all-atom structure validation for macromolecular crystallography.** *Acta Crystallogr D Biol Crystallogr.* Jan;66(Pt 1):12-21. Epub 2009 Dec 21.

Cuneo, MJ., Changela, A., Warren, JJ., Beese, LS., Hellinga, HW. (2006) The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and disaccharide binding sites. *J. Mol. Biol.* 362, 259-270.

Caffrey, M. (2008). On the Mechanism of Membrane Protein Crystallization in Lipidic Mesophases. *Cryst. Growth Des.* 8, 4244-4254.

Chayen, NE. (2004). Turning protein crystallization from an art into a science. *Curr. Opin. Struct. Biol.* 14, 577–583.

Caldas, T., Demont-Caulet, N., Ghazi, A., Richarme, G. (1999) Thermoprotection by glycine betaine and choline. *Microbiology* 145, 2543-2548

Literature

Davidson, AL., Nikaido, H. (1991) Purification and characterization of the membrane associated components of the maltose transport system from *Escherichia coli*. *J. Biol. Chem.* 266, 8946–8951.

Davidson, AL., Dassa, E., Orelle, C., Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol R* 72: 317–364.

Dwyer, MA., Hellinga, HW. (2004). Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr Opin Struct Biol.* Aug;14(4):495-504.

Delcour, AH. (2002). Structure and function of pore-forming betabarrels from bacteria. *J Mol Micro Biotechnol* 4, 1-10

Dean, M., Rzhetsky, A., Allikmets, R. (2001). The human ATP-binding cassette (ABC) transporter superfamily. *Genome Res.* 11, 1156–1166.

Dawson RJP, Locher KP. (2006) Structure of a bacterial multidrug ABC transporter. *Nature.* 443:180–185

Eckey V, Landmesser H, Schneider E. Studying subunit-subunit interactions in a bacterial ABC transporter by in vitro assembly. *Biochim Biophys Acta.* 2010 Jun;1798(6):1250-3. Epub 2010 Mar 10.

Erkens, GB., Berntsson, RP., Fulyan, F., Majsnerowska, M., Vujičić-Žagar, A., Ter Beek, J., Poolman, B., Slotboom, DJ. (2011) The structural basis of modularity in ECF-type ABC transporters. *Nat Struct Mol Biol.* Jun 26;18(7):755-60.

Eitinger T., Rodionov, D.A., Grote, M., Schneider, E. (2011) Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* Apr 23.

Evans P, McCoy A. (2008). An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr.* Jan;64(Pt 1):1-10.

Felder, CB., Graul, RC., Lee, AY., Merkle, HP., Sadee, W. (1999). The Venus Flytrap of Periplasmic Binding Proteins: An Ancient Protein Module Present in Multiple Drug Receptors. *AAPS PharmSci.*; 1 (2): article 2. DOI: 10.1208/ps010202

Fjellstedt, E., Harnevik, L., Jeppsson, JO., Tiselius, HG., Söderkvist, P., Denneberg, T. (2003). Urinary excretion of total cystine and the dibasic amino acids arginine, lysine and ornithine in relation to genetic findings in patients with cystinuria treated with sulfhydryl compounds. *Urol Res.* Dec;31(6):417-25. Epub 2003 Oct 25.

Fukami-Kobayashi, K., Tateno, Y., and Nishikawa, K. (1999) Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* 286, 279-290.

Hollenstein, K., Frei, DC., Locher, KP. (2007). Structure of an ABC transporter in complex

Literature

with its binding protein. *Nature* 446:213–216.

Hollenstein, K., Dawson, R.J., Locher, K.P. (2007) Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol.* Aug;17(4):412-8. Epub 2007 Aug 27.

Higgins, C.F. (1992). ABC transporters: from microorganisms to man. *Annu Rev Cell Biol.* 8:67–113.

Hosie, AHF., Poole, PS. (2001). Bacterial ABC transporters of amino acids. *Res. Microbiol.* 152, 259–270.

Hvorup, R.N., Goetz, B.A., Niederer M., Hollenstein, K., Perozo, E., and Locher, K.P. (2007) Asymmetry in the structure of the ABC-Transporter-binding protein complex BtuCD-BtuF. *Science*, 1145950.

Hung, J., Cooper, D., Turner, MS., Walsh, T., Giffard, PM. (2003). Cystine uptake prevents production of hydrogen peroxide by *Lactobacillus fermentum* BR11. *FEMS Microbiol Lett.* Oct 10;227(1):93-9.

Hung, LW., Wang, IX., Nikaido, K., Liu, PQ., Ames, GF., Kim, SH. (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* Dec 17;396(6712):703-7.

Hunte C, Michel H. (2002) Crystallisation of membrane proteins mediated by antibody fragments. *Curr Opin Struct Biol.* Aug;12(4):503-8.

Griesshammer R, Tate CG. (1995) Overexpression of integral membrane proteins for structural studies. *Quart Rev Biophys* 28:315-422.

Garman, E. F. and Mitchell, E. P. (1996). Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions *J. Appl. Cryst.* 29, 584-587.

Greenfield, NJ. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc.* 1(6):2876-90.

Gerber S, Comellas-Bigler M, Goetz BA, Locher KP. (2008). Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science.* Jul 11;321(5886):246-50.

Chayen NE, Saridakis E. (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods.* Feb;5(2):147-53.

Karpowich NK, Huang HH, Smith PC, Hunt JF. (2003) Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. *J Biol Chem.* Mar 7;278(10):8429-34.

Kadaba N. S., Kaiser J. T., Johnson E., Lee A., Rees D. C., (2008). The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. *Science*, 321, 250– 253.

Kempf, B. and Bremer, E. (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* 170, 319-330

Literature

- Knol, J., Veenhoff, L., Liang, W. J., Henderson, P. J., Leblanc, G., and Poolman, B.** (1996) Unidirectional reconstitution into detergent- destabilized liposomes of the purified lactose transport system of *Streptococcus thermophilus*. *J. Biol. Chem.* 271, 15358–15366
- Kors, CA., Wallace, E., Davies, DR., Li, L., Laible, PD., Nollert, P.** (2009). Effects of impurities on membrane-protein crystallization in different systems. *Acta Crystallogr D Biol Crystallogr.* Oct;65(Pt 10):1062-73. Epub 2009 Sep 16.
- Lebedev, AA., Vagin, AA., Murshudov, GN.** Model preparation in MOLREP and examples of model improvement using X-ray data. (2008) *Acta Crystallogr D Biol Crystallogr.* Jan;64(Pt 1):33-9. Epub 2007 Dec 5.
- Lee, YH., Deka, RK., Norgard, MV., Radolf, JD., Hasemann, CA.** (1999) *Treponema pallidum* TroA is a periplasmic zinc-binding protein with a helical backbone. *Nat Struct Biol.* Jul;6(7):628-33.
- Licht A, Bulut H, Scheffel F, Daumke O, Wehmeier UF, Saenger W, Schneider E, Vahedi-Faridi A.** (2010). Crystal Structures of the Bacterial Solute Receptor AcbH Displaying an Exclusive Substrate Preference for β -d-Galactopyranose. *J Mol Biol.* 2010 Dec 17.
- Locher K.P, Lee A.T, Rees D.C.** (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science.* 296, 1091–1098.
- Locher, K.** (2009) Structure and mechanism of ATP-binding cassette transporters. *Philos Trans R Soc Lond B Biol Sci* 364: 239–245.
- Luecke H, Quioco FA.** (1990). High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature.* Sep 27;347(6291):402-6.
- Linton, KJ., Higgins, CF.** (1998). The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Mol Microbiol* 28: 5–13.
- Linton, KJ.** (2007) Structure and Function of ABC Transporters. *Physiology (Bethesda).* Apr;22:122-30. Review.
- Linton, KJ., Higgins, CF.** (2007). Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch* 453: 555–567.
- Iwata S,** (2003) *Methods and Results in Crystallization of Membrane Protein.* International University Line
- Murray J, Garman E.** (2002) Investigation of possible free-radical scavengers and metrics for radiation damage in protein cryocrystallography. *J Synchrotron Radiat.* Nov 1;9(Pt 6):347-54. Epub 2002 Nov 1.
- Oh BH, Ames GF, Kim SH.** (1994) Structural basis for multiple ligand specificity of the periplasmic lysine-, arginine-, ornithine-binding protein. *J Biol Chem.* Oct 21;269(42):26323-30.

Literature

Oh BH, Pandit J, Kang CH, Nikaido K, Gokcen S, Ames GF, Kim SH. (1993) Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J Biol Chem.* May 25;268(15):11348-55. *Erratum in: J Biol Chem.* Aug 15;268(23):17648-9.

Oh BH, Kang CH, De Bondt H, Kim SH, Nikaido K, Joshi AK, Ames GF. (1994) The bacterial periplasmic histidine-binding protein. structure/function analysis of the ligand-binding site and comparison with related proteins. *J Biol Chem.* Feb 11;269(6):4135-43.

Ozvegy C, Litman T, Szakács G, Nagy Z, Bates S, Váradi A, Sarkadi B., (2001). Functional characterization of the human multidrug transporter, ABCG2, expressed in insect cells. *Biochem. Biophys. Res. Commun.* 285, 111–117

Ostermeier C, Michel H. (1997) Crystallization of membrane proteins. *Current Opinion in Structural Biology*, 7:697–701

Oldham, M.L., Khare, D., Quioco, F.A., Davidson, A.L. and Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**: 515-521.

Oldham, M.L., Davidson, A.L., Chen, J. (2008). Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol.* 18(6):726-33.

Postle, K. and Kadner, R.J. (2003). Touch and go: tying TonB to transport. *Mol Microbiol.* Aug;49(4):869-82

Putman M., van Veen H.W. and Konings W.N., (2000) Molecular properties of bacterial multidrug transporters. *Microbiol. Mol. Biol. Rev.* **64**, pp. 672–693.

Pinkett H.W, Lee A.T, Lum P, Locher K.P, Rees D.C. (2007) An inward-facing conformation of a putative metal-chelate-type ABC transporter. *Science.* 315, 373–377.

Ravelli RB, Garman EF. (2006) Radiation damage in macromolecular cryocrystallography. *Current Opinion in Structural Biology.* Volume 16, Issue 5, October 2006, Pages 624-629

Parcej D, Tampé R. (2007) Caught in the act: an ABC transporter on the move. *Structure.* Sep;15(9):1028-30.

Parcej D, Tampé R. (2009) Solute-Binding Sites in ABC Transporters for Recognition, Occlusion and Trans-Inhibition. *ChemMedChem.* 2009 Jan;4(1):25-8

Parcej D, Tampé R. (2010) ABC proteins in antigen translocation and viral inhibition. *Nat Chem Biol.* Aug;6(8):572-80.

Quioco, F.A. and Ledvina, P.S. (1996) Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* 20, 17-25

Dawson, R.J., Hollenstein, K., Locher, K.P. (2007). Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Mol Microbiol.* Jul;65(2):250-7.

Literature

Rodionov, DA., Hebbeln, P., Eudes, A., ter Beek, J., Rodionova, IA., Erkens, GB., Slotboom, DJ., Gelfand, MS., Osterman, AL., Hanson, AD., Eitinger, T. (2009) A novel class of modular transporters for vitamins in prokaryotes. *J Bacteriol.* Jan;191(1):42-51. Epub 2008 Oct 17.

Rupp, B., 2010. Biomolecular Crystallography Principles, Practice, and Applications to Structural Biology, Garland Science., New York.

Ward, A., Reyes, CL., Yu, J., Roth, CB., Chang, G. (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A.* 2007 Nov 27;104(48):19005-10. Epub Nov 16.

Senior, AE, Bhagat, S. (1998) P-glycoprotein shows strong catalytic cooperativity between the two nucleotide sites. *Biochemistry.* 37, 831–836.

Sutcliffe, IC., and Russell, RR. (1995) Lipoproteins of gram-positive bacteria. *J. Bacteriol.* 177, 1123-1128.

Schiefner, A., Holtmann, G., Diederichs, K., Welte, W., Bremer, E. (2004). Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Biol Chem.* Nov 12;279(46):48270-81. Epub 2004 Aug 11.

Schneider, E., and Hunke, S. (1998) ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol. Rev.* 221-20.

Schneider, E., Eckey, V., Weidlich, D., Wiesemann, N., Vahedi-Faridi, A., Thaben, P., Saenger, W. (2011) Receptor-transporter interactions of canonical ATP-binding cassette import systems in prokaryotes. *Eur J Cell Biol.* May 9.

Schneider, E. (2001). ABC transporters catalyzing carbohydrate uptake. *Res. Microbiol.* 152, 303-310.

Saier, M., H, Jr. (2000) A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.*, 64, 354–411.

Kelly, SM., Jess, TJ., Price, NC., (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* 1751 119-139.

Shultis, DD., Purdy, MD., Banchs, CN., Wiener, MC. (2006) Outer membrane active transport: structure of the BtuB:TonB complex. *Science.* Jun 2;312(5778):1396-9.

Shukla, S., Wu, CP., Ambudkar, SV. (2008) Development of inhibitors of ATP-binding cassette drug transporters: present status and challenges. *Expert Opin Drug Metab Toxicol.* 4(2):205-23.

Sharff, A.J., Rodseth, L.E., Spurlino, J.C., and Quioco, F.A. (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* 31, 10657-10663.

Literature

- Spurlino, J.C., Lu, G.Y., Quioco, F.A.** (1991) The 2.3 Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* 266, 5202-5219.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J. Dubendorff, J.W.** (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185, 60–89
- Szakacs, G., Paterson, J.K., Ludwig, J.A., Booth-Genthe, C., Gottesman, M.M.** (2006) Targeting multidrug resistance in cancer. *Nat Rev Drug Discov*;5(3):219-34
- Trinchieri, A., Dormia, G., Montanari, E., Zanetti, G.** (2004) Cystinuria: definition, epidemiology and clinical aspects. *Arch Ital Urol Androl.* Sep;76(3):129-34.
- Taylor G.L.** (2010) Introduction to phasing. *Acta Crystallogr D Biol Crystallogr.* Apr;66(Pt 4):325-38. Epub 2010 Mar 24.
- Tolosa, L., Ge, X., Rao, G.** (2003). Reagentless optical sensing of glutamine using a dual-emitting glutamine-binding protein. *Anal Biochem.* Mar 15;314(2):199-205.
- Tomii, K., Kanehisa, M.,** (1998). A Comparative Analysis of ABC Transporters in Complete Microbial Genome. *Genome Res.* 8, 1048.
- Ubarretxena-Belandia, I., Stokes, D.L.** (2010). Present and future of membrane protein structure determination by electron crystallography. *Adv Protein Chem Struct Biol.* 81:33-60. Review.
- Vahedi-Faridi, A., Licht, A., Bulut, H., Scheffel F, Keller S, Wehmeier, U.F., Saenger, W., Schneider, E.** (2010). Crystal structures of the solute receptor GacH of *Streptomyces glaucescens* in complex with acarbose and an acarbose homolog: comparison with the acarbose-loaded maltose-binding protein of *Salmonella typhimurium*. *J Mol Biol.* Apr (2010). 2;397(3):709-23. Epub Feb 2.
- Vahedi-Faridi, A., Eckey, V., Scheffel, F., Alings, C., Landmesser, H., Schneider, E., Saenger, W.** (2008) Crystal Structures and Mutational Analysis of the Arginine-, Lysine-, Histidine-binding Protein ArtJ from *Geobacillus stearothermophilus*. Implications for Interactions of ArtJ with its Cognate ATP-binding Cassette Transporter, Art(MP)₂. *J. Mol. Biol.*, 375, 448-459.
- van der Heide T, Poolman B.** (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep.* 3:938–43
- Vasiliou, V., Vasiliou, K., Nebert, D.W.** Human ATP-binding cassette (ABC) transporter family. *Hum Genomics.* 2009 Apr;3(3):281-90.
- Vagin A. and Teplyakov A.** (2010) Molecular replacement with MOLREP *Acta Crystallogr D Biol Crystallogr.* Jan;66(Pt 1):22-5
- Wallin, E., von Heijne, G.** (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 7, 1029–1038.
- Wilkinson, A.J., and Verschueren, K.H.G.,** (2003) Crystal structures of periplasmic solute-binding proteins in ABC transport complexes illuminate their function. In: *ABC proteins: from*

Literature

bacteria to man (Holland, B., Cole, S. P. C., Kuchler, K. and Higgins, C. F., eds.), pp. 187-207, Elsevier, Amsterdam.

Wisedchaisri G, Reichow SL, Gonen T. (2011) Advances in structural and functional analysis of membrane proteins by electron crystallography. *Structure*. Oct 12;19(10):1381-93.

Zhang, Z., Feige, N. J., Chang, A. B., Anderson, I. J., Brodianski, V. M., Vitreschak, A. G., Gelfand, M. S. & Saier, M. H., Jr (2003) A transporter of *Escherichia coli* specific for L- and D-methionine is a prototype for a new family within the ABC superfamily. *Arch. Microbiol.*, 180, 88–100.

Zhang P., Wang J., Shi Y. (2010) Structure and mechanism of the S component of a bacterial ECF transporter. *Nature* 468, 717–720

List of Abbreviations

14. List of Abbreviations

- ABC**, ATP-binding cassette;
- ADP**, adenosine diphosphate;
- AMPcPP**, α , β -methyleneadenosine-5'-triphosphate;
- ATP**, adenosine triphosphate;
- CMC**, Critical micelle concentration;
- Da**, Dalton, unified atomic mass unit (u), 1.660×10^{-24} g;
- DDM**, Dodecyl-b-D-maltopyranoside;
- DM**, Decyl-b-D-maltopyranoside;
- DTT**, dithiothreitol;
- EDTA**, ethylenediaminetetraacetic acid;
- GBP**, glucose/galactose-binding protein;
- GacH**, *Streptomyces glaucescens* maltose/maltodextrin-binding protein;
- IMAC**, immobilized metal affinity chromatography;
- IPTG**, isopropyl- β -D-thiogalactopyranoside;
- LB**, Luria-Bertani broth;
- MALDI**, Matrix-assisted laser desorption/ionization
- MBP**, maltose/maltodextrin-binding protein;
- MPD**, 2-methyl-2, 4-propandiol;
- Min**, minute;
- NBD**, nucleotide binding domain;
- OD**, optical density;
- PAGE**, polyacrylamide gel electrophoresis;
- PEG**, Polyethylene glycol;
- PMSF**, phenylmethylsulfonyl fluoride;
- RMSD**, root mean square deviation;
- SBP**, solute-binding protein;
- SDS**, sodium dodecyl sulphate;
- Tris**, trishydroxymethylaminomethane;
- TMD**, Transmembrane domain;

Acknowledgement

15. Acknowledgement

This thesis arose during past years of research that has been done since I came to Saenger's group. By that time, I have worked with a number of people whose contribution made this thesis possible. It is a pleasure to convey my gratitude to them all in my acknowledgment.

My greatest appreciation goes to Prof. Wolfram Saenger for giving me the opportunity to work on very interesting projects of SFB-499, I am grateful for his tremendous support, encouragement and guidance from the initial to the final level.

Many thank to the members of former Saenger Group especially Dr. Sebastien Moniot for generous helps on structural studies and writing manuscript, Dr. Ardeschir Vahedi-Faridi for sharing a great time full of interesting scientific and non-scientific discussions and Traudy Wandersleben for valuable friendship.

Further I thank Claudia Alings for her support during Crystallization trials, Clemens Langner for his support for protein purification and my colleagues Jacobo Martinez, Dr. Bernhard Loll, Dr. Qingjun Ma, Magdalena Czepnik, Karine dos Santos, Albert Guskov and Sebastian Geibel for creating friendly atmosphere of work in the lab as well as in the offices.

I am grateful to Prof. Erwin Schneider for his invaluable suggestions and fruitful collaboration of his group members especially to Dr. Frank Scheffel for sharing his experiences and plasmids and Anke Licht particularly for her contribution on binding assays.

I also thank Dr. Chris Weise for MALDI-TOF analyses, and the beamline staff of BESSY for the technical support during data collections.

I feel indebted to my great family for their unconditional moral support and endless love and always remembering my father who generated my enthusiasm for science.

16. Curriculum Vitae

For reasons of data protection,
the curriculum vitae is not included in the online version