Biophysical characterization of ABC transporters catalyzing import and export function

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine Universität Düsseldorf

vorgelegt beim Fachbereich

Biochemie

von

Britta Tschapek

aus Viersen

Düsseldorf, Februar 2011

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

1. Gutachter: Prof. Dr. Lutz Schmitt
2. Gutachter: PD Dr. Ulrich Schulte

Tag der mündlichen Prüfung:

Abstract

This work contributes to an understanding of substrate recognition and translocation performed by ABC-transporters. ABC-transporter form one of the largest protein-familiy of membrane proteins. Classically, they consist of two nucleotide binding (NBDs) and two transmembrane domains (TMDs). This architecture facilitates transport of a large variety of substrates using ATP as energy source in all three kingdoms of life. Among these import and export systems are differentiated. Import systems are so far only found in prokaryotes and archaea and require an additional substrate binding protein (SBP), which binds the substrate in the periplasm or extracellular space and delivers it to the transporter. SBPs of ABC-transporter show all a common architecture: They consist of two domains different in size, in between a deep cleft where the substrate binding-site is located. Crystal structures of various SBPs have shown mainly two conformations: one without substrate (open) and one with substrate bound (close) where the substrate is located in the cleft and the two domains have rotated towards each other. Through structural analysis of SBPs substrate specificity, mechanism of substrate binding and release could be illustrated via the example of the ABC-transporter OpuB from B*acillus subtilis* and ProU from thermophilic *Archaeoglobus fulgidus*. The contribution to substrate binding of single amino acids in the binding pocket could be resolved both qualitatively and quantitatively for both proteins. For ProX the SBP of ProU molecular dynamics simulation as well as structural analysis revealed a single amino acid that is responsible for switching the affinity of the binding pocket from a rather low affinity state to a high affinity binding site.

Biophysical investigation of full length ABC-transporter requires application of detergents to keep the membrane protein in solution. Establishment of a method for determination of critical micelle concentration (CMC) and aggregation numbers of diverse detergents in various buffer solutions was also one part of this work. Both parameters depend on pH, salt concentration, temperature and polymer concentration in the solution. Thus a reliable and fast analysis of these characteristics is an important advantage prior to start biophysical and structural analysis on membrane proteins.

With such a "toolbox" in hand establishing purification and functional analysis of single components of Type I secretion system of *Escherichia coli* were the next steps. This system, which translocates its substrate Haemolysin A in a single step across two membranes, is composed of three proteins: The ABC-transporter Haemolysin B (HlyB), a so called membrane fusion protein Haemolysin D (HlyD) and the outer membrane protein TolC. The ABC-transporter has a special feature: an additional domain at its N-terminus, which belongs to the family of C39 peptidase domains. However, in HlyB the catalytic relevant residues are mutated and therefore this C39 domain is no longer able to cleave. Homologous

overexpression and purification of the ABC-transporter HlyB and subsequent functional assays showed that ATP hydrolysis of NBDs can be stimulated by the addition of substrate (HlyA). This increase ATPase activity does not occur due to binding of HlyA to C39 domain but binding of the secretion signal of HlyA to the NBDs of HlyB. To further investigate the role of C39 domain in the transport cycle, the protein was overexpressed separately, purified and for structural analysis successfully crystallized.

Zusammenfassung

Die hier vorliegende Arbeit leistet einen Beitrag zum Verständnis der Substraterkennung und Translokation bei ABC-Transportern. ABC-Transporter verkörpern eine der größten Klassen von Membranproteinen. Ihr klassischer Aufbau aus zwei Nukleotidbindendenden (NBD)- und zwei Transmembrandomänen (TMDs) ermöglicht den Transport unterschiedlichster Substrate unter ATP-Verbrauch in allen Reichen des Lebens. Dabei werden Import und Exportsysteme unterschieden. Importsysteme sind bisher ausschließlich in Prokayoten und Archaeen gefunden worden und benötigen zusätzlich noch ein Substratbindeprotein (SBP), welches das Substrat aus dem Periplasma oder dem extracellulären Raum zum Transporter delegiert. Durch Analyse struktureller Daten dieser Substratbindeproteine wird die Substratspezifität, sowie der Mechanismus der Substratbindung und auch des Entlassen des Substrates in die Translokationspore des Transporters an den Beispielen des ABC-Transporters OpuB aus *Bacillus subtilis* und ProU aus dem thermophilen Archaeon *Archaeoglobus fulgidus* erklärt werden. SBPs von ABC-Transportern zeigen alle einen ähnlichen Aufbau: Sie bestehen aus zwei unterschiedlich großen Domänen, in deren Mitte sich die Substratbindestelle befindet. Kristallstrukturen von verschiedenen SBPs haben gezeigt, dass diese eine offene Konformation ohne Substrat und eine geschlossene Konformation, bei der Substrat in der Mitte gebunden ist und die beiden Domänen zueinander rotiert sind, einnehmen können.

Welche einzelnen Aminosäuren die Substratspezifität der Bindungstasche ausmachen, konnte mittels Analyse der Kristallstrukturen des SBPs OpuBC, das ausschließlich Cholin bindet, und einem Vergleich mit homologen SBPs mit verändertem Substratspektrum gezeigt werden. Für das ProX Protein, das SBP des ABC-Transporters ProU, konnte ein detaillierter Ablauf der Substraterkennung, Bindung und der Entlassung des Substrats in die Translokationspore des Transporters bestimmt werden.

Die biophysikalische Untersuchung gesamter ABC-Transporter erfordert die Anwendung von Detergenzien um diese Membranproteine in wässrigen Lösungen zu stabilisieren. Die Etablierung einer Methode zur Bestimmung der kritischen Mizellen Konzentration sowie der Aggregationszahl unterschiedlicher Detergenzien in verschiedensten Pufferlösungen ist ein weiterer Aspekt dieser Arbeit. Beide Parameter sind stark von pH-Wert, Salzkonzentration, Temperatur und Polymeren in der Lösung abhängig. Daher ist eine zuverlässige und schnelle Analyse dieser Parameter ein entscheidender Vorteil, bevor biophysikalische oder strukturelle Analysen an Membranproteinen durchgeführt werden.

Mit diesem „Werkzeugkoffer" konnte die Reinigung und funktionelle Analyse der einzelnen Komponenten des Typ I Sekretionssystems aus *Escherichia coli* etabliert werden. Dieses System, das das Substrat Hämolysin A in einem Schritt über zwei Membranen transportiert,

besteht aus drei Komponenten: Dem ABC-Transporter Hämolysin B (HlyB), dem Membranfusionsprotein Hämolysin D (HlyD) und dem Außenmembranprotein TolC. Dabei besitzt der ABC-Transporter noch eine zusätzliche N-terminale Domäne, die in Sequenzvergleichen sich als C39 Peptidasedomäne herausstellte. Allerdings sind die katalytisch relevanten Aminosäuren mutiert, so dass sie nicht mehr in der Lage ist ihr Substrat zu spalten. Die homologe Expression und Aufreinigung des ABC-Transporters HlyB, sowie die anschließende funktionale Analyse haben gezeigt, dass die Hydrolyseaktivität der NBDs vom Substrat HlyA stimuliert werden kann. Dabei konnte nachgewiesen werden, dass diese Stimulation nicht auf eine Bindung an die C39 Domäne hervorgerufen wird, sondern es sich um eine Interaktion des Sekretionssignal mit den NBDs handeln muss. Um die Rolle der C39-Domäne zu verstehen, wurde diese separat exprimiert und aufgereinigt und für strukturelle Analysen erflgreich kristallisiert.

1. Introduction

Cellular membranes are composed of a lipid bilayer with a hydrophobic core of about 40Å and are therefore impermeable for ions and only partially permeable for hydrophilic substances (*1*). This cell membrane is essential for bacteria, eukaryotes and archaea since it separates the intracellular space from extracellular environment or in case of eukaryotes for maintaining intra-cellular compartments. It is impossible for amino acids, nutrients, ions and other small compounds to diffuse freely across these membranes. To overcome this barrier prokaryotes, eukaryotes as well as archaea developed distinct transport systems. Referring to the source of energy used one can distinguish channels (*2*), primary and secondary transporters (*3*). Secondary active transporters couple translocation of substrate to electrochemical gradients across the membrane, preferably via a proton or sodium potential. Primary transporters such as the proton pumping ATPases or ATP-binding cassette (ABC)-transporters deploy the free energy released upon ATP hydrolysis (*4*).

1.1 ABC-Transporter

ABC-transporter are found in all three kingdoms of life and form one of the largest membrane protein superfamilies (*5-7*). ABC-transporter share a common structural organization comprising two transmembrane domains (TMDs) that form the translocation pore and two nucleotide binding domains (NBDs) that bind and hydrolyze ATP (*8*). These domains can be arranged as a single fused polypeptide chain, as half-size transporters (one TMD and one NBD fused on a single gene) or as four separately encoded proteins (although this is found solely in prokaryotic ABC importers). Import proteins require an additional component, the substrate binding protein (SBP)(*7*).

Substrates transported by ABC transporters are diverse in nature and size, like for example sugars, amino acids, peptides, vitamins, ions, xenobiotics and even polypeptides, linking these ABC-transporters to various cellular functions that range from energy supply to osmoregulation, detoxification and virulence (*9-12*). The largest substrate identified until today is a 900 kDa Protein which is secreted by *Pseudomonas fluorescens* WCS365 (*13*). This remarkable substrate diversity of ABC-transporters is reflected in sequence variance in the TMD domains, whereas NBDs share rather high sequence homology (see below).

To date, there are a small number of crystal structures of full-length ABC-transporters available (Figure 1). However only for a few of the structural characterized proteins additional biochemical information is available. For most of them even the substrates are unknown. On a structural based view, ABC-transporters can be grouped into 3 classes (*14*). Class I and II

contains ABC-importers, which are distinguished by there arrangement of transmembrane domains and the fact that type I importers contain a substrate binding site in the TMD while type II do not. Type I importer's membrane section is composed of a core of 2 x 5 transmembrane helices (Figure 1), in case of maltose permease one subunit possesses one extra helix while the other has actually two additional helices. Type II ABC-transporters exhibit around 2 x 10 core transmembrane helices and helices from each subunit form distinct and separate domains.
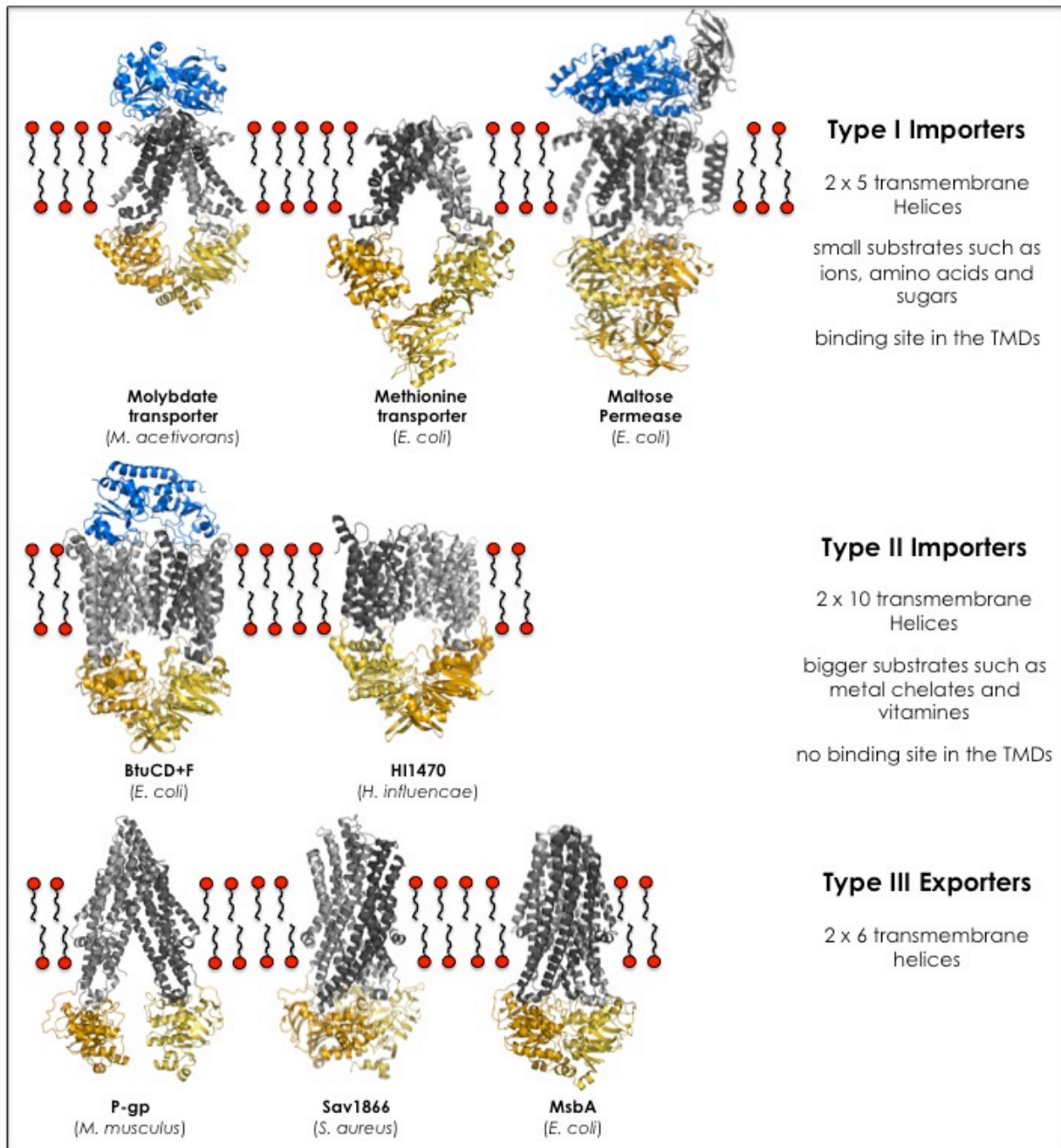

Figure 1: Crystal structures of ABC-transporters. First row: Type I ABC-Importers: ABC-transporter ModBC in complex with its SBP ModA (pdb: 2ONK), putative methionine transporting MetNi (pdb: 3DHW), maltose transporter MalFGK$_2$ in complex with its SBP MalE (pdb: 2R6G); second row: Type II ABC-transporter: BtuCD with BtuF bound (pdb: 2Q19) and HI1470 (pdb: 2NQ2); third row: Type III Exporters: P-glycoprotein from mousse (pdb: 3G60), Sav18660 (pdb: 2ONJ) and *E. coli* MsbA (pdb: 3B5W).

Figure 1 shows crystal structures of full length ABC-transporters known to date. Twining of the TMDs around a hypothetical axis perpendicular to the membrane forms the translocation pathway across the membrane. The NBDs face towards each other thereby the conserved sequence motifs (see below) are positioned at the dimer interface. The similarities in structure support a common mechanism of nucleotide and substrate dependent conformational changes that result in translocation of the substrate across the membrane. The "alternating access" model for transport suggests the presence of a substrate binding site that can alternatively access either the extracellular side or the intracellular side of the membrane, corresponding to the "outward" and "inward" facing conformation of a transporter (*15*).

The best-characterized system is the *E. coli* maltose transporter MalFGK$_2$. For more than 30 years it has been studied in different laboratories. This wealth of knowledge accumulated by genetic, molecular genetic and biochemical means was complemented by crystal structures of the isolated MalK dimer (NBDs) in four conformations as well as two conformational states of the full-length transporter more recently (Figure 2) (*16-18*).
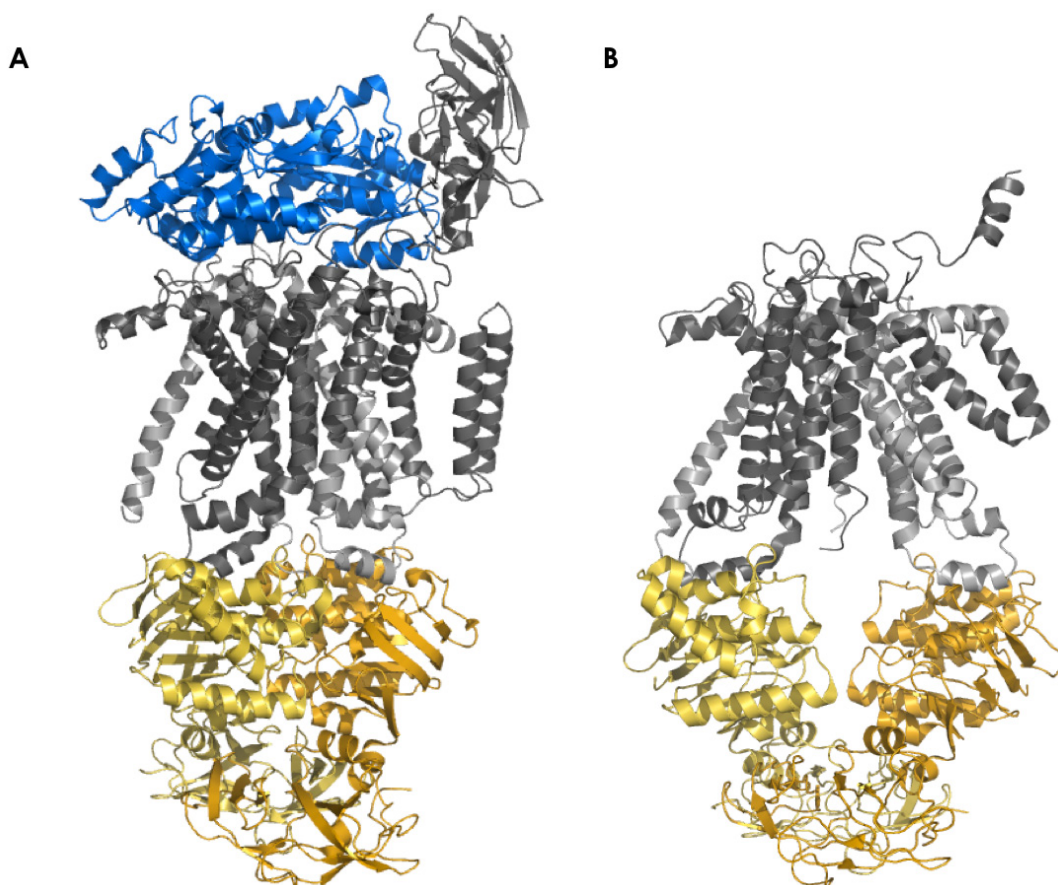

Figure 2: Crystal structures of MalFGK2 in an intermediate state (pdb: 3R6G)(A) and in the apo-state (pdb: 3FH6)(B). Color code is the following: Maltose binding protein (MalE), blue; MalG, light grey; MalF, dark grey; MalK, orange and yellow. In the periplasm maltose is bound and delivered to the transporter by MalE. There it induces a conformational change towards a closed conformation by interacting with the periplasmic loop of MalF (A). This change is communicated to the NBDs and induces ATP-hydrolysis. When ADP is present in the binding pockets of the MalK dimer it opens and forces also the TMDs to change to the cytoplasmic open state (B).

Figure 2 represents two snapshots of the transport cycle. Figure 2A shows the full ABC-transporter including its substrate binding protein in an ATP-bound state. Liganded MalE (SBP) initiates the transport cycle by binding to MalFGK$_2$. It communicates the presents of a substrate to the large periplasmic loop of MalF (Figure 2A). Thereby ATP hydrolysis is initialized by MalK and the formed ADP forces the NBD-dimer to open. This conformational change forces the TMD to change to inward facing conformation (Figure 2B) and substrate can be released into the cytoplasm. Not only crystal structures led to the understanding of this particular transport mechanism also biochemical data on single domains, as well as parts and the full-length transporter supply the required information to interpret these structures ([19-22]).

To analyze the role of the individual domains that comprises an ABC-transporter, I established purification protocols and biophysical analysis tools to get further insights in regulation of transport, requirements for functionality and structural information.

1.1.1 Nucleotide binding domains

The NBDs of ABC-transporters are also called "motor domains" as they fuel the translocation process . NBDs share a conserved architecture among all ABC-transporters, importers as well as exporters. They consist of a RecA-like and an α-helical subunit ([23, 24]), connected by two flexible loop regions, one of which contains a conserved glutamine residue (Q-loop). The RecA-like domain harvests the Walker A and B motifs, indicative of the presence of a nucleotide binding site (Walker A: G-X-X-G-X-G-K-S/T, Walker B: ϕ-ϕ-ϕ-ϕ-D, where ϕ is a hydrophobic residue)([25, 26]). Walker A motif is responsible for binding of the β- and γ-phosphate groups of ADP or ATP, while Walker B motif is arranged as a β-strand, whose terminal aspartate coordinates a Mg^{2+} ion through water needed for ATP hydrolysis ([27]). Moreover, NBDs contain an H-loop with a conserved histidine residue that contacts the γ-phosphate of ATP and is required for hydrolysis ([28]), and finally the famous "signature" motif, the C-loop (L-S-G-G-Q) ([29]), which identifies a amino acid sequence as an ABC-transporter.

NBDs function as dimers in the assembled transport protein, with the nucleotide binding sites of each monomer facing towards the other (head-to-tail arrangement). In this conformation ATP is mainly bound via residues of the Walker A motif of one monomer and the C-loop of the other monomer. One model based on experiments on MalK suggests that upon binding of ATP, the NBD dimer closes in a 'tweezer-like'-fashion ([16]). Hydrolysis of ATP occurs in the tightly dimerized state, which then reopens to release phosphate. The ADP-bound conformer of the NBD dimer resides in an intermediary, semi-open state, which presumably returns to the open apo-state after the dissociation of ADP ([17]). The observed positive cooperativity of ATP-hydrolysis ([28]) makes it highly likely that two molecules of ATP are needed for NBD dimer closure ([7, 16]).

In Chapter I the assayed NBD of the ABC-transporter OpuA from *Bacillus subtilis* differs from most other NBDs but similar to MalK in that it posses an additional C-terminal domain of approx. 150 amino acids. In the 'tweezer-like'-motion a dimer of these accessory domains act as the bottom of the tweezer and keeps the two monomers together during the catalytic cycle. Sequence alignments have shown that the accessory domains of OpuAA from *B. subtilis* as well as its homologue OpuAA from *Lactoccocus lactis* contain two cystathione β-synthetase domains (CBS)(*30*). For OpuAA from *L. lactis* it was observed that these CBS domains are involved in sensing osmotic stress and activate the transport function of OpuA (*31*).

Förster resonance energy transfer was used to analyze the conformational changes taking place during and after ATP-hydrolysis. As a model system the above-mentioned OpuA NBD from *B. subtilis was chosen*. Former publications have shown that it has dynamic monomer-dimer equilibrium, hydrolyses ATP *in vitro* and dissociates in the ADP-bound state (*32*). For fluorescence experiments single labeled OpuAA is required. As OpuAA itself has no cysteine, single mutants were introduced for labeling. Purification and characterization of several mutants resulted in two accessible labeling positions. Both mutants of OpuAA (S45C, G161C) were stable enough to work with and had remaining ATPase activity.

Förster resonance energy transfer is based on overlapping fluorescence and absorption spectra of donor and acceptor fluorophores (Figure 3A). Measuring acceptor fluorescence intensity of different mutants of OpuAA results in distance information between the labeled amino acids (Figure 3B).
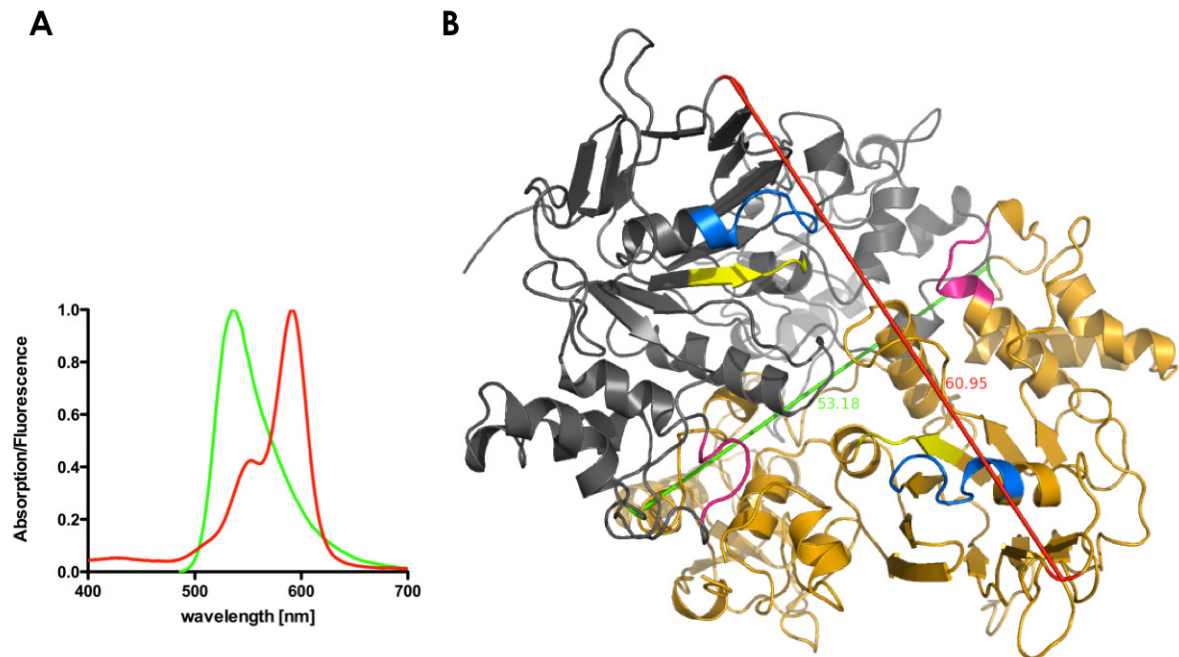
A   **B**

Figure 3: A: Spectral overlap of OregonGreen fluorescence and TexasRed absorption. These are the two fluorophores used in the study presented in Chapter I B: Homology model of OpuAA based on crystal structure of MalK (pdb: 1Q1E). Depicted is the putative ATP-bound dimeric state. Conserved motifs are highlighted and color-coded as follows: Walker A (blue), ABC-signature motif (purple), Walker B (yellow). Cysteine mutations at position S45 and G161 are accentuated in red and green, respectively. Red and green lines indicate the distances between these two residues.

As NBDs of ABC-transporters share rather high homology it was possible to build a homology model based on *E. coli* MalK and to refine it due to the distant restrains obtained with this fluorescence study.

1.1.2   Substrate binding proteins

In prokaryotes, ABC transporter are involved in the uptake of a large variety of substances, including nutrients, osmoprotectants and signal molecules. All ABC import systems require an additional component, a substrate specific receptor or binding protein for function (SBP). These substrate binding-proteins bind their ligands with high affinity either in the periplasm of Gram-negative bacteria, where they can diffuse freely between the inner and outer membrane or linked to the cell surface with a N-terminal lipid moiety in archaea and Gram-positive bacteria (*33*). For some archaea it is also observed that they are inserted in the membrane via a N-terminal hydrophobic helix and even in some bacteria SBPs are directly linked to the TMDs resulting in two or even four substrate binding sites (*34*)(Figure 4).
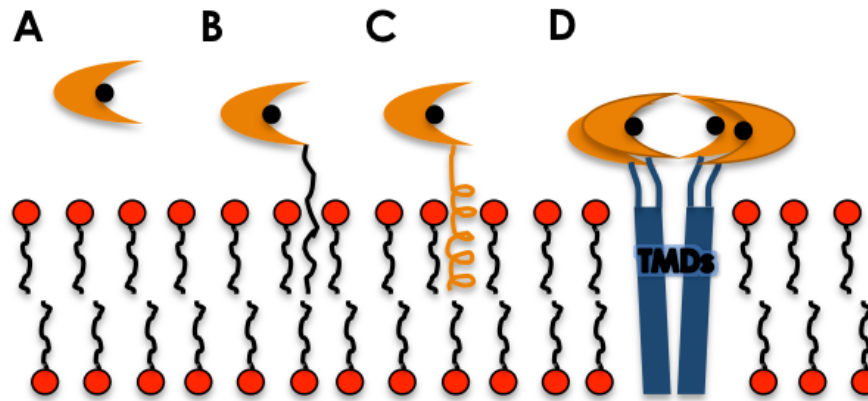
Figure 4: Structural organization of substrate binding proteins in different organisms. A: In Gram-negative bacteria SBP can diffuse freely in the periplasm. B: in Gram-positive bacteria and archeae that have no outer membrane SBPs are fused to the membrane via an N-terminal lipid modification. C: In some archaea SBPs are linked to the membrane via an N-terminal hydrophobic helix. D: In some bacteria and archeae transporters with two or four SBPs directly linked to the TMDs are found.

Despite the absence of significant sequence similarity across SBPs and the diversity of ligands bound, all of the SBPs have a common overall organization. This comprises two globular domains of similar topology, each of which contains a central β-pleated sheet flanked by sets of α-helices (Figure 5)(*35-38*). Shown in Figure 5 is the structure of glycine betaine/proline betaine binding protein ProX from *Archaeoglobus fulgidus*. It illustrates the two-lobe architecture of SBPs. The connection of the two domains is one of the differentiating factors among SBPs and is highlighted in orange.
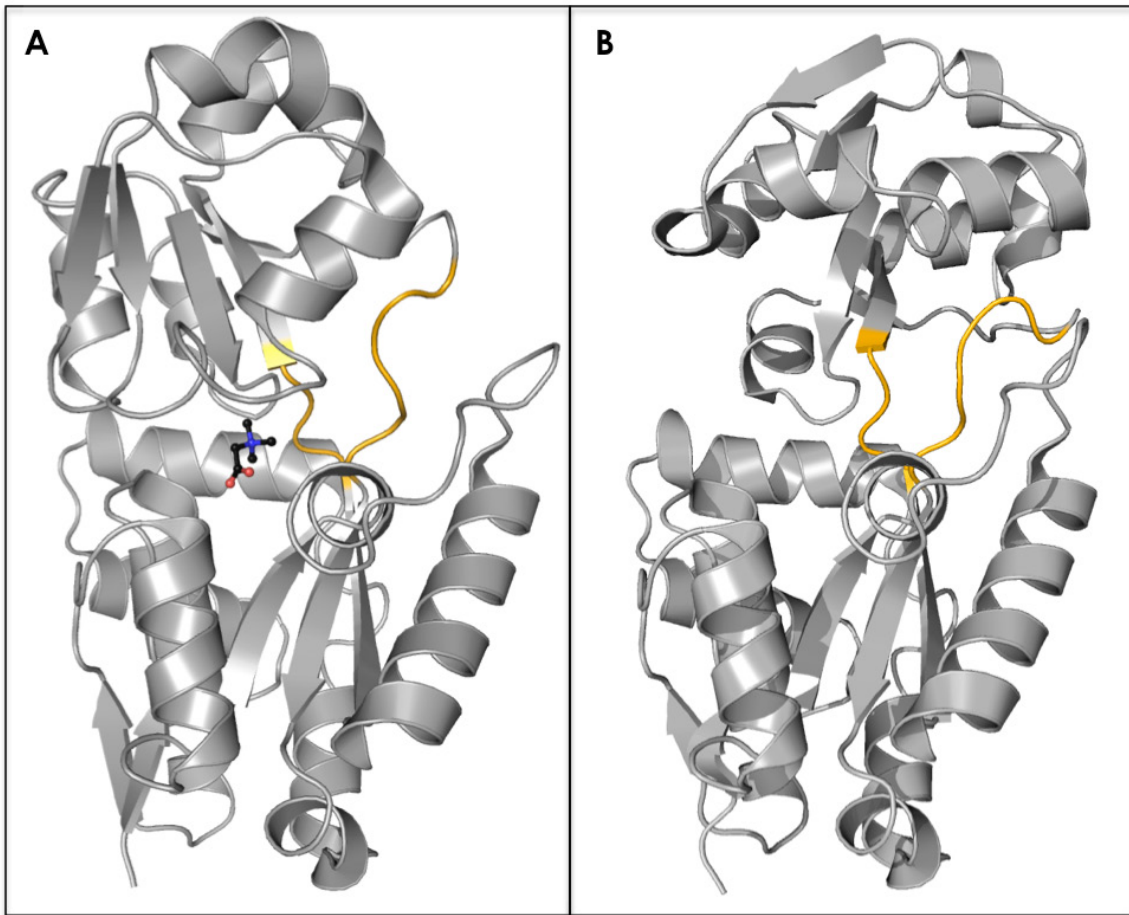
Figure 5: Overall fold of substrate binding proteins. Presented is the structure of ProX from *Archeoglobus fulgidus* in the glycine betaine bound state (A) (pdb: 1SW2) and ligand-free (B) (pdb: 1SW5) (39). The ligand is shown as stick representation. Highlighted in orange are the two loops connecting both domains.

Upon substrate binding, the two lobes twist and close, thereby entrapping the ligand, therefore the ligand binding mechanism is called Venus flytrap (*40*)(Figure 5). For one of the most analyzed SBPs the maltose binding protein (MBP) from *E. coli* it has been shown in NMR studies, that the two domains of MBP are dynamic and fluctuate around an average orientation that is only slightly more closed than this observed in the open conformation found in the crystal structure. This means that the binding of the ligand forces the two domains to close. But how this can be established on a molecular level is still unclear.

In Chapter II crystallization and subsequent structure determination of ProX Y111A from *A. fulgidus* at 2.0 Å resolution is highlighted. Surprisingly the protein adopts a rather open conformation although the substrate (glycine betaine) is bound in the binding pocket. Further analysis of the structure and molecular dynamics simulation enabled us to identify a crucial residue that is involved in triggering domain closure and also communicating with the transmembrane section for release of the substrate. When a ligand enters the binding site a network of amino acid side chains communicates the presence of it towards Arg149, which increases ligand affinity and induces domain closure. These results are corroborated by

molecular dynamics studies and support the view that a single amino acid acts as a switch to induce the high affinity state of the SBP

The observation of highly homologous SBPs with different substrate specificity lead to the question what are the molecular determinants for specificity as well as affinity of SBPs. Therefore crystal structures of OpuBC from *B. subtilis*, which is highly specific for choline betaine (CB), ProX from *A. fulgidus*, that only binds glycine and proline betaine (PB, GB) and structure of OpuCC from *Straphylococcus aureus*, which is rather unspecific and virtually binds all osmolytes, are compared in Chapter III.

1.2   Obstacles to overcome: studying ABC-transporters *in vitro*

It is challenging to analyze these membrane transporters *in vitro*. This is reflected in the rather low number of structures available until today (one archaeal, six bacterial and one eukaryotic ABC-transporters (Figure 1)). For biophysical and structural analysis one has to extract the protein from its natural environment (the lipid bilayer) and keep it "happy" in an artificial solution. Although there are several experimental opportunities to do so, it requires lots of work.

The most commonly employed method to study membrane proteins in aqueous solution is micelle-forming detergents. Detergents are amphiphilic molecules consisting of a hydrophobic and a hydrophilic portion. At concentrations below their specific critical micellar concentration (CMC), detergent molecules exist as monomers in aqueous solution. Above the CMC, multimeric detergent micelles assemble that are in a highly dynamic equilibrium with detergent monomers. A large number of detergents with a wide range of chemico-physical properties are available (for an overview see www.affymetrix.com, www.emdbioscience.com/calbiochem, www.sigmaaldrich.com). However detergents are often not optimal for the activity and stability of a membrane protein (*41*). Micelles are highly dynamic assemblies and their intrinsic instability can result in protein unfolding and aggregation. Additionally, the lateral forces acting on a membrane protein in the roughly spherical detergent micelle are significantly different from those in a native planar membrane (*42*). Summarizing all these aspects leads to the conclusion that a detergent micelle only poorly mimics the architecture and physical properties of a cellular membrane.

In many cases, the presence of detergent molecules interferes with the experimental conditions required for monitoring protein activity by enzymatic, ligand binding or spectroscopic assays. Furthermore, the particular chemico-physical requirements of a given membrane protein are usually only met by a few, if any, out of the bewildering selection of detergents available today. There is no method that could predict a compatible detergent for a given membrane protein a priori. Therefore the search for a suitable detergent that

stabilizes a functional state of the membrane protein is highly empirical as well as time-consuming.
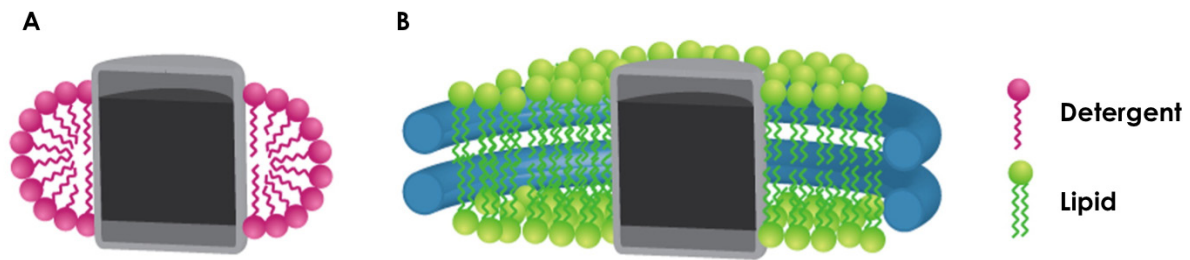

Figure 6: Schematic representation of detergent solubilized (A) and in nanodisc reconstituted membrane protein (B). The membrane protein is shown in black, detergents are colored in pink and lipids in green. Picture is taken from Raschle et al. (43).

For this reason alternative solubilisation techniques are highly desirable and are more and more in use these days. Alternatives that mimic the native lipid bilayer are liposomes and so called nanodiscs. In the case of liposomes, limited solubility, the formation of multilamellar vesicles and the inaccessibility of the vesicle interior may, however, interfere with functional investigations. Recently the self-assembly of a lipid bilayer nanodevice named nanodiscs has been established (*44, 45*). Nanodiscs consist of an assembly of phospholipids arranged as a discoidal bilayer, surrounded by amphipathic apolipoproteins (Figure 6)(*43*). In their native context, apolipoproteins assemble into roughly spherical high density lipoprotein particles, responsible for the reverse transport of cholesterol from the peripheral tissues to the liver (*46*). The native-like environment provided by the nanodisc is likely to support both protein stability and functionality of an incorporated membrane protein.

1.3 Biophysical Methods: Analyzing Detergents

After analyzing components of ABC transport systems, the next step was to investigate full-length transporters and protein complexes containing an ABC-transporter. Therefore the usage of detergents to keep them soluble in water based buffer systems is required (see section 1.2). The amphiphilic detergents cover hydrophobic patches at the surface of the protein while their hydrophilic head groups point towards the buffer containing outside.

For the analysis of membrane proteins and here especially ABC-transporters one starts with a membrane preparation of your expression host, extract the protein of interest with a high amount of detergent and then follows at least one step of protein purification. Finally pure protein can be analyzed biochemically and structurally. For every step the "right" detergent is required. If micelle size is too small, the protein might aggregate, if one uses too much detergent, detergent micelles containing the protein will accumulate maybe not reversible and thus inhibiting protein function, etc. Thus many different kinds of detergents (ionic, non-

ionic, zwitter-ionic, one alkyl chain, two alkyl chains, etc.) are available until today and still this number is increasing (see Figure 7).
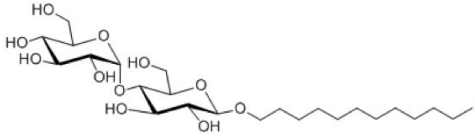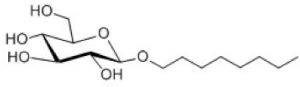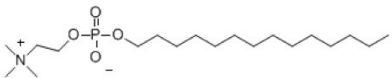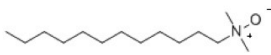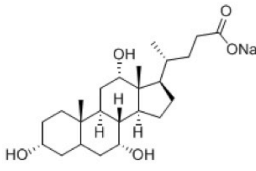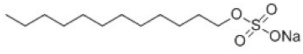
Structure	head group characteristic	Name
	non ionic	N-Dodecyl-β−D-Maltopyranoside (DDM)
	non ionic	N-Octyl-β-D-Glucopyranoside (OG)
	zwitterionic	Fos-Choline14 (FC-14)
	zwitterionic	n-Dodecyl-N,N-Dimethylamine-N-Oxide (LDAO)
	Ionic	Sodium cholate
	ionic	Sodium dodecyl sulfate

Figure 7: A few examples of commonly used detergents (structural representations are taken from www.affymetrix.com).

There are several characteristics of a protein detergent system one must keep in mind during screening for a good alternative for the selected protein. The most famous and important ones are the critical micelle concentration and aggregation number of the detergent.

The CMC is defined as the concentration of a detergent above which micelles are formed spontaneously. The other feature is the number of monomers per micelle, so called aggregation number. In Chapter IV the establishment of a robust an easy to use assay for the determination of both characteristics in a nearly high throughput manner is presented.

1.3.1   CMC determination

As an easy to measure output information the methods described uses the fluorescence of the dye Hoechst 33342, which is applied in biochemical labs for DNA staining *in vitro*. It is cell permeable and emits blue light when it is bound to DNA. For this kind of application these features were used to detect the point when micelles of detergents are formed. Therefore

there is a need for a hydrophobic compound that is soluble in water but partitions greatly and immediately into micelles when they are formed and then it should change some physical properties that can be measured. This is exactly what the fluorescence dye Hoechst 33342 offers: It diffuses greatly into micelles and does almost not emit fluorescence light in hydrophilic media while it emits bright blue light in the micellar phase (Figure 8)(*47*).
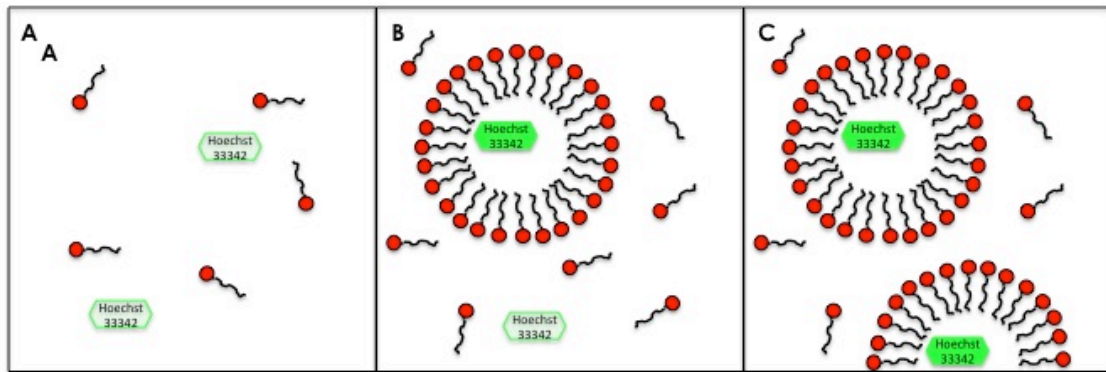
Figure 8: Hoechst Fluorescence in detergent micelles. A: Below the CMC almost no Hoechst 33342 fluorescence is detectable. B: Reaching the CMC detergent micelles form, Hoechst 33342 diffuses into the micelles and due to the hydrophobic environment inside Hoechst emits fluorescent light. C: Increasing detergent amounts lead to formation of more micelles and an increase in fluorescence signal is detectable until all Hoechst is in the micellar phase.

If one measures Hoechst 33342 fluorescence at increasing detergent amounts, one receives a curve like that depicted in Figure 2. There almost no fluorescence of Hoechst is detectable until a concentration of DDM of 0.007 % (w/v) is reached. This infection point indicates the CMC for this specific detergent.
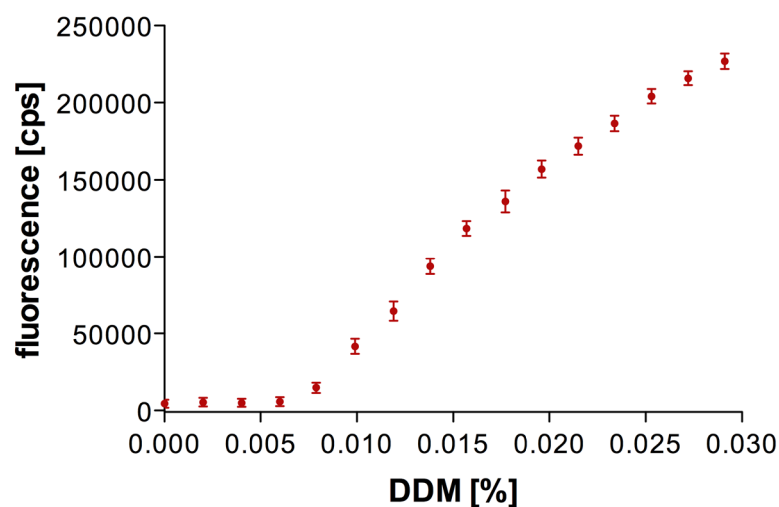
Figure 9: CMC-Determination for DDM. Plotted is the fluorescence signal at 457 nm. Inflection point indicates the formation of micelles (0.007 % DDM).

This kind of measurements can take place in a fluorescence spectrometer but also in a 96well microtiter plate reader that is available in virtually every biochemical lab (see Chapter IV).

1.3.2   Calculating aggregation numbers

The micellar aggregation number (N_{agg}) represents the number of detergent molecules or monomers found in a micelle. Analogous to a molecular weight determination for polymers, the N_{agg} represents an „averaged" value for a collection of micelles, which individually are dynamic structures capable of exchanging monomers among themselves. Several highly sophistical methods are available to determine this micellar parameter (ultracentrifugation, EPR, SAXS, SANS, NMR, X-ray, Light scattering, etc.). The most convenient one for biochemists is obviously fluorescence measurements. Most biochemical labs are equipped with fluorescence detection. Currently, both steady-state and time-dependent fluorescence measurements have been applied to the determination of aggregation numbers (*48*). Steady-state fluorescence is the more common used technique, since it is inexpensive and experimentally simple. They usually rely on the decrease or "quenching" of a fluorescence probe associated with the micelle through the presence of a second species, the "quencher" (Figure 10). How fluorescence quenching varies with increasing quencher concentration is then determined. Several probe-quencher combinations have been investigated (*49-51*). The system worked depends on the pyrene fluorescence quenched through Hoechst 33342, which is described in more detail in Chapter V.
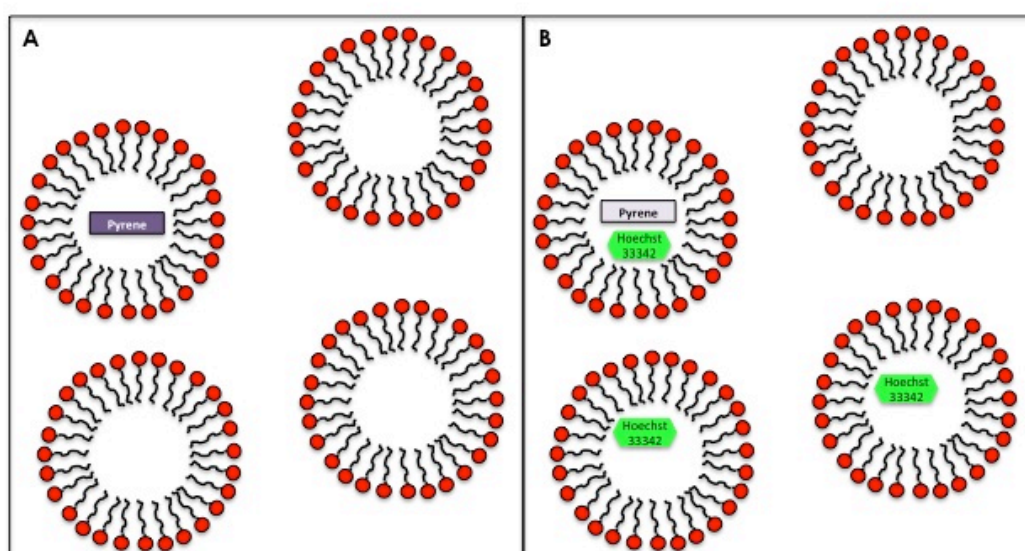

Figure 10: Pyrene quenching upon addition of Hoechst 33342. A: Pyrene emits fluorescence light as no quencher is present. B: Pyrene fluorescence in micelles that posses pyrene and Hoechst 33342 is quenched.

Given several assumptions about probe-quencher interactions and behavior (both are entirely in the micellar phase, the distributions of the quencher and the probe obey Poisson statistics and the probe only fluorescence in the absence of the quencher (*50*), a mathematical expression relating the ratio of fluorescence intensities to the total detergent concentration, free detergent concentration and quencher concentration can be derived. Initially the expression

$$\frac{I}{I_0} = e^{\frac{-[Q]}{[M]}}$$

Equation 1

results, relating the ratio of fluorescence intensities in the absence (I_0) and presence (I) of quencher to the quencher concentration [Q] and the micelle concentration [M]. The micelle concentration is given by

$$[M] = \frac{[Det] - [Det_{free}]}{N_{agg}}$$

Equation 2

There [Det] is the total detergent concentration [Det$_{free}$] is the concentration of detergent monomers not associated with micelles and N$_{agg}$ the average number of detergent monomers in each micelle. Substitution of equation 2 into equation 1 results in:

$$\ln\left(\frac{I}{I_0}\right) = -\frac{N_{agg}[Q]}{[Det] - [Det_{free}]}$$

Equation 3

Thus the aggregation number can be determined from the slope of a linear fit of ln (I/I$_0$) when knowing the detergent concentration in the assay and the CMC.

In Chapter V the practicability of the probe-quencher pair pyrene – Hoechst 33342 is demonstrated and several examples where CMC and aggregation number determination should be considered for biochemical purposes are given. To make it applicable for nearly every biochemical lab we applied our method to 96well plate reader available in virtually every lab.

1.4 Type I secretion in *Escherichia coli*: A multi protein complex including the ABC-transporter HlyB

A Type I secretion machinery in *Escherichia coli* allows the translocation of a 1023 amino acid big protein (Haemolysin A, HlyA) from the cytoplasm to the extracellular medium in a single step, without periplasmic intermediates (*52, 53*). This machinery consists of three proteins localized in the cell membranes. In the inner membrane there is the ABC-transporter HlyB located, which first recognizes the substrate via its secretion signal and is therefore responsible for the specificity of the secretion process. HlyD, a so-called membrane fusion protein consists of a short cytoplasmic domain at the N-terminus followed by a membrane helix and a large periplasmic domain (*54*). It is thought that this protein mediates the connection between the outer membrane tunnel TolC and the ABC-transporter HlyB (see Figure 11 for schematic representation).
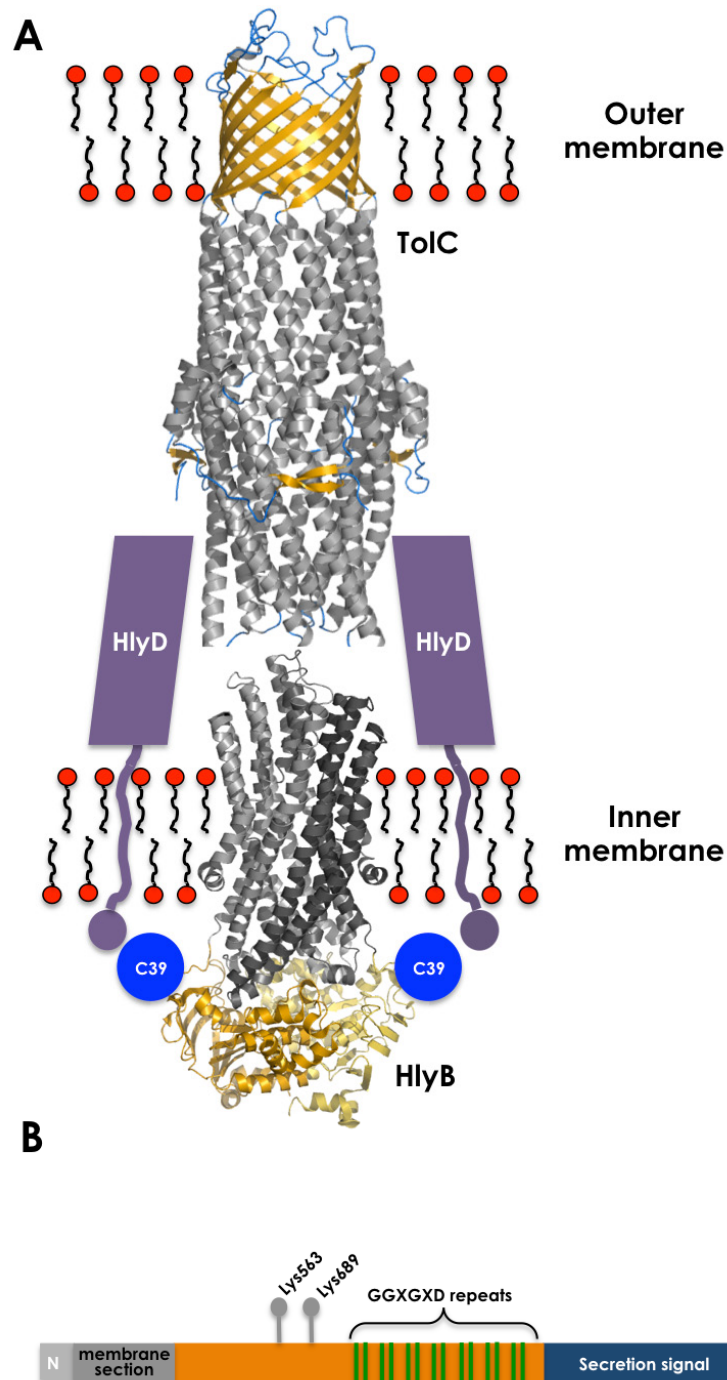
Figure 11: A: Schematic representation of Type I secretion system of *E. coli*. HlyB (shown is the crystal structure of Sav1866, pdb: 2ONJ) resides in the inner membrane. The additional domain (C39) is indicated as blue circle. HlyD is shown in purple. It interacts with the ABC-transporter and the outer membrane protein TolC in the periplasm (pdb: 1EK9), the cytoplasmatic interaction partner is not known so far. B: Schematic drawing of the transported toxin Haemolysin A. Secretion signal of 50-60 amino acids is located at the C-terminus (dark blue). Ca²⁺ -binding motifs are highlighted in green. For its toxity HlyA requires acylation at two indicated lysine residues. In dark grey a putative hydrophobic membrane insertion domain is shown.

In addition to its two TMDs and NBDs, HlyB has an extra domain at its N-terminus of approximately 150 amino acids. Sequence alignments identified these 150 amino acids as a C39 peptidase domain. This family of peptidases cleaves their substrates after a double

glycine motif (*55, 56*). However, the C39 domain of HlyB is degenerated and therefore inactive in terms of cleavage. As secretion of the toxin is completely abolished if the C39 domain is deleted the question arises which role does this domain play in substrate recognition and transport.

By separating the C39 domain from the transporter genetically and subsequent overexpression and purification, it was possible to crystallize the oligomeric C39 protein. The crystals diffract X-rays nicely and a X-ray data set up to a resolution of 2.0 Å was collected of the single domain. With these data structure determination should be possible which would provide us with insights into putative interaction sides, oligomeric state of the domain and its function in the transport process (see Chapter VI).

Another approach to get insights on the business of C39 domain are functional studies on the full-length transporter and a truncation mutant lacking this domain (HlyBΔC39), which is presented in Chapter VII. Therefore I established overexpression and purification of the transporter and performed ATP hydrolysis measurements. Overexpression could be archieved homolougously in *E. coli*. As described above detergent screening is a challenging step on the way to pure, stable and active membrane protein preparation. Membrane extraction in the case of HlyB could only be accomplished with the detergent Foscholine-14. However using it on further steps of purification yield in stable but inactive protein. Thus exchange of detergent during affinity chromatography was tried. A screen of several detergents identified $C_{12}E_8$ (Dodecyl-octaethyleneglycolether) as a suitable detergent to do activity measurements on this protein. ATP concentration dependent analysis did not follow Michaelis-Menten kinetics but fitted very well to a cooperative fit yielding in a Hill coefficient of 2. This is in good agreement with studies on isolated NBDs (*57*)(Figure 12).

	K_m [mM]	V_{max} [nmol/min/mg]
Michaelis-Menten	0.99	24
Hill	0.41	15
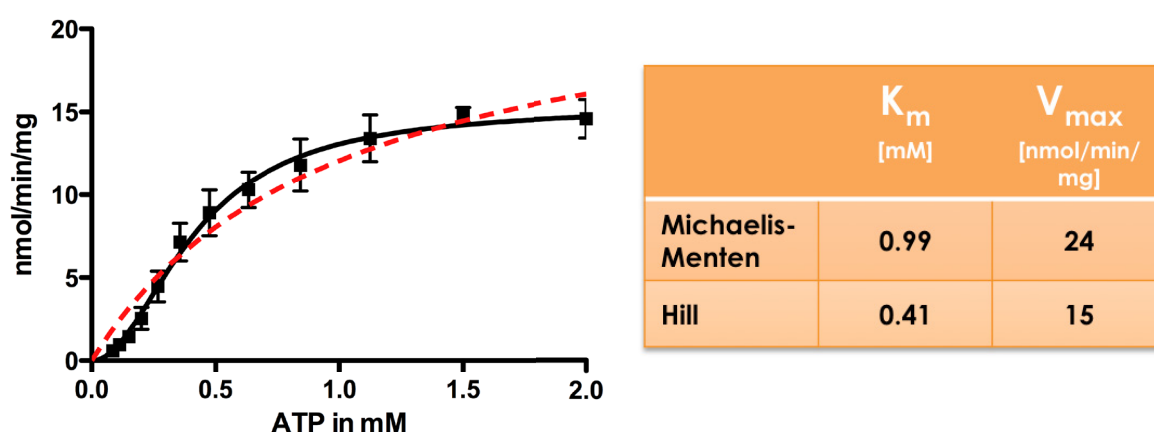
Figure 12: Kinetic data of HlyB wild type. Black straight line belongs to fit function considering cooperativity. Red dashed line shows Michaelis-Menten fit function.

To see whether C39 domain has an influence on ATPase activity we analyzed substrate dependent activity. Thus HlyA was secreted and purified from supernatant. A comparison of wild type and HlyBΔC39 activity is shown in Figure 13.
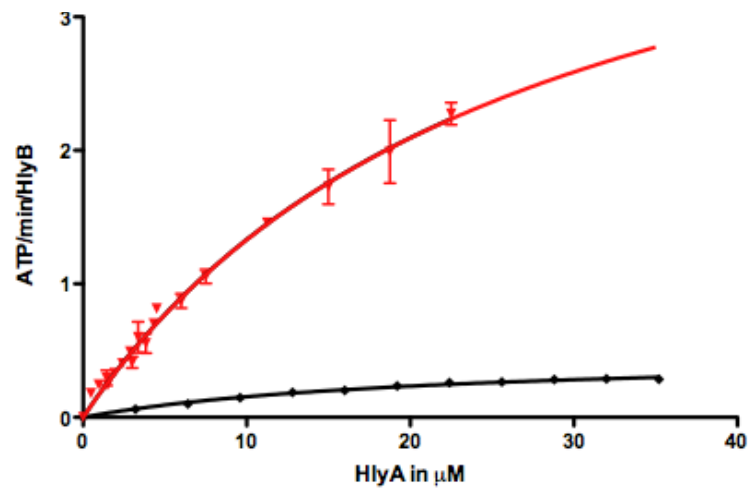
Figure 13: HlyA dependent ATPase activity of wild type HlyB (black) and ΔC39 mutant.

HlyA stimulates ATPase activity of both wild type and ΔC39 protein. However effect on the truncated protein is more than six fold higher. Thus the C39 domain has a regulatory influence in the transport process, making sure that ATP is not wasted. These results are in good agreement with reduced growth of cells overexpressing HlyBΔ39. As no stimulation could be detected for HlyA lacking the C-terminal secretion signal, this part of HlyA must bind to the transporter and induce ATP consumption. In this coherency especially the role of the N-terminal putative cytoplasmic part of HlyD would be interesting to analyze. First experiments provided hints that it interacts with C39 domain. If this domain is lacking ATPase activity is completely abolished after addition of HlyD.

2. Literature

1.   Singer, S. J., and Nicolson, G. L. (1972) The fluid mosaic model of the structure of cell membranes, *Science 175*, 720-731.

2.   Spitzer, N. C. (1979) Ion channels in development, *Annu Rev Neurosci 2*, 363-397.

3.   Abramson, J., Iwata, S., and Kaback, H. R. (2004) Lactose permease as a paradigm for membrane transport proteins (Review), *Mol Membr Biol 21*, 227-236.

4.   Jones, P. M., O'Mara, M. L., and George, A. M. (2009) ABC transporters: a riddle wrapped in a mystery inside an enigma, *Trends Biochem Sci 34*, 520-531.

5.   Holland, I. B., and Blight, M. A. (1999) ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans, *J Mol Biol 293*, 381-399.

6.   Higgins, C. F. (2001) ABC transporters: physiology, structure and mechanism--an overview, *Res Microbiol 152*, 205-210.

7.   Davidson, A. L., Dassa, E., Orelle, C., and Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems, *Microbiol Mol Biol Rev 72*, 317-364, table of contents.

8.   Schmitt, L., and Tampe, R. (2002) Structure and mechanism of ABC transporters, *Curr Opin Struct Biol 12*, 754-760.

9.   Parcej, D., and Tampe, R. (2010) ABC proteins in antigen translocation and viral inhibition, *Nat Chem Biol 6*, 572-580.

10.   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.

11.   Crowley, E., and Callaghan, R. (2010) Multidrug efflux pumps: drug binding--gates or cavity?, *FEBS J 277*, 530-539.

12.   Ernst, R., Kueppers, P., Stindt, J., Kuchler, K., and Schmitt, L. (2010) Multidrug efflux pumps: substrate selection in ATP-binding cassette multidrug efflux pumps--first come, first served?, *FEBS J 277*, 540-549.

13.   Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L., and O'Toole, G. A. (2003) Transition from reversible to irreversible attachment during biofilm formation by Pseudomonas fluorescens WCS365 requires an ABC transporter and a large secreted protein, *Mol Microbiol 49*, 905-918.

14.   Hollenstein, K., Dawson, R. J., and Locher, K. P. (2007) Structure and mechanism of ABC transporter proteins, *Curr Opin Struct Biol 17*, 412-418.

15.   Jardetzky, O. (1966) Simple allosteric model for membrane pumps, *Nature 211*, 969-970.

16.   Chen, J., Lu, G., Lin, J., Davidson, A. L., and Quiocho, F. A. (2003) A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle, *Mol Cell 12*, 651-661.

17.   Lu, G., Westbrooks, J. M., Davidson, A. L., and Chen, J. (2005) ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation, *Proc Natl Acad Sci U S A 102*, 17969-17974.

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

19. Boos, W., and Bohm, A. (2000) Learning new tricks from an old dog: MalT of the Escherichia coli maltose system is part of a complex regulatory network, *Trends Genet 16*, 404-409.

20. Chen, J., Sharma, S., Quiocho, F. A., and Davidson, A. L. (2001) Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport, *Proc Natl Acad Sci U S A 98*, 1525-1530.

21. Daus, M. L., Grote, M., Muller, P., Doebber, M., Herrmann, A., Steinhoff, H. J., Dassa, E., and Schneider, E. (2007) ATP-driven MalK dimer closure and reopening and conformational changes of the "EAA" motifs are crucial for function of the maltose ATP-binding cassette transporter (MalFGK2), *J Biol Chem 282*, 22387-22396.

22. Davidson, A. L., Shuman, H. A., and Nikaido, H. (1992) Mechanism of maltose transport in Escherichia coli: transmembrane signaling by periplasmic binding proteins, *Proc Natl Acad Sci U S A 89*, 2360-2364.

23. Ames, G. F., and Lecar, H. (1992) ATP-dependent bacterial transporters and cystic fibrosis: analogy between channels and transporters, *FASEB J 6*, 2660-2666.

24. Story, R. M., and Steitz, T. A. (1992) Structure of the recA protein-ADP complex, *Nature 355*, 374-376.

25. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold, *EMBO J 1*, 945-951.

26. Azzaria, M., Schurr, E., and Gros, P. (1989) Discrete mutations introduced in the predicted nucleotide-binding sites of the mdr1 gene abolish its ability to confer multidrug resistance, *Mol Cell Biol 9*, 5289-5297.

27. Oswald, C., Holland, I. B., and Schmitt, L. (2006) The motor domains of ABC-transporters. What can structures tell us?, *Naunyn Schmiedebergs Arch Pharmacol 372*, 385-399.

28. Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B., and Schmitt, L. (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB, *EMBO J 24*, 1901-1910.

29. Bianchet, M. A., Ko, Y. H., Amzel, L. M., and Pedersen, P. L. (1997) Modeling of nucleotide binding domains of ABC transporter proteins based on a F1-ATPase/recA topology: structural model of the nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator (CFTR), *J Bioenerg Biomembr 29*, 503-524.

30. Bateman, A. (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein, *Trends Biochem Sci 22*, 12-13.

31. Biemans-Oldehinkel, E., Mahmood, N. A., and Poolman, B. (2006) A sensor for intracellular ionic strength, *Proc Natl Acad Sci U S A 103*, 10624-10629.

32. Horn, C., Bremer, E., and Schmitt, L. (2003) Nucleotide dependent monomer/dimer equilibrium of OpuAA, the nucleotide-binding protein of the osmotically regulated ABC transporter OpuA from Bacillus subtilis, *J Mol Biol 334*, 403-419.

33. Sutcliffe, I. C., and Russell, R. R. (1995) Lipoproteins of gram-positive bacteria, *J Bacteriol 177*, 1123-1128.

34. van der Heide, T., and Poolman, B. (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites?, *EMBO Rep 3*, 938-943.

35. 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.

36.   Wolf, A., Lee, K. C., Kirsch, J. F., and Ames, G. F. (1996) Ligand-dependent conformational plasticity of the periplasmic histidine-binding protein HisJ. Involvement in transport specificity, *J Biol Chem 271*, 21243-21250.

37.   Oh, B. H., Kang, C. H., De Bondt, H., Kim, S. H., Nikaido, K., Joshi, A. K., and Ames, G. F. (1994) The bacterial periplasmic histidine-binding protein. structure/function analysis of the ligand-binding site and comparison with related proteins, *J Biol Chem 269*, 4135-4143.

38.   Olah, G. A., Trakhanov, S., Trewhella, J., and Quiocho, F. A. (1993) Leucine/isoleucine/valine-binding protein contracts upon binding of ligand, *J Biol Chem 268*, 16241-16247.

39.   Schiefner, A., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004) Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon Archaeoglobus fulgidus, *J Biol Chem 279*, 48270-48281.

40.   Mao, B., Pear, M. R., McCammon, J. A., and Quiocho, F. A. (1982) Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model, *J Biol Chem 257*, 1131-1133.

41.   Tate, C. G. (2010) Practical considerations of membrane protein instability during purification and crystallisation, *Methods Mol Biol 601*, 187-203.

42.   Cantor, R. S. (1999) Lipid composition and the lateral pressure profile in bilayers, *Biophys J 76*, 2625-2639.

43.   Raschle, T., Hiller, S., Etzkorn, M., and Wagner, G. (2010) Nonmicellar systems for solution NMR spectroscopy of membrane proteins, *Curr Opin Struct Biol 20*, 471-479.

44.   Bayburt, T. H., and Sligar, S. G. (2002) Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks, *Proc Natl Acad Sci U S A 99*, 6725-6730.

45.   Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., and Sligar, S. G. (2004) Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size, *J Am Chem Soc 126*, 3477-3487.

46.   Schmitz, G., and Grandl, M. (2009) The molecular mechanisms of HDL and associated vesicular trafficking mechanisms to mediate cellular lipid homeostasis, *Arterioscler Thromb Vasc Biol 29*, 1718-1722.

47.   Jumpertz, T., Tschapek, B., Infed, N., Smits, S. H., Ernst, R., and Schmitt, L. (2011) High-throughput evaluation of the critical micelle concentration of detergents, *Anal Biochem 408*, 64-70.

48.   Grieser, F., and Drummond, C. J. (1988) The Physicochemical Properties of Self-Assembled Surfactant Aggregates As Determined by Some Molecular Spectroscopic Probe Techniques, *J. Phys. Chem. 92*, 5580-5593.

49.   Tummino, P. J., and Gafni, A. (1993) Determination of the aggregation number of detergent micelles using steady-state fluorescence quenching, *Biophys J 64*, 1580-1587.

50.   Turro, N. J., and Yekta, A. (1978) Luminescent Probes for Detergent Solutions. A simple Procedure for Determination of Mean Aggregation Number of Micelles, *J. Am. Chem. Soc 100*, 5951-5952.

51.   Infelta, P. P. (1978) Fluorescence quenching in micellar solutions and its application to the determination of aggregation numbers, *Chemical Physics Letters 61*, 88-91.

52.   Holland, I. B., Schmitt, L., and Young, J. (2005) Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review), *Mol Membr Biol 22*, 29-39.

53.   Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B., and Schmitt, L. (2003) A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A, *J Mol Biol 327*, 1169-1179.

54. Johnson, J. M., and Church, G. M. (1999) Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps, *J Mol Biol 287*, 695-715.

55. Havarstein, L. S., Diep, D. B., and Nes, I. F. (1995) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export, *Mol Microbiol 16*, 229-240.

56. Wu, K. H., and Tai, P. C. (2004) Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter, *J Biol Chem 279*, 901-909.

57. Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B., and Schmitt, L. (2005) Functional characterization and ATP-induced dimerization of the isolated ABC-domain of the haemolysin B transporter, *Biochemistry 44*, 9680-9690.

Chapter I

Monitoring conformational changes during the catalytic cycle of OpuAA, the ATPase subunit of the ABC transporter OpuA from *Bacillus subtilis*

Carsten HORN*[1,2], Stefan JENEWEIN*[1], Britta TSCHAPEK*, Werner BOUSCHEN†, Sabine METZGER†, Erhard BREMER‡ and Lutz SCHMITT*[3]

*Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, †Biological and Medical Research Center, Heinrich Heine University, Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, and ‡Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Str., 35032 Marburg, Germany

The ABC transporter (ATP-binding-cassette transporter) OpuA is one of five membrane transport systems in *Bacillus subtilis* that mediate osmoprotection by importing compatible solutes. Just like all bacterial and archaeal ABC transporters that catalyse the import of substrates, OpuA (where Opu is osmoprotectant uptake) is composed of an ATPase subunit (OpuAA), a transmembrane subunit (OpuAB) and an extracellular substrate-binding protein (OpuAC). In contrast with many well-known ABC-ATPases, OpuAA is composed not only of a catalytic and a helical domain but also of an accessory domain located at its C-terminus. The paradigm of such an architecture is MalK, the ABC-ATPase of the maltose importer of *Escherichia coli*, for which detailed structural and functional information is available. In the present study, we have applied solution FRET (Förster resonance energy transfer) techniques using two single cysteine mutants to obtain initial structural information on the architecture of the OpuAA dimer in solution. Analysing our results in detail and comparing them with the existing MalK structures revealed that the catalytic and helical domains adopted an arrangement similar to those of MalK, whereas profound differences in the three-dimensional orientation of the accessory domain, which contains two CBS (cystathionine β-synthetase) domains, were observed. These results shed new light on the role of this accessory domain present in a certain subset of ABC-ATPase in the fine-tuning of three-dimensional structure and biological function.

Key words: ABC (ATP-binding-cassette)-ATPase, conformational changes, dimer, Förster resonance energy transfer (FRET), osmoprotection.

INTRODUCTION

Osmoregulation is of crucial importance for cell viability and survival [1]. *Bacillus subtilis* has to maintain a cell turgor of approx. 20 bar (1 bar = 100 kPa) [2], which is believed to be the driving force of cell expansion. Thus any change in external osmolarity will induce passive influx (hypo-osmotic stress) or efflux (hyperosmotic stress) of water, thereby changing cell turgor [3,4]. Hyperosmotic conditions (an increase in the extracellular osmolarity) result in the uptake of K^+ ions as a first line of defence to maintain cell turgor in *B. subtilis* [5]. However, the accumulation of a large concentration of a charged ion is detrimental to many cellular functions such as protein function and/or protein–DNA interactions. Thus *B. subtilis* switches to a second line of defence and starts to actively take up or synthesize so-called 'compatible solutes' [6]. In the phyla of bacteria, these 'compatible solutes' are small molecules such as sugars, polyols, amino acids, quaternary amines, sulfate esters or small peptides [7]. The common denominator of all these compounds is the fact that they are highly soluble and bear no net charge under physiological conditions. As a consequence, compatible solutes counteract the deleterious effect of changes in the extracellular osmolarity by equilibrating concentration gradients across the cellular membrane and maintaining cell turgor [4]. In addition

to their well-studied role in maintaining an appropriate level of cell water and turgor, compatible solutes also act as protein stabilizers both *in vitro* and *in vivo*. Although the molecular action of compatible solutes as a 'chemical chaperone' is poorly understood, the common opinion is that these molecules act according to the 'volume exclusion model' [8].

Extensive research in the last two decades has identified five transport systems providing osmoprotection in *B. subtilis* through the uptake of compatible solutes [6]. In concert, these five importers allow *B. subtilis* to cope with osmotic stress due to the ability to utilize common compatible solutes such as proline, choline, ectoine, glycine betaine or proline betaine. All of the Opu (osmoprotectant uptake) systems operating in *B. subtilis* belong either to the class of secondary transporters (OpuD and OpuE) or to a subfamily of primary transporters, the ABC (ATP-binding cassette) transporter family (OpuA, OpuB and OpuC; for a recent review see [9]). These membrane transporters ultimately use the energy of ATP for substrate translocation across biological membranes [10]. Although found in all three domains of life and involved in many vital cellular processes, ATP-dependent transporters share a basic blueprint that is composed of four modules: two NBDs (nucleotide-binding domains or subunits) harbouring the ATPase activity and two TMDs (transmembrane domains or subunits) that form the translocation pathway. Next

---

to the substrate, ABC transporters can be classified by the directionality of transport: import or export. OpuA belongs to the family of ABC importers, since it catalyses the import of glycine betaine and proline betaine from the extracellular medium into the bacterial cytosol. Additionally to the NBD (OpuAA) [11] and the TMD (OpuAB) [9], this family always requires another component for functionality: a so-called SBP (substrate-binding protein; OpuAC in the case of the OpuA transporter [12]). Recent insights into the molecular function of these ABC transporters were provided by X-ray crystallography that determined the high-resolution structures of isolated NBDs (for a recent review see [13]) and intact ABC transporters [14–19]. For example, crystal structures of the NBD of the maltose import system (MalK) in the nucleotide-free, ATP- and ADP-bound states have deciphered many of the molecular principles that govern ATP binding and hydrolysis and have resulted in the proposal of a 'tweezer-like motion' that accompanies maltose import [20,21].

OpuAA, the NBD of the OpuA transporter, and MalK, the NBD of the maltose importer, share interesting similarities. Both NBDs not only are composed of a catalytic and a helical domain that harbours all conserved sequence motifs of these ATPases such as the Walker A and B motifs and the hallmark of ABC transporters, the C-loop or ABC-signature motif [10], but also contain an accessory domain, which is composed of approx. 150 amino acids and is located C-terminal to the canonical NBD (catalytic and helical domain). In the case of MalK, this accessory domain interacts with at least two enzymes, MalT [22] and enzyme IIA [23], thereby regulating the expression of the maltose operon. Furthermore, in the 'tweezer-like motion' model, the accessory domains act as the bottom of tweezers keeping two NBDs together during a catalytic cycle, whereas the canonical part (the catalytic and helical domains) undergoes ATP-induced dimerization remaining monomeric in the nucleotide-free and ADP-bound states.

Biochemical studies have established that OpuAA undergoes a dynamic monomer–dimer equilibrium and forms dimers in the nucleotide-free and ATP-bound states but, in striking contrast with MalK, becomes monomeric in the ADP-bound state [11]. The closely related OpuA glycine betaine importer in *Lactococcus lactis* has been intensively studied in the past by biochemical approaches [24]. Like OpuAA from *B. subtilis*, the NBD of the *L. lactis* homologue is composed of a canonical part (catalytic and helical domain) and an accessory domain. Most important was the fact that the accessory domain contains the fingerprint of two CBS (cystathionine β-synthetase) domains [25] and is involved in sensing osmotic stress signals that activate the transport function of the *L. lactis* OpuA system [26].

In the present study, we have applied FRET (Förster resonance energy transfer) studies to analyse the structure of OpuAA in solution by monitoring and determining the conformational changes taking place during the catalytic cycle of the ATPase. After generating suitable single cysteine mutants of OpuAA, static and time-resolved FRET measurements provided a set of distance restrains that were employed to model the overall structures of OpuAA in the nucleotide-free, ATP- and ADP-bound states based on the available crystal structures of MalK [20,21]. Surprisingly, the three-dimensional positioning of the accessory domains with respect to the canonical part of the NBD was different, whereas the overall conformational changes during the catalytic cycle were identical with the ones determined previously for other isolated NBDs [20,27–29]. Thus our results provide novel insights on the architecture of OpuAA from *B. subtilis* and suggest that the accessory domains play an important role in the structural and functional regulation of these proteins.

MATERIALS AND METHODS

Cloning, expression and isolation of OpuAA single cysteine mutants and the F19W mutant

To change native amino acids to cysteine, 5′ and 3′ fragments were amplified by PCR using pBAD33/OpuAA as a template and flanking/mutagenic primer pairs as follows: S45C, 5′ CH7/CH29 and 3′ CH8/CH28; G161C, 5′ CH7/CH41 and 3′ CH8/CH40; F19W, CH47/CH48. The primers are CH7, 5′-GGAATTCCA-TATGAGTGTAGATGAGAAACCAATTA-3′; CH8, 5′-ATATA-TAAAGCTTATTAGTGATGGTGATGGTGATGTTCACCTCC-TGTGCAGAAGGATCTTG-3′; CH28, 5′-ATTAACCCCAA-CGGTACATCCGGTTGCTTTCAG-3′; CH29, 5′-CTGAAAGC-AACCGGATGTACCGTTGGGGTTAAT-3′; CH40, 5′-ATATT-GGTGTTCAAAACATTCCAGTGGAACGAG-3′; CH41, 5′-CTCGTTGGACTGGAATGTTTTGAACACCAATAT-3′; CH47, 5′-ATTCTTTGTTTGTTTCCCCCAAATCTTAGAGAC-3′; and CH48, 5′-GTCTCTAAGATTTGGGGGAAACAAACAAAGA-AG-3′. Secondly, the mutated *opuAA* gene was amplified using both overlapping fragments as a template and the flanking primers CH7 and CH8. The PCR products were digested with NdeI/HindIII (New England Biolabs) and cloned into pBAD33. All gene inserts were sequenced according to the ddNTP method and were free of mutations. Mutant proteins were produced in *Escherichia coli* and purified by IMAC (immobilized metal-ion-affinity chromatography) as described previously [11]. Monomeric OpuAA was isolated by SEC (size-exclusion chromatography) using a Superdex200 HR 26/60 column (GE Healthcare) equilibrated with buffer A [10 mM sodium phosphate and 0.1 mM EDTA (pH 7.5)] supplemented with 1 M NaCl. Protein-containing fractions were pooled and stored at 4 °C until further use.

ATPase activity assays

ATPase activity of monomeric OpuAA was analysed by an enzyme-linked assay as described previously [11] in 10 mM sodium phosphate buffer (pH 7.5) and 1 M NaCl. In brief, absorbance at 340 nm was monitored at 1 min time intervals over a period of 15 min at 22 ± 2 °C. Data points of the whole measuring period or until all NADH was consumed (over a period of at least 7 min) were used to calculate linear slopes $\Delta A / \Delta t_{\text{OpuAA}}$. To account for autohydrolysis of ATP, similar assays were performed in the absence of OpuAA. After background subtraction ($\Delta\Delta A / \Delta\Delta t = \Delta A / \Delta t_{\text{OpuAA}} - \Delta A / \Delta t_{\text{buffer}}$) reaction velocities were calculated according to eqn (1):

$$v = \frac{\Delta\Delta A}{\Delta\Delta t} \cdot \frac{1}{2590 \text{ M}^{-1} \cdot c_0} \tag{1}$$

In this equation, v is the reaction velocity, $\Delta\Delta A / \Delta\Delta t$ denotes the linear slopes of the time-dependent decay of NADH absorbance at 340 nm of OpuAA samples after background subtraction, and c_0 is the final concentration of OpuAA. Reaction velocities represent the average value of two independent measurements.

Fluorophore labelling of OpuAA mutants

For site-specific labelling of mutant OpuAA with fluorophore dyes, cysteine-specific maleinimide chemistry was used. Typically, OpuAA was passed over a PD-10 colum (GE Healthcare) to remove DTT (dithiothreitol) that was added to prevent disulfide bond formation during storage and mixed at a concentration of 10 μM with 25 μM dye [donor dye: OG (Oregon Green® 488 maleinimide; Molecular Probes), 20 mM in dimethyl-formamide; acceptor dye: TR (Texas Red® C2 maleinimide;

Molecular Probes), 4.1 mM in methanol] and incubated overnight at 4°C. Protein and free dye were separated using a PD-10 column (GE Healthcare) equilibrated with buffer A supplemented with 10 mM NaCl. To isolate the labelled mutants in a quantitative fashion, anion-exchange chromatography was used to separate labelled and non-labelled OpuAA. After sample loading on to a 1 ml Q-Sepharose HiTrap column (GE Healthcare) equilibrated with buffer A supplemented with 10 mM NaCl, a 12 column volume gradient (0.01–2.5 M NaCl in buffer A) was applied. Elution was followed at 216 nm, 491 nm and 592 nm respectively.

Labelling efficiency and protein concentration

Protein concentration was determined using the BCA (bicinchoninic acid) assay following the procedure provided by the manufacturer (Pierce) using BSA as a standard. To determine the efficiency of fluorophore labelling, the molar absorption coefficient of the isolated dyes was used as provided by the manufacture (Molecular Probes). However, not only incomplete labelling was evident but also a blue shift of the emission maximum as well as changes in the molar absorption coefficient (up to 2-fold; see Supplementary Figure S2 at http://www.BiochemJ.org/bj/412/bj4120233add.htm). Thus an additional purification step to quantitatively separate labelled and non-labelled protein was employed (see above) to circumvent these problems and avoid spectroscopic determination of the labelling efficiency.

MS

Samples obtained from anion-exchange chromatography (OpuAA S45C–OG or OpuAA S45C–TR) or purified OpuAA S45C were desalted over a C_{18} reversed-phase minicolumn (perfect pure C_{18}; Eppendorf) before mass spectrometric analysis. ESI (electrospray ionization)–MS measurements were performed using a ESI-QqTOF instrument (QSTAR XL, Applied Biosystems) equipped with a nanospray ionsource. The ESI–MS spectra were reconstructed using the deconvolution algorithm included in the analyst software (Applied Biosystems).

FRET sample preparation

To measure FRET efficiencies within homodimeric OpuAA, all donors had to be paired with acceptors in the assembled dimer. Therefore OpuAA–TR and OpuAA–OG were mixed in a 9:1 molar ratio. In brief, species were mixed at conditions of dynamic monomer/dimer equilibrium (buffer A + 100 mM NaCl) and concentrated to 30 μM to force dimer formation. After incubation overnight at 4°C, samples were applied to a Superdex200 HR 10/30 SEC column (GE Healthcare) previously equilibrated with buffer A supplemented with 1 M NaCl to isolate the dimeric species. The dimer fraction was used for the FRET assay, whereas the monomer fraction was used only for control experiments. For ADP/Mg^{2+}-mediated decay the dimer fraction was diluted into buffer A + 100 mM NaCl.

FRET assays

Fluorescence was measured with a Cary-Eclipse spectrometer (Varian) at a λ_{ex} of 492 nm at a temperature (T) of 20 ± 1°C and a slit width of 5 nm in 10 mM sodium phosphate (pH 7.5), 0.1 mM EDTA and 1 M NaCl. In static FRET assays, samples were separately analysed with emission recording from 500 to 650 nm and analysed according to eqn (2) [30]:

$$E_T = \frac{R_0^6}{R_0^6 + R^6} \tag{2}$$

In time-resolved FRET assays, the sample was chased with 10 mM ADP or ADP/Mg^{2+} (pH 7.5) at $t = 0$ to allow dimer dissociation. Fluorescence at time t (F_t) was followed at 519 nm over a period of 3600 s at 5 s intervals and fitted to eqn (3). In this case, fitting parameters were: fluorescence at $t = 0$ ($F_{t=0}$), fluorescence increase at infinite time (ΔF) and half-time (τ).

$$F_t = F_{t=0} + \Delta F \cdot (1 - e^{-t/\tau}) \tag{3}$$

Amplitudes were used to calculate the E_T-time-course according to eqn (4). For $t = 0$ the E_T value for the original, nucleotide-free state was extrapolated and using the Förster-distance [$R_0 = 51$ Å (1 Å = 0.1 nm) for the OG–TR pair] of this donor/acceptor pair, its distance in this state was calculated.

$$E_T(t) = 1 - \frac{F_t}{F_{t=0} + \Delta F} = 1 - \frac{F_t}{F_{t=\infty}} = \frac{R_0^6}{R_0^6 + R(t)^6} \tag{4}$$

Intrinsic tryptophan fluorescence measurements of the F19W mutant of OpuAA

All measurements were performed at 22 ± 1°C on a Cary-Eclipse spectrophotometer employing a slit width of 5 nm. Wild-type OpuAA does not contain any tryptophan residues. The tryptophan residue engineered at position 19 of OpuAA was excited at 290 nm, and fluorescence was collected at the emission maximum (350 nm) at a protein concentration of 2 μM dimeric OpuAA in 10 mM sodium phosphate (pH 7.5), 1 M NaCl and 0.1 mM EDTA. The chase was started by adding 1 mM ADP and 5 mM Mg^{2+}. Data were analysed using a single exponential decay.

Molecular modelling and structural analysis

Molecular modelling was performed with MODELLER 9v1 [31] employing standard settings. As templates, the crystal structures of MalK in the semi-open, nucleotide-free (PDB entry: 1Q1B), the open nucleotide-free (PDB entry: 1Q1E), the ATP- (PDB entry: 1Q12) and ADP-bound (PDB entry: 1AWN) states were used. The quality of the obtained structural models of OpuAA in the corresponding functional states was verified with PROCHECK [32] and compared with the template structure. The initial models were further modified in O after imposing the FRET-derived distance restrains using sgFRET (http://www.mpibpc.mpg.de/groups/grubmueller/start/software/fretsg-1.0/fretsg.html). Structural superimpositions were performed using LSQMAN [33] using the indicated motifs as anchor points. The buried surface areas were calculated using areaimol of the CCP4 suite [33a]. All structural figures were prepared using PyMol (http://www.pymol.org).

RESULTS

Biochemical studies have demonstrated that OpuAA from *B. subtilis* undergoes a dynamic monomer–dimer equilibrium [11]. Interestingly, this equilibrium depends on the presence or absence of nucleotides. To gain further insights into the underlying conformational changes, we applied FRET techniques of single cysteine mutants of OpuAA in solution. Since the crystal structures of MalK from *E. coli* are available in the nucleotide-free and ATP- and ADP-bound states [20,21], they provide a lead structure (homology of 69%) for OpuAA homology modelling and would open up the possibility to translate the distance restrains derived from FRET measurements into a three-dimensional model of OpuAA. A homology model of OpuAA using the crystal structure of MalK in the nucleotide-free state (PDB entry: 1Q1B)
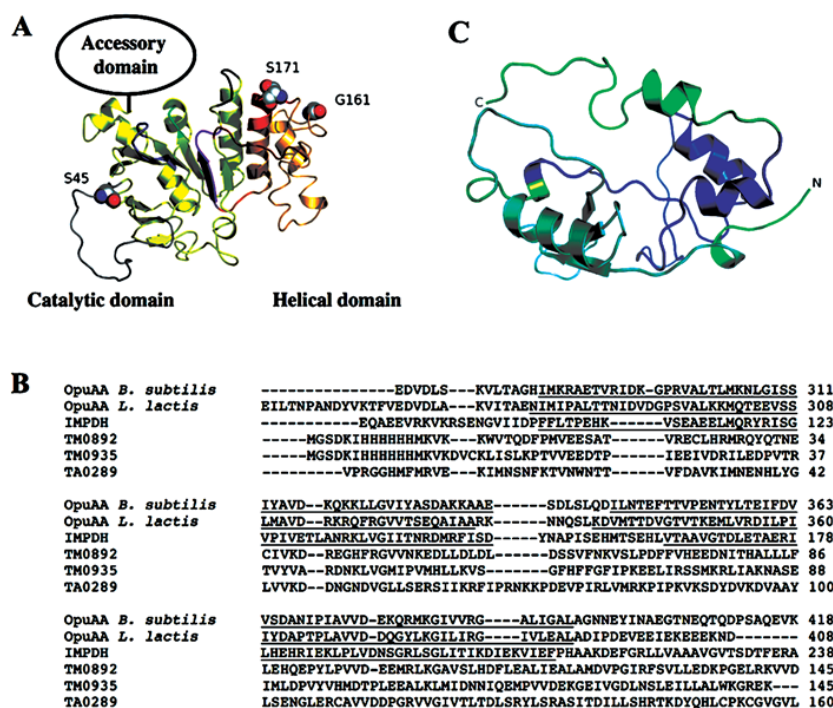
Figure 1   Homology model of OpuAA

(**A**) Homology model of the core domain OpuAA in the nucleotide-free state lacking the accessory domain. The model was generated using MODELLER 9v1 [31] as described in the Materials and methods section. The helical domain is shown in light tan and the catalytic domain in yellow. The OpuAA-specific insertion between β-strands 1 and 2 is highlighted in grey. Conserved motifs are colour-coded as follows: Walker A motif (blue), C-loop or ABC-signature motif (red), Walker B motif (magenta), D-loop (black) and H-loop (green). Positions of the single cysteine mutants are highlighted as spheres and labelled. (**B**) Sequence alignment of the accessory domains of OpuAA from *B. subtilis* (residues 277–418) and *L. lactis* (residues 271–408) containing a tandem CBS domain (residues 283–335 and 342–394 in the case of OpuAA from *B. subtilis*) with selected proteins containing CBS domains, for which a crystal structure is available. The two CBS domains in OpuAA from *B. subtilis* and *L. lactis* and IMPDH [41] are underlined. TM0892, open reading frame 0892 from *Thermotoga maritima* with unknown function (PDB entry: 1VR9); TM0935, open reading frame 0935 from *T. maritima* with unknown function (PDB entry: 1O50); TA0289, open reading frame 0289 from *Thermoplasma acidophilum* with unknown function (PDB entry: 1PVM). (**C**) Homology model of the accessory domain of OpuAA from *B. subtilis*. The tandem CBS domains are coloured cyan (CBS1) and blue (CBS2) respectively. The N- and C-termini, residues 277 and 418 respectively, of OpuAA are indicated. The model was derived from the crystal structure of TA0289 (PDB entry: 1PVM) due to the high sequence identity (25 % identity and 51 % homology) and obtained as described in the Materials and methods section.

was generated. As shown in Figure 1(A), OpuAA indeed adopts a 'MalK-like' structure [20] composed of a helical (coloured tan) and a catalytic (coloured yellow) domain. The catalytic domain of OpuAA contains a 24-amino-acid insertion between β-strands 1 and 2 (highlighted in grey), which is not present in other ABC domains. However, homology modelling of the accessory domain (represented by a circle in Figure 1A) was not straightforward. In clear contrast with MalK, the primary structure of OpuAA does not contain the fingerprint of the CUT1 (carbohydrate uptake transporter 1) subfamily represented by a characteristic GI/VRPE consensus sequence [34]. Rather the sequence indicates the presence of two CBS domains (Figure 1B) as evident by a sequence alignment of the accessory domain of OpuAA from *B. subtilis* with OpuAA from *L. lactis*, IMPDH (inosine-monophosphate dehydrogenase) and the orphan open reading frames from *Thermotoga maritimus* and *Thermoplasma acidophilum* [25]. These open reading frames were chosen, because crystal structures of the CBS modules are available (see the Figure legends for the corresponding PDB entries). Among the four crystal structures, *T. acidophilum* TA0289 had the highest sequence similarity (51 %) and was used for separate homology modelling. The resulting model of the isolated accessory domain of OpuAA is shown in Figure 1(C) displaying the characteristics of the tandem organization of CBS domains.

Since no prior information of the relative orientation of the accessory and the core NBD domain are available, no attempts were made to include the homology model of the CBS domain of OpuAA (Figure 1C) into the MalK-derived homology model of OpuAA (Figure 1A).

With the homology model at hand, we selected three residues of OpuAA within the catalytic and helical domains for changing to a cysteine residue. These residues are highlighted as spheres in Figure 1(A). Ser[45] is located in the catalytic domain at the C-terminal end of the OpuAA-specific insertion between β-strands 1 and 2. Gly[161] is at the N-terminal end of the ABC-signature motif. Ser[171], the conserved serine residue of the ABC-signature motif, was also selected for cysteine mutation to distinguish the composite dimer architecture from the one observed in the MalK crystal structure from *Thermococcus litoralis* [35], which shows a different arrangement of the two monomers in the NBD-dimer. Wild-type OpuAA and the three cysteine mutants were overexpressed in *E. coli* and purified as described in the Materials and methods section. Analysis of the two single cysteine mutants (S45C and G161C) under reducing and oxidizing conditions (see Supplementary Figure S1 at http://www.BiochemJ.org/bj/412/bj4120233add.htm) demonstrated that only negligible intermolecular disulfide bond formation occurred. More importantly for our conformational

Table 1   Kinetic parameters of monomeric OpuAA and the mutants used in the present study

ATPase assays were performed in 10 mM sodium phosphate (pH 7.5) and 1 M NaCl as described in the Materials and methods section. Values are means $\pm$ S.D. of two independent measurements.

Protein	K_m (mM)	k_{cat} (ATP/min)	Catalytic efficiency ($M^{-1} \cdot min^{-1}$)
OpuAA	0.6 ± 0.2	1.1 ± 0.1	1833
OpuAA S45C	0.7 ± 0.1	1.3 ± 0.1	1857
OpuAA G161C	2.8 ± 0.4	0.8 ± 0.1	285
OpuAA S171C	Not determined	0.08 ± 0.01	Not calculated
OpuAA F19W	1.7 ± 0.1	1.5 ± 0.1	882

analysis was the ATPase activity of the mutant proteins (Table 1). The wild-type enzyme hydrolysed ATP at a rate of 1.1 ± 0.1 min^{-1} under the conditions of the assay [11]. In contrast with the S45C and the G161C mutants, which hydrolysed ATP at comparable rates, the S171C mutant displayed hardly any turnover (0.08 ± 0.01 min^{-1}). Owing to this result, the S171C mutant was not used for further studies.

In order to use FRET quantitatively and to extract distance information, it is of importance to obtain a reliable value for the labelling efficiency [36]. The engineered single cysteine mutants of OpuAA allowed us to employ maleimide chemistry and covalently link the individual fluorophores, OG or TR. Surprisingly, labelling efficiencies were rather low (approx. 50% for the G161C and approx. 60% for the S45C mutant for both dyes). To exclude non-specific labelling due to the prolonged reaction time, wild-type OpuAA was treated with the fluorophores under similar conditions. Here, no labelling above background was detected after separating protein and fluorophore (results not shown). Additionally, attachment of the dyes changed their molar absorption coefficient as well as the absorption maximum (12 nm for TR and 8 nm for OR; Supplementary Figure S2). To account for the low labelling efficiency and the changes in the physical properties of the protein-attached fluorophores, we established a second purification step, anion-exchange chromatography, to separate labelled and non-labelled protein (Figure 2). A separation of a labelled and a non-labelled protein species is evident from the absorption properties shown in the representative chromatograms of OpuAA S45C–OG (Figure 2A), OpuAA S45C–TR (Figure 2B), OpuAA G161C–OG (Figure 2C) and OpuAA G161C–TR (Figure 2D).

To investigate further whether the apparent separation was indeed quantitative, we applied ESI–MS of the two fractions obtained by anion-exchange chromatography. ESI–MS analysis of fraction I, which eluted at lower ionic strength, revealed the presence of a single protein species with a molecular mass of 47 174 Da (Figure 3A). The calculated average molecular mass of OpuAA S45C, including a C-terminal hexahistidine tag and a formylated methionine residue, is 47 174.48 Da. The second fraction, which eluted at higher ionic strength, revealed the presence of a single protein species of 47 657 Da in the case of OpuAA S45C–OG (Figure 3B) and 47 921 Da in the case of OpuAA S45C–TR (Figure 3C). These mass differences correspond within experimental error of the MS measurements to the molecular masses of the individual fluorophores [differences of 483 Da for the sodium adduct of OG (expected 486.35 Da) and 747 Da for the sodium adduct of TR (expected 751.83 Da)].

A detailed analysis of the monomer–dimer equilibrium of OpuAA [11], has provided a toolbox to generate monomers or dimers of OpuAA depending on the protein concentration and ionic strength. This allowed us to prepare OpuAA–OG/OpuAA–TR pairs at a molar ratio of 1:9, which ensures that all (>95%) OG-labelled OpuAA molecules are indeed in an OG–TR dimer.
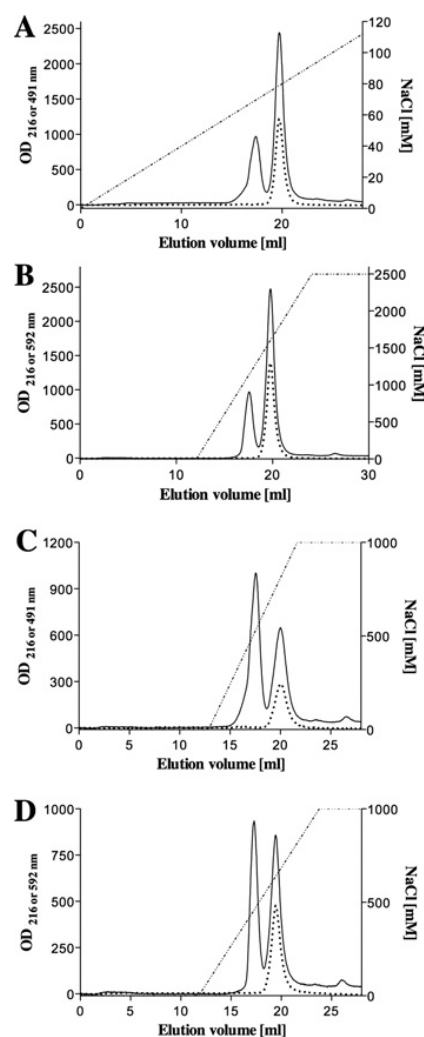
Figure 2   Quantitative separation of labelled and non-labelled single cysteine mutants of OpuAA by anion-exchange chromatography

A (OD) at 216 nm (protein absorption) is shown as a solid line and A at 491 nm (OG absorption; **A** and **C**) and 591 nm (TR absorption; **B** and **D**) is shown as dotted lines respectively. The applied NaCl concentration gradient is shown as a dashed line in the indicated concentration range. (**A**) OpuAA S45C–OG, (**B**) OpuAA S45C–TR, (**C**) OpuAA G161C–OG and (**D**) OpuAA G161C–TR. For further details see the Materials and methods section.

On the other hand, a high molar excess of acceptor might result in an apparent increase in transfer efficiency due to radiation-dependent processes such as acceptor self-absorption.
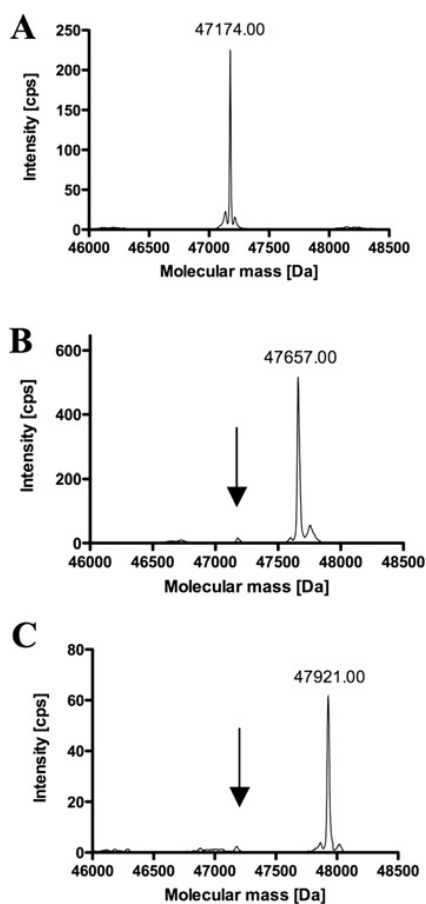
Figure 3   High-resolution MS analysis of OpuAA (A), OpuAA S45C–OG (B) and OpuAA S45C–TR (C)

The arrow in (**B**) and (**C**) indicates the position of unlabelled OpuAA, which is below 2% according to an analysis of the peak area.

However, these processes depend solely on the concentration of the acceptor and a proper choice will eliminate this obstacle. Therefore we measured the fluorescence of a 50 nM sample of OpuAA–OG (donor) in the presence or absence of a 9-fold molar excess of TR-conjugated OpuAA (results not shown) under conditions that abolish FRET. Comparison of the two spectra clearly showed that marginal self-absorption of TR occurred. Consequently, we used a fixed donor concentration of 50 nM in all subsequent experiments.

A representative set of static FRET experiments is shown in Figure 4 for the OpuAA S45C–OG/OpuAA S45C–TR dimer in the nucleotide-free (Figure 4A), the ATP-bound (Figure 4B) and the ADP-bound (Figure 4C) states. The broken lines show the sum of two control experiments, OpuAA S45C–OG and wild-type OpuAA (molar ratio of 1:9) and wild-type OpuAA and OpuAA S45C–TR (molar ratio of 1:9), whereas the solid lines show the fluorescence spectrum of an OpuAA S45C–OG/OpuAA S45C–TR mixture (molar ratio of 1:9). The insets show the corresponding difference spectra. As is evident from these experiments, FRET can be detected for the nucleotide-free and the ATP-bound states of dimeric OpuAA, but not for the ADP-bound state. The FRET efficiencies were calculated to 34% and 39% for the nucleotide-free and ATP-bound state according to
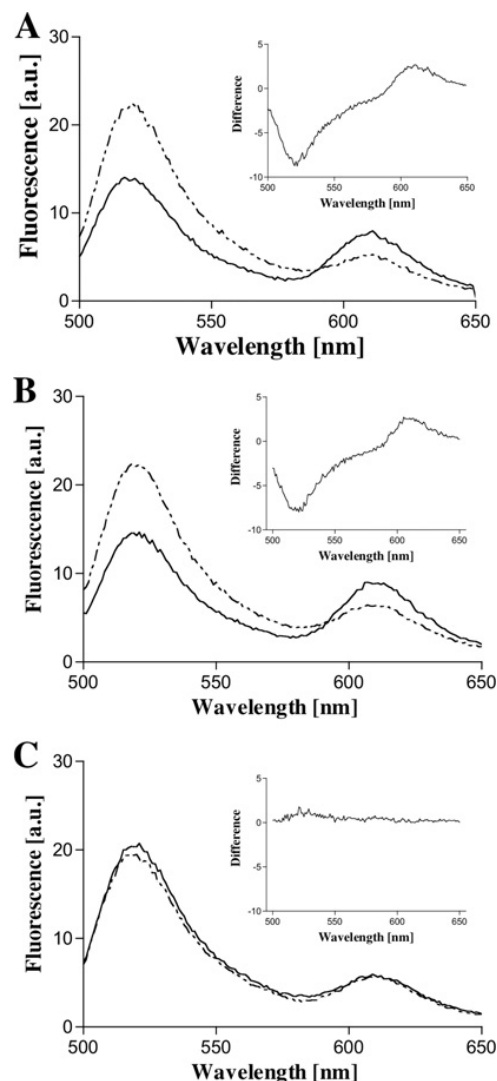
Figure 4   Static FRET measurements of OpuAA S45C in the nucleotide-free (A), ATP-bound (B) and ADP-bound (C) states in the absence of Mg^{2+}

In all three panels, the broken line represents the sum of two individual FRET experiments, OpuAA S45C–OG/OpuAA wild-type (molar ratio of 1:9) and wild-type OpuAA/OpuAA S45C–TR (molar ratio 1:9). The solid line shows the actual FRET experiments of OpuAA S45C–OG/OpuAA S45C–TR (molar ratio of 1:9). Acceptor concentration was always kept at 50 nM, and the corresponding dimers were prepared as described in the Materials and methods section. The inset in each panel shows the difference spectra of the corresponding FRET experiments (OG and TR pairs). The buffer used was 10 mM sodium phosphate (pH 7.5), 0.1 mM EDTA and 1 M NaCl. a.u., absorbance units.

eqn (2). A summary of the FRET efficiencies determined for this and the other combinations is given in Table 2.

Distance determinations through static FRET experiments require an accurate knowledge of the protein concentration. In the present study, concentration determination was complicated by the fact that the molar absorption coefficients of the dyes changed upon covalent attachment to the protein (see Supplementary Figure S2). We determined protein concentrations as outlined in the Materials and methods section and incorporated the changed molar absorption coefficients determined under the experimental conditions, but any error in the determination would

Table 2 **Summary of the static and time-resolved FRET experiments**

The dimeric species of the indicated single cysteine mutants of OpuAA of the nucleotide-free and ATP-bound states and the monomeric species of the indicated single cysteine mutants of OpuAA in the ADP-bound state were analysed as described in the Materials and methods section in the absence of Mg^{2+}. The determined transfer efficiencies at 519 nm (donor emission maximum) or the static FRET experiments were translated in distances according to the Förster equation assuming a Förster radius of the OG/TR pair of 51 Å [30]. In the case of the time-resolved experiments, amplitudes and half times were calculated from the ADP/Mg^{2+}-induced decay of the dimeric forms of labelled OpuAA mutants as described in the Materials and methods section.

Dimer species	Static FRET					Time-resolved FRET		
	Nucleotide free		ATP			Nucleotide-free		
	Transfer efficiency (%)	Distance (Å)	Transfer efficiency (%)	Distance (Å)	ADP/Mg^{2+}	Half time (min)	Transfer efficiency (%)	Distance (Å)
S45C–S45C	34	57	39	55	No transfer	22 ± 0.3	33	59
S45C–G161C	47	52	79	41	No transfer	20.7 ± 0.1	47	52
G161C–G161C	49	51	65	46	No transfer	12.9 ± 0.1	49	51
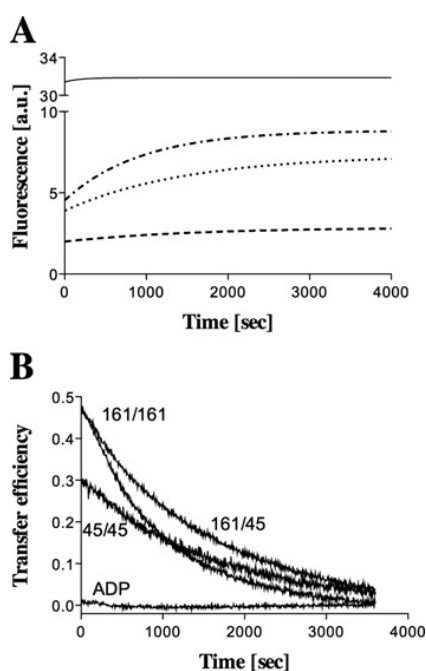
Figure 5 **Time-resolved FRET experiments**

(**A**) Dimers of the nucleotide-free state were prepared as described in the Materials and methods section and were chased by the addition of 10 mM ADP/Mg^{2+}. Fluorescence was measured at 519 nm in 5 s intervals. From bottom to top: dimeric OpuAA S45C–OG/OpuAA S45C–TR (dashed line), dimeric OpuAA G161C–OG/OpuAA S45C–TR (dotted line), dimeric OpuAA G161C–OG/OpuAA G161C–TR (dotted-dashed line) and monomeric OpuAA G161C–OG/OpuAA G161C–TR (solid line). Note that monomeric samples were measured at an approx. 4-fold higher concentration than the dimeric samples. Traces were analysed according to eqn (3). The buffer used was 10 mM sodium phosphate (pH 7.5), 0.1 mM EDTA and 1 M NaCl. (**B**) Time-resolved transfer efficiencies were calculated according to eqn (4) and extrapolated to $t = 0$ to determine the initial transfer efficiencies, which are summarized in Table 2. a.u., absorbance units.

drastically affect the FRET distances extracted from those measurements. Thus we also thought of establishing an experimental approach that was independent of a prior exact knowledge of protein concentration. Here, time-resolved changes in donor fluorescence of monomeric or dimeric OpuAA G161C–OG/OpuAA G161C–TR, dimeric OpuAA G161C–OG/OpuAA S45C–TR and dimeric OpuAA S45C–OG/OpuAA S45C–TR (Figure 5A) provided a tool to countercheck the distance information of the static FRET experiments and to validate these results. Nucleotide-free samples of the indicated combinations
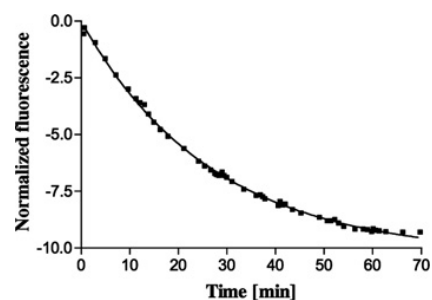
Figure 6 **Time-resolved intrinsic tryptophan fluorescence of the OpuAA F19W mutant**

The nucleotide-free dimeric form of OpuAA was prepared as described for the wild-type ATPase and ADP (1 mM) was added at zero time in the absence of Mg^{2+}. Tryptophan fluorescence was monitored at 350 nm over the indicated time frame and evaluated as described in the Materials and methods section. The buffer used was 10 mM sodium phosphate (pH 7.5), 0.1 mM EDTA and 1 M NaCl.

were incubated with ADP/Mg^{2+}, and fluorescence was measured over the indicated time frame. The observed donor fluorescence increased over time in a strictly mono-exponential fashion for all dimeric probes, although it remained constant for the monomeric probe. Thus the increase in donor fluorescence corresponds to dimer decay, and the decay reaction was evaluated according to eqn (3). This analysis provided the amplitudes and half times of dimer decay and could be used to calculate the transfer efficiency for the individual combinations according to eqn (4) by extrapolation of the data to $t = 0$ (Figure 5B). This extrapolation allowed us to extract the transfer efficiencies without prior knowledge of protein concentration or the exact starting time point of the experiment. A summary of the calculated values is given in Table 2. This demonstrates the agreement between transfer efficiencies determined by static and time-resolved FRET.

Additional support for the ADP-induced dimer decay comes from independent experiments using the F19W mutant of OpuAA (Figure 6). Phe[19], which based on sequence comparisons is interacting with bound nucleotide via π–π stacking, was mutated to a tryptophan residue to use intrinsic tryptophan fluorescence quenching to monitor nucleotide association and dissociation similar to the approach employed for the isolated NBD of HlyB [37]. Kinetic parameters of this mutant are summarized in Table 1. As shown in Figure 6, chasing of dimeric nucleotide-free OpuAA with ADP resulted in a quenching of the detected tryptophan fluorescence. As analysed for the monomeric form of OpuAA, the observed fluorescence reduction

is due to nucleotide binding and $\pi-\pi$ interaction of the adenine moiety of ADP or ATP and Trp[19] (Supplementary Figure S3 at http://www.BiochemJ.org/bj/412/bj4120233add.htm). However, the determined half time of fluorescence decay (18 ± 4 min) is similiar with the half times determined for the time-resolved fluorescence experiments within experimental error (Table 3).

Based on the FRET-derived distance information for the two homo-pairs (S45C/S45C and G161C/G161C) and the hetero-pair (S45C/G161C) in the nucleotide-free and ATP-bound state (Table 2), a molecular model was created using FRETsg as anchor points for a structural alignment in O [38]. Although the Cα–Cα distances obtained from the FRET studies of the homo-pairs [56.9 Å (45–45) and 51.3 Å (161–161)] agreed with the open conformation of MalK (59.9 Å and 51.9 Å respectively) [20], the distance of the S45C/G161C pair was approx. twice the distance measured in MalK (52.1 Å compared with 22.3 Å). Thus the open nucleotide-free dimer of MalK was adjusted accordingly. Here, one monomer was kept fixed while the other was treated as a rigid body. The resulting model for OpuAA satisfying all FRET-derived distance restrains is shown in Figure 7(A). The canonical part of each NBD monomer, composed of the catalytic and helical subdomains, is shown in yellow and light orange respectively, and the accessory domains are in grey. The Walker A and ABC-signature motifs are highlighted in blue and red respectively, whereas the Cα trace of the amino acids, which were mutated to cysteine residues and labelled with the corresponding fluorophores, is coloured green (position 45) and dark grey (position 161). Thin broken lines indicate the distances between the labelled amino acids, which were calculated from the FRET measurements. Interestingly, the closest Cα–Cα distance between opposing Walker A and ABC-signature motifs are 17 and 20 Å respectively. These values agree with the distance determined for the corresponding Cα atoms of the nucleotide-free open form of MalK (17 Å) from *E. coli* [20]. Thus it appears that the overall structure of the canonical part of the NBD is identical between OpuAA and MalK. However, the accessory domains are in a different orientation from that expected from the MalK structure. As a consequence, severe clashes occur between the canonical and accessory part of opposing OpuAA monomers (Figure 7A), indicating that the accessory domain has to adopt a different orientation. Furthermore, the D-loop motifs (SALD in OpuAA and SNLD in MalK) of opposing monomers are in close contact (Cα–Cα distance of 7 Å). As shown in Figure 7(B), structural superimposition reveals that one monomer of the OpuAA model aligns very well with a monomer of the nucleotide-free open structure of MalK [rmsd (root mean square deviation) of 1 Å over 381 Cα atoms]. The two monomers of MalK are coloured magenta and cyan and the Walker A and ABC-signature motifs in blue and red respectively. In clear contrast, the second monomer displayed a translational movement of 9.5 Å as calculated for the conserved lysine residue of the Walker A motifs, which are shown as blue ribbons (Lys[42] in MalK and Lys[72] in OpuAA). As a consequence, the accessory domain of the second monomer of the OpuAA model (light grey) is displaced compared with the MalK structure (light magenta). A structural superimposition of the OpuAA model and the nucleotide-free semi-open form of MalK [20] revealed that one monomer imposes slightly less accurately with OpuAA (rmsd of 1.9 Å over 336 Cα atoms; results not shown). The Cα–Cα distance of opposing Walker A and ABC-signature motifs is also shorter (13 Å) than in the FRET-derived model of OpuAA (17 and 20 Å respectively).

Following the same strategy as outlined for the nucleotide-free state of OpuAA, a model was derived for the ATP-bound dimeric state (Figure 8A) that satisfied the distance restrains derived from
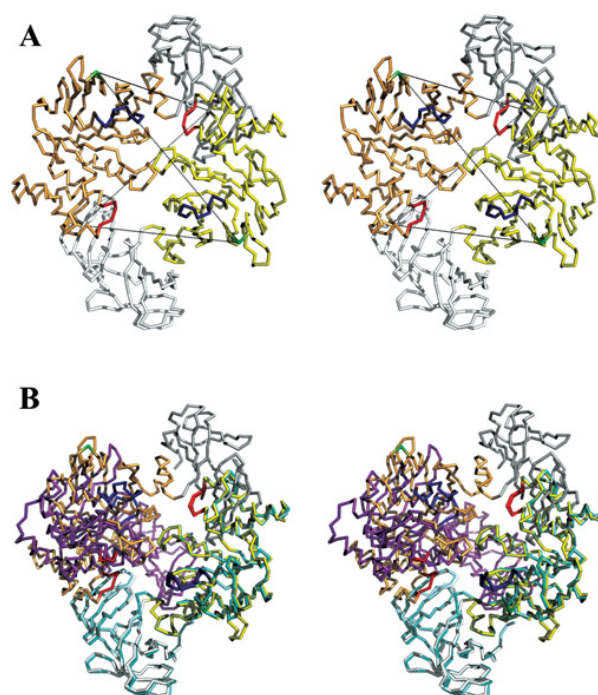
Figure 7   Stereo view of the FRET-derived homology model of nucleotide-free OpuAA

(**A**) The accessory domains are shown in white, whereas the canonical NBDs are given in yellow and light orange respectively. Note that the accessory domain of MalK was used in the homology modelling because proper location of the CBS domain, which is shown in Figure 1(C), is not possible. FRET-derived distances between amino acid residues 45 (green) and 161 (dark grey) are highlighted by black broken lines. The individual distances are given in Table 2. The Walker A motif is highlighted in blue and the ABC-signature motif in red. (**B**) Stereo view of a structural superimposition of the FRET-derived homology model of OpuAA and the open conformation of the nucleotide-free form of MalK (PDB entry: 1Q1E). Colour-coding of OpuAA is as in (**A**). Individual chains of MalK are shown in magenta and cyan, whereas the Walker A and ABC-signature motifs are highlighted in blue and red respectively. For structural superimpositions, the Walker A and B motifs were used. After improvement of the superimposition, the rmsd between monomer A of OpuAA and MalK was 1 Å over 381 Cα atoms.

FRET measurements in solution (Table 2). Here, the Cα–Cα distance between opposing Walker A and ABC-signature motifs shortens to 13 Å compared with the model of the nucleotide-free state. Furthermore, a superposition of the nucleotide-free and the ATP-bound state of the OpuAA model (Figure 8B) revealed that the canonical part of the NBD, coloured cyan and magenta, undergoes an induced fit upon ATP-binding, i.e. a rigid-body motion of the helical domain towards the catalytic domain.

DISCUSSION

The large body of biochemical data [39] and the crystal structures of MalK [20,21], the isolated NBD of the maltose transporter, combined with the recently published crystal structure of the full-length transporter [18], has provided a molecular picture of how substrate import is catalysed by this ABC importer. Thus a 'tweezers-like' motion was proposed for MalK. In this model, the additional domains of MalK form the bottom of the tweezers, keeping two MalK monomers together during the catalytic cycle, while the catalytic and helical domains of the MalK monomers undergo nucleotide-dependent association and dissociation reactions. A similar monomer–dimer equilibrium
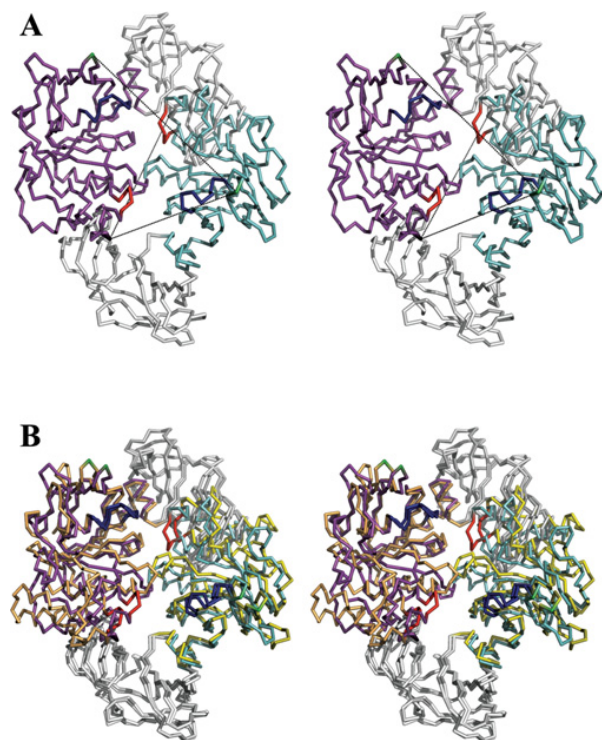
Figure 8   Stereo view of the FRET-derived homology model of OpuAA in the ATP-bound state

(**A**) ATP, which should be located between the Walker A (blue) and ABC signature (red) motifs of opposing monomers is not included in the model, because no structural knowledge of the exact position exists. The accessory domains, which adopt the structure of the MalK model, are depicted in white. The canonical NBDs are shown in cyan and magenta respectively. The anchors points (Ser[45] and Gly[161]), which were used in the FRET studies, are highlighted in green and black respectively and their distances are indicated by broken black lines (exact values are given in Table 2). (**B**) Stereo view of the structural superimposition of the FRET-derived homology models of OpuAA in the ATP-bound state (cyan and magenta) and the nucleotide-free state (yellow and light orange). Colour-coding of the conserved motifs is as in (**A**).

was observed for OpuAA from *B. subtilis* [11]. However, in contrast with MalK, dimeric OpuAA was only observed in the absence of nucleotides or the presence of ATP, although it remained monomeric in the presence of ADP. Moreover, dimeric OpuAA could be converted into the monomeric form in the presence of excess ADP. Such an ADP-induced monomerization of OpuAA, but not of MalK, might be due to differences in the accessory domains. In contrast with MalK, the primary structure of OpuAA indicates the presence of two CBS domains arranged in tandem (Figure 1). CBS domains are found, for example, in voltage-gated chloride channels [40] or IMPDHs [41] and are thought to bind small molecules containing an adenosyl moiety such as AMP, thereby regulating enzyme activity [42,43]. However, ATPase assays of OpuAA in the presence of AMP had no effect on activity and the stoichiometry of ATP (results not shown and [11]). This suggests that the CBS domains of OpuAA do not interact with AMP in a functionally relevant manner. Nevertheless, the core domain of MalK (catalytic and helical domain) is of sufficient homology (36 % identity and 69 % homology) with OpuAA to allow homology modelling (Figure 1), while a homology model of the accessory domain of OpuAA was constructed based on its closest structurally known homologue, TA0289 (Figure 1C). Based on this model, three amino acids were chosen and converted

into cysteine residues (S45C, G161C and S171C) for subsequent fluorophore labelling.

Wild-type and the three mutant proteins could be overexpressed and purified to homogeneity (Supplementary Figure S1). As summarized in Table 1, turnover rates for the wild-type enzyme and the S45C and G161C mutants were identical within experimental error. Also the K_m values for the wild-type enzyme and the S45C were identical, whereas the K_m value for the G161C mutant increased by a factor of 5. The result is somewhat surprising since Gly[161] is a surface-exposed residue and the structural model provides no explanation for this result. Since all subsequent experiments were performed under ATP-saturating conditions, these changes in K_m do not influence the results. In clear contrast, the S171C showed no ATPase activity and was not used in the following analysis.

Comparing this activity with the *in vivo* activity is not possible, because there exist no data on the number of transporters per cell [44]. However, assuming an activity typical for ATP transporters (approx. 1 ATP/s), indicates that the activity of the isolated NBD *in vitro* is 60 times lower than expected. But already in detergent solution, a 10-fold increase of the ATPase activity was observed in the presence of the TMD and the substrate-loaded SBP [9]. This clearly suggests that a properly assembled transporter requires higher activity than displayed by the isolated NBD. However, the results of the re-assembly studies in detergent solution [9] support the notion that the isolated NBD is properly folded and adopts a transport-competent state that can be stimulated in a SBP- and substrate-dependent manner.

The S45C and S161C were labelled employing two thiol-specific maleinimide fluorophores, OG and TR. The use of TR and OG, an efficient FRET pair [30] with a Förster radius of 51 Å, should in principle allow us to extract distance information in the different functional states of the isolated NBD, i.e. nucleotide-free, ATP- and ADP-bound. However, labelling efficiencies were at best 60 %, a result that cannot be explained based on the homology model, because the two residues chosen are surface-exposed. Instead of trying to optimize labelling efficiency, we developed an additional purification step to separate labelled and non-labelled species (Figure 2). Although the ion-exchange chromatogram indicated quantitative separation, we further analysed the identity and homogeneity of the two fractions obtained after ion-exchange chromatography by ESI–MS (Figure 3). Importantly, the amount of non-labelled OpuAA as indicated by an arrow in Figures 3(B) and 3(C) was below 2 % as calculated from the peak areas of the ESI–MS spectra of OpuAA S45C–OG and OpuAA S45C–TR respectively. These two independent sets of experimental evidence demonstrate a quantitative separation of labelled and non-labelled OpuAA, which is a prerequisite for subsequent quantitative FRET measurements.

A process to assemble and de-assemble homodimers of OpuAA has already been described in detail [11] and was readily applied to the labelled single cysteine mutants. However, FRET efficiency strongly depends on a complete donor–acceptor pairing [30]. Statistical mixing of a 1:1 mixture of OG- and TR-labelled OpuAA would in theory result in the formation of 25 % dimers composed of donors only (OG/OG), 25 % dimers composed of acceptors only (TR/TR) and 50 % composed of donor and acceptor (OG/TR). Although such statistical mixing has been verified experimentally for other isolated ABC-NBDs [37,45,46], and additionally a mathematical treatment to correct incomplete assembly is possible in principle, the dynamic monomer–dimer equilibrium of OpuAA would further complicate the evaluation. Therefore we decided to use a molar ratio of donor/acceptor of

1:9 to ensure quantitative 'pairing' of OpuAA–OG with OpuAA–TR. Under these conditions, FRET could be observed for the 45–45, 45–161 and 161–161 pairs in the nucleotide-free and ATP-bound state, but not in the ADP-bound state (Figure 4). Transfer efficiencies are summarized in Table 2. In the light of the biochemical characterization of the isolated NBD of the OpuA transporter [11], these results demonstrate that a dimer is present in the nucleotide-free and ATP-bound state respectively, whereas the protein is monomeric in the ADP-bound state. However, one has to realize that a determination of transfer efficiencies relies on an accurate protein concentration determination. Since the molar absorption coefficient of the dyes changed upon coupling (Supplementary Figure S2), protein concentration determination might become unreliable. To overcome, this problem, we also performed time-resolved FRET experiments (Figure 5). In this case, the decay in FRET signal of nucleotide-free or ATP-bound dimeric OpuAA samples was measured over time in the presence of 10 mM ADP, which induces dimer dissociation. From the determined amplitudes and the half time of dimer decay, transfer efficiencies could be calculated in a protein-concentration-independent way (Table 2). The agreement between these values and the values determined from the static measurements (Table 2) suggests that the distances, which can be derived from the transfer efficiencies, can be used as distance restrains for modelling the dimeric state of OpuAA.

Similarly to the HlyB system, the introduction of a tryptophan residue at position 19 of OpuAA lowered the catalytic efficiency of the enzyme 2-fold (Table 1). However, in our assays, saturating ligand concentrations were used and, more importantly, hydrolysis was prevented by adding EDTA to the buffer so that the enzymatic activity is not as important and the 2-fold reduction can be tolerated. Thus the chase of the nucleotide-free dimer of OpuAA F19W with ADP or ADP/Mg^{2+} displays the same result, as the ADP-induced dissociation of the OpuAA dimer detected by fluorescence, but FRET experiments report changes in distance, i.e. dimer decay directly, the F19W mutant uses nucleotide association/dissociation as the read-out. However, one has to keep in mind that the k_{cat} value of ATPase activity under these conditions is approx. 1 min^{-1} (Table 1 and [11]), whereas the half time of dimer decay was between 12.9 ± 0.1 and 22 ± 0.3 min. Stopped-flow experiments employing changes of the intrinsic tryptophan signal (results not shown) revealed that nucleotide association was faster than the time resolution of the experiment (approx. 1 ms). Thus ATPase activity is either limited by hydrolysis as suggested for the HlyB–NBD [28] or the dissociation of the dimer as suggested for MJ1276 [27]. The current model of ATP hydrolysis within ABC domains assumes that dimer decay occurs every time the two ATPs bound to the composite dimer are hydrolysed. Such a scenario is hard to imagine for OpuAA, because a k_{cat} of 1 min^{-1} and the half time of dimer decay (between 12 and 22 min) are not compatible with this model. However, Gadsby and co-workers [47] have demonstrated that, in CFTR (cystic fibrosis transmembrane conductance regulator), repetitive cycles of ATP hydrolysis can occur before the dimer decays. Clearly, in CFTR the NBD composite sites are unequal, because the degeneration of certain residues critical for catalysis generates one ATP-binding site with no or only little ATP activity, whereas the other site contains the canonical residues and is capable of hydrolysing ATP. Although OpuAA does not contain an asymmetrical ATP-binding site, it is tempting to speculate that the 10–20-fold difference between ATP turnover and dimer decay is due to a similar phenomena. ATP hydrolysis and dimer decay are not strictly coupled, and repetitive cycles can occur in both binding sites before the dimer dissociates. Dimer decay on the other hand would finally terminate ATPase activity. That

would also suggest that a single ATP hydrolysis event and the overall termination of activity (dimer decay) might have different rate-limiting steps. However, this hypothesis requires further verification, and, although appealing, the differences between CFTR and OpuAA in sequence and the resulting functional consequences are obvious, therefore further experimental analysis is necessary before the discrepancy between turnover and dimer decay can be resolved.

Based on the distance restrains derived from the static and time-resolved FRET measurements (Table 2), the dimeric homology models of OpuAA (nucleotide-free and ATP-bound) that were based on the dimeric crystal structures of MalK, were modified to satisfy the restrains (Figures 7 and 8). As shown, in Figures 7(A) and 8(A), the FRET-derived distances can only be incorporated into the homology models if the accessory domains are not in close contact in OpuAA as they are in MalK. Furthermore, the catalytic and helical domains in the second monomer are translationally shifted by 9 Å if comparing MalK with OpuAA (Figure 7B). This excludes that a simple rotation of the linker between core NBD and the CBS-containing accessory domain is sufficient to generate a 'MalK-like' arrangement. However, one has to stress that the accessory domain of the OpuAA model shown in the present study contains the MalK-fold, whereas sequence analysis clearly revealed the presence of CBS domains in OpuAA that adopt a fold similar to the one shown in Figure 1(C). The low structural homology between these two folds prohibited any further homology modelling. Analysis of the buried surface area of the OpuAA model revealed a value of 130 Å^2 compared with 1900 Å^2 in the ATP-bound state of MalK [20]. Clearly, the calculated value of the buried surface area is too small to explain the dimeric nature of nucleotide-free OpuAA in solution [11]. However, the accessory domains, as shown in Figure 7(A), do not have any contact, so the only stabilization is provided by interactions of the canonical part (helical and catalytic domain) of OpuAA. Thus the FRET-derived model strongly suggests that the CBS-containing accessory domains stabilize the dimeric nature of this ATPase subunit, but that the relative orientation of canonical NBD and the accessory domain is different in OpuAA and MalK. This finding is similar to the situation observed in the crystal structure of CysA [48], the NBD of a putative sulfate-import system. In this case, the orientation of the accessory domain in one of the two dimers is also different from that of MalK and suggests that the relative three-dimensional orientation of the core NBD and the accessory domain is different in different NBDs and might be a consequence of functional requirements imposed on the transporter (nature of substrate, interacting molecules/proteins, etc.). On the other hand, a rigid-body inward motion is observed for the helical domain of the ATP-bound model compared with the nucleotide-free state (Figure 8B). This rigid-body motion of induced fit upon ATP-binding was first proposed for MJ1276 [49] and subsequently verified in the crystal structures of other isolated NBDs in the ATP-bound state [20,27–29]. Here, the inward motion ranged from 16° to 25°. In our model, the helical domain of the nucleotide-free state (coloured yellow and light orange in Figure 8B) undergoes an inward motion of approx. 17°. This demonstrates that the basic motion in the core NBD is also preserved in OpuAA.

In summary, although the orientation of the accessory domain is not identical in OpuAA and MalK, the basic principles of structural changes occurring within the helical domain upon ATP-binding that have been elaborated for many systems are also conserved in OpuAA. It is also tempting to speculate that the OpuAA dimer adopts the open conformation of MalK in solution in the absence of nucleotides and not the semi-open state, but with a different arrangement of the accessory domain

that very probably stabilizes the dimer and also performs a tweezer-like function as it was proposed for MalK [20]. Further studies, such as additional FRET or cysteine-based cross-linking experiments, are necessary to determine the exact orientation of the CBS domains to each other, but a definitive answer will only become available upon determining the three-dimensional crystal structure of OpuAA in different functional states. However, in the light of the proposed role of this module in regulating transport function, it is important to realize that not only the primary structure, but also the quaternary arrangement is important for *in vivo* function.

REFERENCES

1 Wood, J. M., Bremer, E., Csonka, L. N., Kraemer, R., Poolman, B., van der Heide, T. and Smith, L. T. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria. Comp. Biochem. Physiol. A Mol. Integr. Physiol. **130**, 437–460

2 Whatmore, A. M. and Reed, R. H. (1990) Determination of turgor pressure in *Bacillus subtilis*: a possible role for K$^+$ in turgor regulation. J. Gen. Microbiol. **136**, 2521–2526

3 Miller, K. J. and Wood, J. M. (1996) Osmoadaptation by rhizosphere bacteria. Annu. Rev. Microbiol. **50**, 101–136

4 Wood, J. M. (1999) Osmosensing by bacteria: signals and membrane-based sensors. Microbiol. Mol. Biol. Rev. **63**, 230–262

5 Ventosa, A., Nieto, J. J. and Oren, A. (1998) Biology of moderately halophilic aerobic bacteria. Microbiol. Mol. Biol. Rev. **62**, 504–544

6 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

7 Welsh, D. T. (2000) Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. FEMS Microbiol. Rev. **24**, 263–290

8 Minton, A. P. (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J. Biol. Chem. **276**, 10577–10580

9 Horn, C., Bremer, E. and Schmitt, L. (2005) Functional overexpression and *in vitro* assembly of OpuA, an osmotically regulated ABC-transporter from *Bacillus subtilis*. FEBS Lett. **579**, 5765–5768

10 Schmitt, L. and Tampé, R. (2002) Structure and mechanism of ABC-transporters. Curr. Opin. Struct. Biol. **12**, 754–760

11 Horn, C., Bremer, E. and Schmitt, L. (2003) Nucleotide dependent monomer/dimer equilibrium of OpuAA, the nucleotide-binding protein of the osmotically regulated ABC transporter OpuA from *Bacillus subtilis*. J. Mol. Biol. **334**, 403–419

12 Horn, C., Sohn-Bosser, L., Breed, J., Welte, W., Schmitt, L. and Bremer, E. (2006) Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. J. Mol. Biol. **357**, 592–606

13 Oswald, C., Holland, I. B. and Schmitt, L. (2006) The motor domains of ABC-transporters: what can structures tell us? Naunyn Schmiedebergs Arch. Pharmacol. **372**, 385–399

14 Dawson, R. J. and Locher, K. P. (2006) Structure of a bacterial multidrug ABC transporter. Nature **443**, 180–185

15 Locher, K. P., Lee, A. T. and Rees, D. C. (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. Science **296**, 1091–1098

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

17 Hollenstein, K., Frei, D. C. and Locher, K. P. (2007) Structure of an ABC transporter in complex with its binding protein. Nature **446**, 213–216

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

19 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 **317**, 1387–1390

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

21 Lu, G., Westbrooks, J. M., Davidson, A. L. and Chen, J. (2005) ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. Proc. Natl. Acad. Sci. U.S.A. **102**, 17969–17974

22 Panagiotidis, C. H., Boos, W. and Shuman, H. A. (1998) The ATP-binding cassette subunit of the maltose transporter MalK antagonizes MalT, the activator of the *Escherichia coli* mal regulon. Mol. Microbiol. **30**, 535–546

23 Bluschke, B., Volkmer-Engert, R. and Schneider, E. (2006) Topography of the surface of the signal-transducing protein EIIA(Glc) that interacts with the MalK subunits of the maltose ATP-binding cassette transporter (MalFGK2) of *Salmonella typhimurium*. J. Biol. Chem. **281**, 12833–12840

24 van der Heide, T. and Poolman, B. (2000) Glycine betaine transport in *Lactococcus lactis* is osmotically regulated at the level of expression and translocation activity. J. Bacteriol. **182**, 203–206

25 Bateman, A. (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein. Trends Biochem. Sci. **22**, 12–13

26 Biemans-Oldehinkel, E., Mahmood, N. A. and Poolman, B. (2006) A sensor for intracellular ionic strength. Proc. Natl. Acad. Sci. U.S.A. **103**, 10624–10629

27 Smith, P. C., Karpowich, N., Millen, L., Moody, J. E., Rosen, J., Thomas, P. J. and Hunt, J. F. (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Mol. Cell **10**, 139–149

28 Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B. and Schmitt, L. (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. EMBO J. **24**, 1901–1910

29 Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B. and Schmitt, L. (2006) A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. EMBO J. **25**, 3432–3443

30 Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, Kluwer Academic Publishers, New York

31 Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. **29**, 291–325

32 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291

33 Kleywegt, G. J. (1996) Use of non-crystallographic symmetry in protein structure refinement. Acta Crystallogr. Sect. D Biol. Crystallogr. **52**, 842–857

33a Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **50**, 760–763

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

35 Diederichs, K., Diez, J., Greller, G., Muller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W. and Welte, W. (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. EMBO J. **19**, 5951–5961

36 Berney, C. and Danuser, G. (2003) FRET or no FRET: a quantitative comparison. Biophys. J. **84**, 3992–4010

37 Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B. and Schmitt, L. (2005) Functional characterization and ATP induced dimerization of the isolated ABC-domain of the haemolysin B transporter. Biochemistry **44**, 9680–9690

38 Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. Sect. A Found. Crystallogr. **47**, 110–119

39 Davidson, A. L. and Chen, J. (2004) ATP-binding cassette transporters in bacteria. Annu. Rev. Biochem. **73**, 241–268

40 Meyer, S. and Dutzler, R. (2006) Crystal structure of the cytoplasmic domain of the chloride channel ClC-0. Structure **14**, 299–307

41 Zhang, R., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A. and Collart, F. R. (1999) Characteristics and crystal structure of bacterial inosine-5′-monophosphate dehydrogenase. Biochemistry **38**, 4691–4700

42 Bennetts, B., Rychkov, G. Y., Ng, H. L., Morton, C. J., Stapleton, D., Parker, M. W. and Cromer, B. A. (2005) Cytoplasmic ATP-sensing domains regulate gating of skeletal muscle ClC-1 chloride channels. J. Biol. Chem. **280**, 32452–32458

43 Scott, J. W., Hawley, S. A., Green, K. A., Anis, M., Stewart, G., Scullion, G. A., Norman, D. G. and Hardie, D. G. (2004) CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. J. Clin. Invest. **113**, 274–284

44 Kempf, B. and Bremer, E. (1995) OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. J. Biol. Chem. **270**, 16701–16713

45 Janas, E., Hofacker, M., Chen, M., Gompf, S., van der Does, C. and Tampe, R. (2003) The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p. J. Biol. Chem. **278**, 26862–26869

46   Nikaido, K. and Ames, G. F. (1999) One intact ATP-binding subunit is sufficient to support ATP hydrolysis and translocation in an ABC transporter, the histidine permease. J. Biol. Chem. **274**, 26727–26735

47   Vergani, P., Lockless, S. W., Nairn, A. C. and Gadsby, D. C. (2005) CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains. Nature **433**, 876–880

48   Scheffel, F., Demmer, U., Warkentin, E., Hulsmann, A., Schneider, E. and Ermler, U. (2005) Structure of the ATPase subunit CysA of the putative sulfate ATP-binding cassette (ABC) transporter from *Alicyclobacillus acidocaldarius*. FEBS Lett. **579**, 2953–2958

49   Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y. R., Dai, P. L., MacVey, K., Thomas, P. J. and Hunt, J. F. (2001) Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. Structure **9**, 571–586

SUPPLEMENTARY ONLINE DATA

Monitoring conformational changes during the catalytic cycle of OpuAA, the ATPase subunit of the ABC-transporter OpuA from *Bacillus subtilis*

Carsten HORN*[1,2], Stefan JENEWEIN*[1], Britta TSCHAPEK*, Werner BOUSCHEN†, Sabine METZGER†, Erhard BREMER‡ and Lutz SCHMITT*[3]

*Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, †Biological and Medical Research Center, Heinrich Heine University, Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, and ‡Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Str., 35032 Marburg, Germany

Figure S1   Purification of wild-type and single cysteine mutants of OpuAA

Left-hand panel: Coomassie-Blue-stained SDS/PAGE (15 % gel) of OpuAA wild-type, S45C and G161C. Right-hand panel: Coomassie-Blue-stained SDS/PAGE (15 % gel) under reducing (Red.) and oxidizing (Oxid.) conditions of OpuAA S45C and G161C. Molecular masses in kDa of the marker proteins are given on the left-hand side of each panel.
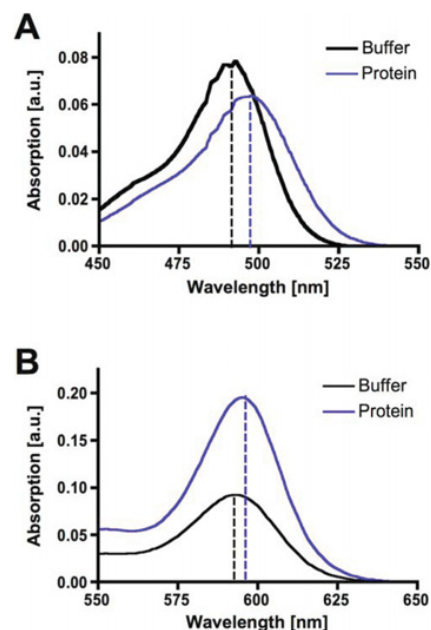
Figure S2   Absorption spectra of identical concentrations of OG (A) and TR (B) in buffer [10 mM sodium phosphate (pH 7.5) and 1 M NaCl; black lines) and after covalent attachment to OpuAA [10 mM sodium phosphate (pH 7.5) and 1 M NaCl; blue lines]

The vertical broken lines indicate the change in the absorption maxima. As depicted, the absorption coefficent at the absorption maxima increased 2-fold for TR but decreased 1.2-fold for OG after the labelling reaction. a.u., absorbance unit.
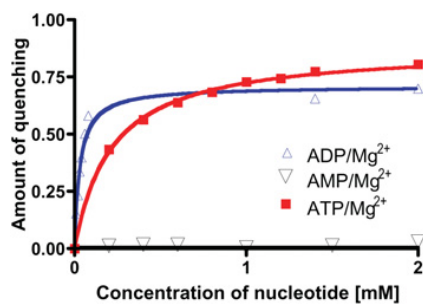
Figure S3   Nucleotide-induced tryptophan-quenching of the F19W OpuAA mutant

Data were recorded as described in the Materials and methods section of the main text at $20 \pm 1\,^{\circ}$C in 10 mM sodium phosphate (pH 7.5), 1 M NaCl and 2 mM Mg^{2+} using monomeric OpuAA. Data were analysed using a 1:1 Langmuir binding isotherm and K_d values were calculated to be $215 \pm 9\ \mu$M (ATP/Mg^{2+}) and $30 \pm 0.4\ \mu$M (ADP/Mg^{2+}). The control AMP/Mg^{2+}, shows the specificity of the binding reaction. This specificity clearly demonstrates the involvement of residue 19 of OpuAA in nucleotide binding.

Own proportion to this work: 10%

Published in: Biochemical Journal 2008

Impact factor: 5.155

Chapter II

Publication in preparation

Osmoregulatory ABC-transporter OpuB from *Bacillus subtilis*: crystal structure of the solute receptor OpuBC in complex with choline

Marco Pittelkow[2†], Britta Tschapek [1†], Sander H. J. Smits[1], Lutz Schmitt[1]* and Erhard Bremer[2]*

[1] Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, D-402325 Düsseldorf, Germany

[2] Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Str. 8, D-35032 Marburg, Germany,

For correspondence during the reviewing and editorial process:
Lutz Schmitt; Phone: 49(0)211-81-10773 Fax: 49(0)211-81-15310. E-mail: lutz.schmitt@uni-duesseldorf.de

Running Title: Choline binding protein OpuBC

*Corresponding author. Mailing address: Lutz Schmitt, Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany. Phone: 49(0)211-81-10773 Fax: 49(0)211-81-15310. E-mail: lutz.schmitt@uni-duesseldorf.de

*Corresponding author. Mailing address: Erhard Bremer, Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Str. 8, D-35032 Marburg, Germany. Phone: (49)-6421-2821529. Fax: (49)-6421-2828979. E-mail: bremer@staff.uni-marburg.de.

† Both authors contributed equally to this work.

Abstract

Bacillus subtilis synthesizes the compatible solute glycine betaine from an exogenous supply of the precursor choline as an osmoprotectant. Import of choline is mediated by two osmotically inducible ABC-transport systems, OpuB and OpuC. OpuB is specific for choline, whereas OpuC mediates the import of various osmoprotectants. OpuBC is the substrate binding protein of the OpuB transporter and we have analyzed the affinity of the OpuBC/choline complex by intrinsic tryptophan fluorescence (K_d approximately 30 µM). The X-ray crystal structure of the OpuBC/choline complex revealed a fold typical for class II substrate-binding proteins with the ligand trapped in a deep cleft formed by the two domains of OpuBC. The positively charged trimethylammonium head-group of choline is wedged into an aromatic cage formed by four tyrosine residues and is bound via cation-pi interactions. The alcohol group of choline protrudes from this aromatic cage and makes a single direct interaction with Gln-19. The substitution of this residue by Ala decreases choline binding-affinity by approximately 15-fold. A water network stabilizes choline within its substrate-binding site and promotes indirect interactions between the two lobes of OpuBC. Disruption of this intricate water network by site directed mutagenesis of selected residues in OpuBC either strongly reduces choline-binding affinity (between 18-fold and 25-fold), or abrogates ligand binding altogether. A comparison of the OpuBC/choline structure with the crystal structure of the glycine betaine binding protein ProX from *Archaeoglobus flulgidus* revealed why glycine betaine cannot be bound stably by OpuBC despite the fact that the trimethylammonium head group of this ligand would fit correctly into the aromatic cage of OpuBC.

Introduction

Changes in the osmotic conditions of the environment are key factors that influence growth and survival of microorganisms in their varied ecological niches (*1*). Since most microorganisms possess a turgor, an intracellular hydrostatic pressure considered to be essential for cell expansion (*2*). increases or decreases in the external osmolarity inevitably trigger osmotically driven water fluxes across the semi-permeable cytoplasmic membrane. The direction and magnitude of these water fluxes is determined by the differential in the osmotic potential between the interior of the cell and the surrounding (*3, 4*).

The slowing of cell growth and the eventual growth arrest under severe high osmolarity conditions is correlated with the diminished content of free water in the bacterial cell (*5*). To counteract the efflux of water and to promote a physiologically adequate level of hydration and turgor, microorganisms initiate a cellular adjustment process that aims at an increase in the osmotic potential of the cytoplasm through the amassing of water-attracting solutes (*3, 4*). In many microorganisms, this is typically a multiphasic process in which potassium ions are initially imported as an immediate stress reaction. This initial phase is then followed by the amassing of a selected class of organic osmolytes, the compatible solutes, for a sustained adjustment to high osmolarity surroundings (*1-4*). The accumulation of compatible solutes allows the cell to export the initially imported potassium ions and thereby reduce the ionic strength of the cytoplasm without compromising its osmotic potential (6).

Compatible solutes possess special biophysical and biochemical properties that make them highly congruous with cellular functions so that cells can amass them to exceedingly high levels (*7, 8*). These organic osmolytes have been selected in the course of evolution as effective cytoprotectants not only by members of the *Archaea* and *Bacteria* but also by *Eukarya* (*1, 9, 10*). Microbial cells accumulate compatible solutes either through synthesis or uptake for environmental resources (*3, 4, 11, 12*). Glycine betaine, a tri-methylated derivative of the amino acid glycine, is without any doubt the most widely used compatible solute in nature (*1, 9, 10, 13*). Most microbes that have been studied in the context of their adaptation to high osmolarity environments have been found to possess osmotically stimulated high-affinity uptake systems for glycine betaine (*2-4, 11, 12*).

Frequently, microorganisms can also synthesize glycine betaine and they can do so either *via* de novo synthesis or by oxidizing the precursor choline. The de novo synthesis of glycine betaine is accomplished by a step-wise methylation of the amino acid glycine with *S*-adenosylmethionine as the methyl donor (*14*). This enzymatic process is energetically very expensive due to the required regeneration of the *S*-adenosylmethionine cofactor (*14*). Consequently, most microorganisms synthesize glycine betaine by oxidizing choline, an energetically more favorable process. For this latter biosynthetic route, choline is first imported from the environment through osmotically inducible transporters and this precursor can then be converted into glycine betaine by a combination of various enzymes (*15-18*). It should be

noted in this context, that choline has no osmoprotective property *per se* in microorganisms (*19*).

Glycine betaine plays a key role in the defense of the soil-dwelling Gram-positive bacterium *Bacillus subtilis* against high osmolarity challenges that are caused by the drying of the upper layers of the soil (*20*). *B. subtilis* possesses three high-affinity uptake systems for this osmoprotectant: the ABC-transporters OpuA and OpuC and the BCCT-type transporter OpuD (Fig. 1)(*21-23*). The expression of the structural genes for these transport systems is up-regulated in high-salinity stressed *B. subtilis* cells (*24, 25*). thereby providing more glycine betaine transport capacity to the cell under osmotic stress conditions (*21, 22*). Glycine betaine can also be synthesized by *B. subtilis* from the precursor choline (*19*). For this purpose, choline is imported with high affinity through the substrate-specific OpuB ABC-transporter and the promiscuous ABC-transporter OpuC (Fig. 1)(*20, 23*). The imported choline is then oxidized to glycine betaine by the concerted actions of the alcohol dehydrogenase GbsB and the glycine betaine aldehyde dehydrogenase GbsA (Fig. 1)(*15*). The choline to glycine betaine biosynthetic pathway provides *B. subtilis* with a considerable degree of osmotic tolerance (*19, 23*).

Where does the choline for glycine betaine synthesis come from in the main natural habit of *B. subtilis*, the soil? Choline is an important component of the eukaryotic cell membrane and hydrolytic enzymes released by microorganisms liberate the choline moiety from lipids of decaying cells, primarily from rotting plant material (*19, 26*). Since the choline content of the soil is very low and variable (*26*), high affinity uptake systems are requited to effectively scavenge it. The kinetic parameters of the *B. subtilis* OpuB and OpuC transporters with K_m-values in the low μM range, make these two ABC systems suitable for such a task (*23*).

The structural genes for the OpuB and OpuC ABC importers are located only a few kilo-base pairs apart on the *B. subtilis* genome and the high degree of amino acid sequence identity of their components (up to 70 % in the SBPs) suggests that these uptake systems have evolved from a gene duplication event (*23*). The OpuB transporter is highly specific for the uptake of choline, whereas OpuC not only imports choline but also various other compatible solutes (Fig. 1)(*20*). Gene duplication events are an important driving force in microbial evolution since they generate gene copies onto which adaptive evolution can act without impairing the physiological function of the original gene (*27*). Hence, a possible scenario that can be envisioned is the initial presence of the substrate-specific OpuB choline uptake system that worked in conjunction with the GbsA and GbsB glycine betaine biosynthetic enzymes (Fig. 1). The *opuB* operon was then duplicated and the resulting second copy of the transporter subsequently evolved into an uptake system with a broad substrate specificity, OpuC. In this way, the spectrum of compatible solutes that could be used by *B. subtilis* cells as osmoprotectants was significantly expanded (Fig. 1)(*20*) since some of the known osmoprotectants of *B. subtilis* (e.g. L-carnitine, choline-*O*-sulfate, ectoine) enter the cell exclusively via the OpuC transporter (*28-30*).

If the above described scenario for the evolution of the *B. subtilis* OpuB and OpuC transporters is correct, then the question arises what are the structural determinants of the OpuB and OpuC ABC transport systems that govern their substrate specificity? It is the generally accepted view that the substrate specificity of a given microbial binding-protein-dependent ABC transporter is primarily determined by the substrate specificity of the cognate ligand binding protein (*31, 32*). However, it is known from the genetic and structural analysis of the maltose/maltodextrin ABC transporter from *Escherichia coli*, that a low-affinity substrate binding site is present in one (MalF) of the two permease components (MalF and MalG) of the transporter and that this binding site contributes significantly to substrate import and specificity (*33-35*).

Since *B. subtilis* is amenable to sophisticated genetic techniques, the sequence-related OpuB and OpuC transporters are attractive model systems to study the evolution of the substrate specificity of microbial binding-protein-dependent ABC transporters. The substrate binding proteins (OpuBC and OpuCC) of the OpuB and OpuC ABC transporters are lipoproteins that are tethered on the outer surface of the cytoplasmic membrane of *B. subtilis* (Fig. 1)(*23*). The amino acid sequence of the OpuBC and OpuCC proteins are 25 % identical, but judging from the substrate profile of their cognate transporters (Fig. 1), their substrate specificity must be very different.

To further characterize the OpuB and OpuC ABC transport systems, we report here a high-resolution crystal structure of the OpuBC ligand binding protein in complex with its natural ligand choline. This crystal structure does not only reveal the molecular determinants for high-affinity choline binding by OpuBC but it also provides a rational explanation why glycine betaine, a substrate for the OpuC transporter but not for the OpuB system (Fig. 1), cannot be stably bound by the OpuBC ligand binding protein.

Results and Discussion

Purification and functional characterization of the *B. subtilis* OpuBC protein

The OpuBC protein is initially synthesized in *B. subtilis* as a precursor protein with a 22 amino acid long signal sequence to direct its export across the cytoplasmic membrane. During protein translocation, the OpuBC signal sequence is proteolytically removed by a signal sequence-specific peptidase and the newly formed mature N-terminus of the OpuBC protein (Cys-23) is lipid modified (Fig. 1)(*23*). We constructed an expression system for the *opuBC* gene that allowed the recombinant production of the *B. subtilis* OpuBC protein in *E. coli* and its purification via *Step*-tag-II affinity chromatography. The OpuBC protein used for the subsequent biochemical and crystallization experiments (Fig 2A) did not possess the originally present *Strep*-Tag-II affinity peptide but it carried a Gly residue (instead of Cys-23) at its N-terminus (see Materials and Methods) in order to prevent its lipid modification.

Substrate-binding activity of OpuBC

The OpuBC protein contains three Trp and 17 Tyr residues spread throughout its amino acid sequence. Using the intrinsic Trp/Tyr fluorescence of the purified OpuBC protein, we measured an emission spectrum between 290 and 400 nm and detected a maximum of fluorescence emission at about 340 nm (Fig 2B). The addition of 100 μM choline, the physiological substrate of the OpuB transporter (*23*) to the OpuBC protein solution led to an increase in the intensity of the intrinsic fluorescence (Fig 2B), consistent with the expected binding of choline to the OpuBC solute receptor. In contrast, the addition of 1 mM glycine betaine to the OpuBC protein solution did not trigger such a change in the fluorescence intensity (data not shown). Since the properties of ligand-binding proteins typically determine the substrate specificity of a given microbial ABC importer (*31, 32*) the recombinantly produced OpuBC protein exhibited a substrate-binding specificity for choline expected from previously reported osmoprotection assays and transport experiments conducted with intact *B. subtilis* cells (*23*).

Equilibrium binding experiments of OpuBC with its ligand choline

The change in the intrinsic Trp/Tyr fluorescence of OpuBC observed after addition of choline to the purified protein (Fig 2B) was used to determine the dissociation constant for the OpuBC:choline complex. For these experiments, we kept the OpuBC concentration constant (1 μM) and varied the choline concentration from 1 μM to 200 μM. From the resulting spectroscopic data we calculated a K_D value for choline of 30.5 ± 3.2 μM (Fig 2C). This K_d value is very similar to that for the OpuAC:glycine betaine complex (approximately 20 μM) determined by the same method for the OpuAC ligand binding protein (Fig 1) from the *B. subtilis* OpuA ABC-transport system (*36, 37*).

Crystal structure of OpuBC from *Bacillus subtilis* in complex with choline

As noted previously by Schiefner *et al.* (*38*), the amino acid sequence of the *B. subtilis* choline-binding protein OpuBC is related (29 % and 48 % amino acid sequence identity and similarity, respectively) to the glycine betaine binding protein ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus* (Fig 5). In particular, four Tyr residues that are part of the glycine betaine ligand-binding site of ProX (*38*) are conserved in the amino acid sequences of the OpuBC protein (Fig 4).

We took advantage of the available 2.1 Å crystal structure of the ProX:glycine betaine complex (*38*) to solve the crystal structure of the *B. subtilis* OpuBC solute receptor in complex with choline by molecular replacement. Crystals of the OpuBC protein were grown in the presence of choline and a native data set of the OpuBC:choline complex was collected at beamline ID-23 (ESRF, Grenoble) and scaled using XDS (*39*). Initial phases were obtained by molecular replacement using the program PHASER (*40*) with the crystal structure of ProX from *A. fulgidus* as a template (PDB entry: 1SW2)(*38*). The final structure was refined to a resolution of 2.0 Å, which allowed the modeling of water molecules. Data and refinement statistics as well as model content are summarized in Table 1.

The OpuBC protein was crystallized in the presence of choline and yielded a closed ligand bound conformation. The overall structure of the OpuBC:choline complex is depicted in Figure 3A. It shows the typical bi-lobal fold of a ligand binding protein functioning in the context of bacterial binding-protein-dependent ABC-importers (*31, 32, 41*). The OpuBC protein possesses two globular domains that are connected by two loops. A single choline molecule is trapped in a deep cleft formed by the two domains of the OpuBC protein, a location of the ligand that is characteristic for microbial substrate-binding proteins (*41*). Domain I of the OpuBC protein comprise residue 4 - 107 and 213 - 275 (depicted in dark gray) and domain II consists of residues 108 - 212 (depicted in orange) (Fig 3A). The two domains are formed by two central five stranded beta-sheets flanked by several alpha-helices. Residues 104 - 108 and 212 - 215 form the two hinge regions connecting both domains of OpuBC. When one compares domain I with domain II of the OpuBC protein, it becomes apparent that the latter domain is not only smaller in size but also exhibits differences in secondary structure (Fig 3A).

We found that the *B. subtilis* OpuBC protein belongs to cluster F-III of the newly developed crystal structure-based classification scheme for substrate binding proteins (SBPs)(*41*). Cluster F solely contains class II SBPs, binding proteins that possess two hinge regions connecting the two domains of the SBP. A characteristic feature of the binding proteins grouped into cluster F-III is the presence of a long hinge segment that in all likelihood permits a more flexible domain movement between the substrate-free (open) and substrate-bound (closed) conformations (*41*). In the OpuBC protein, this longer hinge segment comprises 10 residues and extends from Gln216 to Ser224; an additional, but shorter hinge region, spans from Asp114 to Tyr117 (Fig 3A). Although the OpuBC and *Af*ProX proteins exhibit

only a modest degree of amino acid sequence identity (29 %) (Fig 5), an overlay of the OpuBC:choline crystal complex with that of the ProX:glycine betaine crystal complex reveals a high degree of similarity that is reflected in the RMSD between both crystal structures of 2.8 Å for 206 C-alpha atoms. The overall fold of the OpuBC protein from the mesophile *B. subtilis* (optimal growth temperature 37° C) and the *Af*ProX protein from the hyperthermophile *A. fulgidus* (optimal growth temperature 83° C) is basically the same with only a few minor exceptions in some of the outer loop regions (data not shown).

The SBPs grouped into cluster F-III not only have a very similar overall structure, but they also possess ligand-binding sites with similar architectures (*41*). The current functionally characterized members of cluster F-III are the glycine betaine/proline betaine binding proteins ProX from *E. coli* (*42*) and *A. fulgidus* (*38*) the OpuAC glycine betaine/proline betaine binding proteins from *B. subtilis* (*36, 37*) and *Lactococcus lactis* (*43*) and the choline/acetylcholine binding protein ChoX from *Sinorhizobium meliloti* (*44*). The ligands of these solute receptors all possess bulky, positively charged trimethyl- or dimethylammonium headgroups that are positioned and held within the ligand binding site via cation–pi interactions (*45*) with aromatic amino acid side chains forming a hydrophobic cage of varying geometries. As described below, the architecture of the substrate-binding site of OpuBC and the coordination of the choline ligand within the binding site conforms to the features established for all members of cluster F-III of the structure-based binding protein classification scheme (*41*).

Architecture of the choline-binding site in OpuBC

The ligand for the OpuBC protein, choline, is trapped deep in the cleft formed by the two lobes of the OpuBC protein (Fig 3A). A closer inspection of the ligand-binding site revealed that residues originating from both domain I and from domain II (Fig 3A) are used to form part of the choline binding site: the side chains of Tyr-71 and Tyr-221 (both contributed by domain I) and of Tyr117 and Tyr-197 (both contributed by domain II) together form an aromatic cage (Fig 3B). This aromatic cage accommodates and coordinates the positively charged trimethylammonium moiety of choline. Within the positively charged head group of choline, the charge is not fixed on the nitrogen atom of the choline molecule but is instead delocalized over the three methyl groups thereby forming a bulky cation. The binding of this cation is established by interactions with the aromatic side chains of Tyr-71, Tyr-117, Tyr-197 and Tyr-221 via cation-pi interactions as was previously found in ligand binding proteins for the compatible solutes glycine betaine and proline betaine from *E. coli* (*42*), *B. subtilis* (*37*), *L. lactis* (*43*) and *A. fulgidus* (*38*) and in the choline/acethylcholine binding protein ChoX from *S. meliloti* (*44*). Distances between the carbon atom in the methyl groups of the choline head group and the four aromatic Tyr side chains within the substrate binding site of OpuBC are in the range of ~3.5 Å, consistent with the van der Waals radii of these atoms. In addition to these cation-pi interactions, the trimethylammonium head group of choline forms hydrogen

bonds with Asn-115 side chain and backbone, respectively. The hydroxyl tail of choline protrudes form the aromatic cage and possesses only a single strong direct interaction with the OpuBC protein. The alcohol function of choline is coordinated by Gln-19 a residue that originates from domain I of OpuBC and has a distance of 2.7 Å to the choline ligand (Fig 3B).

In addition to the ligand choline that interacts with residues originating from both domains of the OpuBC protein (Fig 3B), a domain-domain interaction is observed between residues Asp-74 and Tyr-117 that is mediated via a water molecule depicted in blue spheres in Figure 3B. This water-mediated interaction between the two lobes of OpuBC is further stabilized via the interactions of an additional water molecule with Leu-155. Via these two water molecules, Leu-155 (located in domain II) interacts indirectly with Asp-74 (located in domain I). These water-mediated interactions are, in addition to the interactions mediated by the trimethylammonium head group of choline, the only interactions observable between domains I and II of the OpuBC protein (Fig 3B).

Mutational analysis of the choline-binding site in OpuBC

The architectures of the aromatic cage that coordinates the bulky cationic head groups of choline in OpuBC (Fig 3B) and glycine betaine in *Af*ProX (*38*) via cation-pi interactions are super-imposable (Fig 4). The spatial orientation of the side chains of Tyr-71, Tyr-117, Tyr-197 and Tyr-221 from OpuBC overlay perfectly with the four Tyr residues (Tyr-63, Tyr-111, Tyr-190 and Tyr-214) present in the *Af*ProX glycine betaine-binding site. As a consequence of this architectural arrangement, the trimethylammonium head groups of choline and glycine betaine are coordinated identically in both crystal structures. The ProX protein from the hyperthermophile A. *fulgidus* binds its ligand glycine betaine with a K_d of about 50-60 nM (Tschapek *et. al.* submitted). The replacement of single Tyr-residues within the aromatic cage of *Af*ProX by Ala-residues causes a drop in ligand binding, depending on which Tyr-residues is mutated. Hence, a mutant aromatic cage formed by only three Tyr-residues still affords ligand binding by the *Af*ProX solute receptor with a reasonable affinity. However, the simultaneous replacement of two Tyr-residues by Ala-residues (regardless in which combination) generates *Af*ProX mutant proteins that are no longer proficient in glycine betaine binding (Tschapek *et. al.* submitted).

Since the aromatic cages present in the ligand binding sites of ProX and OpuBC have an identical architecture (Fig 4), we have not repeated a mutational analysis of the aromatic cage of the OpuBC protein. We surmise that the data obtained on ligand binding for the *Af*ProX protein can be projected onto the ligand-binding attributes of the OpuBC protein. Instead, we conducted a mutational analysis of the residue coordinating the alcohol function of choline and those residues that contributes to domain-domain interactions of the two lobes of the OpuBC protein and the build-up the water network that stabilizes the ligand within the binding site.

Therefore, we focused on Gln-19, Asp-74, Asp-149 and Leu-155; each of these residues was replaced by an Ala-residue and each of the purified mutant OpuBC proteins was analyzed by intrinsic Trp/Tyr fluorescence. Replacement of Gln-19 by Ala reduced the binding affinity by about 14-fold (Table 2), highlighting the importance of the direct interaction of the Gln-19 side-chain with the alcohol group of choline (Fig 3B) for high affinity ligand binding. However, those residues that are involved in setting the water network within the ligand-binding site were (at least) of equal importance to the contribution of Gln-19 to ligand binding. The separate replacement of Asp-149 and Leu-155 with Ala residues let to a 18-fold and 25-fold drop in ligand binding activity, respectively (Table 2). The simultaneous substitution of these two residues with Ala residues let to an OpuBC variant with a K_d of about 1132 μM, a 38-fold drop in ligand affinity exhibited by the wild type OpuBC protein (K_d of about 30 μM (Table 2). A key residue within the ligand-binding site is Asp-74. The side-chain of Asp-74 does not make any direct contacts to the choline ligand, but it catalyzes via water-mediated interactions between Gln-19, Leu-155 and Asp-149 the closure of the two domains of the OpuBC protein. It is thereby critical for the stable entrapment of the choline ligand (Fig 3B). Replacement of Asp-74 by an Ala-residue generates an OpuBC mutant protein that is unable to bind choline, even when the ligand is present at a concentration of 50 mM in the binding assay (Table 2). Likewise, the simultaneous replacement of the residue that makes a direct contact to the choline ligand (Gln-19) and Leu-155, a key residue for the formation of the water network (Fig 3B), with Ala residues leads to an OpuBC protein that no longer can bind choline (Table 2). Collectively, this mutational assessment of the ligand-binding site of the OpuBC solute receptor underscores the functional roles of those residues that were deduced from the crystal structure of OpuBC for both direct (Gln-19) and water-mediated interactions (Asp-74, Asp-149, Leu-155) with the choline ligand and the two lobes of the OpuBC protein.

It is generally thought that the interaction of the ligand with residues in the substrate-binding site induces a hinge-bending (and sometimes also a rotational motion) movement of the two domains of the binding protein. As a result, the substrate-binding site is dynamically formed with contributions of residues from both domains of the binding protein that then lead to the entrapment of the ligand and the formation of the closed conformation of the binding protein. The hinge segments that connect the two lobes play a critical role for these domain movements (*41, 44, 46*). Asn-115 is part of one of the hinge segments present in OpuBC and its replacement by Ala destroys the ability of OpuBC to bind its ligand (Table 2).

Comparison of the choline-binding site of OpuBC with the glycine betaine-binding site present in ProX from *A. fulgidus*

The architectures of the aromatic cage for the coordination of the trimethylammonium headgroups of choline in OpuBC (Fig 3B) and glycine betaine in *Af*ProX(*38*) via cation-pi interactions are super-imposable (Fig 4). Despite the identical spatial

orientation of the side chains of the four Tyr residues forming the aromatic cages, it should be recalled that the *B. subtilis* OpuBC protein does not bind glycine betaine (data not shown). Conversely, the ProX protein from *A. fulgidus* does not bind choline (Tschapek *et. al.* submitted). We traced these differences in ligand-specificity to the way in which the *Af*ProX and OpuBC proteins coordinate the carboxylate group of glycine betaine and the alcohol group of choline, respectively.

An important role for the fixation of the choline ligand within the substrate-binding site of OpuBC plays the side chain of Gln-19, since it interacts directly with the alcohol group of the choline molecule (Fig 3B). At the equivalent position in the glycine betaine-binding site of *Af*ProX, a lysine residue (Lys-13) is present that interacts with one oxygen of the carboxy tail of glycine betaine (Fig 4). The other hydroxyl group of glycine betaine is coordinated through interactions with Arg-149 (*38*). As described above, the interaction between both domains of OpuBC is mediated via a water network into which Leu-155 is critically involved (Fig 3B) (Table 2). Strikingly, in the overlay of the OpuBC and *Af*ProX crystal structures, a water molecule is located exactly at the position of the N3 atom of the glycine betaine-contacting Arg-149 residue of *Af*ProX (Fig 4). Hence, the strong interaction partners for the carboxylic tail of glycine betaine present in the ProX protein from *A. fulgidus* are absent in the OpuBC protein from *B. subtilis* (Fig 4). As a result, glycine betaine cannot be bound effectively despite the fact that its trimethylammonium head group would in all likelihood fit without any problems correctly into the aromatic Tyr-cage of the OpuBC protein.

Our analysis of the crystal structures of the OpuBC and *Af*ProX proteins therefore suggests that glycine betaine can only fleetingly interact with the open conformation of the OpuBC protein and is unable to induce a closure of the two domains of OpuBC to entrap it stably and thereby form the closed conformation of the solute receptor. This conformation, however, is required for the productive interactions of the substrate-loaded binding proteins with the membrane-embedded components of the cognate ABC transport system (*35, 47*) to release the ligand into the translocation pathway for ATP-hydrolysis-dependent import of the substrate into the microbial cell (*31, 32*). Hence, the crystal structure of the OpuBC protein in complex with its natural ligand choline provides a rational for the observed choline-specificity of the OpuB ABC transporter and its inability to catalyze the import of glycine betaine into the *B. subtilis* cell (*23*).

Analysis of the crystal structure of OpuCC from *Staphylococcus aureus*, a ligand binding protein related to the OpuBC and OpuCC proteins from *B. subtilis*

In the brookehaven protein databank (www.pdb.org), a crystal structure (1.9 Å) of a substrate binding protein (OpuCC) from a *St*aphylococcus *aureus* strain has been deposited (PDB code 3O66). This structure has been determined by a structural genomics consortium but has not been published yet; the OpuCC protein is annotated in the PDB as a glycine betaine/carnitine/choline binding protein but no functional data are available for this protein

to assess its true substrate specificity. However, a sequence-related OpuCC protein from the *S. aureus* strain RN6390 is part of a functionally characterized ABC-importer for choline.*(48)* Due to the high degree of amino acid sequence identity (>60 % identity and >80 % homology) towards the *B. subtilis* OpuBC protein (Fig 5), we compared the crystal structures of the OpuCC$_{S.a}$-3O66 and OpuBC proteins. The superimposition of both structures (Fig 6) revealed that the OpuCC$_{S.a}$-3O66 structure has been crystallized in the open unliganded conformation. This open conformation is stabilized by two triethylene glycol molecules bound between both domains of the OpuCC$_{S.a}$-3O66 protein, a position where the physiological relevant ligand is expected to be localized (*41, 46*). Presumably, the two triethylene glycol molecules present in the OpuCC$_{S.a}$-3O66 structure have likely been captured from the crystalization condition.

Conclusion

The crystal structure of choline bound OpuBC from *B. subtilis* expand our knowledge on substrate binding and the determinants for differences in affinity of SBPs involved in osmoprotection. The excessive mutation of amino acids involved in ligand binding of *Af*ProX and OpuBC and subsequent affinity measurements resulted in the conclusion that the aromatic box which interacts via cation-π-interactions with the substrate's trimethyl ammonium head group has a rather low influence on affinity to the substrate. However, most important are amino acids involved in domain-domain interactions, that stabilize the closed conformation of the SBP. At this point the nature of interaction seem to determine somehow the quantitative value of affinity. In our case *Af*ProX shows direct interactions via hydrogen bonds of residues of domain I and II and therefore wild type protein has a K_D of 50-50nM. This is rather high when compared to OpuBC: Here only water mediated interactions between domain I and II can be identified and measured K_D for wild type is more than 600 fold lower than affinity constant of ProX. Thus microorganisms the binding site in SBP binding osmolytes needs to be mutated at one or two amino acids positions to adjust affinity to match their requirements.

Materials and Methods

Chemicals

Choline, glycine betaine, anhydrotetracyline and desthiobiotin were purchased from Sigma-Aldrich (Munich, Germany). The antibiotic ampicillin was purchase from Carl Roth (Karlsruhe, Germany) and gentamycin was obtained from Serva (Heidelberg, Germany). The matrix (*Strep*-Tactin Sepharose) for affinity purification of *Strep*-tag-II marked recombinant proteins were purchased from IBA (Göttingen, Germany).

Bacterial strains and growth media

Chromosomal DNA of the *B. subtilis* wild type strain JH642 (*trpC pheA1*; BGSC 1A96; a kind gift of J. Hoch, Scripps Research Institute, CA, USA) was used as the donor for the cloning of the *opuBC* gene. Strain JH642 was propagated aerobically in Luria-Bertani liquid medium at 37° C (49). The *Escherichia coli strain* DH5α (Fermentas, St. Leon-Rot, Germany) was used for all cloning experiments. For the overproduction of the recombinant *B. subtilis* OpuBC protein and its mutant derivatives, we used the *E. coli* strain BL21 [Artic express (DE) RIL] (Stratagene; La Jolla, CA, USA). Overproduction of OpuBC and its derivatives was carried out in a slightly chemically defined minimal medium (MMA)(49) supplemented with 1mg/L thiamin, 1mM MgSO₄, 0.2% casamino acids (w/v) and 0.5% glucose (w/v) as the carbon source and 100 µg/ml ampicillin to select for recombinant plasmids carrying the *opuBC* gene or its mutant derivatives.

Cloning and site-directed mutagenesis of the *B. subtilis opuBC* gene

The OpuBC choline binding protein from *B. subtilis* is initially synthesized with a cleavable 22-residue N-terminal signal sequence for export via the Sec-pathway from the *B. subtilis* cell; the signal sequence is proteolytically processed and the newly formed N-terminus of the mature protein (Cys–23) is lipid-modified (23). The coding region of *opuBC* was cloned without the DNA segment that encodes the signal sequence and the codon for Cys-23 was replaced by a codon encoding a Gly residue. A PCR-reaction was carried out with the forward (5'-CGTCTCCGCGCGGCTCGCTTCCTGGTCTCAGCGC-3') and backward (5'-CGTCTCGTATCTCACGATTCGAAATAGCGATGTTTTTC-3') primers on chromosomal DNA isolated from the *B. subtilis* wild type strain JH642. The resulting 877 bp *opuBC* DNA fragment was then cleaved with the restriction enzyme *Esp*3I (recognition sequences in the primer sequence are underlined) and then cloned into the expression vector pASK-IBA6 (IBA, Göttingen, Germany) which had been cut with *Bsa*I. The DNA sequence of one of the resulting plasmids, pMP31, was determined (eurofins MWG operon; Ebersberg, Germany) to ensure that no undesired mutations had been introduced into the *opuBC* coding sequence during the cloning procedure. In plasmid pMP31, the coding region of *opuBC* is inserted into the pASK-IBA6 vector in such a way that it is fused to an OmpA signal sequence followed by

the *Strep*-tag-II affinity peptide and a cleavage site for the Factor Xa protease. Expression of the hybrid *opuBC* gene in plasmid pMP31 is under the control of the anhydrotetracyclin (AHT)-responsive TetR repressor that is encoded by the pASK-IBA6 expression vector.

Mutant derivatives of the *opuBC* gene were generated with the QuickChange site-directed mutagenesis kit (Statagene, La Jolla, CA) using plasmid pMP31 as the DNA template and custom-synthesized mutagenic DNA primers (purchased from Sigma-Aldrich, Munich, Germany). The following mutant *opuBC* plasmids were created: pMP49 [Gln-19 →Ala (CAA/GCA)], pMP51 [Asp-74 →Ala (GAC/GCC)], pMP52 [Asn-115 →Ala (AAT/GCT)], pMP53 [Leu-155 →Ala (CTC/GCC)] and pMP65 [Asp-149 →Ala (GAT/GCT)]. Selected mutant plasmids with single amino acid substitutions in the OpuBC protein were re-mutagenized to generate the following double *opuBC* derivatives: pMP62 [Gln-19 →Ala (CAA/GCA) and Leu-155 →Ala (CTC/GCC)] and pMP66 [Asp-149 →Ala (GAT/GCT) and Leu-155 →Ala (CTC/GCC)]. The entire coding region of each of the mutant *opuBC* alleles was sequenced to ensure the presence of the desired mutation and the absence of unwanted DNA alterations.

Overexpression and purification of recombinant OpuBC proteins in *E. coli*

Plasmid pMP31 (resistant to ampicillin) was transformed into competent cells of the *E. coli* strain BL21 [Artic express (DE) RIL] (resistant to gentamycin) by simultaneously selecting for ampicillin (100 μg/ml) and gentamycin (20 μg/ml) resistant transformants on LB agar plates. One antibiotic resistant colonies was then picked and used to inoculate a pre-culture (100 ml modified MMA containing the antibiotics ampicillin and gentamycin at the above indicated concentrations) that was then propagated over-night at 37°C in a 1 L Erlenmeyer flask in an aerial shaker set at 200 rpm. For the main culture, we used a Minifors fermenter system (Infors AG, Bottmingen, Switzerland) with a 5 L reaction vessel filled with 4 L modified MMA (see above) containing 100 μg/ml ampicillin. The medium present in the fermenter vessel was inoculated with the pre-culture of strain BL21 [Artic express (DE) RIL] to an OD_{578} of about 0.1. The cells were then grown at 37°C first as batch culture (up to an OD_{578} of about 3) and subsequently grown as a fed-batch culture up to an OD_{578} of about 15. Subsequently, the temperature of the culture was cooled-down to 20°C and the expression of the TetR-controlled *opuBC* gene present on plasmid pMP31 was induced by the addition of anhydrotetracyclin (0.2 μg anhydrotetracylin/ml culture) to the culture. The bacterial cells were then propagated under *opuBC*-inducing conditions for additional 16 h at 20°C and were subsequently harvested by centrifugation (30 min, 3 000 x g at 4°C). The cells (10 g wet weight) were re-suspended in 20 ml buffer W [100 mM Tris-HCl (pH 7.5), 150 mM NaCl] and then were disrupted by passing them five-times through a French-pressure cell (1000 psi) at 4°C. Unbroken cells and cellular debris were removed by ultracentrifugation (1 h, 35 000 x g at 4°C) and the supernatant was then loaded onto a 10-ml *Strep*-Tactin column (IBA, Göttingen, Germany) pre-equilibrated with 10 bed volumes of buffer W [100 mM Tris-HCl (pH

7.5) 150 mM NaCl]. After washing the column with 20 bed volumes of buffer W, proteins bound to the affinity matrix were eluted with three bed volumes of buffer E [100 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2.5 mM desthiobiotin]. Two forms of the recombinantly produced OpuBC protein were found: (i) the non-processed OmpA-*Strep*-Tag-II-OpuBC protein and (II) the processed *Strep*-Tag-II-OpuBC protein.

All fractions containing OpuBC proteins were pooled and concentrated using Vivaspin 4 concentrator columns (exclusion size, 10 kDa) (Vivascience, Hannover, Germany) with a simultaneous exchange of the butter to 10 mM Tris-HCl (pH7.5) and 10 mM NaCl. The non-processed OmpA-*Strep*-Tag-II and the processed *Strep*-Tag-II portions of the OpuBC protein were then treated with the factor Xa protease (Merck, Darmstadt, Germany) to remove the OmpA signal sequence and the *Strep*-Tag-II affinity peptide. Proteolytic cleavage of the recombinant OpuBC protein with Factor Xa was carried out at room temperature over-night [0.5 U per 50 μg of OpuBC protein in the cleavage buffer provided by Merck (Darmstadt, Germany)]. The completeness of the proteolytic cleavage of the two forms of the recombinant OpuBC protein was assessed by SDS-gel electrophoresis and found to typically yield a single OpuBC band on a 12% SDS-polyacrylamide gel (see Fig. 2A). The OpuBC protein was then applied to a 5 ml ion exchange column (HiTrap Q FF 5ml; GE Healthcare, Munich, Germany) equilibrated with two bed volumes of buffer A [25 mM Tris-HCl (pH 7.5), 20 mM NaCl]. We observed that a approximately 50% of the OpuBC protein passed right though the ion exchange column and this portion of the purified OpuBC protein was subsequently found to be inactive in ligand binding experiments using fluorescence spectroscopy (see below). Functional OpuBC protein (as assayed by subsequent fluorescence spectroscopy) was eluted from the HiTrap Q FF 5ml column by applying a gradient of buffer B [25 mM Tris-HCl (pH 7.5), 1 M NaCl]. From a typical overproduction (from 4 l of culture) and purification experiment, we obtained approximately 20 mg of pure and biologically active OpuBC protein. For crystallization trials and ligand binding experiments, the purified OpuBC protein was concentrated to about 10 mg/ml with Vivaspin 4 concentrator columns (exclusion size, 10 kDa) (Vivascience, Hannover, Germany) with a simultaneous exchange of the butter to 10 mM Tris-HCl (pH7.5) and 10 mM NaCl. All mutant derivatives of OpuBC were overproduced and purified as described above for the wild type protein.

Determination of the dissociation constant of the OpuBC-choline complex

The dissociation affinity of the OpuBC-choline complex was determined by intrinsic tyrosine and tryptophan fluorescence spectroscopy as described for the glycine betaine binding protein OpuAC from *B. subtilis* (36, 37). The OpuBC protein possesses three Trp and 17 Tyr residues and the intrinsic fluorescence of these residues was monitored from 295 nm to 400 nm using a Cary Eclipse fluorescence spectrometer (Varian, Surry, United Kingdom). The excitation wavelength was set to 285 nm, the slit width was 5 nm, and the temperatures during the assays was maintained at 25°C with a cooling/heating device. One ml sample of 1

μM OpuBC protein (in 10 mM Tris-HCl; pH 7.5) was equilibrated for 15 min in the measuring cuvette. Subsequently, different amounts of choline (1 μM to 200 μM) were titrated to the OpuBC protein solution and the fluorescence was recorded. Upon choline binding, an increase in the fluorescence intensity of OpuBC was detected (Fig. 2B), and this change in the emission spectrum was used to determine the apparent K_d value for the OpuBC protein for choline binding by plotting the change in fluorescence intensity at 339 nm against the choline concentration used in the assays. The recorded data were analyzed by assuming a standard one-site-binding model. All given K_d values of the wild type OpuBC protein and its mutant derivatives (Table 2) represent the average of three independent measurements, with the standard deviations given as errors.

Crystallization of the OpuBC-choline complex, data collection and model refinement

Crystals of OpuBC were obtained using the hanging-drop vapor diffusion technique at 12°C. Purified OpuBC was concentrated to 10mg/ml and incubated with 15 μM choline chloride prior to crystallization. 1 μl of protein was mixed with 1μl of reservoir solution containing 200 mM potassium acetate and 20% PEG3350. Crystals appeared after 3 days and were harvested after 2 weeks in 200 mM potassium acetate, 20% PEG3350 and 30% ethylenglycol and flesh frozen in liquid nitrogen.

A native dataset of the OpuBC-choline complex was collected at beamline ID23-eh1 (ESRF, Grenoble, France) and scaled using XDS (*39*). Initial phases were obtained by molecular replacement using the program PHASER (*40*) with the crystal structure of *Archaeoglobus fulgidus* glycine betaine binding protein ProX as a template (PDB entry 1SW2)(*38*). Data refinement statistics and model content are summarized in Table 1. Model building and refinement were performed using coot and REFMAC5 (*50*).

Sequence alignments and Structure alignments

Amino acid sequence alignments of proteins related to the *B. subtilis* OpuBC protein were performed with ClustalW (*51*).
Structure alignments were performed using LSQMAN employing standard settings.

Figure preparation

Figures of protein structures were prepared using PYMOL (http://pymol.sourceforge.net/).

Protein Data bank accession number

The coordinates of the OpuBC-choline complex will be deposited in the RCSB Protein Data Bank.

Literature

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

2. Wood, J. M., Bremer, E., Csonka, L. N., Kraemer, R., Poolman, B., van der Heide, T., and Smith, L. T. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria, *Comp Biochem Physiol A Mol Integr Physiol 130*, 437-460.

3. Bremer, E., and Krämer, R. (2000) Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes., In *Bacterial Stress Responses* (Storz, G., and Hengge-Aronis, R., Eds.), pp 79-97, ASM Press, Washington, DC.

4. Wood, J. M. (2011) Osmotic stress, In *Bacterial Stress Responses* (Storz, G., and Hengge, R., Eds.), pp 133-156, ASM Press, Washington, DC.

5. Cayley, S., and Record, M. T., Jr. (2003) Roles of cytoplasmic osmolytes, water, and crowding in the response of *Escherichia coli* to osmotic stress: biophysical basis of osmoprotection by glycine betaine, *Biochemistry 42*, 12596-12609.

6. Whatmore, A. M., Chudek, J. A., and Reed, R. H. (1990) The effects of osmotic upshock on the intracellular solute pools of *Bacillus subtilis*, *Journal of General Microbiology 136*, 2527-2535.

7. Street, T. O., Bolen, D. W., and Rose, G. D. (2006) A molecular mechanism for osmolyte-induced protein stability, *Proc Natl Acad Sci U S A 103*, 13997-14002.

8. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant, *Proc Natl Acad Sci U S A 103*, 13357-13361.

9. Burg, M. B., and Ferraris, J. D. (2008) Intracellular organic osmolytes: function and regulation, *J Biol Chem 283*, 7309-7313.

10. Yancey, P. H. (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses, *J Exp Biol 208*, 2819-2830.

11. Ziegler, C., Bremer, E., and Kramer, R. (2010) The BCCT family of carriers: from physiology to crystal structure, *Mol Microbiol 78*, 13-34.

12. Poolman, B., Spitzer, J. J., and Wood, J. M. (2004) Bacterial osmosensing: roles of membrane structure and electrostatics in lipid-protein and protein-protein interactions, *Biochim Biophys Acta 1666*, 88-104.

13. Le Rudulier, D., Strom, A. R., Dandekar, A. M., Smith, L. T., and Valentine, R. C. (1984) Molecular biology of osmoregulation., *Science 224*, 1064-1068.

14. Nyyssölä, A., Kerovuo, J., Kaukinen, P., von Weymarn, N., and Reinikainen, T. (2000) Extreme halophiles synthesize betaine from glycine by methylation, *J Biol Chem 275*, 22196-22201.

15. Boch, J., Kempf, B., Schmid, R., and Bremer, E. (1996) Synthesis of the osmoprotectant glycine betaine in *Bacillus subtilis*: characterization of the *gbsAB* genes, *Journal of Bacteriology 178*, 5121-5129.

16. Rosenstein, R., Futter-Bryniok, D., and Götz, F. (1999) The choline-converting pathway in *Staphylococcus xylosus* C2A: genetic and physiological characterization, *J Bacteriol 181*, 2273-2278.

17. Lamark, T., Kaasen, I., Eshoo, M. W., Falkenberg, P., McDougall, J., and Strom, A. R. (1991) DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*, *Mol Microbiol 5*, 1049-1064.

18. Finnegan, S., Yuan, H., Wang, Y. F., Orville, A. M., Weber, I. T., and Gadda, G. (2010) Structural and kinetic studies on the Ser101Ala variant of choline oxidase: catalysis by compromise, *Arch Biochem Biophys 501*, 207-213.

19. Boch, J., Kempf, B., and Bremer, E. (1994) Osmoregulation in *Bacillus subtilis*: synthesis of the osmoprotectant glycine betaine from exogenously provided choline, *Journal of Bacteriology 176*, 5364-5371.

20. Bremer, E. (2002) Adaptation to changing osmolarity, In *Bacillus subtilis and its closest relatives* (Sonenshein, A. L., Hoch, J. A., and Losick, R., Eds.), pp 385-391, ASM Press, Washington, DC.

21. Kempf, B., and Bremer, E. (1995) OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*, *JBC 270*, 16701-16713.

22. Kappes, R. M., Kempf, B., and Bremer, E. (1996) Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD, *Journal of Bacteriology 178*, 5071-5079.

23. Kappes, R. M., Kempf, B., Kneip, S., Boch, J., Gade, J., Meier-Wagner, J., and Bremer, E. (1999) Two evolutionarily closely related ABC-transporters mediate the uptake of choline for synthesis of the osmoprotectant glycine betaine in *Bacillus subtilis.*, *Molecular Microbiology 32*, 203-216.

24. Hahne, H., Mäder, U., Otto, A., Bonn, F., Steil, L., Bremer, E., Hecker, M., and Becher, D. (2010) A comprehensive proteomics and transcriptomics analysis of *Bacillus subtilis* salt stress adaptation, *J Bacteriol 192*, 870-882.

25. Steil, L., Hoffmann, T., Budde, I., Völker, U., and Bremer, E. (2003) Genome-wide transcriptional profiling analysis of adaptation of *Bacillus subtilis* to high salinity, *J. Bacteriol. 185*, 6358-6370.

26. Kortstee, G. J. J. (1970) The aerobic decomposition of choline by micrroganisms., *Arch. Microbiol. 71*, 235-244.

27. Andersson, D. I., and Hughes, D. (2009) Gene amplification and adaptive evolution in bacteria, *Annu Rev Genet 43*, 167-195.

28. Nau-Wagner, G., Boch, J., Le Good, J. A., and Bremer, E. (1999) High-affinity transport of choline-O-sulfate and its use as a compatible solute in *Bacillus subtilis*, *Appl Environ Microbiol 65*, 560-568.

29. Kappes, R., and Bremer, E. (1998) Response of *Bacillus subtilis* to high osmolarity: uptake of carnitine, crotonobetaine and g-butyrobetaine via the ABC transport system OpuC, *Microbiology 144*, 83-90.

30. Jebbar, M., von Blohn, C., and Bremer, E. (1997) Ectoine functions as an osmoprotectant in *Bacillus subtilis* and is accumulated via the ABC-transport system OpuC, *FEMS Microbiology Letters 154*, 325-330.

31. Davidson, A. L., Dassa, E., Orelle, C., and Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems, *Microbiol Mol Biol Rev 72*, 317-364.

32. Eitinger, T., Rodionov, D. A., Grote, M., and 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.

33. Oldham, M. L., Davidson, A. L., and Chen, J. (2008) Structural insights into ABC transporter mechanism, *Curr Opin Struct Biol 18*, 726-733.

34. Bordignon, E., Grote, M., 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.

35.   Cui, J., Qasim, S., and Davidson, A. L. (2010) Uncoupling substrate transport from ATP hydrolysis in the *Escherichia coli* maltose transporter, *J Biol Chem 285*, 39986-30993.

36.   Smits, S. H., Höing, M., Lecher, J., Jebbar, M., Schmitt, L., and Bremer, E. (2008) The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, site-directed mutagenesis, and crystallographic studies, *J Bacteriol 190*, 5663-5671.

37.   Horn, C., Sohn-Bosser, L., Breed, J., Welte, W., Schmitt, L., and Bremer, E. (2006) Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine, *J Mol Biol 357*, 592-606.

38.   Schiefner, A., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004) Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*, *J Biol Chem 279*, 48270-48281.

39.   Kabsch, W. (1993) Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants, *J. Appl Crystallogr 40*, 658-674.

40.   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.

41.   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.

42.   Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004) Cation-pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*, *J Biol Chem 279*, 5588-5596.

43.   Wolters, J. C., Berntsson, R. P., Gul, N., Karasawa, A., Thunnissen, A. M., Slotboom, D. J., and Poolman, B. (2010) Ligand binding and crystal structures of the substrate-binding domain of the ABC transporter OpuA, *PLoS One 5*, e10361.

44.   Oswald, C., Smits, S. H., Hoing, M., Sohn-Bosser, L., Dupont, L., Le Rudulier, D., Schmitt, L., and Bremer, E. (2008) Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unliganded-closed states, *J Biol Chem 283*, 32848-33259.

45.   Dougherty, D. A. (1996) Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp, *Science 271*, 163-168.

46.   Quiocho, 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.

47.   Shilton, B. H. (2008) The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems, *Biochim Biophys Acta 1778*, 1772-1780.

48.   Kiran, M. D., Akiyoshi, D. E., Giacometti, A., Cirioni, O., Scalise, G., and Balaban, N. (2009) OpuC--an ABC transporter that is associated with *Staphylococcus aureus* pathogenesis, *Int J Artif Organs 32*, 600-610.

49.   Miller, J. H. (1992) *A Short course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

50.   Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr 53*, 240-255.

51.   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 22*, 4673-4680.

52.   Ziegler, C., Bremer, E., and Krämer, R. (2010) The BCCT family of carriers: from physiology to crystal structure, *Mol Microbiol 78*, 13-34.

53.   Kempf, B., Gade, J., and Bremer, E. (1997) Lipoprotein from the osmoregulated ABC transport system OpuA of *Bacillus subtilis*: purification of the glycine betaine binding protein and characterization of a functional lipidless mutant, *J Bacteriol 179*, 6213-6220.

Figure Legend

Figure 1:

The osmoprotectant uptake systems of *B. subtilis* and the synthesis pathway of glycine betaine from the precursor choline. The OpuA, OpuB and OpuC osmoprotectant uptake (Opu) systems are members of the ABC family (*21, 23*), the OpuD osmoprotectant uptake system is a member of the BCCT family (*22, 52*). The OpuA, OpuB and OpuC transporters possess extracellular solute receptors (OpuAC, OpuBC, OpuCC) that are each tethered to the outer face of the cytoplasmic membrane through a lipid modification of an N-terminal Cys-residue (*23, 53*). The OpuA transporter possesses a single integral membrane component (OpuAB) that likely functions as a homodimer; the integral membrane components of the OpuB transporter (OpuBB, OpuBD) and that of the OpuC transporter (OpuCB, OpuCD) are likely to function as a heterodimer. Each of the OpuA, OpuB and OpuC ABC transporters possesses a single ATPase (OpuAA, OpuBA, OpuCA) protein that probably function as dimers (*31, 32*). The substrate specificities of the various Opu transporters have been assessed though osmoprotection growth assays and transport studies with high salinity challenged *B. subtilis* cells (*21-23, 28-30*). The synthesis pathway of glycine betaine from the precursor choline has been elucidated by Boch *et al.* (*15, 19*). Glycine betaine production in *B. subtilis* is catalyzed by the sequential enzymatic reactions of an alcohol dehydrogenase (GbsB) and a glycine betaine aldehyde dehydrogenase (GbsA).

Figure 2:

A) SDS-PAGE of purified OpuBC used for binding experiments and crystallography. B) Ligand binding experiments with Choline. Red line: Fluorescence emission without ligand, blue line: Fluorescence after addition of $100\mu M$ Choline. C) Equilibrium binding experiments of wild type OpuBC with CB. Calculated binding constant is $30\mu M$.

Figure 3: Structure of OpuBC crystallised in space group C2

A) Overall fold of OpuBC from *B. subtilis*. Two domain architecture is indicated in dark grey and orange. Bound ligand is represented as sticks. B) View of the ligand-binding pocket of OpuBC –Choline complex. Interactions of the hydroxy tail of CB are highlighted with red dashed lines. For simplicity both representations show the same orientation of the protein.

Figure 4: Overlay of OpuBC and ProX from Archaeoglobus fulgidus (pdb: 1SW2)

Residues belonging to OpuBC are color coded as described in Figure 3. ProX is shown in marine. Residues that are identical in OpuBC and ProX binding site are shown as lines. Amino acid involved in binding the hydroxy /carboxy tail are highlighted as stick representation.

Figure 5: Sequence alignment of OpuBC from B. subtilis, ProX from A. fulgidus and OpuCC from *Staphylococcus aureus*.

Figure 6: Superimposition of OpuBC from B. subtilis and OpuCC from *Staphylococcus aureus*.

A) Overall representation of both proteins. OpuBC is shown in dark grey and orange (see Fig 3). OpuCC is depicted in light blue. B) Overlay of OpuBC binding site with residues of OpuCC that align in Fig 5. Residues of domain II overlap while outward rotation of domain I of OpuCC compared to OpuBC indicates an open conformation of OpuCC.
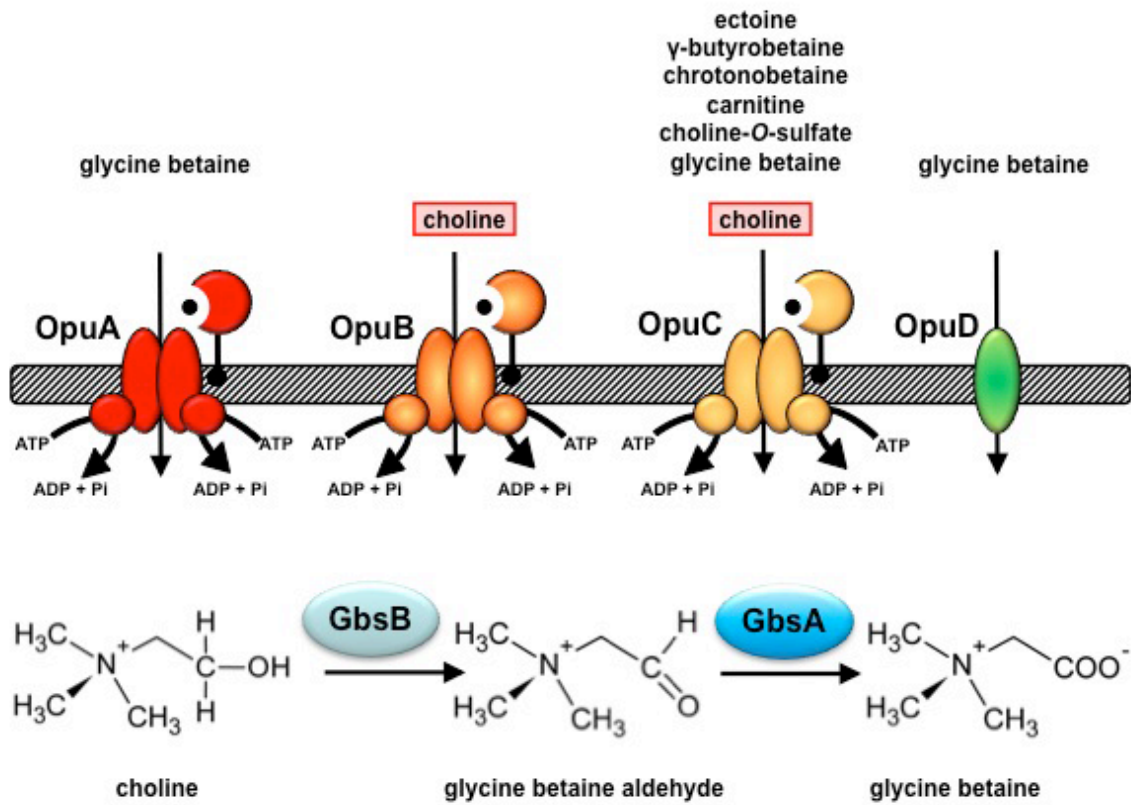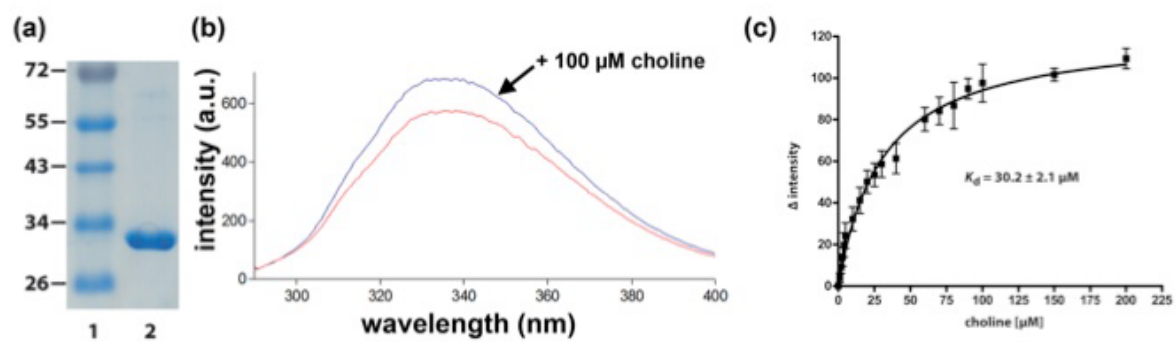
Figure 1:

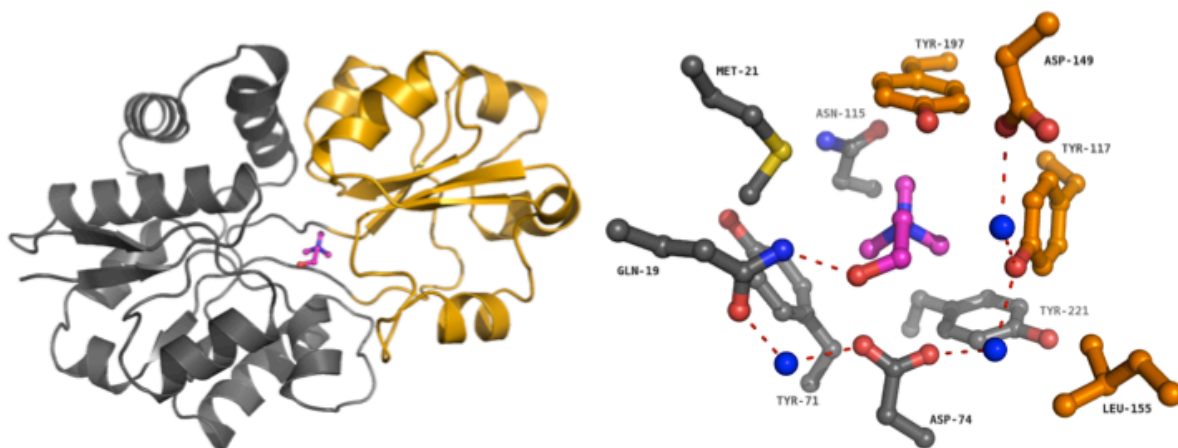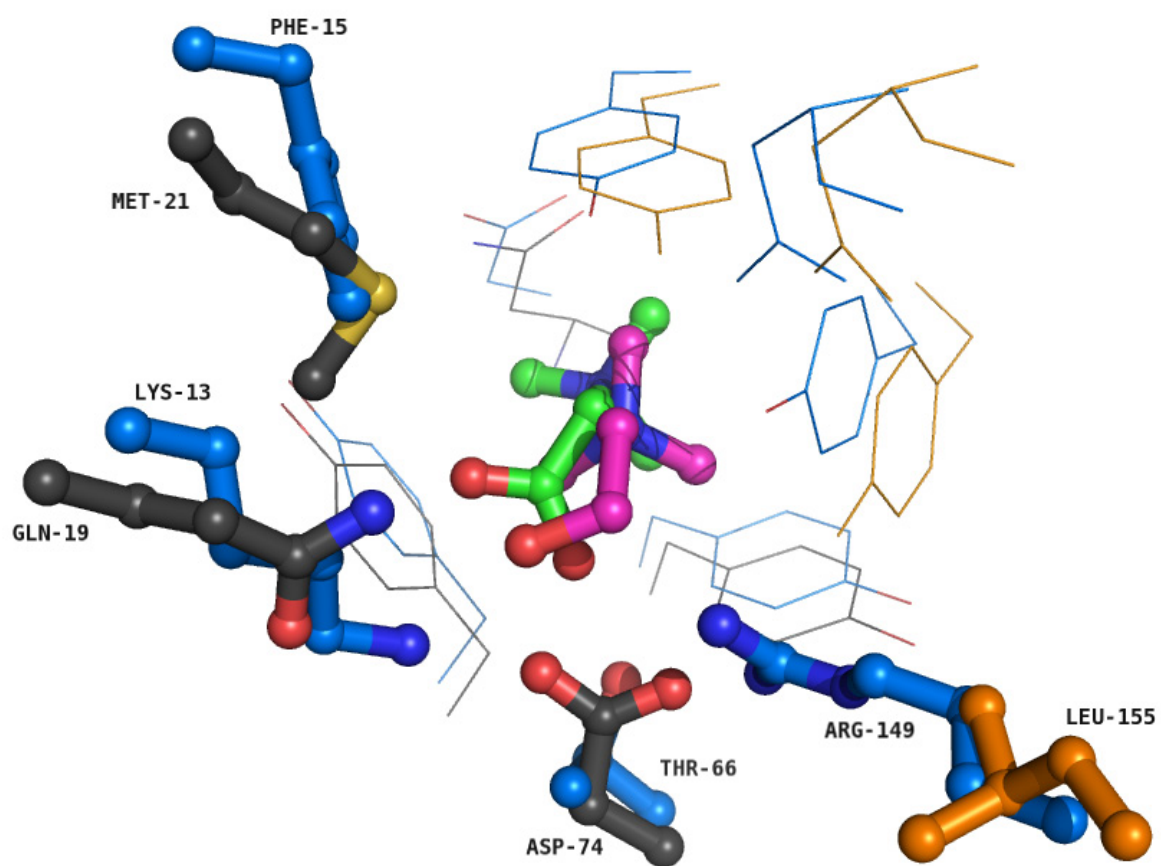**Figure 2:**

Figure 3:

Figure 4:

Figure 5:

```
OpuBC    CSLPGL-SAAADQTIKIGAQSMSESEIIASMLGQLIEHHTD--LKTTTIKNLGSNAVQQQ 57
OpuCC    CSLPGLGSKSTKNDVKITALSTSESQIISHMLRLLIEHDTHGKIKPTLVNNLGSSTIQHN 60
ProX     -------CSQSSERVVIGSKPFNEQYILANMIAILLEENGY---KAEVKEGLGGTLVNYE 50
              .  :.: : * : . .*. *:: *:  *:*..      *.   :.**.. :: :

OpuBC    ALMNGEIDIAATRYTGDALTGTLRMEP--EKDPDKALALTQREFKKRYDLKWYDSYGFDN 115
OpuCC    ALINGDANISGVRYNGTDLTGALKEAP--IKDPKKAMIATQQGFKKKFDQTFFDSYGFAN 118
ProX     ALKRNDIQLY-VEYTGTAYNVILRKQPPELWDQQYIFDEVKKGLLEADGVVVAAKLGFRD 109
         ** ..: ::  ..*.*   . *: *    * . :  .:: : :  .     . ** :

OpuBC    TYAFTVSKELADQYHLETVSDVKKWAPQLKLGVDNYWMKLKGNGYQDFTKTYGMTFGGTY 175
OpuCC    TYAFMVTKETAKKYHLETVSDLAKHSKDLRLGMDSSWMNRKGDGYEGFKKEYGFDFGTVR 178
ProX     DYALAVRADWAEENGVEKISDLAEFADQLVFGSDPEFASRP-DGLPQIKKVYGFEFKEVK 168
          **: *   : *.:  :*.:**: : : :* :* *  : .    :*   :.* **: *  .

OpuBC    PMQIGLVYDAVKSGKMDIVLAYSTDGRIKSYGLKMLKDDKQFFPPYDCSPVVPEKVLKEH 235
OpuCC    PMQIGLVYDALNTEKLDVALGYSTDGRIAAYDLKVLKDDKQFFPPYAASAVATNELLRQH 238
ProX     QMEPTLMYEAIKNKQVDVIPAYTTDSRVDLFNLKILEDDKGALPPYDAIIIVNGNTAKDE 228
          *:  *:*:*::. ::*:  .*:**.*:  :.**:*:*** :*** . . :.  :  ::.

OpuBC    PELEGIIKKMLGKIDTATMQELNYEVDGNLKEPSVVAKEYLEKHRYFES------ 284
OpuCC    PELKTTINKLTGKISTSEMQRLNYEADGKGKEPAVVAEEFLKKHHYFDKQKGGHK 293
ProX     -KLISVLKLLEDRIDTDTMRALNYQYDVEKKDAREIAMSFLKEQGLVK------- 275
          :*    :: : .:*.*  *: ***: * : *:.  :* .:*::: ..
```

Figure 6:
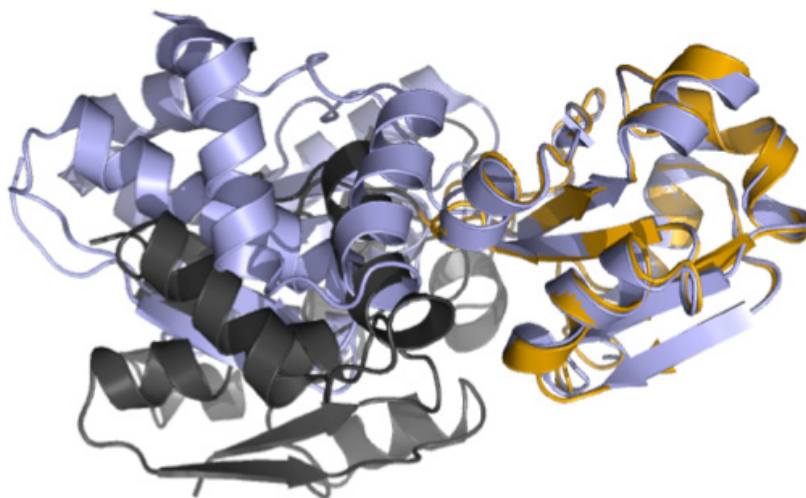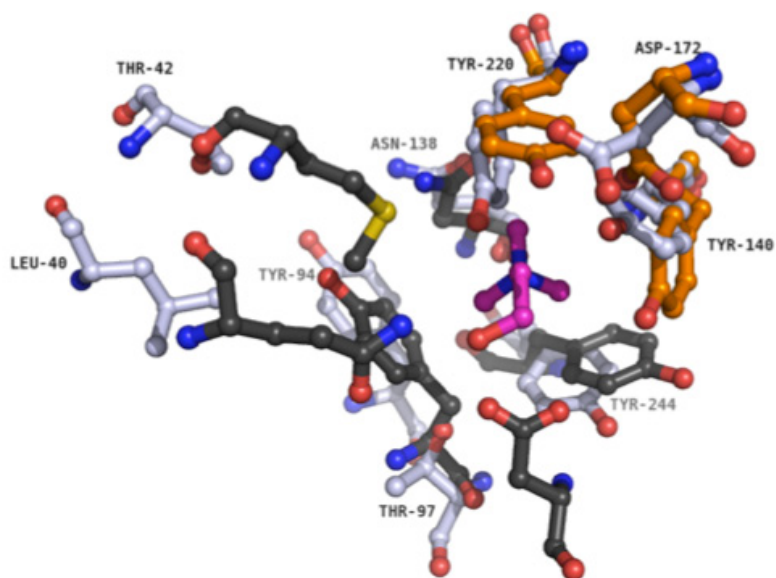
Table 1

Space group	C2
Unit cell parameters	
a, b, c (Å)	104.0, 42.5, 60.5
α, β, γ (deg)	90.0, 97.6, 90.0
Data collection and processing	0.97625
wavelength (Å)	20-1.8
resolution (Å)	2.68 (2.64)
mean redundancy	22957
completeness (%)	93.5 (96.4)
I/σ	23.97 (6.73)
Refinement	
RFc (%)	18.07
Rfree (%)	23.30
RMSD from ideal	
Bond length (Å)	0.006
Bond angles (deg)	1.027
Ramachandran plot	
most favored (%)	96.92
allowed (%)	2.69
disallowed (%)	0.38
Model content	
monomers/ASU	1
protein residues	270
ligand	1

Table 2:

Mutant	K$_D$ choline [µM]
wild type	30,5 ± 3,5
Q19A	434,3 ± 55,4
M21A	97,6 ± 14,3
D74A	n.b.*
N115A	n.b.*
D149A	556,6 ± 32
L155A	767,6 ± 77
Q19A / L155A	1450 ± 61,3
D149A / L155A	1132 ± 96

Own proportion to this work: 30%

In preparation

Chapter III

A single amino acid defines the high affinity state in the glycine betaine binding protein ProX from the hyperthermophile *Archaeoglobus fulgidus*

Britta Tschapek[1][†], Marco Pittelkow[3][†], Linda Sohn-Bösser[3], Gudrun Holtmann[3], Sander H. J. Smits[1], Holger Gohlke[2], Erhard Bremer[3], and Lutz Schmitt[1]*

[1]Institute of Biochemistry, [2]Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Duesseldorf, Universitaetsstr. 1, D-40225 Duesseldorf, Germany.

[3]Laboratory for Microbiology, Department of Biology, Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35032 Marburg, Germany.

Running title: Arg149 seals the high affinity state of the glycine- and proline betaine binding protein *AfProX*

[†]Both authors contributed equally

_____

*: To whom correspondence should be addressed:

Lutz Schmitt

Institute of Biochemistry

Heinrich Heine University Duesseldorf

Universitaetsstr. 1

40225 Duesseldorf, Germany.

Phone: +49-211-81-10773

Fax: +49-211-81-15310

Email: lutz.schmitt@hhu.de

Abstract

The substrate binding protein (SBP) ProX from *A. fulgidus* is highly selective for glycine betaine (GB) and proline betaine, which are required for thermoprotection to counterbalance temperature fluctuation in the habitat. A detailed mutational analysis of the substrate-binding site revealed the contribution of individual amino acids for ligand binding. Interestingly, exchange of a single amino acid, Arg149, displayed the largest impact. To gain further insights, the structure of a mutant was solved in the liganded-open conformation. Here, GB is bound differently compared to the liganded-closed structure. A network of amino acid side chains communicates the presence of GB towards Arg149, which increases ligand affinity and induces domain closure. These results are corroborated by molecular dynamics studies and support the view that a single amino acid acts as a switch to induce the high affinity state of the SBP.

Introduction

Many microorganisms import a variety of chemical compounds from environmental sources through high-affinity transport systems, many of which belong to the family of ATP Binding Cassette (ABC) transporters (*7*). These systems depend strictly on a so-called substrate binding protein (SBP) that captures the ligand with high affinity and subsequently delivers it to its cognate transporter (*4*). The SBP thereby determines the directionality of the overall transport reaction, and its interaction with the transmembrane domain of the transporter regulates the ATPase activity of the nucleotide-binding domain of the transport complex (*4, 47*).

Structural studies of a substantial number of SBPs revealed a common fold with a bilobal organization connected via a linker region (*48, 49*). In the ligand-free, open conformation, the two lobes or domains are separated from each other, thereby forming a deep, solvent exposed cleft, which harbors the substrate-binding site. Upon ligand binding, both domains of the SBP move towards each other through a hinge-bending motion or rotation, which results in the so-called liganded-closed conformation. As a consequence of this movement, residues originating from both domains generate the ligand-binding site and trap the ligand deeply within the SBP (*48*). In the absence of a ligand, unliganded-open and unliganded-closed states of the SBP are in equilibrium, and the ligand solely shifts this equilibrium towards the liganded-closed state. This sequence of events has been coined the "Venus-fly trap mechanism" (*29, 50, 51*), which is supported by numerous crystal structures in the absence and presence of a ligand(*52, 53*) and other biophysical techniques(*48*).

For the maltose binding protein (MBP) from *Escherichia coli*(*4*), the paradigmatic SBP, it has been shown that both domains are dynamically fluctuating around an average orientation in the absence of the ligand (*54*). NMR spectroscopy of MBP in solution revealed that the ligand-free form of MBP consists of a predominantly open species (95%) and a minor species (5%) that corresponds to a partially closed state; both forms co-exist in rapid equilibrium(*55*). The open form of MBP observed by NMR is similar to the crystal structure of the unliganded-open conformation (*56*). However, the partially closed species detected by NMR (*55*) does not correspond to the ligand-bound, fully closed form found in crystallographic studies. Instead, it represents an intermediate, partially closed conformation (*57*), suggesting that the substrate is required to reach the final, liganded-closed conformation.

In line with this suggestion, Millet *et al.* have shown that every degree of domain closure in the unliganded MBP requires a conformational deformation energy of approximately 212 cal/mol. Since MBP closes by about 20°, this would lead to an energetic cost of ~ 4 kcal/mol when forming the unliganded-closed conformation of MBP (*58*). This unliganded-closed state, first observed by X-ray crystallography, is therefore highly unfavorable, and resembles a loaded spring (*58*). Changing the stiffness of this region by introducing amino acids with bulky

side chains changes the degree of closure of MBP, which directly correlates with maltose binding affinity (*58*),(*59*). As a consequence, mutations in the hinge region likely alter the equilibrium between the open and partially closed states of MBP (*55*). Upon maltose binding, a further shift in this equilibrium occurs towards the fully liganded-closed conformation, which becomes accessible because maltose forms favorable interactions with the protein.

Generally, the hinge region provides MBP with a large degree of flexibility. This allows MBP to bind a variety of malto-oligosaccharides, some of which lead to domain closure (*56, 60*). Although this highlights a distinct role of the substrate in domain closure, the exact nature of the substrate-induced closing mechanism of MBP and SBPs in general is unknown. Furthermore, SBPs bind their substrates with high affinity (in the low μM or even nM range) (*48*) but it is not fully understood how the substrate is released to its cognate ABC importer. One scenario could be that interactions between the ABC transporter and the SBP during the transport cycle lead to a modulation of the SBP's affinity to its substrate(*47*).

In the sulfate-reducing thermophilic Archaeon *Archaeglobus fulgidus (Af)*, the ABC importer ProU has been annotated as a specific uptake system for the compatible solutes glycine betaine (GB) and proline betaine (PB) (*61*). The corresponding SBP (*Af*ProX) has been crystallized in different conformations: a liganded-closed conformation in complex either with GB or PB as well as in an unliganded-open conformation (*62*). Here, we present the structure of *Af*ProX in the liganded-open conformation together with a systematic mutational analysis of the binding site of *Af*ProX and molecular dynamics simulations of liganded *Af*ProX, using starting structures of different conformational states. Based on our results, residue Arg149 is crucial for the switch from a low-affine open structure of *Af*ProX to a high-affine closed state. Furthermore, we elucidate the role of the substrate in this process, whose presence is communicated through a network of amino acids located in both domains of *Af*ProX, ultimately triggering domain closure. Our data implies that the "Venus-fly trap" mechanism is overlaid with distinct molecular events to ensure the specific biological function of SBPs.

Results

Thermoprotection of the hyperthermophilic archaeon *A. fulgidus* by GB and PB

*Af*ProX possesses striking substrate specificity: from all of the osmoprotectants tested only GB and PB are substrates (Figure 1A). Both, GB and PB, typically function as protectants in microorganisms especially during changes in the salinity, which lead to osmostress. However, we observed that these compounds had no osmoprotective effect for *A. fulgidus* (Supplementary Figure 1). Therefore, the uptake of GB or PB must serve another physiological role. Since GB has been shown previously to serve as a thermoprotectant in different bacteria [63], we explored a possible thermoprotective effect of GB and PB for *A. fulgidus*. Neither GB nor PB had any effect on the growth of this archaeon when it was cultivated at its optimal growth temperature of 83°C (Supplementary Figure 1). In contrast, both compounds exerted a striking thermoprotective effect on cell growth at an elevated growth temperature of 90°C. Without GB or PB, *A. fulgidus* was unable to grow at that temperature (Figure 1B). To our knowledge, such thermoprotection has only been observed for organisms living at moderate temperatures but never for a hyperthermophile.

Mutational analysis of the ligand-binding site in *Af*ProX

*Af*ProX has been crystallized with GB or PB at a resolution of 2.1 Å and 1.9 Å, respectively (62). The positively charged trimethlyammonium head group of GB and the dimethylammonium head group of PB are wedged into an aromatic cage that is formed by the main chain carbonyl of Asp109 at the bottom and by four Tyr residues (Tyr63, Tyr111, Tyr119, Tyr214) forming the four faces of a box. Here, cation-π interactions are key determinants for coordinating the head group. The carboxylic tails of GB and PB protrude from this partially negatively charged cage and form two salt bridges and one hydrogen bond with Lys13, Thr66, and Arg149, respectively (62). We have measured the binding of GB and PB to *Af*ProX using fluorescence spectroscopy and determined a K_D of 60 ± 10 nM and 50 ± 10 nM, respectively (Supplementary Table 1). To further support this high but not unusual affinity of SBPs (49, 64), we also determined the binding constant of GB by isothermal calorimetry (ITC) and revealing a K_D of 100 ± 30 nM (Supplementary Figure 2). This value is within experimental error identical to the K_D determined by fluorescence spectroscopy.

To assess the contribution of individual residues within the substrate-binding site, we carried out a systematic mutagenesis study. The individual replacement of the four Tyr residues forming the aromatic cage reduced the binding affinity between 60-fold and 2400-fold for GB and between 60-fold and 4500-fold for PB (Supplementary Table 1). The simultaneous

replacement of two Tyr residues by Ala, regardless of the combination, abolished ligand binding, highlighting the importance of the aromatic cage. The individual substitution of Lys13, Thr66, and Arg149 by Ala also resulted in a decreased ligand binding affinity. The largest decrease, however, was observed for the Arg149Ala mutation, which exhibited a K_D of 320 ± 60 µM for GB (Supplementary Table 1), a 5000-fold reduction in binding affinity, while no binding of PB could be detected. Therefore, Arg149 is of high importance in *Af*ProX.

Crystal structure of Tyr111Ala *Af*ProX

The Tyr mutations of the aromatic box had an unexpectedly large effect on the affinity of GB. Therefore, we crystallized one of these mutants, Tyr111Ala, and solved its structure in the presence of GB at 2.0 Å. Tyr111Ala has an affinity of 76 ± 4 µM towards GB (Supplementary Table 1). Crystallization conditions are given in Materials and Methods and data statistics are shown in Supplementary Table 2. The structure of Tyr111Ala *Af*ProX consists of two domains, the ligand GB, and a HEPES molecule captured from the crystallization solution. A comparison with the unliganded-open and a liganded-closed structure revealed that the structure of Tyr111Ala *Af*ProX adopts a liganded-open conformation, with the hinge region positioned between residues 109-111 and 213-215 (Figure 2). A comparison between the liganded-open and unliganded-open structure revealed that only one loop (amino acid 142 to 153) differs. This loop contains Arg149 that is part of the substrate-binding site in the liganded-closed structure and will be referred to as "Arg-loop" henceforth.

The substrate-binding site of the *Af*ProX Tyr111Ala variant

In the *Af*ProX Tyr111Ala mutant, the trimethylammonium moiety of GB is located at the same position as in the liganded-closed structure (Figure 3). Although GB lacks interactions with Tyr111, it is bound by Tyr63, Tyr190, and Tyr214 of the aromatic cage and further stabilized by interactions with Thr66, Asp109, and the piperazinyl ring of the bound HEPES. Notably, when compared to the closed structure of the wild type protein, the carboxyl tail of GB is rotated by almost 180° in the *Af*ProX Tyr111Ala mutant. Here, GB interacts through a water molecule with Thr66 and through two water molecules with the side chain of Tyr190. As the carboxyl group of GB neither forms a direct nor a water-mediated interaction with the HEPES molecule, the orientation of GB should not be influenced.

The architectures of the binding sites in the liganded-closed, liganded-open, and unliganded-open structures of *Af*ProX are closely related because Tyr63, Tyr214, Lys13, Thr66, and Asp109 are superimposable in all three structures (Figures 4A and 4B). However, the side chain of Tyr190 is flipped in the closed structure by 130° when compared to both open structures. Interestingly, the position of the Tyr111 side chain in the liganded-closed structure is

located at the position of the side chain of Tyr190 in the unliganded-open and liganded-open structure (Figures 4A and 4B). These observations imply that during the opening and closing movements of *Af*ProX, the binding site does not undergo a major conformational change; rather, it is largely pre-formed in the unliganded-open structure and, thus, pre-dispositioned to capture GB. Still, the side chains of Tyr111 and Tyr190 need to reorient. Furthermore, the largest movement is observed for the Arg-loop, which moves by approximately 10Å, starting from an outward-rotated orientation. This leads to side chain interactions of Arg149 with GB, Tyr111, and Thr66, thereby complementing and locking the binding site.

The role of Arg149

In addition to the direct interaction of Arg149 with GB and residues that are part of the substrate-binding pocket (Tyr111, Thr66), Arg149 is a major determinant in domain-domain interactions in the closed structure (Supplementary Table 3). As such, Arg149 interacts with Val70 (domain I) and Asp151 (domain II), thereby acting as a linking element between the two domains enforcing a stable domain closure. These interactions complement those mediated by Pro172 of domain II, where Pro172 interacts via its C_α-atom and a water molecule with Glu155 of domain II. Together, this provides a further explanation for the crucial role of Arg149 for the stability of the liganded-closed state. This crucial role is manifested by a drop in affinity towards GB and abolished binding of PB (Supplementary Table 1) when replacing Arg149 by Ala.

Partial domain closure of *Af*ProX occurs even in the absence of GB

To study the role of Arg149 with respect to the domain closure of *Af*ProX in further detail, we performed molecular dynamics (MD) simulations starting from the liganded-open *Af*ProX structure with bound GB. Here, Ala111 was mutated *in silico* to Tyr in the Tyr111Ala *Af*ProX crystal structure. After 50 ps of MD simulation, GB reoriented from the configuration observed in the *Af*ProX Tyr111Ala crystal structure to the configuration seen in the closed structure (Supplementary Figure 4). In the initial configuration, only water-mediated interactions are formed between the carboxyl moiety of GB and *Af*ProX. After the reorientation, the carboxyl moiety of GB forms a hydrogen bond with Thr66 and a salt bridge with Lys13. Furthermore, after 4 ns of MD simulation, GB dissociates from *Af*ProX and does not re-associate during another 196 ns of simulation time (Supplementary Figure 5). Notably, after 50 ns of MD simulation, unliganded *Af*ProX undergoes a drastic conformational change: domain II comes as close as 2 Å rmsd to the configuration found in the closed *Af*ProX structure (Supplemental Figure 5). During the remaining 150 ns of MD simulation, domain II repeatedly opens again but always returns to the configuration akin to the closed structure. In summary, MD simulations

started from the open *Af*ProX structure reveal that a tendency to domain closure exists even in the absence of GB. Such motions are inline with the "Venus-fly trap mechanism" (*48*) and have been experimentally verified for MBP (*55*).

Coupled structural reorganization in the binding site upon domain closure

In order to determine in atomic detail the coupling of conformational changes in the ligand-binding site that occurs upon domain closure in the presence of GB, MD simulations starting from the liganded-open *Af*ProX structure were repeated. However, this time GB was restrained to the ligand-binding site to prevent dissociation. Again, a domain closure was observed, with domain II coming as close to the ligand-bound bound state as 4.2 Å rmsd (Figures 5A and 6A). Notably, this state is reached after 11 ns, about five times faster than in the absence of GB (Figure 5A). Although single events that occur only once during a MD simulation must be interpreted with caution, these results indicate that binding of GB facilitates conformational changes occurring during the transition from the open to the closed state. Furthermore, domain closure is accompanied by a reorientation of GB (Figure 5A). In parallel, the Tyr cage that coordinates the binding of the trimethyl ammonium moiety of GB via cation-π interactions starts to adopt the conformation observed in the closed structure (Figure 5B). In this process, Tyr111 shifts about halfway between the conformations of the open and the closed structures, whereas the conformation of Tyr190 is still close to the one observed in the open structure (Figure 6B). For Tyr111 to reach its final position, Tyr190 must reorient. In addition, the minimal distance between any Oe atom of Glu145 and any NH atom of Arg149 fluctuates between 2.8Å and ~ 6Å during the first 8 ns of simulation time (Figure 5B), indicating the temporary formation of a salt bridge. This salt bridge consolidates during the next 2 ns, after which it is almost permanently formed for another 2 ns. Subsequently, upon domain opening after about 12 ns of simulation time, the interaction between Glu145 and Arg149 breaks up, and Glu145 and Arg149 remain fluctuating between a salt bridge-formed state and complete separation for the remainder of the simulation time. Thereby, maxima of the domain II-rmsd curve (Figure 5A) are significantly paralleled by large Glu145-Arg149 distances (Figure 5B), again demonstrating in a remarkable way the coupling of events that lead to domain closure and opening of the *Af*ProX protein.

In order to further investigate the central role of Arg149 for the domain closure of *Af*ProX, unrestrained MD simulations started from liganded-closed wild type and Arg149Ala *Af*ProX structures were performed. Over the course of 50 ns of simulation time, the wild type structure starts to open up once but immediately returns to the closed conformation (Figure 7). In contrast, opening of the domains is observed earlier, more frequently, and more pronounced in the Arg149Ala structure. Most notably, after opening, the Arg149Ala structure never regains the closed conformation again.

Substrate binding triggers domain closure to a closed state with a "loaded spring"

The above-described experiments unambiguously established the importance of Arg149 for the high-affine binding of GB (or PB) to *Af*ProX. Furthermore, *Af*ProX is essential during thermal fluctuations in the habitat of *A. fulgidus* (see above), and the response to such fluctuations must be fast. Thus, binding and release of GB (or PB) to and from *Af*ProX must be fast, too. This raises the question if an initial binding of GB (or PB) can trigger a domain closure leading to a closed conformation stabilized by Arg149. To identify such a possible triggering mechanism, we analyzed interactions of the bound substrate in the liganded-open structure with Arg149. Already in this state, the presence of GB is communicated to Arg149 through interactions of the side chains of Tyr190, Tyr111, and Phe146; that is, GB and Arg149 are connected with each other via a direct side-chain network. Interestingly, this network does not change from the open to the closed structure as judged by unaltered distances between these amino acids in both states.

Together with the information from MD simulations, this suggests the following substrate-induced triggering mechanism in *Af*ProX. As highlighted in Figure 8A, the open conformation appears to be stable and ready to bind GB (or PB) due to the largely preformed ligand-binding site. The next state, after GB binding, is represented by the liganded-open structure of *Af*ProX presented here (Figure 8B). In this state, GB interacts with Tyr190 (distance between the C_α atom of GB and the C_{e1} atom of Tyr190: 3.8 Å). Immediately after binding, GB reorients around its trimethylammonium group to the configuration seen in the closed structure as observed in the MD simulation. This reorientation leads to GB attracting the side chain of Tyr190 and Tyr111. The additional cation-π interactions between Tyr111 and Tyr190 and GB's trimethylammonium group are energetically favorable; their formation may thus be the cause why the domain closure of *Af*ProX is facilitated in the ligand-bound state compared to the unliganded structure as also observed in the MD simulations. The Tyr111 and Tyr190 movements are relayed to Arg149 via the above-described interactions. Ultimately, the reorientation of GB triggers domain closure and induces the high-affine liganded-closed conformation in which Arg149 finally complements and locks the binding site and stabilizes the domain closure (Figure 8C).

During the domain closure, part of the Arg-loop changes from a stable α-helical conformation in the open structure to an unwound conformation in the liganded-closed structure. This secondary structure change leads to a "loaded spring" in this region, which is restrained by a salt bridge of Arg149 (domain II) with the carboxyl group of GB. GB itself forms additional salt bridges with Lys13 and Thr66 of domain I. If this Arg149-GB interaction is disrupted, the Arg-loop will adopt its energetically favored, "relaxed" helical conformation. This leads to the opening of *Af*ProX and, hence, the release of the substrate. As a consequence, *Af*ProX should not bind solutes that lack a carboxyl group, e.g. choline. Furthermore, since the exact positioning of the carboxyl group is important, compatible

solutes with increased size, e.g. crotonobetaine, homobetaine, carnitine, dimethylsulphonio acetate, and ectoine should not induce a stable closed conformation. This is what we observed in competition assays with radiolabeled GB (Figure 1A). Hence, our analysis of the crystal structures of the *AfProX* protein provides a rational explanation for the observed substrate specificity.

Discussion

High-affinity ligand binding by the *Af*ProX SBP

The crystal structures of wild-type *Af*ProX in complex with GB and PB and an open unliganded configuration were previously reported (*62*). The latter structure is thought to represent the open conformation of *Af*ProX as suggested by the "Venus fly trap model" (*48*). *Af*ProX exhibits high affinities for its substrates GB (K_D = 60 nM ± 10 nM) and PB (K_D = 50 nM ± 10 nM), and it is highly substrate-specific for these two compatible solutes (Figure 1A). GB and PB serve as excellent thermoprotectants (Figure 1B). To our knowledge, thermoprotection of a hyperthermophilic archaeon by GB or PB has never been described before, although both compatible solutes are known to possess thermoprotective properties for mesophilic bacteria (*63*).

We probed the contribution of individual residues within the substrate-binding site of *Af*ProX to ligand binding and found that single amino acid substitutions within the Tyr cage reduced ligand binding between 60-fold (Tyr214/Ala; K_D = 3.5 µM ± 0.7 µM) to 2400-fold (Tyr63/Ala; K_D = 149 µM ± 17 µM). Mutating any two Tyr residues in the aromatic cage to Ala abrogates ligand binding completely (Supplementary Table 1). Similar effects have been observed for the ectoine binding protein EhuB from *Sinorhizobium mellilotti*(*65*), GB/PB-specific OpuAC from *Bacillus subtilis* (*66*), and ProX from *E. coli* (*67*). Here, cation-π interactions also contribute strongly to ligand binding.

Switching the *Af*ProX SBP from low to high affinity

In *Af*ProX, Arg149 plays a distinct role in substrate binding as its substitution by an Ala residue results in a loss of binding affinity from K_D ≈ 50-70 nM to K_D ≈ 300–400 µM. In the liganded-open structure, Arg149 is located ~10 Å away from its final position in the liganded-closed structure. Hence, the ligand-binding site is laid open, and the substrate has free access to the binding pocket. Our data suggest that Arg149 is the final amino acid to interact with the substrate and, thereby, captures GB or PB in the high-affine closed state of *Af*ProX.

Domain closure triggered by substrate binding

The "Venus fly trap" model describes the opening and closing of SBPs in solution where the equilibrium between these two conformations is shifted towards the closed state upon substrate binding (*68*). Many crystal structures of SBPs have been solved in the unliganded-

open, liganded-closed, and, more rarely, in the unliganded-closed state (*48, 53, 62, 69*). In striking contrast, only a few proteins have been crystallized in a liganded-open conformation. Examples are the leucine-isoleucine-valine binding protein (LIV), where the substrate was soaked into preformed unliganded protein (*68*) and the liganded-open form of MBP (*56*). In the case of MBP, the non-natural substrate b-cyclodextrin is bound, which is not transported, however (*56*). In the liganded-open structures of LIV or MBP, the binding site is completely open, and the substrate is bound by amino acids that originate from only one domain of the SBP. In the case of MBP, it was therefore postulated that, first, maltose interacts with the aromatic groups of the C-terminal domain to form a complex with the open form (*70*). Then, the complex converts into the stable closed conformation of MBP. However, molecular dynamics studies of MBP show that the unliganded protein can adopt a wider variety of conformations than the maltose-bound state, indicating that ligand binding not only affects the relative orientation of the two domains but also attenuates their movement relative towards each other (*71*). In contrast to the situation described for MBP, we found that the substrate-binding site of *Af*ProX is pre-formed in the open conformation. Thus, only a slight reorientation has to occur to optimize the binding site towards its closed conformation. Still, although the open conformation contains a largely assembled binding site, it exhibits only low binding affinity towards GB or PB. Our mutational analysis of *Af*ProX revealed a distinct role of Arg149 in substrate binding: the movement of this residue, which complements the binding site in the closed *Af*ProX structure, changes *Af*ProX from a low- to a high-affine state by locking the binding site and mediating domain-domain interactions, thus finally trapping the ligand.

Comparison with other SBP

The triggering mechanism that we have observed in *Af*ProX is intriguing and might apply to other SBPs. Therefore, we carefully compared open and closed structures of MBP. Domain I was chosen as anchor point for the super-positioning of both conformations because it shows almost no conformational change in going from the open to the closed state.

The comparison reveals that Met330 undergoes only a relative small movement of 2.5 Å from the unliganded-open to the liganded-closed structure. Like Arg149 in *Af*ProX, Met330 is connected via a network of side chain interactions to all amino acids participating in substrate binding. First, this is achieved through an interaction network of Met330-Tyr155-Glu153-Arg344. Second, Tyr155 interacts with Trp230 and Phe156. The distance between these side chains does not change in the open and closed structures, identical to the observation in *Af*ProX. Hence, the movement of Met330 allows triggering a pushing event towards all amino acids involved in substrate binding in the second MBP domain. In addition to this network, Met330 also interacts with Phe258, which is located at the beginning of a beta

sheet. It has been shown that this beta sheet is important for stabilizing the closed conformation as it contains Gly260, which is part of a salt bridge that is only observed in the closed conformation of MBP (*71*).

Similar observations have been made when comparing the open and closed structures of the ribose binding protein from *E.coli* (*72*) as well as the binding protein of the *N*-Acetyl-5-neuraminic acid TRAP transporter from *H. influenza (73)*.

Substrate release by external modulation of Arg149 interactions

The Arg-loop of the *Af*ProX protein has a α-helical conformation in the open structure. In the liganded-closed structure, however, this helix is unwound. The energetic cost of unwinding this region is diminished by multiple interactions of Arg149 with domain II of *Af*ProX and the GB ligand. Furthermore, in the liganded-closed structure of *Af*ProX, the side chain of Arg149 is interacting with Glu145, whereas no such interaction occurs in the open structure. Finally, salt-bridge formation between Arg149 and Glu145 upon domain closure is also observed during the MD simulations (see above). Interestingly, Glu145 is exposed at the membrane-facing side of *Af*ProX.

Taken together, our data leads us to suggest a model where a conformational change in the membrane components (ProW-1, ProW-2) and the ATPase (ProV) of the ProU ABC transporter from *A. fulgidus* results in the manipulation or even the disruption of the Glu145-Arg149 salt-bridge. This event is expected to loosen the interactions of Arg149 with the GB and PB substrates, which, in turn, would induce a re-formation of the helical conformation of the Arg-loop. Finally, the opening of the binding site of the *Af*ProX protein would be triggered. Consequently, the affinity of the substrate-binding site would change from low nM to medium µM range, which would facilitate release of the ligand into the substrate-translocation pathway for ATP-dependent import into the *A. fulgidus* cell.

In summary, our data suggest that Arg149 acts as a trigger to fulfill the role of *Af*ProX in thermoprotection: here, a fast response of the SBP is required to ensure a fast and efficient ligand capture. A preformed ligand-binding site that only requires a single amino acid to switch the system from a low- to a high-affine state is a perfect solution to environmental restraints. Our data implies that distinct molecular events such as single amino acid triggers are overlaid with the "Venus-fly trap" mechanism to ensure the specific biological function of SBPs.

Materials and Methods

Crystallization

*Af*ProX crystals were obtained using the hanging drop method at 293 K against a reservoir solution of 100 mM HEPES pH 7.0, 15-25% PEG6000 and 15-25% PEG 8000. *Af*ProX was incubated with 10 mM glycine betaine prior to crystallization. Crystals grew slowly and appeared after one month. They were flash frozen in liquid nitrogen with mother liquor, supplemented with 20-25% ethylenglycerol as cryoprotectant.

Data collection, refinement and structure analysis

*Af*ProX crystals diffracted X-rays beyond 1.6 Å. However for refinement purposes the dataset was truncated at a resolution of 2.0 Å. The dataset of *Af*ProX was collected at the ID23-EH2 beamline at the ESRF and processed with XDS (*74*). *Af*ProX in the ligand free form (PDB code: 1SW5) (*62*) was used as a template to obtain initial phases using PHASER (*75*). The structure was further refined using REFMAC5 (*76*) and COOT (*77*). Dataset and refinement statistics are listed in Table 2 in the Supplementary Information. As analyzed with Procheck (*78*), the Ramachandran plot of *Af*ProX shows one residue in the disallowed region. However, this residue interacts with the bound HEPES molecule, which explains the rather unusual conformation. Figures of protein molecules were prepared using PyMol (www.pymol.org).

Molecular dynamics simulations

Molecular dynamics simulations were performed with the AMBER 10 suite of programs (*79*) together with the force field as described by Cornell *et al*. (*80*) using modifications suggested by Simmerling *et al*. (*81*). In total, four different MD simulations were performed: I) liganded-open wt *Af*ProX. The starting structure was generated by mutating Ala111 in the crystal structure of Tyr111Ala *Af*ProX to Tyr. The simulation length is 200 ns; II) liganded-open wt *Af*ProX. In contrast to I), a distance restraint by means of a harmonic potential was applied between the ammonium nitrogen of GB and the N of Asp109 of *Af*ProX. The simulation length is 25 ns; III) liganded-closed wt *Af*ProX. The starting structure was taken from PDB code 1SW4 (*62*). The simulation length is 50 ns; IV) liganded-closed Arg149Ala *Af*ProX. The starting structure was generated from the liganded-closed wt *Af*ProX structure by mutating Arg149 to Ala. The simulation length is 50 ns.

In all cases, the starting structure was placed into an octahedral periodic box of TIP3P water molecules (*82*). The distance between the edges of the water box and the closest atom of the protein was at least 11 Å, resulting in a system of ~43000 atoms. The system was minimized by 50 steps of steepest descent minimization followed by 450 steps of conjugate

gradient minimization. The particle mesh Ewald (PME) method (*83*) was used to treat long-range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained using SHAKE (*84*). The time-step for all MD simulations was 2 fs, with a direct-space, non-bonded cutoff of 8 Å. Applying harmonic restraints with force constants of 5 kcal mol^{-1} Å^{-2} to all solute atoms, canonical ensemble (NVT)-MD was carried out for 50 ps, during which the system was heated from 100 K to 300 K. Subsequent isothermal isobaric ensemble (NPT)-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to zero during 100 ps of NVT-MD. The following NVT-MD at 300 K with a time constant of 10 ps for heat-bath coupling was used for analysis, with conformations extracted every 20 ps. Atomic charges for the GB ligand were generated following the RESP procedure (*85*); force field parameters for GB were taken from GAFF (*86*).

For analyzing the trajectories, conformations were superimposed with respect to the C_α atoms of domain I. This resulted in an almost perfect overlay in all cases. Root mean-square deviations with respect to the open or closed starting structures were then determined for C_α atoms of domain II or the atoms of GB.

ACKNOWLEDGEMENTS

Literature

1. Kempf, B. & Bremer, E. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch Microbiol* **170**, 319-30 (1998).

2. Davidson, A.L., Dassa, E., Orelle, C. & Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* **72**, 317-64, table of contents (2008).

3. Cui, J., Qasim, S. & Davidson, A.L. Uncoupling substrate transport from ATP hydrolysis in the Escherichia coli maltose transporter. *J Biol Chem* **285**, 39986-93 (2010).

4. Wilkinson, J. & Verschueren, K.H.G. Crystal structures of periplasmic solute-binding proteins in ABC transport complexes illuminate their function. in *ABC proteins: from bacteria to man* (eds. Holland, I.B., Cole, S.P.C., Kuchler, K. & Higgins, C.F.) 187-208 (Academic Press (Elsevier Science), London, 2003).

5. Berntsson, R.P., Smits, S.H., Schmitt, L., Slotboom, D.J. & Poolman, B. A structural classification of substrate-binding proteins. *FEBS Lett* **584**, 2606-2612 (2010).

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

7. Mao, B., Pear, M.R., McCammon, J.A. & Quiocho, F.A. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J Biol Chem* **257**, 1131-3 (1982).

8. Sack, J.S., Saper, M.A. & Quiocho, F.A. Periplasmic binding protein structure and function. *J Mol Biol* **206**, 171-191 (1989).

9. Oh, B.H. et al. Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J Biol Chem* **268**, 11348-55 (1993).

10. Oswald, C. et al. Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from Sinorhizobium meliloti in the liganded and unliganded-closed states. *J Biol Chem* **283**, 32848-59 (2008).

11. Shilton, B.H. The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems. *Biochim Biophys Acta* **1778**, 1772-80 (2008).

12. Tang, C., Schwieters, C.D. & Clore, G.M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **449**, 1078-82 (2007).

13. Sharff, A.J., Rodseth, L.E. & Quiocho, F.A. Refined 1.8-A structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein. *Biochemistry* **32**, 10553-9 (1993).

14. Spurlino, J.C., Lu, G.Y. & Quiocho, F.A. The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem* **266**, 5202-19 (1991).

15. Millet, O., Hudson, R.P. & Kay, L.E. The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. *Proc Natl Acad Sci U S A* **100**, 12700-5 (2003).

16. Marvin, J.S. & Hellinga, H.W. Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nat Struct Biol* **8**, 795-8 (2001).

17. Duan, X., Hall, J.A., Nikaido, H. & Quiocho, F.A. Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. *J Mol Biol* **306**, 1115-26 (2001).

18. Klenk, H.P. et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364-70 (1997).

19. Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Biol Chem* **279**, 48270-81 (2004).

20. Holtmann, G. & Bremer, E. Thermoprotection of Bacillus subtilis by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *J Bacteriol* **186**, 1683-93 (2004).

21. Vahedi-Faridi, A. et al. 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**, 448-59 (2008).

22. Hanekop, N. et al. Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J Mol Biol* **374**, 1237-50 (2007).

23. Smits, S.H. et al. The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, site-directed mutagenesis, and crystallographic studies. *J Bacteriol* **190**, 5663-71 (2008).

24. Schiefner, A. et al. Cation-pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J Biol Chem* **279**, 5588-96 (2004).

25. Sack, J.S., Saper, M.A. & Quiocho, F.A. Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J Mol Biol* **206**, 171-91 (1989).

26. Oswald, C., Smits, S.H., Hoing, M., Bremer, E. & Schmitt, L. Structural analysis of the choline-binding protein ChoX in a semi-closed and ligand-free conformation. *Biol Chem* **390**, 1163-70 (2009).

27. Diez, J. et al. The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic Archaeon Thermococcus litoralis at 1.85 A. *J Mol Biol* **305**, 905-15 (2001).

28. Stockner, T., Vogel, H.J. & Tieleman, D.P. A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys J* **89**, 3362-71 (2005).

29. Bjorkman, A.J. et al. Probing protein-protein interactions. The ribose-binding protein in bacterial transport and chemotaxis. *J Biol Chem* **269**, 30206-11 (1994).

30. Muller, A. et al. Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. *J Biol Chem* **281**, 22212-22 (2006).

31. Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J Appl Cryst* **26**, 795-800 (1993).

32. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

33. Murshudov, G., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallography D* **53**, 240-255 (1997).

34. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biological Crystallography* **60**, 2126-2132 (2004).

35. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**, 283-291 (1993).

36. Case, D.A. et al. The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668-1688 (2005).

37.   Cornell, W.D. et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* **117**, 5179-5197 (1995).

38.   Simmerling, C., Strockbine, B. & Roitberg, A.E. All-atom structure prediction and folding simulations of a stable protein. *J Am Chem Soc* **124**, 11258-9 (2002).

39.   Jorgensen, W.L., Chandrasekhar, J., Madura, J. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926-935 (1983).

40.   Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald - A Nlog(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092 (1993).

41.   Ryckaert, J.P., Ciccotti, G. & Berendsen, H.J.C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular Dynamics of n-Alkanes. *J Comput Phys* **23**, 327-341 (1977).

42.   Bayly, C.I., Cieplak, P., Cornell, W.D. & Kollman, P.A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model. *J Phys Chem* **97**, 10269-10280 (1993).

43.   Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. & Case, D.A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-74 (2004).

Figures

Figure 1: *Af*ProX mediated thermoprotection of *A. fulgidus*

A) Substrate specificity of *Af*ProX. The substrate specificity of ProX was investigated in a competitive radioactive binding assay. 5 µM *Af*ProX in 100 µl 10 mM Tris-HCl (pH 7.5) were incubated at 85°C with 5 µM [1-14C] glycine betaine and forty-fold excess of different unlabeled compatible solutes. As control the same experiment was done with no competitive substrate and that value was set to 100 %. B) Influence of GB and PB on the growth of *A. fulgidus* at elevated temperature. Pre-warmed media lacking or containing 2 mM GB or 2 mM PB were inoculated with 5 % exponentially growing *A. fulgidus* cells and incubated on a rotary shaker at 90°C. Squares: *A. fulgidus* cells grown in the absence of either GB or PB; diamonds: *A. fulgidus* cells grown in the presence of 2 mM GB; triangles: *A. fulgidus* cells grown in the presence of 2 mM PB.

Figure 2: The binding site of Tyr111Ala *Af*ProX

The GB-binding site of Tyr111Ala *Af*ProX is shown in sticks as is Arg149 with its neighboring residues Asp145 and Glu151. Dashed lines depict distances < 3.5 Å between Arg149 and Asp145 or Glu151. Bound GB is highlighted in blue.

Figure 3: Overlay of Tyr111Ala *Af*ProX with the liganded-closed and the unliganded-open structure of *Af*ProX.

A) Overall structure of *Af*ProX highlighted as cartoons. B) Overlay of the Tyr111Ala *Af*ProX structure with the liganded-closed structure of *Af*ProX (pink, PDB code 1SW2). C) Overlay of the Tyr111Ala *Af*ProX structure with the unliganded-open structure of *Af*ProX (light green, PDB code 1SW5). The overlays were calculated using LSQMAN.

Figure 4: Overlay of the Tyr111Ala *Af*ProX binding sites

A) The binding site of Tyr111Ala *Af*ProX (shown in blue) is overlayed with the open unliganded structure (orange). Residues involved in substrate binding are highlighted.

B) The binding site of Tyr111Ala *Af*ProX (shown in blue) is overlayed with the liganded-closed structure (white). Residues involved in substrate binding are highlighted. For simplicity, Asp109 was not included in A and B. We note that this residue superimposes perfectly in both structures.

Figure 5: MD simulation started from the open-liganded wild type *Af*ProX structure tethering GB in the binding site.

A) Rmsd of domain II with respect to the closed-liganded structure (red); rmsd of GB with respect to the starting structure (green); B) rmsd of Tyr 111 and Tyr190 with respect to the closed structure (red); minimal distance between $O\varepsilon$ atoms of Glu145 and NH atoms of Arg149 (green). The dashed line indicates the simulation time at which the conformation shown in Figure 6 was extracted.

Figure 6: Overlay of the open structure (orange), the closed structure (dark blue), and the conformation extracted after 11 ns from the MD simulation started from the open-liganded wild type *Af*ProX structure, in which GB was tethered in the binding site (cyan).

A) Domain II of the extracted conformation comes as close as 4.2 Å rmsd to the closed structure. B) Close-up view of the tyrosine pocket.

Figure 7: MD simulation started from the closed-liganded *Af*ProX structure for wild type *Af*ProX (red) and *Af*ProXArg149Ala (green).

Rmsd values of domain II with respect to the open *Af*ProX structure are shown.

Figure 8: Proposed mechanism of *Af*ProX during substrate capturing and high affinity binding.

A) The open conformation of the binding site of *Af*ProX is shown. B) The ligand GB is captured and bound in the binding site C) Binding of the substrate triggers closure of the two domains via the Tyr190-Tyr111-Phe146 network. In the closed structure Arg149 interacts with GB (red, dashed line). All highlighted distances (dashed lines) are between 3.2 and 3.8Å
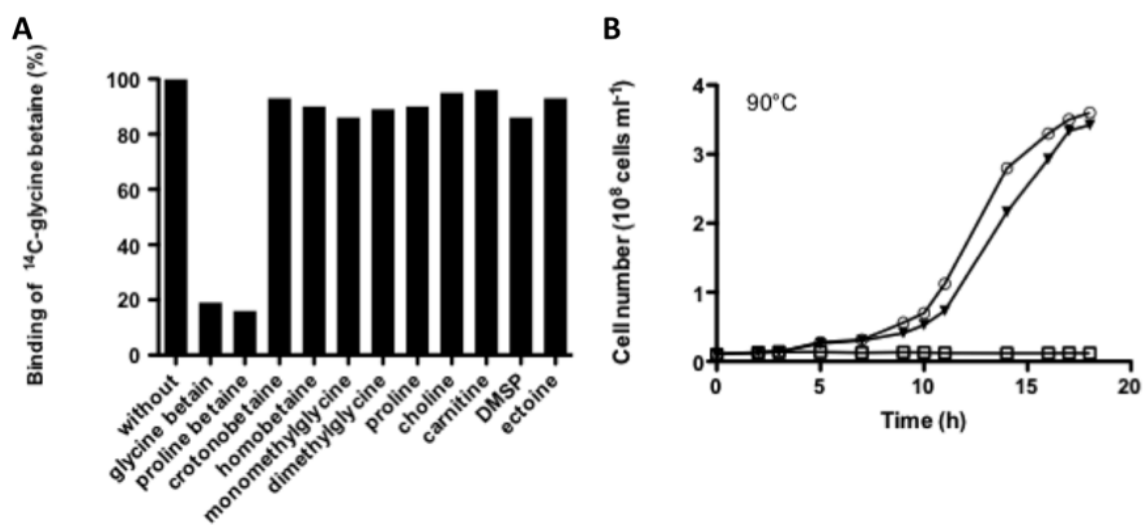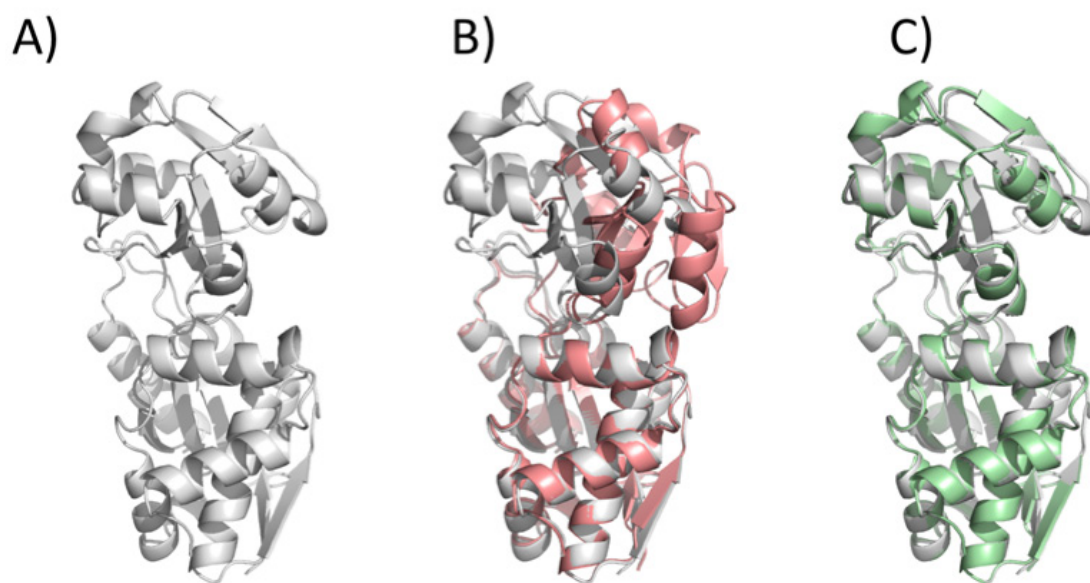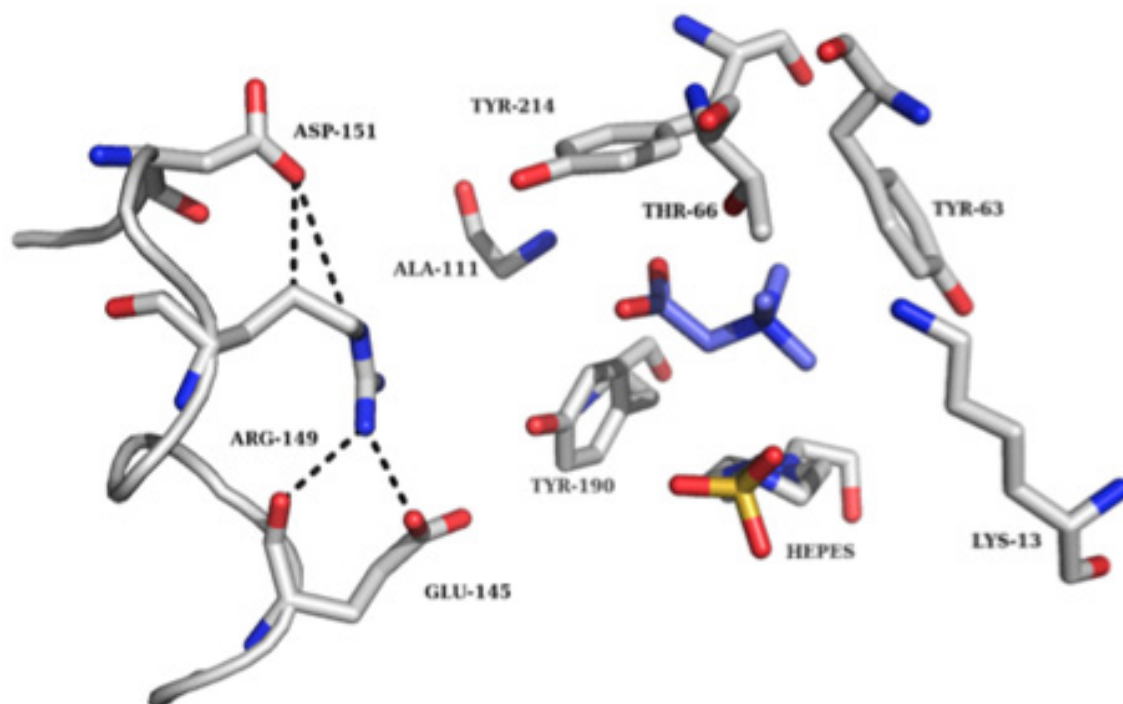
Figure 1:

Figure 2:

Figure 3:

Figure 4:

A)

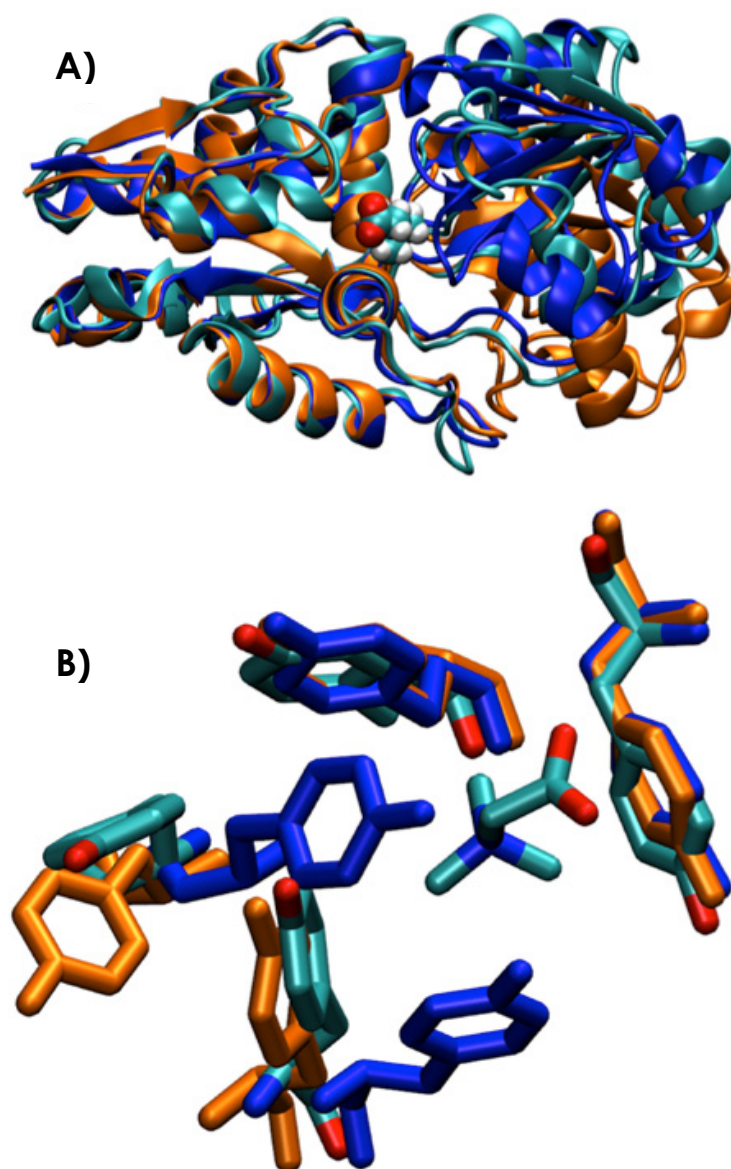**B)**
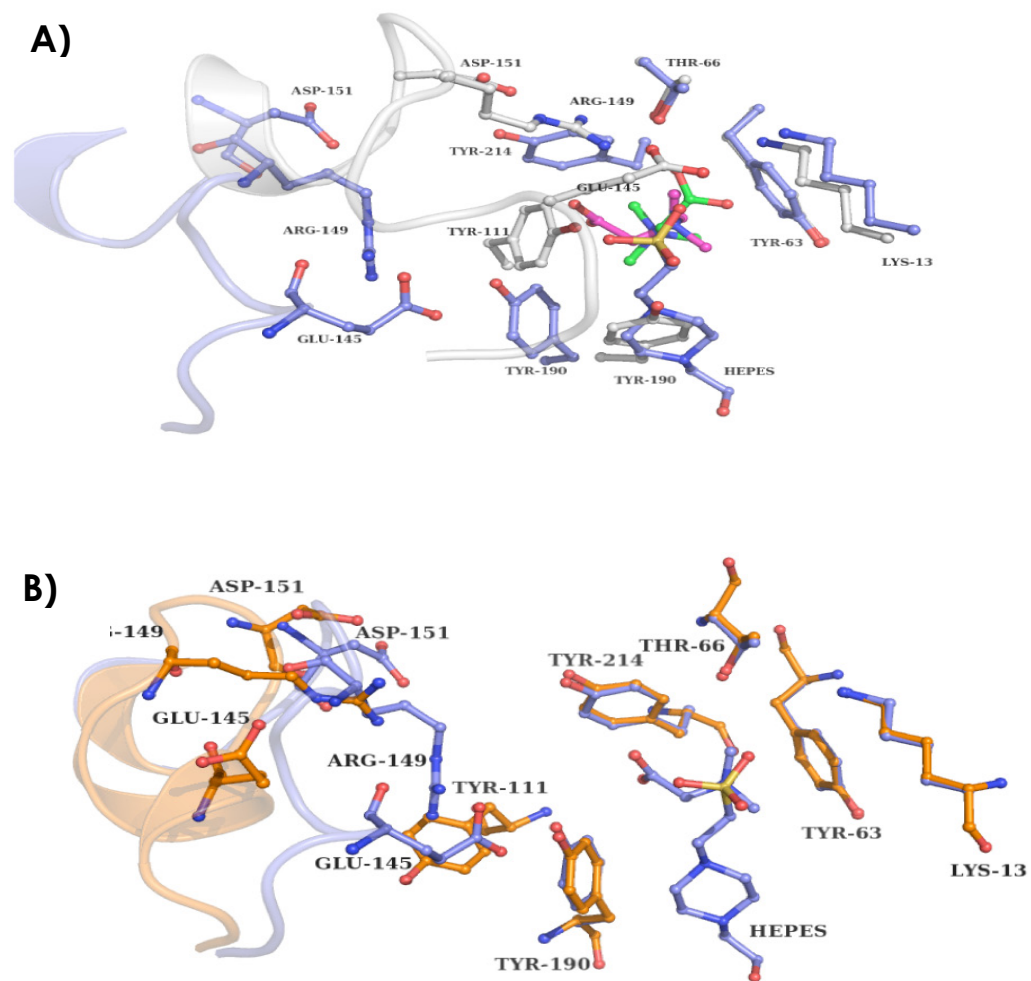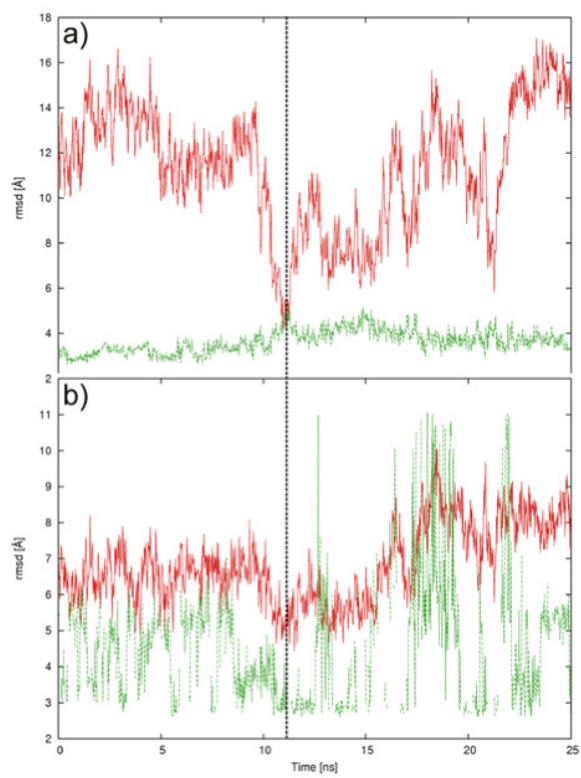
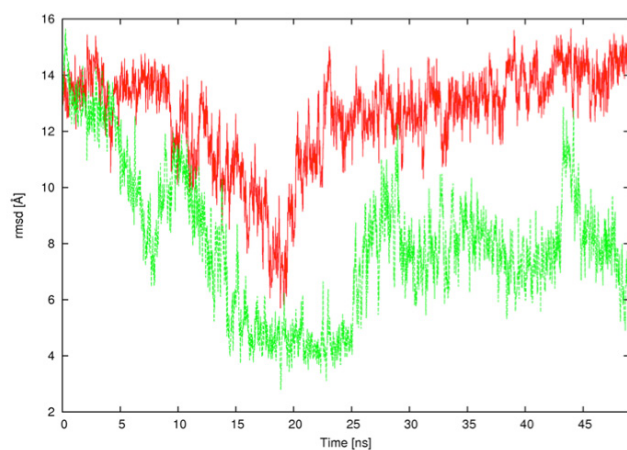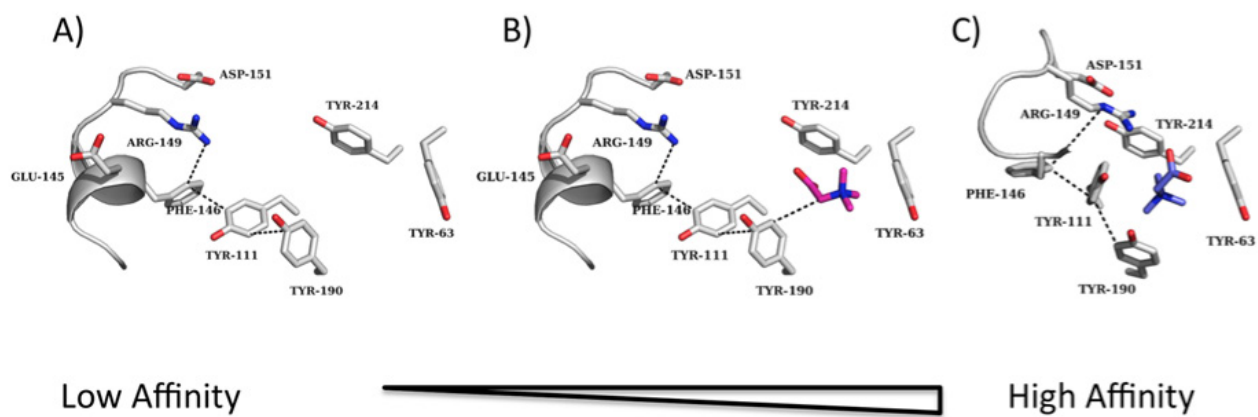**Figure 5:**

Figure 6:

Figure 7:

Figure 8:

Supplementary Information

A molecular switch changes the low to the high affinity state in the substrate binding protein *Af*ProX

Britta Tschapek[1], Marco Pittelkow[3], Linda Sohn-Bösser[3], Gudrun Holtmann[3] Sander H. J. Smits[1], Holger Gohlke[2], Erhard Bremer[3], and Lutz Schmitt[1]*

[1]Institute of Biochemistry, [2]Institute of Pharmaceutical and Medicinal Chemistry,
Heinrich-Heine-University Duesseldorf, Universitaetsstr. 1,
D-40225 Duesseldorf, Germany.

[3]Laboratory for Microbiology, Department of Biology,
Philipps-University Marburg, Karl-von-Frisch Str. 8,
D-35032 Marburg, Germany.

Thermoprotection of *A. fulgidus* by glycine betaine and proline betaine.

In addition to the well-established role of compatible solutes as osmostress protectants (*87-89*), recent experiments attribute a thermoprotective function to these compounds. Thermoprotection by glycine betaine (GB) has been observed both in bacteria (*90, 91*) and in plants (*92*). The hyperthermophilic archaeon *A. fulgidus* can grow in a temperature range from 64°C to 92°C, with an optimal growth temperature of 83°C; this growth behavior was investigated in media containing 0.1% yeast extract (*93*). Since yeast extract is known to contain GB (*94*), we wondered whether GB and proline betaine (PB) might serve as thermoprotectants for *A. fulgidus*. The presence of either 2 mM GB or 2 mM PB in a chemically defined medium lacking yeast extract did not have any influence on the growth of *A. fulgidus* at its optimal growth temperature of 83°C (Supplemental Fig. 1). However, this picture changed dramatically when we investigated the growth of *A. fulgidus* at 90°C. In the absence of a compatible solute, *A. fulgidus* was unable to grow at this elevated temperature. In contrast, growth was rescued to a large extent by the presence of either 2 mM GB or 2 mM PB (Supplemental Fig. 1). Consequently, both compatible solutes function as thermoprotectants for the hyperthermophilic archaeon *A. fulgidus*.

Growth at high temperature (90°C) had a noticeable effect on the cellular level of the *Af*ProX protein. In comparison with cells that were grown at 83°C, the amount of *Af*ProX was increased approximately four-fold in *A. fulgidus* cells that were propagated at 90°C .

Determination of the affinity constants of wildtype and mutants of *Af*ProX for glycine betaine and proline betaine.

Binding of GB and PB to purified *Af*ProX influenced the spectroscopic properties of *Af*ProX. Substrate binding of these compatible solutes leads to an increase of the intrinsic tryptophan fluorescence. We exploited this feature to determine the affinity constants (K_D) for both substrates. *Af*ProX binds GB and PB with high affinity, 60 ±10 and 50 ± 10 nM, respectively (Supplemental Table 1). According to the structure of *Af*ProX (62), it is evident that the substrate binding pocket is built up by four tyrosines (Tyr63, Tyr111, Tyr190, and Tyr214) as well as Lys13, Thr66, and Arg149. To further understand the residues' roles for substrate binding, they were mutated to alanine, and the substrate affinity was measured (Supplemental Table 1). Mutation of any of the four tyrosines reduced the affinity, ranging from 3.5 ± 0.7 µM for Y214A to 149 ± 17 µM for Y63A, respectively (Supplemental Table 1). Any one of the double Tyr to Ala mutants was not able to bind GB at all.

The K13A and T66A mutants also showed a lower affinity, 1.8± 0.2 µM and 107±20 µM, respectively (Supplemental Table 1). The largest decrease in affinity in the case of a single mutation was observed for the Arg149Ala mutation. Here, the K_D drops to 320± 60 µM for GB, and no binding could be observed for PB (Supplmental Table 1). This suggests a specific role of Arg149 in the substrate binding to *Af*ProX.

Crystallographic analysis and overlay

Crystals were grown as described in Material and Methods. A dataset of the *Af*ProX Tyr111Ala in complex with GB was collected at beamline ID-23 (ESRF, Grenoble, France) and processed using XDS (74). Initial phases were obtained by molecular replacement using the program PHASER (75) with the unliganded-open *Af*ProX structure as template (PDB entry: 1SW5) (62). The final structure was refined to a resolution of 2.0 Å. Data, refinement statistics, and model content are summarized in Supplemental Table 2.

Tyr111Ala *Af*ProX adopts the same conformation as the unliganded-open structure, as revealed by a perfect overlay (Figure 2). The RMSD between two open and a closed liganded structure was analyzed by the DYNDOME server (95). An overall RMSD value of 4.75 Å for all C-α atoms was calculated for the Tyr111Ala structure with respect to the closed-liganded structure (PDB code 1SW2) whereas no significant structural difference was found with respect to the unliganded-open structure (PDB code 1SW5) (RMSD < 0.2 Å).

Material and Methods

Microorganisms. The *Archaeoglobus fulgidus* strain VC16 (DSMZ 4304), whose genome sequence is known (96), was kindly provided by K. O. Stetter (University of Regensburg, Germany). *A. fulgidus* cells were grown anaerobically at 83°C or 90°C in sealed 100-ml serum bottles on a rotary shaker set at 120 rpm. We used the chemically defined medium described by Stetter et al. (93) with slight modifications for the propagation of *A. fulgidus*. The medium contained (per liter) 100 ml salt solution, 10 ml trace element solution, 5.5 g $NaHCO_3$, 2 mg $(NH_4)_2Fe(SO_4)_2$ and 10 mmol L-lactate as carbon and energy source. The salt solution contained (per liter) 74 g $MgSO_4 \cdot 7H_2O$, 3.4 g KCl, 27.5 g $MgCl_2 \cdot 6H_2O$, 1.4 g $CaCl_2 \cdot 2H_2O$, 1.4 g K_2HPO_4, 180 g NaCl, and 5 ml Resazurin (0.2% w/v). The trace element solution contained (per liter) 1.5 g Titriplex I, 3 g $MgSO_4 \cdot 7H_2O$, 0.01 g $NaWO_4$, 0.01g H_3BO_3, 0.01 g $Na_2MoO_4 \cdot 2H_2O$, 0.025 g $NiCl_2 \cdot 6H_2O$, 0.3 g $Na_2SeO_3 \cdot 5H_2O$, 0.45 g $MnSO_4 \cdot 1H_2O$, 1 g NaCl, 0.1 g $FeSO_4 \cdot 7H_2O$, 0.18 g $CoSO_4 \cdot 7H_2O$, 0.1 g $CaCl_2 \cdot 2H_2O$, 0.18 g $ZnSO_4 \cdot 7H_2O$, 0.01 g $CuSO_4 \cdot 5H_2O$, and 0.02 g $KAl(SO_4)_2 \cdot 12H_2O$. The pH was adjusted to 6.5 with KOH. We omitted the yeast extract from the originally described growth medium (93) because it is known to contain compatible solutes (e.g. GB or PB) (94); the presence of undefined concentrations of compatible solutes in the growth medium would interfere with the growth assays we performed with *A. fulgidus*. Instead, we used a vitamin solution that had the following composition per liter: 2 mg biotin, 2 mg folic acid, 10 mg pyridoxine-HCl, 5 mg thiamine-HCl·2 H_2O, 5 mg riboflavin, 5 mg nicotinic acid, 5 mg D-Ca-pantothenate, 0.1 mg vitamine B_{12}, 5 mg p-aminobenzoic acid, and 5 mg lipoic acid. The medium was adjusted to pH 6.8 with H_2SO_4, evacuated and gassed with N_2:CO_2 (80:20, v/v). After heat sterilization, the medium was reduced by the addition of 0.5 ml 0.5 M Na_2S and 0.5 ml $Na_2S_2O_4$. The osmolality of the medium was adjusted by the addition of solid NaCl to the indicated final concentrations. For growth experiments, the pre-warmed (83°C) medium was inoculated with 5 % of exponentially growing cells of an *A. fulgidus* pre-culture, and growth of the cells was monitored by determining the total cell number with a Helber counting chamber under a phase contrast microscope.

Chemicals. GB, choline, carnitine, proline, monomethylglycine, and dimethyglycine were purchased from Sigma (Deisenhofen, Germany). PB (stachydrine-HCl) was obtained from Extrasynthese (Genay, France). Ectoine was purchased from BIOMOL (Hamburg, Germany), and dimethylsulfoniumpropionate (DMSP) was a kind gift from T. Hansen (University of Groningen, the Netherlands). Radiolabeled [1-^{14}C] glycine betaine (55 mCi mmol^{-1}) was purchased from American Radiolabeled Chemicals Inc. (St. Louis, Mo., USA). Factor Xa was obtained from New England Biolabs (Schwalbach, Germany). Desthiobiotin was purchased from Sigma (Deisenhofen, Germany).

Methods used with nucleic acids and plasmid construction. To construct an expression plasmid carrying the *Af*ProX gene with an N-terminal Strep-Tag affinity peptide, we amplified the *proX* gene with the oligonucleotides 5´-AAAAAA<u>GGTCTC</u>AGCGCGGTTCCCAATCATCTGAGAG-3´ and 5´-AAAAAA<u>GGTCTC</u>ATATCATTTAACGAGTCCCTGCTCCT-3´ by PCR using chromosomal DNA of *A. fulgidus* VC16, kindly provided by S. Shima (Max-Planck Institute for Terrestrial Microbiology, Marburg, Germany). The PCR-product was digested with *BsaI* and cloned into the expression vector pASK-IBA6 (IBA, Göttingen, Germany). This construction removed the putative signal sequence of *Af*ProX and also mutated the N-terminal cysteine residue into a glycine residue. The sequence of the *A. fulgidus ProX* gene in pGH26 was verified by DNA sequence analysis.

Heterologous expression and purification of *Af*ProX. *Af*ProX was heterologously overexpressed using the *E. coli* strain BL21 Codon Plus RIL strain. A ten-liter glass flask containing five liters of MMA supplemented with 150 µg ml^{-1} of ampicillin, 30 µg ml^{-1} chloramphenicol, 0.5% glucose and 0.2% casaminoacids were inoculated to an OD$_{578}$ of 0.1 from an overnight culture. Cells were grown at 37°C with vigorous stirring until the culture had reached OD$_{578}$ = 0.6 - 0.7. Then expression was induced by the addition of anhydrotetracycline (final concentration 0.2 µg ml^{-1}). Cells were grown for an additional two hours and were subsequently harvested by centrifugation (10 min, 3000 x g). To release the periplasmic proteins the cell paste was resuspended in 50 ml ice cold buffer P [100 mM Tris-HCl (pH 8), 500 mM sucrose and 1 mM EDTA]. After a 30-minute incubation on ice, the periplasmic protein extract was harvested by centrifugation (15 min, 21000 x g at 4°C), followed by an ultra-centrifugation step (30 min, 120000 x g at 4°C) to remove insoluble material. The supernatant was then loaded onto a 10-ml Strep-tactin column (IBA, Göttingen, Germany) equilibrated with buffer W [100 mM Tris-HCl (pH 8)]. The column was washed with five column volumes of buffer W, and bound proteins were released from the affinity resin by washing the column with buffer E [100 mM Tris-HCl (pH 8), 2.5 mM desthiobiotin]. Since a portion of the pro-OmpA signal sequence was not proteolytically removed from the ProX protein by the *E. coli* cells, the purified ProX protein was cleaved with factor Xa (1 µg factor Xa per 200 µg ProX) for 16 hours at room temperature in a buffer

containing 100 mM Tris-HCl (pH 8), 100 mM NaCl, and 1 mM CaCl$_2$. This proteolytic cleavage removed the OmpA signal sequence and the Strep-Tag sequence, thereby resulting in *Af*ProX containing no N-terminal extensions. To remove factor Xa from the solution, the protein solution was diluted 1:2 with deionized water and loaded on a UnoQ6 column (BioRad, Muenchen, Germany) equilibrated with 50 mM Tris-HCl (pH 8) (buffer A). The column was washed with 20 ml buffer A, and the protein was eluted with a linear NaCl gradient. *Af*ProX eluted at 250 mM NaCl. *Af*ProX was subsequently dialyzed overnight against five liters of 10 mM Tris-HCl (pH 7.5) at 4°C and stored until further use at 4°C. In general, 2 mg of pure *Af*ProX protein were obtained per one liter of overproducing cells.

Substrate binding of purified *Af*ProX.

To study substrate binding properties of *Af*ProX, we used the ammonium sulfate precipitation technique of Richarme and Kepes (97). 5 µM of *Af*ProX were incubated at 85°C (or at other temperatures as indicated) with 5 µM [1-^{14}C] glycine betaine in a volume of 100 µl 10 mM Tris-HCl (pH 7.5) for 5 minutes. *Af*ProX was then precipitated by the addition of 900 µl of an ice-cold saturated ammonium sulfate solution. After a 5-minute incubation on ice, the mixture was sucked through cellulose filters (0.45 µM pore size, Scheicher and Schuell, Dassel, Germany), and the radioactivity retained by the filters was determined by liquid scintillation counting. To determine the substrate specificity of *Af*ProX, the protein was incubated at 85°C with 5 µM radiolabeled glycine betaine and a forty-fold excess of different unlabeled compatible solutes and was further treated as described.

Determination of the binding affinity (K$_D$) of *Af*ProX using fluorescence spectroscopy.

All spectra were obtained at room temperature with a VARIAN CARY Eclipse fluorometer (VARIAN, Darmstadt, Germany). Emission scans were collected at an excitation wavelength of 280 nm from 290 nm to 450 nm. Measurements were carried out in 1 ml 10 mM Tris-HCl (pH 7.5). 5 µg of *Af*ProX protein were added to the buffer, and the solution was mixed. To titrate the protein, 5 µl substrate was added from a 40 µM GB or PB stock solution. After substrate addition, the sample was mixed in the cuvette and the change in the fluorescence was

recorded. This step was repeated until a stable fluorescence signal was obtained. Assuming one binding site per monomer in *Af*ProX, the K_D for substrate binding was determined by nonlinear least-squares fitting of the data to the following equation: $F=F_0+(DF/2P_0)[(K_D+P_0+L_0)-((K_D+P_0+L_0)^2-4L_0P_0)^{1/2}]$. F is the measured fluorescence; F_0 the fluorescence of *Af*ProX; DF is the change of fluorescence at saturation; P_0 and L_0 are the total concentrations of protein and substrate.

Isothermal titration calorimetry. ITC was performed using an ITC200 calorimeter from Microcal Inc. (GeHealthcare) at 50°C. Protein concentration was 10μM in 20mM Tris pH 7.0, 20mM NaCl. The titration syringe was loaded with 50 μM glycine betaine prepared from the same solution as the protein. Titrations were usually carried out using 72 injections of 0.5 μl each at 160 s intervals. Stirring was set to 350 rpm. Data were corrected for heats of dilution by subtracting a blank titration, acquired by injecting an equal amount of ligand into buffer.

Substrate triggered domain closure in other SBPs

In the ribose binding protein (RBS) the substrate is bound by several amino acid of both domains (72)..Here again the superimposition of the open and the closed structure revealed a perfect overlay of the amino acids located in domain I. In domain II however the AA undergo a significant movement. Here a network was identified via the side chain of Gln235, which during closing undergoes a small conformational change of 1.5Å. The network can be described as follows: Gln235-Phe15-Phe16-Asn64-Arg90-Asp89. The Gln235 is located directly at the hinge between both domains which explains the small movement.

Siap is the binding protein of the N-Acetyl-5-neuraminic acid TRAP transporter from *H. influenza* (73). Here a specific role can be assigned to Arg127. Besides interacting with the substrate in the closed conformation a side chain interaction network can be observed via interactions with Pro149 and Asn187. Asn187 itself is in contact with Phe170 and Arg147, which all together build up the substrate-binding site in SIAP of domain II. Here as well as in MBP and RBS the distances in this network do not change in the open and closed form of the structures. In SIAP Arg127 has been mutated to alanine as well as Lysine. In elisa competitions

assays the function of SIAP has altered for the mutants. Furthermore no binding of substrate to the R127A mutant could not be observed anymore by using ITC, which is explainable by the proposed specific role of Arg127 in domain closure (*73, 98*).

Legends

Table 1: Mutational analysis of the binding site of *Af*ProX.

The affinity of wild type *Af*ProX and *Af*ProX mutants are listed for GB and PB. No binding for any of the double mutants was observed; hence, these mutants are not listed in the table. n.b. indicates that no binding was observed under this experimental setup.

Table 2: Crystallographic parameters.

Crystal parameters and data collection statistics are derived from XDS (99). Refinement statistics were obtained from REFMAC5 (76). Ramachandran analysis was performed using PROCHECK (100, 101). [a] R_{sym} is defined as $R_{sym} = \sum_{hkl} \sum_{i} \left| I_i(hkl) - \langle I(hkl) \rangle \right| \Big/ \sum_{hkl} \sum_{i} I_i(hkl)$ and [b] R_F as $R_{sym} = \sum_{hkl} \sum_{i} \left| I_i(hkl) - \langle I(hkl) \rangle \right| \Big/ \sum_{hkl} \sum_{i} I_i(hkl)$. R_{free} is calculated as R_F but for 5% randomly chosen reflections that were omitted from all refinement steps.

Table 3: Summary of the interactions of domain I and domain II in the closed form of *Af*ProX.

Suppl Figure 1: Influence of glycine betaine on the growth yield of *A. fulgidus*

A) *A. fulgidus* was grown in a defined minimal medium lacking yeast extract without (•) or with 2 mM GB (**o**) at different salt concentrations. Growth yields of the various cultures were monitored after 20 hours by counting the total cell number in a Helber counting chamber under a phase contrast microscope.

B) Pre-warmed media (83°C) lacking or containing 2 mM GB or 2 mM GB were inoculated with 5 % exponentially growing *A. fulgidus* cells and incubated on a rotary shaker at the indicated temperatures. Growth was monitored over time by counting the cell number with a Helber counting chamber under a phase contrast microscope. Squares: *A. fulgidus* cells grown in the absence of both GB and PB; diamonds: *A. fulgidus* cells grown in the presence of 2 mM GB; triangles: *A. fulgidus* cells grown in the presence of 2 mM PB.

Suppl Figure 2: Glycine betaine-binding to purified *Af*ProX measured with ITC.

ITC measurements were performed at 10 μM of purified *Af*ProX, and the integrated heat peaks were fit to a one site-binding model (black line).

Suppl Figure 3: Overall structure of *Af*ProX Tyr111Ala.

The overall structure of *Af*ProX is shown in white. The substrate glycine betaine (blue) and the bound HEPES molecule are shown in sticks representation.

Suppl Figure 4: Reorientation of the ligand within the first 50 ps of unrestrained MD simulation started from the open-liganded wild type *Af*ProX structure.

Orange: starting structure; blue: closed-liganded wt *Af*ProX structure; cyan: conformation after 50 ps.

Suppl Figure 5: Unrestrained MD simulation started from the open-liganded wild type *Af*ProX structure.

The rmsd of GB with respect to the starting structure is shown in green; the rmsd of domain II with respect to closed-liganded wt *Af*ProX structure is shown in red.

Table 1:

Mutant	K$_D$ GB [μM]	K$_D$ PB [μM]
Wild type	0.06 ± 0.01	0.05 ± 0.01
K13A	107 ± 20	101 ± 12
Y63A	149 ± 17	288 ± 26
T66A	1.8 ± 0.2	18 ± 1.4
Y111A	76 ± 4	148 ± 28
E145A	2.7 ± 0.5	23.2 ± 4
F146A	6.6 ± 0.6	4.0 ± 0.5
Y190A	67 ± 9	19 ± 5
R149A	320 ± 60	n.b.
Y214A	3.5 ± 0.7	3.5 ± 0.7
Double tyrosine mutations	n.b.	n.b.

Table 2:

Space group	P$_1$
a, b, c (Å) α, β, γ (°)	33.6, 36.9, 57.8; 83.8, 80.5, 95.9
Data collection and processing	
Wavelength (Å)	0.8726
Resolution (Å)	20 – 2.0
Mean redundancy	4.2 (4.3)
Unique reflections	17040
Completeness (%)	92.6 (93.6)
I/s	47.7 (29.0)
R$_{merge}$	2.1 (4.3)
Refinement	
R$_F$^{c} (%)	17.8
R$_{free}$^{d} (%)	24.4
Overall B factor from Wilson scaling (Å^2)	20.9
RMSD from ideal	
Bond lengths (Å)	0.023
Bond angles (°)	0.956
Average B factors (Å^2)	7.6
Ramachandran plot	
Most favored (%)	93.4
Allowed (%)	6.2
Disallowed (%)	0.4
Model content	
Monomers/ASU	1
Protein residues	6-275
Ligands	
Glycine Betaine	1
HEPES	1
H$_2$O	300

Table 3

| Amino acid	Atom	Interaction partner	Atom	Distance (Å)
Pro 172	Cδ	Pro 14	O1	3.4
	Cη		O1	3.3
	Cβ	Tyr 190	OH	3.6
	Cη		OH	3.6
Glu 155	NE1	Water-Pro76	CG	3.3
			Cα	3.3
Glu 145	Oε1	Thr 45	N	2.8
	Oε2	Thr 45	Cη	2.5
Asp 151	OD2	Tyr 214	CZ	2.6
Arg 149	NH1	Tyr 111	OH	3.5
	NH1	Betaine	O2	3.1
	NH1	Glu 145	OE2	3.0
	NH2	Tyr 111	OH	3.8
	NH2	Thr 66	OG1	2.8
	NH2	Thr 66	Cβ	3.5
	NH2	Thr 66	Cα	3.6
	CZ	Val 70	Cη2	3.6
	NE	Val 70	Cη2	3.3
	NE	Asp 151	OD1	2.8
	NH2	Tyr 111	OH	3.8

Suppl Figure 1:
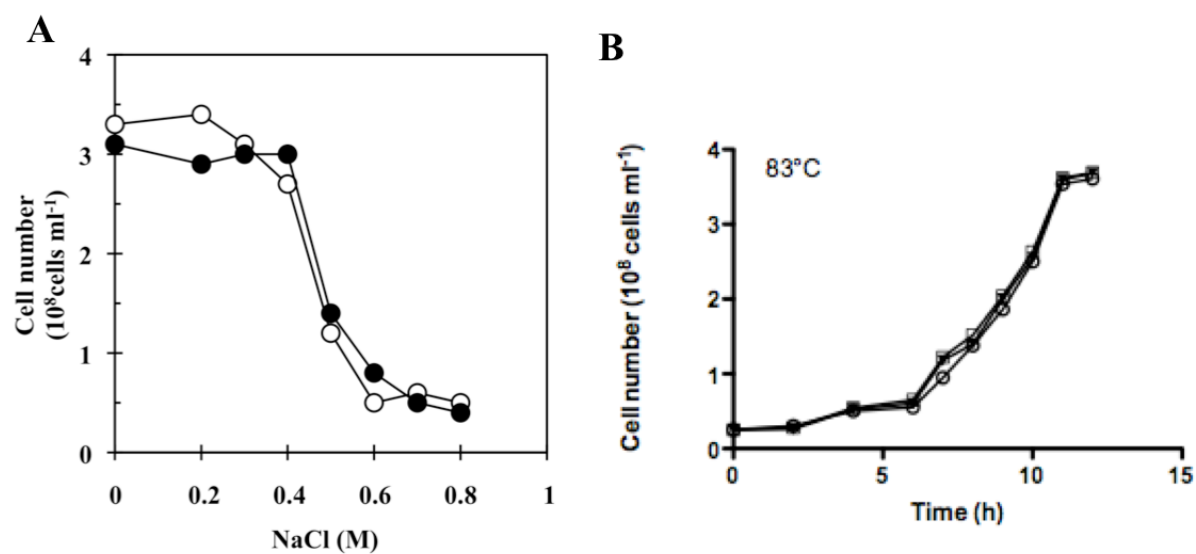
Suppl Figure 2:
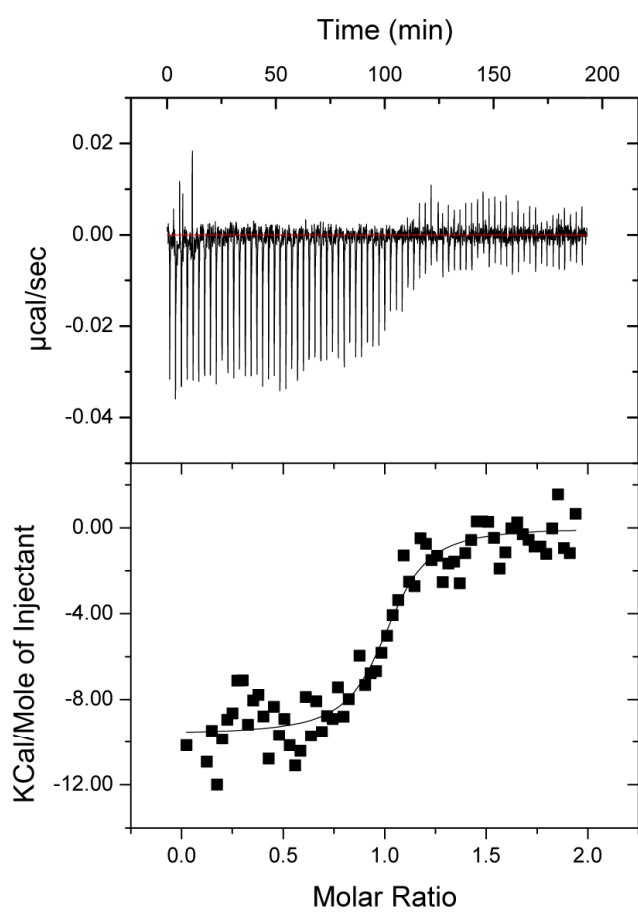
Suppl Figure 3:
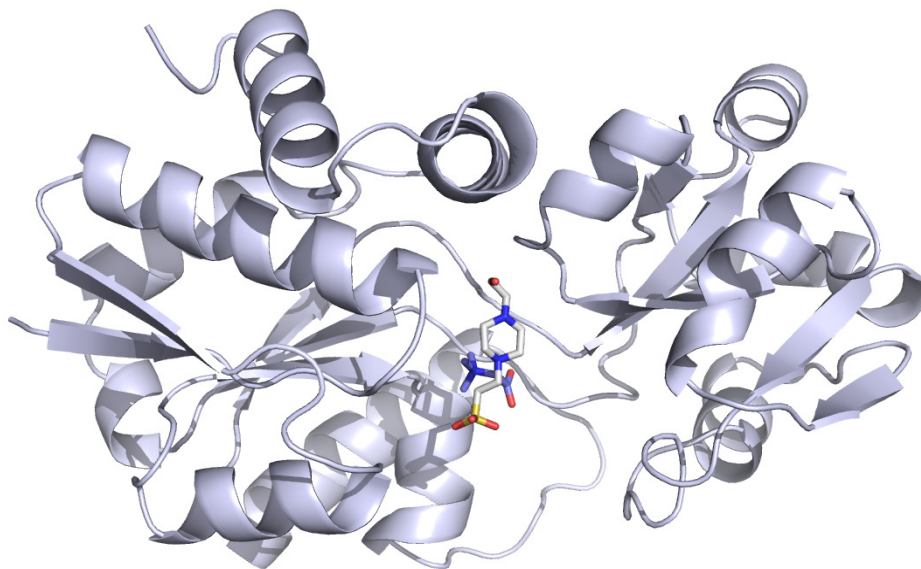
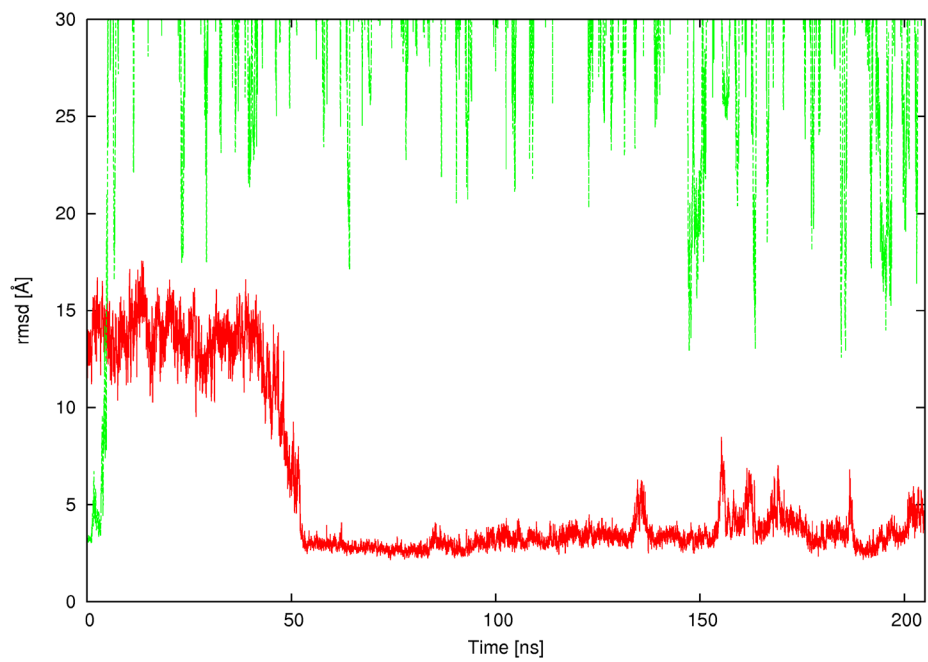**Suppl Figure 4:**
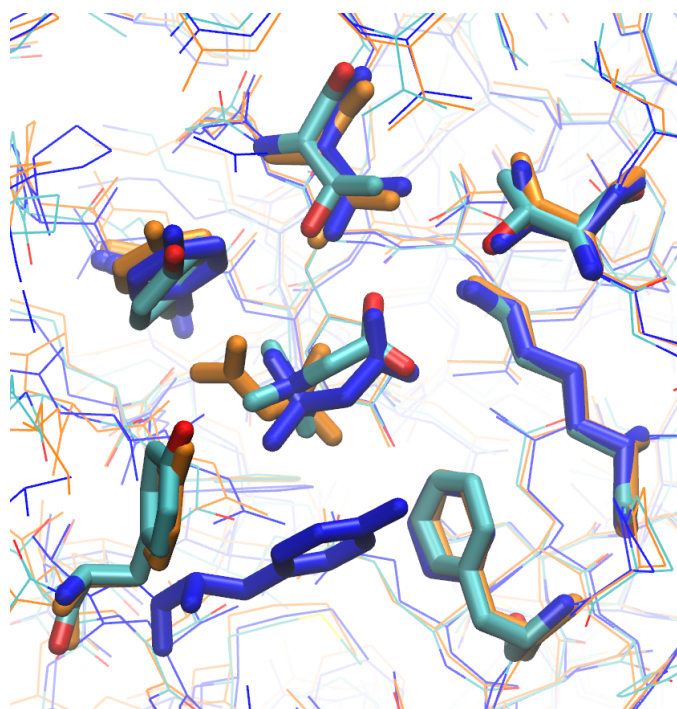
Suppl Figure 5:

References

1. Galinski, E.A. & Trüper, H.G. Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiology Reviews* **15**, 95-108 (1994).

2. Bremer, E. & Krämer, R. Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes in bacteria. in *Bacterial stress responses* (eds. Storz, G. & Hengge-Aronis, R.) 79-97 (ASM, Washington, D.C., 2000).

3. da Costa, M.S., Santos, H. & Galinski, E.A. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv Biochem Eng Biotechnol* **61**, 117-53 (1998).

4. Caldas, T., Demont-Caulet, N., Ghazi, A. & Richarme, G. Thermoprotection by glycine betaine and choline. *Microbiology* **145**, 2543-8. (1999).

5. Canovas, D., Fletcher, S.A., Hayashi, M. & Csonka, L.N. Role of trehalose in growth at high temperature of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **183**, 3365-71. (2001).

6. Alia, Hayashi, H., Sakamoto, A. & Murata, N. Enhancement of the tolerance of *Arabidopsis* to high temperatures by genetic engineering of the synthesis of glycinebetaine. *Plant J* **16**, 155-61 (1998).

7. Stetter, K.O., Lauerer, G., Thomm, M., Neuner, A. Isolation of extremely thermophilic sulfate reducers evidence for a novel brach of archaebacteria. *Science* **236**, 822-824 (1987).

8. Dulaney, E.L., Dulaney, D.D. & Rickes, E.L. Factors in yeast extract which relieve growth inhibition of bacteria in defined medium of high osmolarity. *Dev. Ind. Microbiol.* **9**, 260-269 (1968).

9. Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Biol Chem* **279**, 48270-81 (2004).

10. Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J Appl Cryst* **26**, 795-800 (1993).

11. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

12. Hayward, S. & Lee, R.A. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J Mol Graph Model* **21**, 181-3 (2002).

13. Klenk, H.P. et al. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364-70 (1997).

14. Richarme, G. & Kepes, A. Study of binding protein-ligand interaction by ammonium sulfate-assisted adsorption on cellulose esters filters. *BBA* **742**, 16-24 (1983).

15. Bjorkman, A.J. et al. Probing protein-protein interactions. The ribose-binding protein in bacterial transport and chemotaxis. *J Biol Chem* **269**, 30206-11 (1994).

16. Cuneo, M.J., Tian, Y.J., Allert, M. & Hellinga, H.W. The backbone structure of the thermophilic *Thermoanaerobacter tengcongensis* ribose binding protein is essentially identical to its mesophilic *E-coli* homolog. *Bmc Structural Biology* **8**, - (2008).

17. Muller, A. et al. Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. *J Biol Chem* **281**, 22212-22 (2006).

18. Johnston, J.W. et al. Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable Haemophilus influenzae strain 2019. *J Biol Chem* **283**, 855-65 (2008).

19. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. in *Meth. Enzymol.*, Vol. 276 (eds. Carter, C.W. & Sweet, R.M.) (Academic Press, London, 1997).

20. Murshudov, G., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallography D* **53**, 240-255 (1997).

21. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291 (1993).

22. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**, 449-57 (2005).

Own proportion to this work: 25%

Submitted to: Structure 2011

Impact factor: 5.9

Chapter IV

High-throughput evaluation of the critical micelle concentration of detergents

Thorsten Jumpertz [1,2], Britta Tschapek [1], Nacera Infed, Sander H.J. Smits, Robert Ernst [3], Lutz Schmitt *

Institute of Biochemistry, Membrane Transport Group, Heinrich Heine University Duesseldorf, 40225 Duesseldorf, Germany

ARTICLE INFO

ABSTRACT

Determination of the critical micelle concentration (CMC) value of detergents routinely used in biological applications is necessary to follow possible changes due to different buffer compositions (e.g., temperature, pH) such as those in solutions that are used for protein activity assays or crystallization. Here we report a method to determine the CMC values of detergents through a fast and robust assay that relies on the fluorescence of Hoechst 33342 using a 96-well plate reader. Furthermore, this assay provides the possibility and sensitivity to measure the CMC of detergent mixtures. The examples described here emphasize the potential and applicability of this assay and demonstrate that analysis of the physico-chemical parameters of detergents can now be investigated in virtually every laboratory.

© 2010 Elsevier Inc. All rights reserved.

Facing the fact that the "-omics" era has reached nearly every branch of natural, life, and medical sciences, new techniques often must prove their potential for handling an enormous number of samples or experimental conditions that need to be screened during the quest to find the right conditions for one-pot synthesis, drug–protein interactions, crystallizing proteins, and the like.

Genome projects estimate that roughly 30–40% of all proteins encoded in the human genome are membrane-embedded or membrane-associated proteins (MPs)[4] [1,2]. In addition, MPs account for approximately 70% of all drug targets [3]. Therefore, it is of great importance to describe MPs functionally and structurally. A major obstacle is to solubilize MPs from their native environment, the lipid bilayer, in a functional form. To overcome this intrinsic problem, a huge variety of biologically important detergents have emerged, and new ones are being constantly synthesized to match the ambitious needs of biochemists and medicinal pharmacologists [4]. The solubilization of MPs is not the only critical procedure, also subsequent steps

such as purification, dialysis, concentration, biochemical assays, and crystallization are performed in the continuous presence of detergents and numerous additives specific for the desired experiment [5–7]. These additional substances (e.g., salt, glycerol, buffer, metal ions, sugars) as well as physical parameters (e.g., pH, temperature) influence the critical micelle concentration (CMC) of detergents. Because detergents prevent MPs from aggregation, precise knowledge of the detergent concentration and its CMC at a given experimental condition may become crucially important. Most MPs lose their activity when the detergent concentration drops below the CMC, and because too high detergent concentrations can induce protein unfolding and disrupt protein–protein interactions, this can hinder protein activity and crystallization. Therefore, it can be crucial to minimize the detergent concentration, which requires detailed knowledge of how the experimental conditions affect the detergent's CMC.

Different methods to determine the CMC and concentrations of detergents already exist, including fluorescence techniques [8], infrared spectroscopy [9], thin layer chromatography [10], colorimetric assays [11], and refractive index measurements [12]. However, most of these experimental setups are either not accessible to laboratories working with MPs or not very well suited for higher throughput screening if one considers the amount of material that is needed.

Therefore, we developed an easy-to-use, robust, and fast assay for the determination of the CMC value of detergents routinely used in biological applications that makes it possible to determine the CMC of different buffer compositions such as those in solutions that are used for protein crystallization. The assay is based on the dye Hoechst 33342, which displays literally no fluorescence in hydrophilic environments but emits efficiently in hydrophobic environments.

---

* Corresponding author. Fax: +49 211 8115310.
E-mail address: Lutz.Schmitt@uni-duesseldorf.de (L. Schmitt).
[1] Both of these authors contributed equally to this work.
[2] Present address: Institute of Neuropathology, Molecular Neuropathology Group, Medical School Duesseldorf, 40225 Duesseldorf, Germany.
[3] Present address: Max Planck Institute for Molecular Cell Biology and Genetics, 01307 Dresden, Germany.

[4] *Abbreviations used:* MP, membrane protein; CMC, critical micelle concentration; DDM, dodecyl maltoside; FC-14, 1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1′-rac-glycerol); FC-16, 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phospho-(1′-rac-glycerol); ddH$_2$O, double-distilled H$_2$O; PFO, perfluorooctanoic acid; MALS, multiple angle light scattering; UV, ultraviolet; DM, decyl maltoside; β-OG, β-octylglucoside; ELISA, enzyme-linked immunosorbent assay; C$_{12}$E$_8$, polyoxyethylene(8)dodecyl ether; ANS, 8-anilinonaphthalene-1-sulfonate; PEG, polyethylene glycol; MPD, 2-methyl-1,3-propanediol; MWCO, molecular weight cutoff; NMR, nuclear magnetic resonance; SAXS/NS, small angle X-ray/neutron scattering.

Materials and methods

Materials: Chemicals and detergents

All chemicals were purchased from Sigma–Aldrich at the highest purity available. Dodecyl maltoside (DDM) was obtained from Glycon Biochemicals (Luckenwalde, Germany). FosCholine 14 [FC-14: 1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1′-rac-glycerol)] and FosCholine 16 [FC-16: 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phospho-(1′-rac-glycerol)] were purchased from SynphaBase (Pratteln, Switzerland). All other detergents used in this study were obtained from Anatrace (Maumee, OH, USA).

Methods

Fluorescence spectroscopy

To determine the CMC, various detergents were titrated into a fluorescence cuvette from a stock solution. The cuvette contained either double-distilled H$_2$O (ddH$_2$O, Milli-Q grade, 0.45 μm filter sterilized) or buffer (50 mM Hepes [pH 7.2] or 50 mM Na–phosphate supplemented with 50 mM NaCl and 5% [w/v] glycerol) supplemented with 7 μM Hoechst 33342. The detergent stock solution also contained 7 μM Hoechst 33342 to circumvent dilution of the fluorescent dye. The resulting fluorescence was measured on a Horiba–Jobin–Yvon Fluorolog 3 spectrophotometer with λ_{ex} = 355 nm and λ_{em} = 457 nm (λ_{em} = 487 nm for perfluorooctanoic acid [PFO]), and the slit widths were 2 and 1 nm, respectively. Data points were recorded until a mean standard deviation below 0.5% was reached. To transfer the method of CMC determination via Hoechst 33342 to a less laborious and time-consuming setup, we used 96-well plates suitable for fluorescence spectroscopy (Greiner Bio-One, FIA plate, black, flat bottom, medium binding), pipetted up to 80 different concentrations for each detergent (this step can also be automated by a pipetting robot) into 96-well plates, and recorded emission spectra (filters used: λ_{ex} = 355 ± 10 nm, λ_{em} = 460 ± 80 nm, BMG Labtech, Offenburg, Germany). Data recorded from the Fluorolog 3 and plate reader gave virtually identical results (Table 1). Fluorescence data were background corrected and analyzed according to Eq. (1) using Prism (version 5, GraphPad, San Diego, CA, USA):

$$F = \frac{F_{max}([Det] - CMC)}{(K_{0.5} - CMC) + ([Det] - CMC)} \tag{1}$$

where F is the measured fluorescence, F_{max} is the maximal fluorescence, [Det] is the absolute detergent concentration, $K_{0.5}$ is the midpoint of the fit function, and CMC is the critical micelle concentration of the detergent, which is determined by the fit.

dn/dc Measurements

We determined the dn/dc values for certain detergents (Triton X-100, DDM, and PFO) and compared these values with data found in the literature (see, e.g., http://www.affimetrix.com/estore). In brief, 1 ml of a stock solution of 0.25 g/ml Triton X-100, 1 g/ml DDM, and 0.25 g/ml PFO was injected in appropriate dilutions via the direct injection mode into an optiLAB rex (Wyatt Technologies, Santa Barbara, CA, USA). Refractive index measurements were used to calculate dn/dc values from peak area versus sample concentration at fixed sample volume (optiLAB rex). The dn/dc values were used to calculate micelle size/mass from subsequent light scattering experiments.

Multiple angle light scattering

To measure micelle size, we used a multiple angle light scattering (MALS) setup miniDAWN Treos/optiLAB rex (Wyatt Technologies) connected to an Äkta Purifier (GE Healthcare, Freiburg; Germany) using either a Superose 6 or a Superdex 200 10/300 analytical size exclusion column. Here 200 μl of a detergent solution (5% [w/v] DDM, 2% [w/v] FC-16, 5% [w/v] Triton X-100, and 30% [w/v] PFO) was applied onto the column. The flow rate was 0.2 ml/min. Ultraviolet (UV) detection was at 280 nm, light scattering was detected at angles of 0°, 90°, and 107°, and the obtained values were averaged. MALS data were evaluated using the program ASTRA (Wyatt Technologies). For PFO, the micelle size could not be determined reliably. Due to its surface properties inherent to perfluorinated substances, PFO interacted intensively with the column material. Using an asymmetric flow field–flow fractionation setup (Eclipse AF4, Wyatt Technologies) connected to a MALS detector, no conclusive results could be produced because the detergent seemed to interact with the membrane and the Teflon fittings.

Measurements of detergent mixtures

To measure the CMC values of detergent mixtures, we applied the same methodology used for the determination of CMC values of single detergents. In the examples presented here, the wells of a 96-well plate were preloaded with FC-16, decyl maltoside (DM), and β-octylglucoside (β-OG) at half the CMC and then were titrated with either DDM, DM, or β-OG, and the fluorescence of Hoechst 33342 was measured as described above.

Results

Determination of the CMC and comparison of CMC value derived from plate reader and Fluorolog measurements

In ddH$_2$O, aqueous buffers, and buffers containing detergents below the CMC, no (or only marginal) Hoechst 33342 fluorescence can be detected. On constant increase of the concentration of detergent, the CMC value will be reached and micelles start to form. As soon as this point is passed, micelles begin to incorporate Hoechst 33342, which can be monitored by an increase of the fluorescence signal (shown schematically in Fig. 1). This point represents the CMC of the detergent (see Ref. [13] and references cited therein). The assay described here is optimized for 96-well fluorescence plates, and data can be evaluated with a plate reader (i.e., enzyme-linked immunosorbent assay [ELISA] reader) available in nearly every laboratory working with membrane proteins.

Formation of micelles was followed by the emitted fluorescence of Hoechst 33342 (emission wavelength 457 ± 10 or 487 ± 10 nm [only in the case of PFO] and 460 ± 40 nm for the Fluorolog 3 and plate reader setups, respectively). We obtained the CMC values for the detergents shown in Table 1 in 50 mM Hepes (pH 7.2), 50 mM NaCl, and 5% (w/v) glycerol except for PFO. This detergent precipitated in Hepes buffer and was dissolved in NaP$_i$ or ddH$_2$O. The higher sensitivity of the Fluorolog 3 setup allowed us to thoroughly compare the data with data from the plate reader. Table 1 and Fig. 2 clearly indicate that both approaches yielded highly similar CMC values. Therefore, we propose that the 96-well plate reader setup is perfectly suited to accurately determine the CMC value of detergents.

At saturating concentrations, different detergents result in different fluorescent signals (e.g., Triton X-100 and DDM in Fig. 2). This difference most likely results from the different environments that the dye encounters in the core of the micelles and could be used to characterize detergent micelles even further.

Table 1
Comparison of CMC values of different detergents obtained by the Hoechst 33342 or ANS assay and literature values.

Detergent	CMC value Hoechst 33342 (%, w/v) (Plate reader)	CMC value Hoechst 33342 (%, w/v) (Fluorolog)	CMC value ANS (%, w/v) (Plate reader)	CMC value (%, w/v) (Literature)	Aggregation number (Literature)	Aggregation number (MALS)
$C_{12}E_8$	n.d.	0.0039 ± 0.0001	0.0044 ± 0.0003	0.0048	123	87
CyMal7	n.d.	0.011 ± 0.002	n.d.	0.0099	150	n.d.
DM	0.093 ± 0.002	0.092 ± 0.001	0.127 ± 0.013	0.087	69	n.d.
DDM	0.0081 ± 0.0002	0.0083 ± 0.0002	0.013 ± 0.002	0.0087	78–149	148
FC-14	0.0055 ± 0.0001	0.0067 ± 0.0002	n.d.	0.0046	108	n.d.
FC-16	0.00054 ± 0.00005	0.00055 ± 0.00003	0.00135 ± 0.00019	0.00053	178	196
PFO	0.80 ± 0.02	0.81 ± 0.02	n.d.	1.18	106–179	n.d.
	0.20 ± 0.01[a]					
Triton X-100	0.018 ± 0.001	0.015 ± 0.0001	n.d.	0.010–0.016	75–165	136

Note. Literature values were taken from http://www.affimetrix.com/estore. Errors reported are standard deviations determined by Eq. (1). n.d., not determined; CyMal7, 7-cyclohexyl-1-heptyl-β-D-maltoside.
   [a] Measured in ddH_2O.

Salt dependency of CMC

It is known that ionic strength can influence the CMC of certain detergents (see, e.g., Ref. [14]). Therefore, we performed experiments to address the question of whether or not our assay can be used to analyze the influence of ionic strength on the CMC value. For DDM, a detergent routinely used for MP purification and crystallization, we showed that micelle formation depends on salt concentration (Fig. 3A). In addition, a great impact on the CMC value in the presence of buffer/salt is shown for PFO (0.8% in ddH_2O vs. 0.2% in buffer). Increasing the NaCl concentration from 0 to 1 M decreased the CMC of DDM nearly fourfold (Fig. 3A). The largest change (twofold) was observed from 0 to 200 mM NaCl, whereas a nearly linear decrease of the CMC value was detected in the case of further increases up to 1 M NaCl (overall twofold change). For PFO, we compared the change of the CMC between ddH_2O and a buffer (50 mM NaP_i, 50 mM NaCl, and 5% [w/v] glycerol) (Fig. 3B). Again, a nearly fourfold reduction of the CMC was observed. The slope of the linear regression of the CMC versus NaCl concentration reflects changes in the strength of self-interaction of the detergent molecules [15]. For the anionic detergent PFO, the observed slope is steeper than that observed for DDM, a nonionic detergent. This indicates that the self-association of PFO depends stronger on changes of the ionic strength than that of DDM. In addition, the CMC of DDM was not strictly linearly dependent on the NaCl concentrations. This is likely due to different

interactions of the hydrophilic and hydrophobic parts of the detergent in the absence or presence of salt. However, a more detailed analysis would be required to determine the exact nature of this deviation. Nevertheless, these results demonstrate the capability of our assay to easily analyze the influence of certain physicochemical parameters on the CMC of the detergent.

dn/dc values and MALS

We determined the dn/dc values (data not shown) and micelle sizes for the detergents shown in Fig. 2 for cross-validation of our results. From light scattering experiments, the molecular mass of detergent micelles was calculated to be 88 kDa for Triton X-100, 76 kDa for DDM, 80 kDa for FC-16, and 47.4 kDa for polyoxyethylene(8)dodecyl ether ($C_{12}E_8$) (Fig. 4). These numbers are in good agreement with literature values (Table 1).

Mixed micelles

An interesting observation is that crystal quality of MP can improve dramatically when a second detergent is added. A recent article nicely described the purification and crystallization strategies of the Na^+–H^+-antiporter NhaA from *Escherichia coli,* and the authors assumed that the mixture of detergents used improved crystal contacts, allowing a subsequent structure determination [16]. Employing our assay, we showed that detergent mixtures have different CMCs (Fig. 5). For example, when titrating DDM to a solution containing FC-16 at half CMC, micelles formed at a concentration of DDM below the CMC of 'pure' DDM. This is indicative of the formation of mixed micelles [17–19]. In addition, we observed that identical concentrations of FC-16 resulted in higher fluorescence in the presence of DDM, likely due to the formation of a higher number of (mixed) micelles. A similar behavior could be observed for a mixture of detergents with a rather detergents with large CMC value, such as DM and β-OG, or a mixture of detergents with large and small CMCs (DDM and DM [data not shown]).

After solubilization of a MP from a lipid bilayer, detergent micelles mimic the hydrophobic environment of its transmembrane region. Because the transmembrane regions of MPs from different organelles/species can differ greatly in their hydrophobic length [20], the simultaneous use of two detergents with different aliphatic chains is more likely to overcome the problems due to a hydrophobic mismatch such as the lack of crystallization [16] or loss of protein activity. This highlights the importance of a systematic analysis of detergent mixtures.
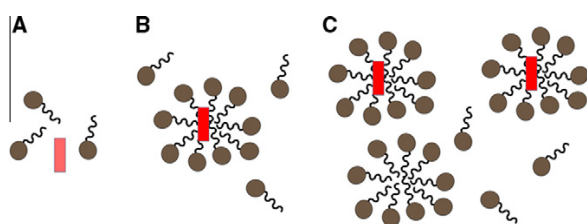

Fig. 1. Model of micelle formation and the Hoechst 33342-dependent assay. Below the CMC, only detergent monomers are present in solution and no fluorescence of Hoechst 33342 (pale red square) is measured (A). When the detergent concentration increased and reaches the CMC, micelles form and provide a hydrophobic environment for Hoechst 33342, resulting in its fluorescence (deep red square). This point can easily be determined using the assay presented in this study (B). Increasing the detergent concentration will increase the concentration of micelles while the concentration of detergent monomers remains constant. Under saturating conditions, newly formed micelles are not saturated with Hoechst 33342 anymore, and no further increase in fluorescence can be detected (C). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
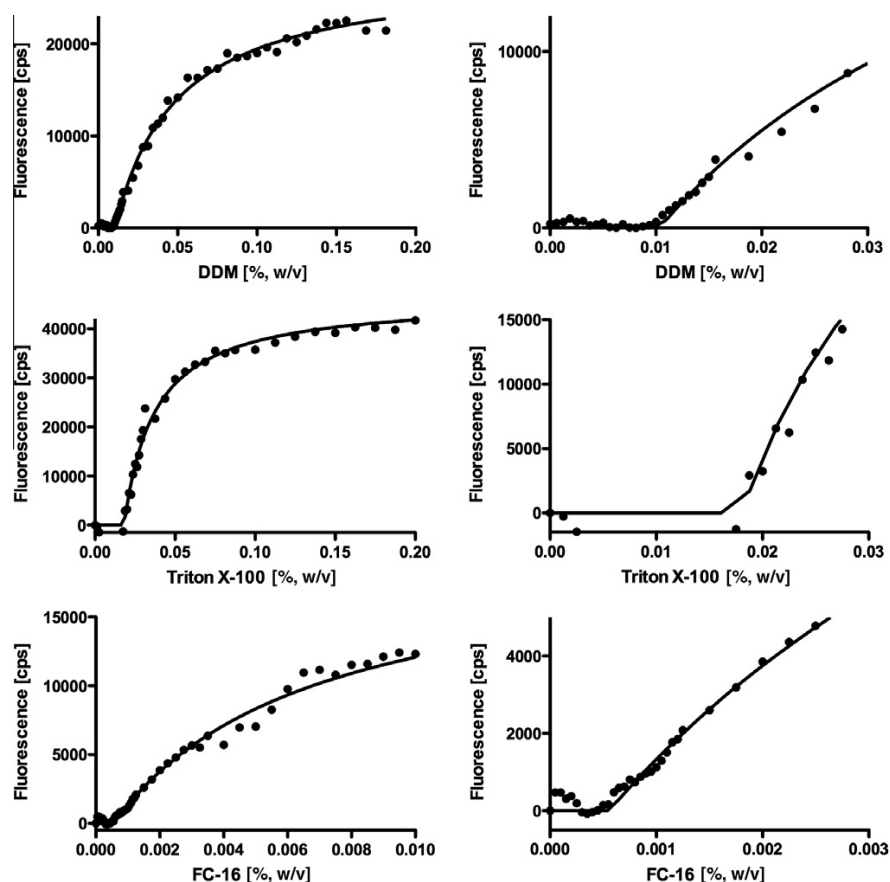
Fig. 2. CMC determination of DDM, Triton X-100, and FC-16 via a 96-well plate reader setup (left panels). Data were analyzed according to Eq. (1) (see Materials and methods). The determined CMC values for these and further detergents are listed in Table 1. The difference in maximum fluorescence between for example DDM and FC-16 is likely due to local environmental effects on the Hoechst 33342 fluorescence such as local pH, water penetration, and hydrophobicity of the detergent. A zoom-in of the CMC region of the detergents is shown in the right panels.

ANS measurements

We also performed the experiments using 8-anilinonaphthalene-1-sulfonate (ANS). ANS has been used extensively as a reporter dye to examine the CMC (see, e.g., Ref. [8]). Our results obtained by ANS measurements vary in the range of 1.3–2.0 times the literature values for the CMC with the exception of $C_{12}E_8$, which fit nicely to the literature values (Table 1). We stress that our analysis with ANS is not sufficient to provide an explanation for this behavior. However, the data clearly demonstrate that the assay is not restricted to Hoechst 33342, and if needed other dyes specifically adapted to particular detergent characteristics (e.g., pH, redox sensitivity) might be employed.

Influence of major precipitants on CMC

The crystallization of proteins depends strongly on precipitant solutions that, in most cases, are directly mixed with the protein/detergent solution. Therefore, we also examined the influence of some precipitants routinely used to screen for crystallization conditions (e.g., polyethylene glycol [PEG], 2-methyl-1,3-propanediol [MPD]) on the CMC with the Hoechst 33342 assay. The CMC value increases in the presence of precipitants (Table 2). At higher concentrations of precipitants, the CMC can no longer be determined because the fluorescence values seem to

scatter around a midpoint. In addition, the maximum fluorescence is decreased in comparison with a buffer–detergent mixture that contains no precipitants.

Protein–detergent mixtures

From the fluorescence intensity, it is also possible to determine the detergent concentration of a certain sample so long as the buffer composition does not differ from the control measurement (see experiments shown in Fig. 2). One issue during the purification of MPs is that filtration techniques that concentrate the MPs can also result in a potentially detrimental increase in detergent concentration. We purified a MP via affinity purification (details of expression and purification of the MP will be provided elsewhere). Prior to the measurements, the MP was further purified by size exclusion chromatography in a buffer containing 0.02% DDM. This detergent–MP mixture was then split into three identical aliquots. The first aliquot was mixed with 7 µM Hoechst 33342, and the resulting fluorescence was directly determined (Fig. 6, left bar). The other two aliquots were concentrated by a factor of 2 using either a 30-kDa or 100-kDa molecular weight cutoff (MWCO) membrane and subsequently diluted back to their original volume with the original buffer and, hence, back to the initial MP concentration. Then 7 µM Hoechst 33342 was added to both samples, and the corresponding fluorescence values were
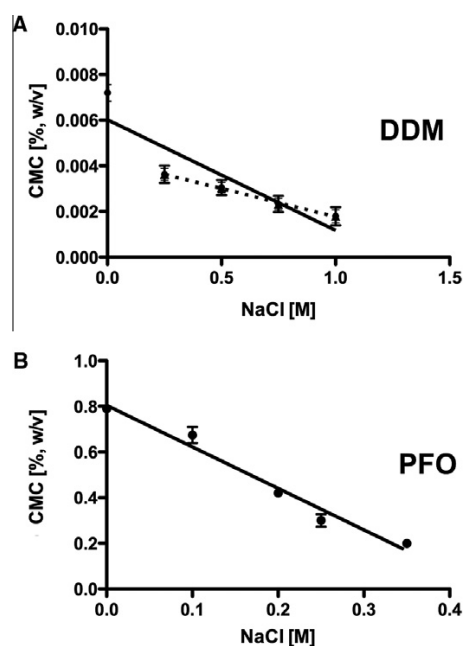
Fig. 3. Salt dependency of the CMC of DDM (A) and PFO (B). Decreasing concentrations of DDM or PFO are required for the formation of micelles if the ionic strength of the solution is increased. Straight lines show the linear regression of the data. The dependence of the CMC of PFO in dependence of ionic strength shows a clear linear behavior (black line), whereas a discrepancy is visible in DDM (black line). However, for ionic strengths nonequal to zero, a linear dependency is present (dashed line).

measured (Fig. 6, middle and right bars). An increase of fluorescence as compared with the original mixture would indicate that the detergent-to-protein ratio increased upon concentration; in other words, more detergent molecules are present per MP molecule in the concentrated sample. The back-diluted concentrate from the 30-kDa MWCO membrane exhibited a significantly higher fluorescence than the back-diluted 100-kDa concentrate or the nonconcentrated sample. This indicates that even empty micelles are concentrated on a 30-kDa MWCO concentrator. Thus, our assay can be used to choose suitable concentrators and, thereby, to avoid a potentially deleterious increase of detergent concentration on protein concentration.

Discussion

We have developed an easy-to-use and robust assay to determine the CMC of detergents by Hoechst 33342 fluorescence (Fig. 2). Different kinds of influences on the CMC of detergents can be investigated without the need for highly sophisticated experimental setups. The behavior of the detergents under various conditions, including increasing or decreasing salt concentrations (Fig. 3), high or low pH values, temperature, and additives such as precipitants (Table 2), ligands, and cofactors, can be easily followed and may give valuable hints for adjusting the experimental conditions.

Highly sophisticated and reliable approaches, such as nuclear magnetic resonance (NMR), small angle X-ray/neutron scattering (SAXS/NS), and contact angle measurements, will result in similar values to those with the method described here. But most of the time, relative changes of the CMC due to changes in temperature, buffer composition, or additives are the focus. Our method provides a fast and robust way to analyze such changes.
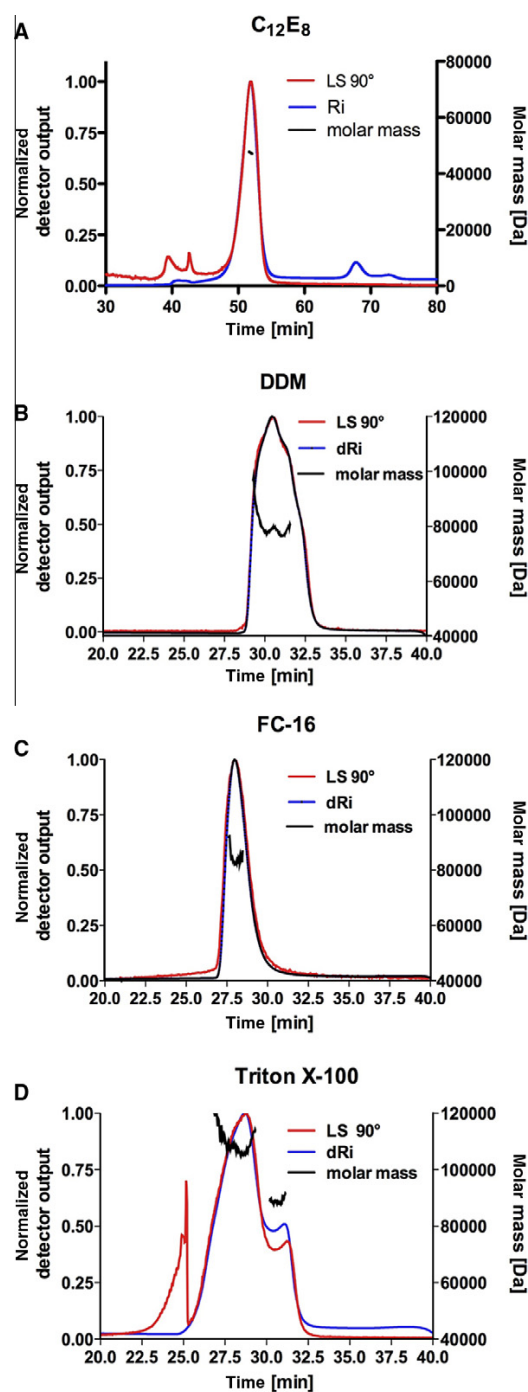

Fig. 4. MALS combined with size exclusion chromatography for $C_{12}E_8$ (A), DDM (B), FC-16 (C), and Triton X-100 (D) solutions. To determine the molecular masses of the corresponding detergent micelles, dn/dc values were determined by refractive index measurements (see Materials and Methods for further details).

We also investigated the potential application of the assay on MP preparations that helps to choose suitable protein concentrators and to avoid a potentially harmful increase of detergent concentration. It must be stressed that the MP by itself very likely affects the fluorescence of Hoechst 33342; therefore, an absolute
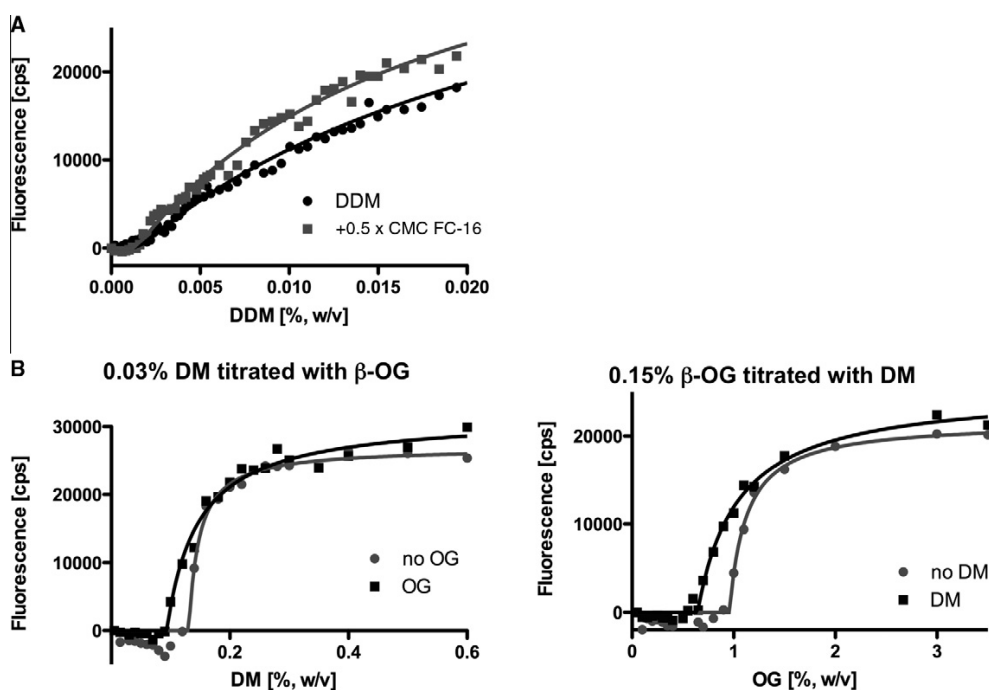
Fig. 5. Analysis of detergent mixtures. (A) FC-16 was supplied at a concentration of $0.5 \times$ CMC, and DDM was added stepwise (gray squares). The resulting fluorescence curve and the CMC determined from the titration experiments showed small but significant changes compared with that of pure DDM (black squares). (B) When using detergents with a higher CMC (DM and β-OG), mixed micelles were also observed. The CMC value of pure DM (0.087%) is shifted toward larger values in the presence of β-OG (left panel). The same behavior was observed for the inverse experiment (right panel).

Table 2
Influence of precipitants on CMC values of β-OG and DDM.

Detergent	Precipitant	CMC value (%, w/v) (Plate reader)
DDM	Buffer	0.0078
DDM	Jeffamine–M600 (10%)	0.0089
DDM	MPD (5%)	0.0087
DDM	MPD (10%)	0.0055
DDM	PEG400 (10%)	0.0082
DDM	PEG400 (20%)	n.d.
DDM	PEG8000 (7%)	0.0099
DDM	PEG8000 (14%)	n.d.

Note. n.d., not determined (for high concentrations of high-molecular-weight precipitants such as PEG4000, a determination of the CMC was not possible due to random scattering of the obtained fluorescence values).

Fig. 6. Analysis of relative changes in the detergent-to-protein ratio in MP preparations (here shown for DDM). Left bar: Hoechst 33342 fluorescence at an MP concentration of 0.5 mg/ml in 0.02% DDM; middle bar: Hoechst 33342 fluorescence after the preparation was concentrated twofold with a spin concentrator using a 30-kDa MWCO membrane and subsequently diluted twofold with buffer containing 0.02% DDM to the original MP concentration (0.5 mg/ml); right bar: Hoechst 33342 fluorescence after the preparation was concentrated twofold with a spin concentrator using a 100-kDa MWCO membrane and subsequently diluted twofold with buffer containing 0.02% DDM to the original MP concentration (0.5 mg/ml). Error bars represent the standard deviations of five independent measurements.

quantification of the detergent concentration in the presence of an MP is not possible. The Hoechst 33342 fluorescence of a detergent–MP mixture arises from three different sources that cannot be separated: binding to empty detergent micelles, to MP–detergent micelles, and directly to the MP. However, changes in detergent concentration are detectable when the concentration of the MP is kept constant (Fig. 6).

In summary, the described method has proven to be a valuable tool for addressing questions arising in all areas of research dealing with detergents. Transfer into a 96-well plate format and the striking simplicity make this method suitable to follow (bio)physical questions that were not easy to address before. We are convinced that on the basis of this assay, further developments will be achieved to investigate different parameters of detergent/detergent–protein complexes and will add an auxiliary biochemical "tool" for membrane protein research.
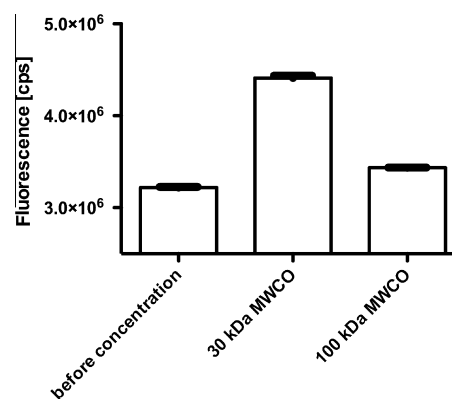
References

[1] E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (1998) 1029–1038.
[2] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.

[3] J. Drews, Drug discovery: a historical perspective, Science 287 (2000) 1960–1964.

[4] R.M. Garavito, S. Ferguson-Miller, Detergents as tools in membrane biochemistry, J. Biol. Chem. 276 (2001) 32403–32406.

[5] M.C. Wiener, A pedestrian guide to membrane protein crystallization, Methods 34 (2004) 364–372.

[6] L.E. Fisher, D.M. Engelman, J.N. Sturgis, Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain, J. Mol. Biol. 293 (1999) 639–651.

[7] L.E. Fisher, D.M. Engelman, J.N. Sturgis, Effect of detergents on the association of the glycophorin A transmembrane helix, Biophys. J. 85 (2003) 3097–3105.

[8] E.B. Abuin, E.A. Lissi, A. Aspee, F.D. Gonzalez, J.M. Varas, Fluorescence of 8-anilinonaphthalene-1-sulfonate and properties of sodium dodecyl sulfate micelles in water–urea mixtures, J. Colloid Interface Sci. 186 (1997) 332–338.

[9] C.J. daCosta, J.E. Baenziger, A rapid method for assessing lipid:protein and detergent:protein ratios in membrane–protein crystallization, Acta Crystallogr. D 59 (2003) 77–83.

[10] L.R. Eriks, J.A. Mayor, R.S. Kaplan, A strategy for identification and quantification of detergents frequently used in the purification of membrane proteins, Anal. Biochem. 323 (2003) 234–241.

[11] A. Urbani, T. Warne, A colorimetric determination for glycosidic and bile salt-based detergents: applications in membrane protein research, Anal. Biochem. 336 (2005) 117–124.

[12] P. Strop, A.T. Brunger, Refractive index-based determination of detergent concentration and its application to the study of membrane proteins, Protein Sci. 14 (2005) 2207–2211.

[13] A. Helenius, K. Simons, Solubilization of membranes by detergents, Biochim. Biophys. Acta 415 (1975) 29–79.

[14] A. Chattopadhyay, K.G. Harikumar, Dependence of critical micelle concentration of a zwitterionic detergent on ionic strength: implications in receptor solubilization, FEBS Lett. 391 (1996) 199–202.

[15] P. Mukerjee, Salt effects on nonionic association colloids, J. Phys. Chem. 69 (1965) 4038–4040.

[16] E. Screpanti, E. Padan, A. Rimon, H. Michel, C. Hunte, Crucial steps in the structure determination of the Na^+/H^+ antiporter NhaA in its native conformation, J. Mol. Biol. 362 (2006) 192–202.

[17] P. Sehgal, J.E. Mogensen, D.E. Otzen, Using micellar mole fractions to assess membrane protein stability in mixed micelles, Biochim. Biophys. Acta 1716 (2005) 59–68.

[18] R. Nagarajan, Molecular theory for mixed micelles, Langmuir (1985) 331–341.

[19] E.A.G. Aniansson, S.N. Wall, On the kinetics of step-wise micelle association, J. Phys. Chem. 78 (1974) 1024–1030.

[20] H.J. Sharpe, T.J. Stevens, S. Munro, A comprehensive comparison of transmembrane domains reveals organelle-specific properties, Cell 142 (2010) 158–169.

Own proportion to this work: 35%

Published in: Analytical Biochemistry 2011

Impact factor: 3.29

Chapter V

High Throughput analysis of mixed detergent micelles

Britta Tschapek[1], Sander H. J. Smits[1] and Lutz Schmitt[1]*

[1]Institute of Biochemistry, Heinrich Heine University Duesseldorf, Universitaetsstrasse 1, 40225 Duesseldorf, Germany

*Corresponding author. Phone: +49 211 8110773, Fax: +49 211 8115310, E-mail address: Lutz.Schmitt@hhu.de

Abstract

A simple, reliable and fast method for the analysis of detergent containing buffer systems is presented. It enables the determination of CMC values and aggregation numbers of detergents as well as detergent mixtures in 96well plate standard device relying on steady state fluorescence and fluorescence quenching. To prove the universal applicability we demonstrate DDM and mixtures at different ratios with six other detergents.

Abbreviations

CMC: critical micellar concentration; DDM: Dodecyl-maltoside; DM: decyl-maltoside; FC-14: 1-myristoyl-2-hydroxy-sn-glycero-3-phospho-(1'-rac-glycerol); MP: membrane protein; OG: octylglucoside; LDAO: lauryldimethylamine oxide; SDS: Sodium dodecyl-sulfate; SAXS: small angle X-ray scattering; SANS: small angle neutron scattering

Introduction

Analyzing membrane proteins (MP) *in vitro* requires usage of detergents. Initially, MPs are solubilized from cellular membranes and further purified until biochemical, biophysical or even structural studies can be performed. Every of these steps requires a stable and homogenous protein in detergent solution. Often not a single detergent is used for all intermediate steps, in some cases the detergent is exchanged during purification or even detergent mixtures are used to keep the MP in solution. Another pitfall lies within the physical behavior of detergents in buffers or under ionic strength (*36, 102*). For example, the critical micellar concentration (CMC) differs for DDM in a buffer with no salt compared to a 250 mM sodium chloride solution. Although and in comparison to other detergents, the change of the CMC for DDM from 0.12 to 0.098 mM is rather small (*36*), the effect of salt on a solution containing more than one detergent is even more complicated and can *a priori* not been estimated.

How crucial the choice of the "right" detergent can be, shows the example of Ca^{2+}-ATPase from *Sarcoplasmic Reticulum*. For this membrane protein more than 30 detergents have been tested for preserving the activity of this MP. The analysis focused on chain length, influence of double bonds and character of the head group. A reduced activity was observed for shorter chain length (<10 C-Atoms) and bulky head groups just to mention two examples of this study (*103*).

Optimal detergent (or detergent mixture) conditions cannot be predicted and an analysis is needed for every single MP or purification step individually. Due to the hydrophilic and hydrophobic part of a membrane protein the exact size of the detergent micelle can play a crucial role for any experiment scheduled after the MP has been purified.

An example for such a screening procedure was described for the crystallization of the Na^+/H^+ antiporter NhaA from *E. coli* (*104*). Here, 14 different detergents were screened in the crystallization process yielding different crystal forms resulting in varying quality of X-ray diffraction of the corresponding crystals. Finally, a mixture of α-dodecylmaltoside (DDM) and β-octyl-glucoside (OG) yielded the best diffracting crystals. The author's assumption is that

the addition of β-OG to NhaA purified in α-DDM reduced the micelle size thereby allowing the formation of stable crystal contacts (*104*). This phenomena has been suggested as a general strategy for the crystallization of membrane proteins (for other examples see also (*105, 106*)).

Here, we present a simple 'tool-box' combining different, already published procedures (*36, 38-40*) to determine the CMC values as well as the aggregation numbers of detergent mixtures. As representative detergent we choose DDM as this is the most widely used detergent in membrane protein purification, analysis and crystallization.

Several studies have shown that pyrene quenching is suitable to determine aggregation numbers of detergents (*38, 39, 107-109*). However, in most biochemical laboratories this is not commonly used, because these methods are time consuming and require high sophisticated fluorescence equipment. With our method - using 96well standard microtiter plates - it is possible to analyze most properties of the detergent micelles in any buffer system. This can be very valuable information for example when screening for a reduced micelle size, or the amount of free detergent in only a couple of hours.

Materials and Methods

Triton X-100, LDAO and Pyrene were purchased from Sigma Chemicals. All other detergents used in this study were received from Affymetrix.

Fluorescence Measurements

CMC-determination

CMCs were determined with the set-up described earlier (*36*). In brief, Hoechst 33342 is highly fluorescence in hydrophobic environment whereas it is not in aqueous solution. Therefore, Hoechst 33342 fluorescence was monitored at increasing detergent concentrations. When fluorescence starts to increase (first inflection point), micelles start to form and interpolation leads to the CMC value.

Partitioning Coefficient determination

The partitioning coefficient of Hoechst 33342 into the micellar phase is given by equation 1:

$$P_{m(Hoechst33342)} = \frac{\left[Hoechst33342\right]_m}{\left[Hoechst33342\right]_w} \qquad \text{Equation 1}$$

As described earlier (*38*), one can calculate the partitioning coefficient by measuring the fluorescence intensity of the dye at different α_m values. α_m corresponds to the ratio of micelle volume to total volume of solution. If $F_m >> F_w$, where F is the fluorescence of Hoechst33342, one obtains equation 2:

$$\frac{F}{F_\infty} = \frac{\alpha_m \left[Hoechst33342\right]_m}{\alpha_m \left[Hoechst33342\right]_m + (1-\alpha_m)\dfrac{\left[Hoechst33342\right]_m}{P_m}} \qquad \text{Equation 2}$$

F_∞ corresponds to the fluorescence when all Hoechst33342 molecules are located in micelles. A plot of inverse fluorescence versus inverse α_m yields in a linear function, which provides the partitioning coefficient through the intercept divided by the slope (equation 3).

$$\frac{1}{F} = \frac{1}{F_\infty} - \frac{1}{P_m F_\infty} + \left(\frac{1}{P_m F_\infty} \bullet \frac{1}{\alpha_m}\right)$$
Equation 3

Steady State Fluorescence Quenching Studies

Detergent concentrations used in the assay were five fold above the CMC, determined with the method mentioned above, except for detergents with a very high CMC (≥5mM). In these cases, the detergent concentration above CMC was adjusted to 2mM. For example pure SDS aggregation number was measured at 10mM (CMC = 8mM). Pyrene concentration was adjusted to at least 10 fold below the concentration of micelles present as calculated from literature values for the aggregation number to avoid excimer formation. The quencher concentration varied from 0 to 20µM.

Steady-state fluorescence quenching experiments were carried out either in a Fluorolog-3 (Horiba Jobin Yvon GmbH, Germany, Unterhaching) or in a 96well plate reader (Fluorostar, BMG labtech, Offenburg, Germany). Excitation wavelength was set to 335 nm (filter: 330 ± 10 nm), emission was monitored from 365 to 450 nm or with a filter of 372 ± 10 nm. Fluorescence data were corrected for background fluorescence and analyzed according to equation 4:

$$\ln\frac{I}{I_0} = -\frac{N[Q]}{(C_t - cmc)}$$
Equation (4)

Here, *I* and I_o are the intensities of the first vibronic peak of the fluorescence emission spectra of pyrene in the presence and absence of several quencher concentrations. C_t is the total detergent concentration, N the aggregation number and *cmc* is the critical micellar concentration (*107*).

Theoretical treatment of mixed detergent systems: Ideal Mixing

To analyze whether it is really necessary to measure the CMC of the mixed detergent system or if one can rely on theoretical data we compared our data with the ideal mixing model (*110*). This model relates the ideal mixed CMC, CMC*, with the experimental CMC of the pure components by using the following equations:

$$c_1 = x_1 CMC_1 = \alpha_1 CMC^*$$
<div align="right">Equation (5)</div>

and

$$c_2 = x_2 CMC_2 = (1 - \alpha_1) CMC^*$$
<div align="right">Equation (6)</div>

Equation (7) can be obtained after combining Equations (5) and (6):

$$\frac{1}{CMC} = \frac{\alpha_1}{CMC_1} + \frac{1 - \alpha_1}{CMC_2}$$
<div align="right">Equation (7)</div>

where α_1 and x_1 are the mole fractions of detergent 1 (for example DDM) in the mixed solute and in the mixed micelle, respectively. CMC_1 and CMC_2, c_1 and c_2 are the CMCs and the overall concentrations of components 1 and 2, respectively. If both detergents do not mix well one has to add the micellar activity coefficient f:

$$\frac{1}{CMC} = \frac{\alpha_1}{f_1 CMC_1} + \frac{(1 - \alpha_1)}{f_2 CMC_2}$$
<div align="right">Equation (8)</div>

For a more detailed description of this theory, see Sarmoria et al. and Holland et al (*111, 112*).

Results

Here, we established a method to measure the CMC and aggregation numbers of detergents in a nearly high throughput manner. Due to the information gained with this assay it is possible to adjust buffer conditions of the purification and/or subsequent activity assay, in a controlled manner. For example it will be possible to maintain or regain MP activity and to stabilize the MP in a more controlled fashion (*113, 114*).

As recently described, fluorescence of Hoechst 33342 in detergent solutions can be used to calculate CMCs for different detergent-water or detergent-buffer systems (*36*). Hoechst 33342 is highly fluorescence in hydrophobic environments but shows nearly no emission in water-based solutes. In addition to the method to measure the CMC of detergent (*36*), we further developed an approach allowing the investigation of the influence of salts or pH on the CMCs and aggregation numbers of detergent mixtures .

To calculate aggregation numbers of detergents, pyrene quenching has been used extensively (*38-40*). Hoechst 33342 absorption spectra and pyrene fluorescence emission spectra overlap. This implicates that Hoechst 33342 can quench pyrene fluorescence in micelles as described in detail before (*38, 39*). Figure 1 shows fluorescence spectra of pyrene in the absence and presence of increasing Hoechst33342 concentrations. Here, the DDM concentration was 5 fold above CMC (1mM) and the pyrene concentration used was adjusted to an expected micelle to pyrene ratio of 10:1. Under these conditions no fluorescence peak at 470 nm was observed, which would have been present if pyrene excimers were formed.

To validate our method exercisable to determine aggregation numbers we needed to exclude possible errors introduced from poor solubility of Hoechst 33342 in different detergents. Therefore, we first analyzed the partitioning of Hoechst 33342 into different micelles. Partitioning coefficients (Pm) were calculated according to Equation 3 (see Materials and Methods). Hoechst 33342 greatly migrates into the micelles as shown by the Pm values listed in Table 1. The Pm value for DDM is 2500 demonstrating, that almost all Hoechst 33342 molecules are in the micellar phase under the conditions used in the assay. To phrase it

differently, 2500 fold excess of Hoechst 33342 is present in micelles compared to the aqueous phase. The same holds true for the other detergent tested (see Table 1).

As already observed earlier, micelle size is not strictly constant but varies slightly. With higher amounts of detergent, the micelle size in most cases also increases (*38, 115, 116*). Our assay is sensitive enough to reliably measure this increase. Pyrene quenching by Hoechst 33342 was analyzed measuring the emission peak of pyrene at 373 nm at different quencher concentrations. The intensity of pyrene fluorescence decreases with increasing quencher concentration (Figure 2). A plot of lnF/lnFo vs. quencher concentration results in a straight line from which the aggregation number can be calculated using the slope. Here an exact knowledge of the CMC and detergent concentration are an absolute prerequisite (see equation 4). Pyrene quenching was measured for three different DDM concentrations. Increasing DDM concentrations resulted in a decrease in the slope of the logarithmic fluorescence ratio, which is expected as the presence of more micelles decreases the probability to pair a donor and an acceptor molecule in the same micelle. The calculated aggregation numbers of 104, 135 and 154, respectively, correspond very well with literature values (see Table 2). To support our interpretation and the ease of our assay, we like to note that increasing aggregation numbers with increasing detergent concentrations were also observed with other techniques for example SANS, NMR, and SAXS (table 2) (*38, 115-117*).

As a second experiment, we looked at the influence of salt and pH on the behavior of the detergent in terms of CMC and aggregation number. Information about the influence of salt, pH, denatured reagents or further parameters on the physicochemical nature of a detergent micelle can be very useful to optimize buffer conditions especially to reduce the amount of free detergent or empty detergent micelles, when using MPs. Therefore, we analyzed the influence of increasing ionic strength on the aggregation number of SDS. DDM is a non-ionic detergent and shows only small changes in aggregation number under increasing salt concentrations as seen with other methods before (*118*). Therefore, we illustrated a more drastic example using SDS. In Figure 3, we measured SDS aggregation numbers without salt and after addition of 250mM sodium chloride. With increasing salt concentrations, the aggregation number increased from 68 for water to 86 (Figure 3). These

data are in good agreement with data obtained by the more sophisticated methods like ultracentrifugation (*119*), time-resolved fluorescence quenching (*120*), light scattering (*121*) and EPR (*122*), but we like to stress that this values were obtained in a 96well plate reader.

Often large micelles might stabilize a protein but interfere with activity or for example with the formation of crystal contacts (*106*). With the pyrene quenching method it is possible to measure several detergent mixtures in a 96-well plate reader in a couple of hours. Figure 4 shows exemplary mixtures of DDM with other detergents. DDM is a very widely used detergent for MP purification and activity measurements. The different ratios of DDM with the added detergent where analyzed for their influence on CMC and aggregation number and plotted against the molar ratio of added detergent. Most of the calculated CMCs show almost ideal behavior according to equation (7), which is highlighted by a straight line. Here, we refer to the pseudo phase thermodynamic model described by Clint and Motomura (*110, 123*). In our study, little deviation from ideal behavior is observed only for example for high amounts of SDS in DDM micelles. As seen in Figure 4, the measured values for CMCs (triangles), when more than half of the detergent present is SDS do not increase with the same intensity as the straight line indicative for ideal behavior does. For example, for 90% SDS in the detergent mixture, we observed a CMC of 1mM whereas theoretically it should be 1.54 mM. The same deviation for this mixture was observed with surface tension measurements by Zhang et al. (*124*). Therefore, one has to be careful when mixing detergents with dramatic differences in CMC and different head groups (nonionic vs. ionic). The reason for this deviance might be poor mixing in the micelle of the two components. This is an important information when choosing to work with a mixture of detergents.

The aggregation numbers plotted in Figure 4 seems to behave linear in most mixtures with DDM, which is in good agreement with data obtained by SAXS studies on mixed micelles (*125*). Only SDS and Triton X-100 show unexpected curve progression. SDS seems to have little impact on DDM as also seen in CMC behavior (see above). The aggregation number stays roughly 100 until 100% SDS is reached, where a value of 50 monomers per micelle was calculated. For Triton X-100 the observation is the opposite: here aggregation number increases already after the addition of 25% Triton X-100 to DDM from 108 to 189.

These findings strongly recommend investigating the CMC and aggregation number of the detergent mixture in the buffer used before the experiment is started.

Discussion

Quenching of pyrene fluorescence by Hoechst 33342 can be used as a simple and rapid method to determine aggregation numbers. Together with the usage of a 96well plate reader ten or more detergent mixtures can be measured in less than an hour. Pyrene was chosen due to its long fluorescence lifetime (≈400ns)(*126*), but also because it is relatively cheap and widely available.

Hoechst 33342 partitions greatly into the micellar phase and shows an overlap with the pyrene fluorescence spectrum. Furthermore, it does almost not fluorescence in the region of pyrene fluorescence peak maximum (273 nm). Hoechst 33342 fluorescence can be used to determine CMC values (36). Combining both characteristics gives the opportunity to determine aggregation numbers with equipment that is literally available in every biochemical lab. Of course other combination of fluorophore and quencher are possible and might be investigated. Furthermore, the physical behavior of detergent mixtures can be measured using basically every imaginable ratio of detergent. We show here the proof of principle using DDM as the first detergent. To the best of our knowledge this method is applicable for any detergent.

Acknowledgements.

Figure 1: Corrected emission fluorescence spectra of a 5.0 x 10^{-7} M solution of pyrene in an aqueous solution of 1mM DDM at 25°C. The black curve was collected in the absence of the quencher Hoechst 33342; grey and blue curves in the presence of increasing (5 to 20 μM) quencher concentrations.

Figure 2: Quenching of Pyrene fluorescence by Hoechst 33342 at three different detergent concentrations. The logarithm of the fluorescence intensity of the first vibronic peak of pyrene is shown as a function of Hoechst 33342 concentration. Pyrene was excited at 335 nm and emission monitored at 372 nm. Aggregation numbers were calculated from the slope of these three measurements and are 104, 135 and 154, respectively.

Figure 3: Salt dependency of aggregation number of SDS. Increasing salt concentrations result in a swelling of the micelle. Aggregation numbers calculated from this plots resulted in a value of 68 and 86 monomers per micelle, respectively.

Figure 4: CMC and aggregation number determination of mixtures of DDM with different detergents via the 96-well plate reader set-up. Triangles show the behavior of the CMC at different ratios of a second detergent. Quadrangles indicate aggregation number measured for different detergent mixtures.

Table 1: Calculated partitioning coefficients and literature values for CMC and aggregation number taken from www.affymetrix.com unless otherwise indicated.

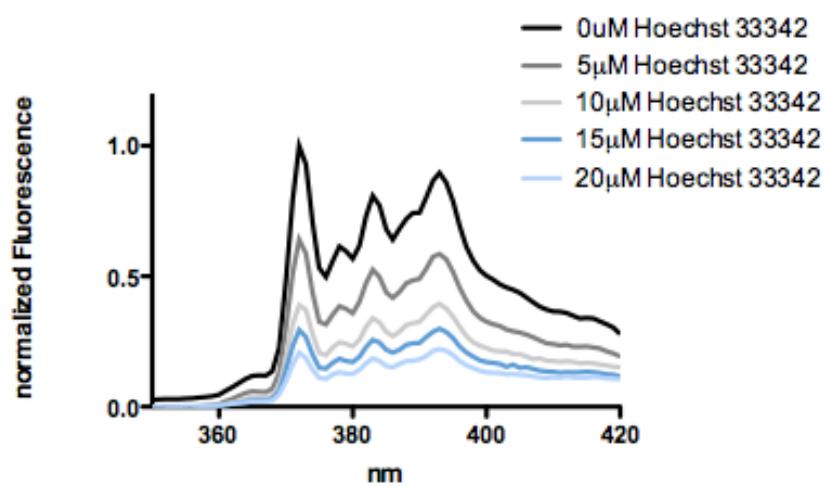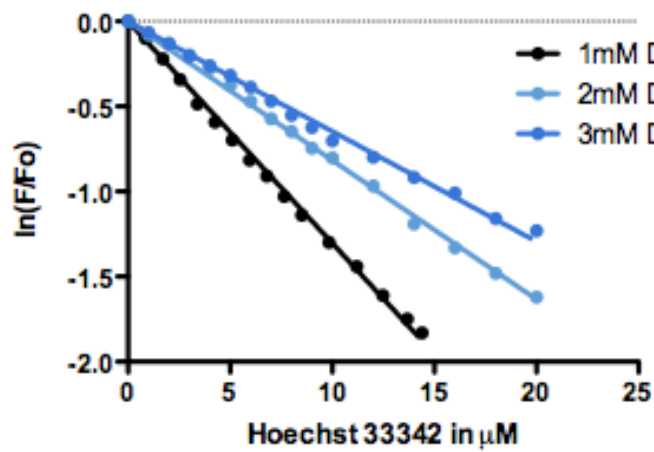**Table 2:** Concentration dependency of aggregation numbers.
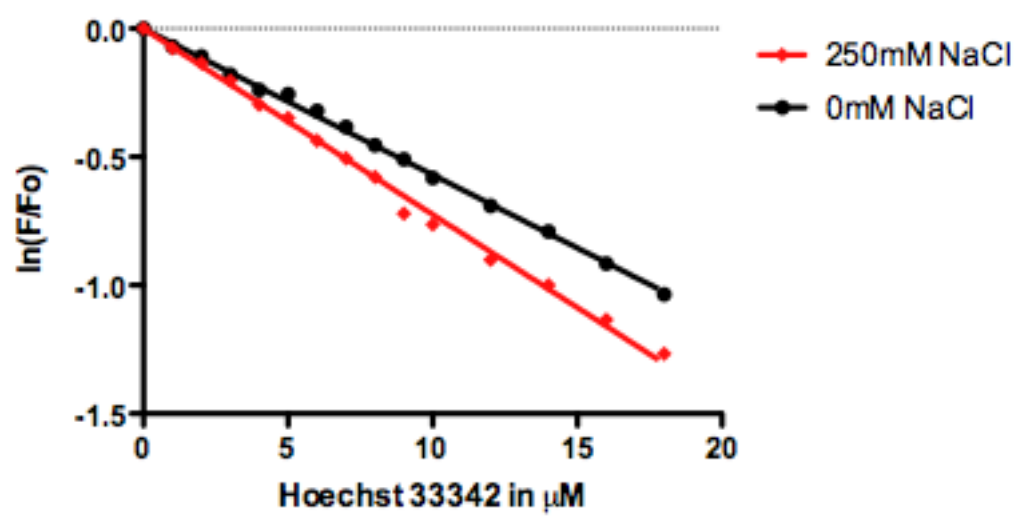
Figure 1:

Figure 2:

Figure 3:

Figure 4:
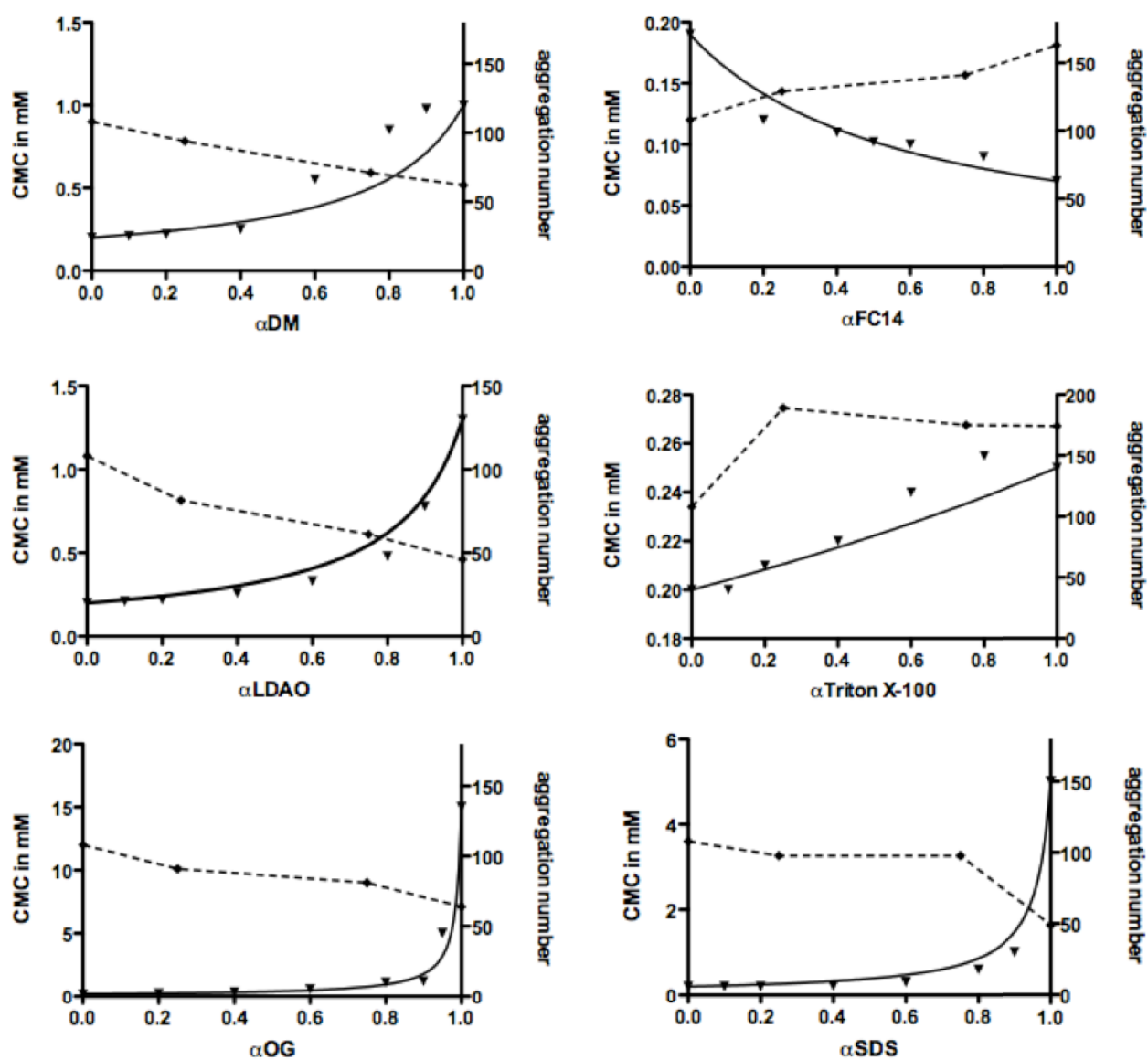
Table 1:

Detergent		Pm	n_{agg} Lit.	CMC Lit. (mM)
DDM	Nonionic	2500	78-149	0.17
DM	Nonionic	390	69	1.8
Triton X-100	Nonionic	3500	75-165	0.15-0.24
FC-14	Zwitterionic	-	108	0.12
LDAO	Zwitterionic	1352	74[$]	1-2[§]
OG	Nonionic	100	27-100	18-20
SDS	Anionic	80	62-101	8

§: (*127*), $: (*128*)

Table 2:

Detergent	Concentration [mM]	Aggregation number	Method	Literature
DDM	1 2 3	104 135 154	Steady state fluorescence quenching	This study
DDM	5 10	136 140	SAXS	(*115*)
SDS	100 200 300 400 500	69 78 85 90 95	NMR	(*117*)
	30 70 100 125 160 250	84 100 113 122 129 142	SANS	(*116*)
Triton X-100	1.6 3.2 4.8 6.4	111 125 140 156	Steady state fluorescence quenching	(*38*)
OG	50 75 100 150	83 99 111 120	SAXS	(*115*)

References

1. Jumpertz, T., Tschapek, B., Infed, N., Smits, S. H., Ernst, R., and Schmitt, L. (2011) High-throughput evaluation of the critical micelle concentration of detergents, *Anal Biochem 408*, 64-70.

2. Paradies, H. H. (1980) Shape and size of a nonionic surfactant micelle. Triton X-100 in aqueous solution, *J. Phys. Chem. 84*, 599-607.

3. Lund, S., Orlowski, S., de Foresta, B., Champeil, P., le Maire, M., and Moller, J. V. (1989) Detergent structure and associated lipid as determinants in the stabilization of solubilized Ca2+-ATPase from sarcoplasmic reticulum, *J Biol Chem 264*, 4907-4915.

4. Screpanti, E., Padan, E., Rimon, A., Michel, H., and Hunte, C. (2006) Crucial steps in the structure determination of the Na+/H+ antiporter NhaA in its native conformation, *J Mol Biol 362*, 192-202.

5. Lemieux, M. J., Song, J., Kim, M. J., Huang, Y., Villa, A., Auer, M., Li, X. D., and Wang, D. N. (2003) Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily, *Protein Sci 12*, 2748-2756.

6. Michel, H. (1983) Crystallization of membrane proteins, *Trends Biochem. Sci. 8*, 56-59.

7. Tummino, P. J., and Gafni, A. (1993) Determination of the aggregation number of detergent micelles using steady-state fluorescence quenching, *Biophys J 64*, 1580-1587.

8. Turro, N. J., and Yekta, A. (1978) Luminescent Probes for Detergent Solutions. A simple Procedure for Determination of Mean Aggregation Number of Micelles, *J. Am. Chem. Soc 100*, 5951-5952.

9. Infelta, P. P. (1978) Fluorescence quenching in micellar solutions and its application to the determination of aggregation numbers, *Chemical Physics Letters 61*, 88-91.

10. Pastor, O., Junquera, E., and Aicart, E. (1998) Hydration and Micellization Processes of n-Octyl-beta-D-Glucopyranoside in Aqueous Solution. A Thermodynamic and Fluorimetric Study in the Absence and Presence of Salts, *Langmuir 14*, 2950-5957.

11. Atik, S. S., and Singer, L. A. (1978) Nitroxyl Radical Quenching of the Pyrene Fluorescence in Micellar Environments. Development of a kinetic model for Steady-State and Transient Experiments, *Chem. Phys. Letts 59*, 519-524.

12. Hansson, P., Jönsson, B., Ström, C., and Södermann, O. (2000) Determintion of Micellar Aggregation Numbers in Dilute Surfactant Systems with the Fluorescence Quenching Method, *J. Phys. Chem. 104*, 3496-3506.

13. Clint, J. H. (1974) Micellization of Mixed Nonionic Surfacre Active Agents, *J. Chem. Soc. Faraday. Trans. 71*, 1327-1334.

14. Sarmoria, C., Puvvada, S., and Blankschtein, D. (1992) Prediction of Critical Micelle Concentrations of Nonideal Binary Surfactant Mixtures, *Langmuir 8*, 2690-2697.

15. Holland, P. M., and Rubingh, D. N. (1983) Nonideal Multicomponent Mixed Micelle Model, *J. Phys. Chem. 87*, 1984-1990.

16. Vinogradova, O., Sonnichsen, F., and Sanders, C. R., 2nd. (1998) On choosing a detergent for solution NMR studies of membrane proteins, *J Biomol NMR 11*, 381-386.

17. Bowie, J. U. (2001) Stabilizing membrane proteins, *Curr Opin Struct Biol 11*, 397-402.

18. Lipfert, J., Columbus, L., Chu, V. B., Lesley, S. A., and Doniach, S. (2007) Size and shape of detergent micelles determined by small-angle X-ray scattering, *J Phys Chem B 111*, 12427-12438.

19. Bezzobotnov, V. Y., Borbély, S., Cser, L., Faragó, B., Gladkih, I. A., and Ostanevich, Y. M. (1988) Temperature and Concentration Dependence of Properties of Sodium

Dodecyl Sulfate MIcelies DetermIned from Small-Angle Neutron Scattering Experiments, *J. Phys. Chem. 92*, 5738-5743.

20. Gangabadage, C. S., Najda, A., Bogdan, D., Wijmenga, S. S., and Tessari, M. (2008) Dependence of the Size of a Protein–SDS Complex on Detergent and Na+ Concentrations, *J. Phys. Chem. 112*, 4242-4245.

21. Bucci, S., Fagotti, C., Degiorgio, V., and Piazza, R. (1991) Small-angle neutron-scattering study of ionic-nonionic mixed micelles, *Langmuir 7*, 824-826.

22. Doughty, D. A. (1979) Isopiestic compositions of aqueous ionic surfactant systems as a measure of preferential interactions. Application to the determination of micelle aggregation numbers by equilibrium ultracentrifugation, *J. Phys. Chem. 83*, 2621-2628.

23. Lianos, P., and Zana, R. (1980) Use of Pyrene Excimer Formation To Study the Effect of NaCl on the Structure of Sodium Dodecyi Sulfate Micelles, *J. Phys. Chem. 84*, 3339-3341.

24. Emerson, M. F., and Holtzer, A. (1967) On the ionic strength dependence of micelle number.II, *J. Phys. Chem. 71*, 1898-1907.

25. Bales, B. L., Messina, L., Vidal, A., and Peric, M. (1998) Precision Relative Aggregation Number Determinations of SDS Micelles Using a Spin Probe. A Model of Micelle Surface Hydration, *J. Phys. Chem. 102*, 10347-10358.

26. Motomura, K., Yamanaka, M., and Aranato, M. (1984) *Colloid & Polymer Science 262*, 948.

27. Zhang, R., Zhang, L., and Somasundaran, P. (2003) Study of mixtures of n-dodecyl-β-D-maltoside with anionic, cationic, and nonionic surfactant in aqueous solutions using surface tension and fluorescence techniques, *Journal of Colloid and interface Science 278*, 453-460.

28. Columbus, L., Lipfert, J., Jambunathan, K., Fox, D. A., Sim, A. Y., Doniach, S., and Lesley, S. A. (2009) Mixing and matching detergents for membrane protein NMR structure determination, *J Am Chem Soc 131*, 7320-7326.

29. Turro, N. J., Graetzle, M., and Braun, A. M. (1980) Photophysical and Photochemical Processes in Micellar Systems, *Angew Chem. Int. Ed. Engl. 19*, 675-696.

30. Herrmann, K. W. (1962) Non-Ionic-Cationic Micellar Properties of Dimethyldodecylamine Oxide, *J. Phys. Chem. 66*, 295-300.

31. Herrmann, K. W. (1966) Micellar Properties of some zwitterionic surfactants, *Journal of Colloid and interface Science 22*, 352-359.

Own proportion to this work: 60%

149

Submitted to: Biochemistry

Impact factor: 3.226

Chapter VI

Crystallisation and preliminary X-ray crystallographic studies of an oligomeric species of a refolded C39 peptidase-like domain of the *Escherichia coli* ABC transporter Haemolysin B.

Christian K.W. Schwarz[1], Britta Tschapek[1], Thorsten Jumpertz[1], Stefan Jenewein[1], Justin Lecher[2], Dieter Willbold[2,3], Santosh Panjikar[4], I. Barry Holland[5], Sander H.J. Smits[1], and Lutz Schmitt[1][#]

[1]Institute of Biochemistry, [2]Institute of Physical Biology, Heinrich Heine University Duesseldorf, Universitaetsstr. 1, 40225 Duesseldorf, Germany

[3] Institute for Structural Biochemistry (ICS-6), Research Centre Juelich GmbH, Juelich, Germany

[4]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany.

[5] Institute of Microbiology and Genetics, University of Paris-Sud, Orsay, France

[#] To whom correspondence should be addressed:

Lutz Schmitt

Institute of Biochemistry

Heinrich Heine University

Universitätsstraße 1

40225 Duesseldorf, Germany.

Phone: +49-211-81-10773

Fax: +49-211-81-15310

E-mail: lutz.schmitt@hhu.de

Abstract

The ABC transporter haemolysin B (HlyB) from *E. coli* is part of a Type 1 secretion system that translocates a 110 kDa toxin in one step across both membranes of this Gram-negative bacterium in an ATP-dependent manner. Sequence analysis indicates that HlyB contains a C39 peptidase-like domain (C39-like) at its N-terminus. C39 domains are thiol-dependent peptidases that cleave their substrates after a GG motif. Interestingly, the catalytically invariant cysteine is replaced by a tyrosine in the C39 domain of HlyB. Here, we describe overexpression, purification and crystallization of the isolated C39 as a first step to obtain structural insights into this domain and eventually answer the question concerning the function of a degenerated C39 in the ABC transporter HlyB.

Introduction

The ATP-binding cassette (ABC)-transporter HlyB is a central element of the Type I secretion machinery of *Escherichia coli (E. coli)*. Together with the inner membrane protein HlyD which interacts with the outer membrane protein TolC, HlyB forms a continuous channel-tunnel from the cytoplasm into the exterior to secrete the 110 kDa toxin HlyA without any periplasmic intermediates (*42, 129-132*).

In general, a functional ABC transporter consists of two nucleotide-binding domains (NBD) and two transmembrane domains (TMD). However, and in contrast to this general blueprint, HlyB contains an additional N-terminal domain of about 150 amino acids that is conserved in bacteriocin ABC transporters (*45, 133*). A sequence alignment classifies the N-terminal region of HlyB as a C39 domain. This family of cysteine peptidases cleaves their substrate after a GG-motif using a catalytic dyad consisting of a conserved cysteine and a histidine. In HlyB, however, the catalytic dyad is degenerated; the highly conserved cysteine residue is mutated to tyrosine (*45, 133*). Thus, this domain is very likely inactive giving rise to the name C39-like domain.

Curiously, the C39-like domain of HlyB appears to be essential for secretion of the toxin HlyA. In a system where HlyB lacks the C39-like domain, secretion of HlyA is completely abolished (data not shown) and therefore the question arises what specific role this defective protease may play during substrate secretion. We expressed and purified the C39-like domain in *E. coli* and with the objective determine its structure using X-ray crystallography. Here, we describe the preliminary results of the crystallization of C39 of HlyB from *E.coli*.

Material and Methods

Cloning, expression and purification

E. coli strain BL21 (DE3) pLysS was transformed with the vector pET28b harboring the HlyB-C39 gene for overexpression of the C39-like domain (residues 2-144 of HlyB). Since the original sequence of the C39-like domain contains no methionine, two isoleucines were exchanged for methionine to facilitate the possibility the introduction of selenomethionines to provide a tool for structure determination by anomalous diffraction. Cells were grown at 37°C and 180 rpm in the rich medium (2xYT) until an OD_{600} of 0.6 before expression of the C39-like domain was induced by addition of 1mM isopropyl β-D-thiogalactopyranoside (IPTG). All following steps were performed at 4°C. Cells of 4 L bacterial cell culture were harvested 5 h after induction by centrifugation and re-suspended in 60 mL 50 mM HEPES pH 7.0, 150 mM NaCl, 10 % (v/v) glycerol. After addition of DNAse (0.1 mg/mL) cells were incubated for 20 min and disrupted three times with 2.5 kbar using a cell disrupter (Constant Systems). The homogenate was centrifuged for 30 min at 30,000 x g. The expression of the C39-like domain resulted in the formation of inclusion bodies. The pellet of inclusion bodies was washed once with 50 mM Tris-HCl pH 7.0, 0.5 % Triton X-100, 1 mM dithiotreitol (DTT), 0.05 % NaN_3 and once with 50 mM Tris-HCl pH 7.0, 1 mM EDTA, 0.05 % NaN_3. Following the final centrifugation step, inclusion bodies were resuspended in 20 mL buffer (6 M guanidinium hydrochlorid, 10 mM Tris-HCl pH 7.0, 100 mM KCl, 10 mM DTT) until the pellet was dissolved. The protein suspension was centrifuged for 1 h at 50.000 x g and the protein concentration of the supernatant was determined using a NanoDrop device (PeqLab Biotechnology GmbH). The resulting supernatant was diluted to 2 mg/mL with resuspension buffer and dialyzed overnight in a 1:125 ratio against renaturation buffer containing 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 10 % (v/v) glycerol, 2 mM DTT, 0.5 % (w/v) PEG 3350 using dialysis tubing with a molecular weight cut-off of 6-8 kDa (Spektra/Por, Spectrum Laboratories, CA, USA). After dialysis any precipitate was separated from soluble protein by centrifugation for 10 min at 21,000 x g and the supernatant was applied to a cation-exchange (cIEX) chromatography (5 mL SP HP column, GE Healthcare) pre-equilibrated with buffer A (10 mM HEPES pH 7.0, 1 mM DTT). After washing with 5 column

volumes (CV) of buffer A, proteins were eluted with a linear gradient of 40 CV to 100 % buffer B (buffer A containing 500 mM NaCl). Proteins eluting at 190 mM NaCl (38% buffer B) and 375 mM NaCl (75 % buffer B) were concentrated using an Amicon Ultra Centrifugal Filter device (10 kDa MWCO, Millipore), and analyzed by SDS-PAGE and Western blotting with an antibody recognizing the His-Tag (according to the manufacturer, Qiagen). However, the protein fraction that eluted at 375 mM NaCl from the column could not be concentrated above 2 mg mL^{-1}. Nevertheless, both species were applied to a size exclusion chromatography (SEC) using a Superdex-200 16/60 column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT.

Buffer Screen

A buffer screen was performed with the C39-like domain fraction eluting at 375 mM NaCl during cIEX chromatography. The protein was concentrated immediately after elution to 2 mg mL^{-1}. 300 nL of the C39-like domain were mixed in a 1 : 1 ratio with commercially available crystalization buffers (Basic Crystallography Kit, Sigma-Aldrich) in 96-well plates using a pipette robot and incubated at 18°C for 24 h. Buffer conditions with no visible precipitations were analyzed further. 1 mL of the C39-like domain (2 mg mL^{-1}) was dialyzed against 1 L of each buffer positively evaluated in the 96-well format using Slide-A-Lyzer dialysis cassettes (10 kDa MWCO, ThermoScientific) and subsequently concentrated via filter devices (see above) until precipitations occurred. After purification and dialysis against 50 mM citrate pH 6.5 and 20 mM LiCl, the C39-like domain could be concentrated to 20 mg mL^{-1}.

SEC-MALS

SEC-MALS (SEC - multi-angle light scattering) experiments were performed using an analytical Superdex200 10/300 column (GE Healthcare) equilibrated with 25 mM Tris pH 8.0 and 150 mM NaCl on an Äkta-Purifier (GE Healthcare) connected to a triple-angle light scattering detector (miniDAWN™ TREOS, Wyatt Technology) followed by a differential refractive index detector (Optilab® rEX, Wyatt Technology). 100µL of purified C39-like domain was loaded at a protein concentration of 1 mg mL^{-1} on the analytical Superdex200 10/300 column. Data were analyzed with the ASTRA software package (Wyatt Technology).

Crystallization and preliminary X-ray analysis

Crystallization conditions were screened via the sitting-drop vapor-diffusion method in combination with commercially available buffers (Sigma basic screen I and II) using the monomeric as well as oligomeric form of the C39-like domain at a protein concentration of 20 mg/ml. Crystals appeared when samples of protein were mixed with an equal volume of the precipitant solution containing 3.6–4 M formiate. Using the "No-fail cryoprotection" method (http:/capsicum.colgate.edu/chwiki/) 37.5% glucose in mother liquor was added gradually to the crystal solutions to a final concentration of 20% glucose before flash cooling of the crystals in liquid nitrogen. A high resolution dataset was collected from a single crystal at 100 K at a wavelength of 0.87260 Å on a MAR225 detector (MAR Research) at the ESRF beamline id23eh1 using a 0.75° rotation per frame. The dataset was processed using the XDS Package (*74*) and scale using XSCALE.

Results and discussion

The C39-like domain of HlyB was expressed as inclusion bodies in *E. coli*. After preparation and crude purification of the inclusion bodies, proteins were denatured and refolded. A cIEX step was used to purify the refolded C39-like domain (see methods). During elution with a salt gradient, two species eluted at 190 mM NaCl (38% Buffer B) and 375 mM NaCl (75% Buffer B), respectively (Figure 1). SDS-PAGE and Western Blot analysis using an anti-His-tag antibody showed that both eluted fractions correspond to the C39-like domain (data not shown). Therefore, both species were concentrated and applied to a SEC column. However, the resulting elution chromatograms were different. The low-salt species (190 mM NaCl) gave two peaks (61 mL and 93 mL) corresponding to an oligomeric and a monomeric species. Based on a calibration curve the molecular weights were calculated to approximately 200 kDa and 20 kDa, respectively. For further investigation of the molecular weight of the low salt oligomeric species more precise MALS experiments were performed and the molecular weight was determined to be 233.4 ± 1.4 kDa, corresponding to a dodecamer of the C39-like domain (Figure 2). Analysis by size exclusion chromatography of the high-salt species (375 mM NaCl) revealed that species that corresponded only to a monomeric C39-like domain.

The monomeric species precipitated at a relatively low protein concentration of 2 mg mL^{-1}. In our hands, this concentration was too low to obtain any protein crystals since more than 90% of the drops in the crystallization experiments did not show any precipitation at such low protein concentration. Therefore, a buffer screen was performed to search for conditions with enhanced solubility and stability of the C39-like domain using commercially available crystallization screens. As read-out we chose to rely on visible aggregation, which was evaluated one day after setting up the buffer-protein mixtures. The C39-like domain in buffers resulting in no visible precipitations was spun down using ultracentrifugation and the concentration of the C39-like domain was determined in the supernatant. This method revealed a buffer condition in which the protein was more stable; 20 mM citrate pH 6.5, and 20 mM NiCl$_2$. Monomeric C39-like domain, originally precipitating at 2 mg mL^{-1}, could be concentrated above 20 mg mL^{-1} and was stable for weeks when dialyzed against this buffer

directly after refolding. Unfortunately, despite extensive efforts no well diffracting crystals of the C39-like domain could be obtained.

Therefore, we went back to the initial oligomeric species and crystallized this form of the C39-like domain. Crystals were observed in various screening conditions. The best diffracting crystals (Figure 3) were obtained after 30 d at 12°C in 3.6 - 4 M Na-formiate. For diffraction analysis, crystals were cryoprotected using a "No-fail cryoprotection" method (http:/capsicum.colgate.edu/chwiki/). Here, crystal-containing drops were substituted step by step with small volumes of 37% glucose dissolved in mother liquor to a final glucose concentration of 20%. With this method well-diffracting crystals could be obtained and a native dataset was collected.

Preliminary analysis of the dataset using the XDS package revealed that the crystals belong to the premitive orthorhombic space group (either $P2_12_12$ or $P2_12_12_1$) with a large unit cell axis (a=117.6, b=134.2 c=139.4 Å) for a protein of this size. At the moment, we cannot distinguish between both space groups, since the systematic absences observed are not conclusive. The overall dataset was cut off at a resolution of 2.1 Å with an R_{merge} value in the highest resolution shell below 30% (see Table 1 for data statistics). Consideration of the unit-cell volume (2.2×10^6 Å^3) and the molecular weight of 19,460 Da suggests that the asymmetric unit contains either one copy of 12mer corresponding to a V_M value of 2.36 Å^3 Da^{-1} with an estimated solvent content 47.61% (*134*).

This suggests that the initially observed oligomeric species is likely the state in which the C39-like domain is in a sufficiently stable and homogenous state to crystallize under the conditions used in this study.

An attempt to solve the structure of C39-like taking advantage of the introduced selenomethionine residues for usage of multiple anomolous dispersion (MAD) phasing is currently under way. The result of this structural analysis should provide an insight into the function of C39-like of HlyB, and its role in the secretion of HlyA.

Acknowledgments

References:

1.	Balakrishnan, L., Hughes, C. & Koronakis, V. (2001). J Mol Biol 313, 501-510.

2.	Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. & Schmitt, L. (2003). J Mol Biol 327, 1169-1179.

3.	Gray, L., Mackman, N., Nicaud, J. M. & Holland, I. B. (1986). Mol Gen Genet 205, 127-133.

4.	Ishii, S., Yano, T., Ebihara, A., Okamoto, A., Manzoku, M. & Hayashi, H. (2010). J Biol Chem 285, 10777-10785.

5.	Kabsch, W. (2010). Acta Cryst. D66, 125-132.

6.	Mackman, N., Baker, K., Gray, L., Haigh, R., Nicaud, J. M. & Holland, I. B. (1987). EMBO J 6, 2835-2841.

7.	Mackman, N., Nicaud, J. M., Gray, L. & Holland, I. B. (1986). Curr Top Microbiol Immunol 125, 159-181.

8.	Matthews, B. W. (1968). J. Mol. Biol 33, 491-497.

9.	Wu, K. H. & Tai, P. C. (2004). J Biol Chem 279, 901-909.

Table 1:

Data collection statistics. Native dataset were collected at the ESRF beamline

ID23eh1. Statistics for the highest resolution shell are given in parentheses.

Figure 1:

cIEX column elution profile of the refolded C39-like domain. Two peaks are observed which correspond to the oligomeric dodecameric state (*) and the monomeric state (#) of C39.

Figure 2:

The oligomeric state of the C39-like domain was analyzed by SEC-MALS. The elution profile from a Superdex200 10/300 column is shown. Straight line: normalized UV-Signal; Dashed Line: normalized LS-(90°)-Signal; Dotted line: refractive index signal. Filled cycles indicate the corresponding molar mass.

Figure 3:

Crystal of the C39-like domain

Table 1:

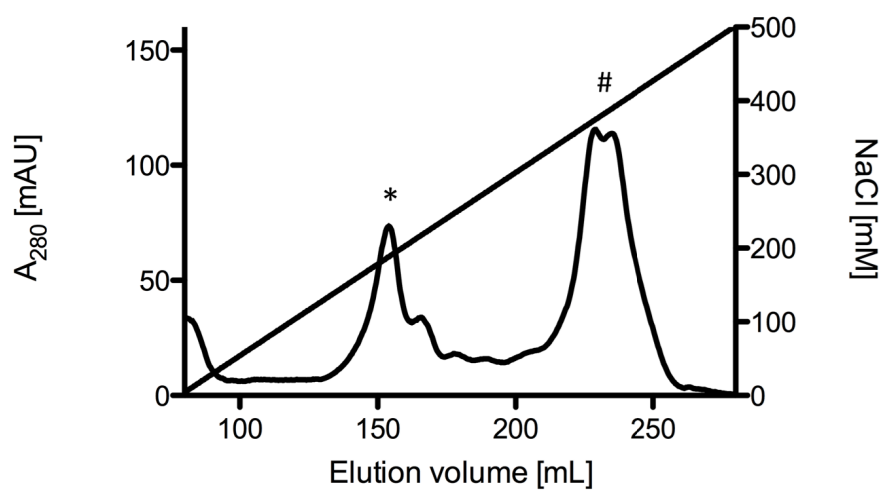| Radiation source	ESRF
Beamline	ID23eh1
Detector	MAR mosaic 225 mm
Space group	$P2_12_12$ or $P2_12_12_1$
Unit cell parameters (Å)	
a	117.64
b	134.15
c	139.44
$\alpha=\beta=\gamma$ (°)	90.0
Resolution range (Å)	20-2.1
Highest resolution shell (Å)	2.2-2.1
Wavelength (Å)	0.870260
Total reflections	720832
No. of unique reflections	240680
Completeness (%)	95.6 (97.3)
R_{merge} (%)	9.0 (21.9)
$I/\sigma(I)$	9.4 (4.6)

Figure 1:
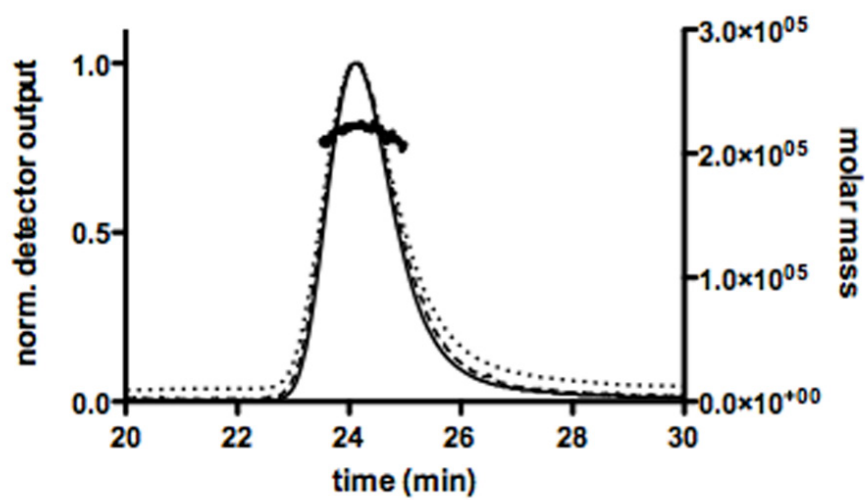

Figure 2:

Figure 3:

Own proportion to this work: 15%

165

Submitted to: Acta Crystallography Section F 2011

Impact factor: 0.55

Chapter VII

Publication in preparation

Investigating the ATPase activity

of the *E. coli* transporter HlyB *in vitro*

Britta Tschapek[1], Christian KW Schwarz[1], Thorsten Jumpertz[1+], Stefan Jenewein[1], Barry Holland[2], Sander HJ Smits[1] and Lutz Schmitt[1*]

[1]Institute of Biochemistry, Heinrich Heine University Duesseldorf, Universitaetsstrasse 1, 40225 Duesseldorf, Germany

[2]Institute of Genetics and Microbiology, University of Paris-Sud, Orsay, France

[+]present address. Institute of Neuropathology, Molecular Neuropathology Group, Medical school Duesseldorf, 40225 Düsseldorf, Germany

[*]Corresponding author. Phone: +49 211 8110773, Fax: +49 211 8115310, E-mail address: Lutz.Schmitt@hhu.de

Introduction

Type-1 secretion in *Escherichia coli* or other Gram-negative bacteria allows the translocation of a wide variety of substrates from cytoplasm to the extracellular space in a single step, without periplasmic intermediates (*1*). In contrast to Sec-dependent translocations, information required and sufficient for secretion, the secretion signal, is located at the very C-terminus of the substrate and not cleaved after translocation.

The paradigm of a Type I secretion system is the Haemolysin A (HlyA) secretion system. The translocation complex is composed of three different proteins: An ABC-Transporter, HlyB in the inner membrane, the membrane fusion protein HlyD, that is supposed to seal the translocation pore of HlyB and the outer membrane protein TolC. HlyD spans the membrane a single time with the first 60, N-terminal amino acids being located in the cytosol. However, the role of this putative cytoplasmic domain remains unclear although deletion abrogates substrate secretion (*2*). The ABC-transporter HlyB is supposed to act as a dimer and therefore contains two transmembrane domains (TMDs), two nucleotide binding domains (NBDs) and additionally to "classic" ABC-transporters, a C39 peptidase domain at its N-terminus. In general C39 domains cleave their substrates after a canonical double glycine motif. However, and as mentioned before HlyA is not degraded during translocation (see Chapter V!). This is due to the fact that the C39 domain of HlyB is degenerated, i.e. the amino acids required for cleavage activity, a cysteine and an aspartate are missing. As this domain is on the other hand essential for toxin secretion it must play a role in substrate recognition, stabilization during transport or communication with other components of the secretion complex.

Despite the fact that the ATP hydrolysis cycle of the NBDs of HlyB is rather well analyzed and understood (*3-6*), little is known about interaction and regulation of the whole transport complex. Therefore we purified the inner membrane components of the secretion system and investigated the interplay of HlyB and the substrate HlyA using ATPase activity measurements as a read-out.

Materials and Methods

Materials

Fos-choline14 was obtained through SynphaBase (Pratteln, Swizerland). Octaethylene glycol monododecyl ether ($C_{12}E_8$) was supplied from Affymetrix (Santa Clara, USA). All other chemicals were purchased from Sigma-Aldrich at the highest purity available.

Strains and Plasmids

pLG814 (*7*) encoding HlyB and HlyD was used as template for HlyB and the HlyB sequence lacking the C39 peptidase domain (HlyBΔC39). By standard PCR methods nucleotides encoding amino acids 1-707 (HlyB) where amplified. Primers were as follows: HlyBfor 5'GCTATCCATGGC<u>GAATT</u>CTGATTCTTGTCATAAAATTGATTATGGG3'; HlyBrev 5' CGTGTCTAGATTACA<u>GGATCCC</u>GTCTGACTGTAACTGATATAAGTAACTG3'. Underlined sequences encoded restriction sites for cloning purposes (EcoRI/BamHI). The HlyB sequence was cloned into pET401 vector using compatible, cohesive ends. Introduction of the affinity tag with a protease cleavage site (factor Xa) was performed postcloning by addition of an oligonucleotide pair at the 5' end (N-terminus; Nhis10Eco_for: 3'AATTCTCATCACCATCACCATCACCATCACCATCATAGCATCGAAGGGCGC5', Nhis10Eco_rev: 3'AATTGCGCCCTTCGATGCTATGATGGTGATGGTGATGGTGATGGTGATGAG5'). The final gene was further subcloned into pBAD vector yielding pSJ027.

HlyBΔC39 plasmid is based on pSJ027 and gained with In-Fusion kit from Clontech (*8*). The gene of interest was amplified using the following primer sequences: CATCACCATCACCATCACCATCATGAGAATCTTTATTTTCAGGGCGCTTCCCGTTCTTCTGTTGC and GGCGGCCGCTCTAGATTACAG. The vector was linearized using the following primers: TCTAGAGCGGCCGCCACCG and ATGGTGATGGTGATGGTGATGAGAATTCGCCATGGT, respectively. Correctness of all constructs was verified by sequencing.

For HlyA secretion experiments same constructs as described in Bakkes et al. were used (*9*).

Protein overexpression and purification:

E. coli strain BL21(DE3) was transformed with plasmid pSJ027 coding wild type or mutant (H662A, ΔC39) His_{10}-tagged HlyB protein. Cultures in 2litres of 2YT medium were grown at 37°C until the OD_{600} reached 0.8, induced by the addition of arabinose (final concentration of 10mM) and were grown for an additional hour. Cells were harvested by centrifugation; cell pellet was resuspended in buffer R (50mM Tris/HCl pH 8.0, 150mM NaCl and 10% glycerol) and

lysed by three passages through a cell disruption machine at 2.5kbar (IUL systems). The cell lysate was centrifuged at 18,000xg for 30min before membranes were isolated by ultracentrifugation (125,000 g, 90min). The resulting membrane pellets were homogenized in buffer M (50mM Tris/HCl, 250mM NaCl, 10% Glycerol). Membrane vesicles were lysed with 1% Fos-choline 14 and HlyB was purified using HiTrap Chelating column loaded with Zn^{2+} (GE Healthcare). The buffer employed was 50mM Tris/HCL, 250mM NaCl, 10 or 200mM Imidazol and 0.012% (w/v) $C_{12}E_8$. His-tagged protein was eluted with 150mM Imidazol and subsequently loaded on a gelfitration column (Superose6 10/300 GL, GE Healthcare). Protein fractions were analyzed by SDS-PAGE according to Laemmli (*10*), using a 10% acrylamide gel, followed by staining with Coomassie Blue.

HlyA secretion

Plasmids encoding for HlyB/HlyD (pSJ037) and HlyA (pSUI) were transformed in BL21 DE3 cells. HlyA secretion was performed for 6h until cells were separated by centrifugation and subsequently supernatant was filtered through 0.25µm filter. Supernatant was concentrated and loaded on a Superdex200 16/60 gelfiltration column (GE Healthcare). Elution fractions containing HlyA were analyzed by SDS-PAGE, shock frozen in liquid nitrogen and stored at -80°C until further use.

Static light scattering

SEC-MALS (SEC and multi-angle light scattering) experiments were performed using an analytical Superose6 10/300 column (GE Healthcare) equilibrated in 50mM Tris pH 8.0, 150mM NaCl run on an Äkta-Purifier (GE Healthcare) connected to a triple-angle light scattering detector (miniDAWN™ TREOS, Wyatt technology) followed by differential refractive index detector (Optilab® rEX, Wyatt technology). Purified HlyB was loaded at 1.5 mg mL^{-1} with a sample size of 100ul. Data were analyzed with the ASTRA software package (Wyatt Technology).

ATPase activity assay

The purified HlyB was assayed for ATP hydrolysis using the Malachite Green method, which depends upon release of P_i and performed as described in (*11*). Assays were performed at 22°C in a temperature controlled 96-well plate spectrometer (Fluostar, BMG labtec). Reaction mixture contained 10mM $MgCl_2$ and increasing amounts of ATP or HlyA.

Results

Isolation and Purification of HlyB:

N-terminally His$_{10}$-tagged *E. coli* HlyB was overexpressed in *E. coli* BL21 DE3 cells and purified by affinity chromatography followed by gelfiltration. As the choice of the detergent used had a great impact on oligomeric status and activity of HlyB detergent was exchanged during affinity chromatography. Nonionic, zwitterionic and ionic headgroups as well as different chain length of detergents were tried. Best results in terms of activity were obtained using nonionic detergents of rather low CMC and more than 100 monomers per micelle (for example DDM and C$_{12}$E$_8$).

Purified HlyB in C$_{12}$E$_8$ buffer solution was >95% pure (Fig 1). Minor contamination at 50 kDa corresponds to a degradation product, which could be confirmed by western blot analysis using polyclonal antibody against the NBD of the transporter (data not shown). A typical preparation from 4 liters of bacterial culture yielded 5 mg of pure membrane protein.
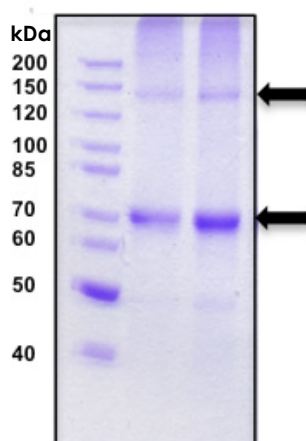

Figure 1: Purification of HlyB. The HlyB dimer is observed at a molecular weight of around 140kD, the monomer runs at 70kD.

Both HlyB mutants (H662A, HlyBΔC39) could be obtained in same quality and amounts following this procedure.

Secretion and purification of HlyA

Highly pure secreted HlyA was further purified by size exclusion chromatography (Fig 2) After concentrating the supernatant residual medium components could be removed with this step.
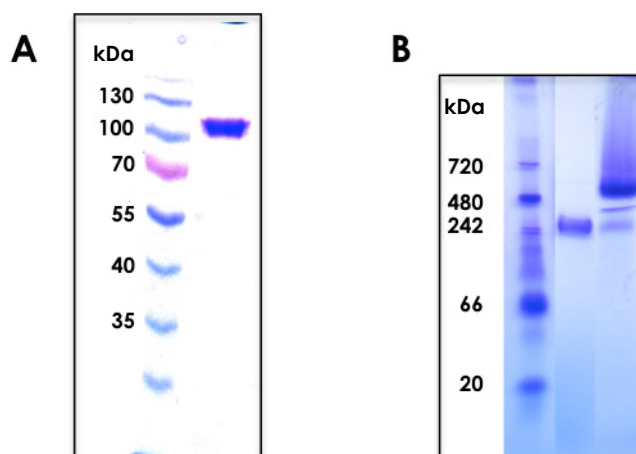
Figure 2: A: 10% SDS-PAGE of purified HlyA. Observed band of a molecular weight of 110kDa belongs to HlyA. B: Blue Native Page of purified HlyA. First lane shows HlyA after incubation at 65°C with SDS, second lane shows HlyA running without denaturizing agents.

Blue Native PAGE analysis was tried to identify the oligomeric state of HlyA after secretion. However, comparison of unfolded and native protein is rather delicate as the unfolded protein (treated with SDS and incubated at 90°C) runs at a higher molecular mass as expected (approxx. 250kD). Therefore the slower migration of the folded protein indicates a dimeric form of the protein but clearly needs further analysis for verification.

Static light scattering to verify the oligomeric state

Online multi-angle light scattering experiments were performed to analyze the oligomeric state of HlyB in detergent solution.
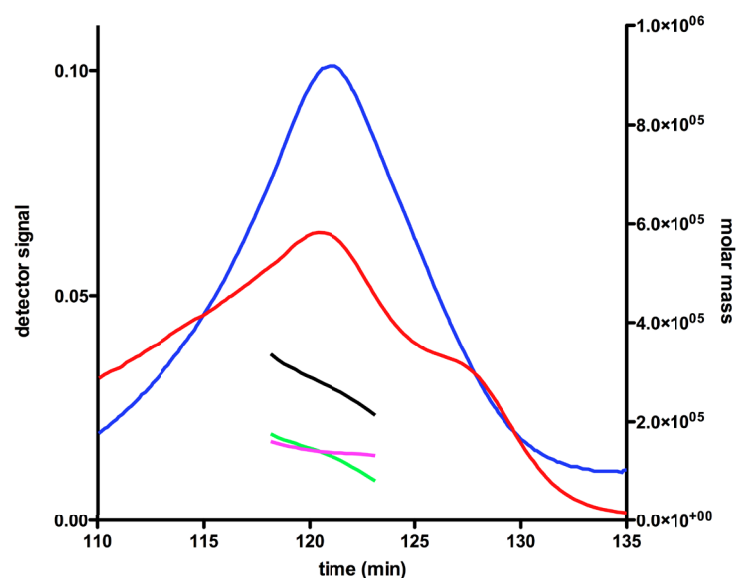

Figure 3: Gelfiltration profile of HlyB on a Superose6 column (GE Healthcare). Blue line: UV-Absoption at 280nm, Red line: Light scattering signal at an angle of 90°, Black line: molar mass, purple line: molar mass of the protein fraction, Green line: molar mass of the detergent fraction.

The results in Figure 3 confirm the dimeric status of purified HlyB in detergent solution as expected for an ABC-transporter. HlyB dimer accumulates roughly the same mass of detergent to stay in solution. The whole protein detergent complex has a calculated molar mass of 301kD, while protein fraction is 155kD. The mass decay in Figure 3 clearly indicates that our preparation is not homogeneous in terms of detergent-protein ratio. However, as we can show specific activity for HlyB, this issue is neglected.

ATPase activity of isolated HlyB

ATPase activity of wildtype HlyB was monitored over 20min and at ATP concentration ranging from 0 to 2mM. Resulting data did not obey Michealis-Menten kinetics, while a nearly perfect fit was obtained using the Hill equation. Such cooperativity has been also observed for the isolated HlyB-NBD but not for example for HisP, MalK or Mdl1p NBD (*12-14*). Kinetic data are summarized in **Table 1**. It was shown previously that mutation of H662 to Ala abolished ATPase activity for the isolated NBD (*3*). Therefore this mutation was introduced in full length HlyB by site-directed mutagenesis and also assayed for ATPase activity. As expected no activity could be detected under the conditions tested for the wild-type protein (Fig 4)
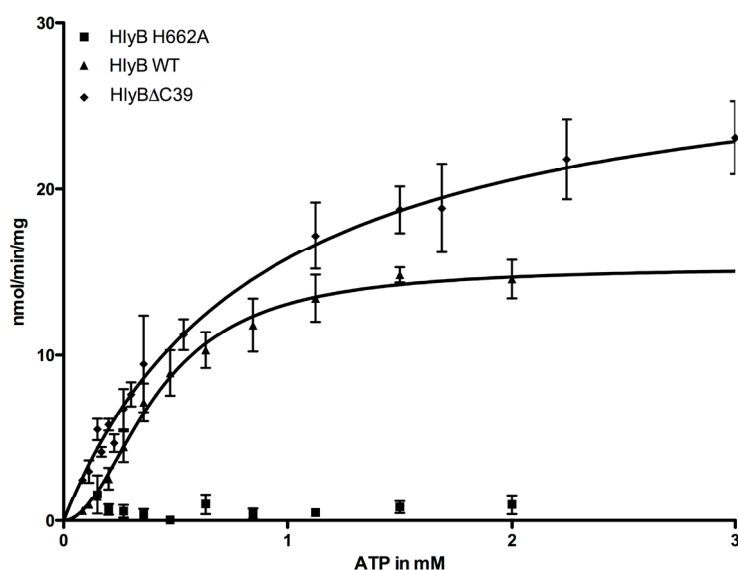

Figure 4: ATPase activity measurements of HlyB in detergent solution. Triangles show HlyB wt activity, squares: HlyB H662A mutant, rotated squares: HyBΔC39.

Table 1: Kinetic data of HlyB ATPase activity

	K_M in mM	v_{max} in nmol/min/mg	h	K_d in µM
HlyB wt	0.42	15.3	2	21.01
HlyB ΔC39	2.00	51.0	-	26.80
HlyB_NBD (3.6µM)	0.26*	600*	1.35*	4[#]
(1.1µM)	0.36*	202	1.31*	

* Zaitseva et al. 2005 (*3*)   [#]Benabdelhak et al. 2003 (*15*)

For some ABC-transporter stimulation of ATPase activity after addition of substrate could be detected (*16*). Interaction of HlyB and HlyA were substantiated by two individual methods (pulldown assays with the separated C39 domain (see Chapter VI) and C-terminal fragment of HlyA, Biacore measurements with HlyB-NBD and HlyA fragment (*15*) as well as C39 domain and this fragment (own lab, unpublished data). Therefore we analyzed the impact of HlyA on the ATPase activity of HlyB (Fig 5) and investigated the impact of the C39 domain. To see if the C39 domain is involved in regulation of the transporter, we deleted the first 150 amino acids of the full-length transporter and measured the remaining ATPase activity (Fig 5) HlyBΔC39 shows more than four fold higher ATP turnover than wild type protein (Fig 4). As stability is increased for this truncation mutant we assume that this effect is not physiologically relevant.

However measurement of activity in presence of increasing amounts of HlyA showed a stimulation of the ATP-hydrolysis rate for the full-size transporter as well as for the ΔC39 version (Fig 5)
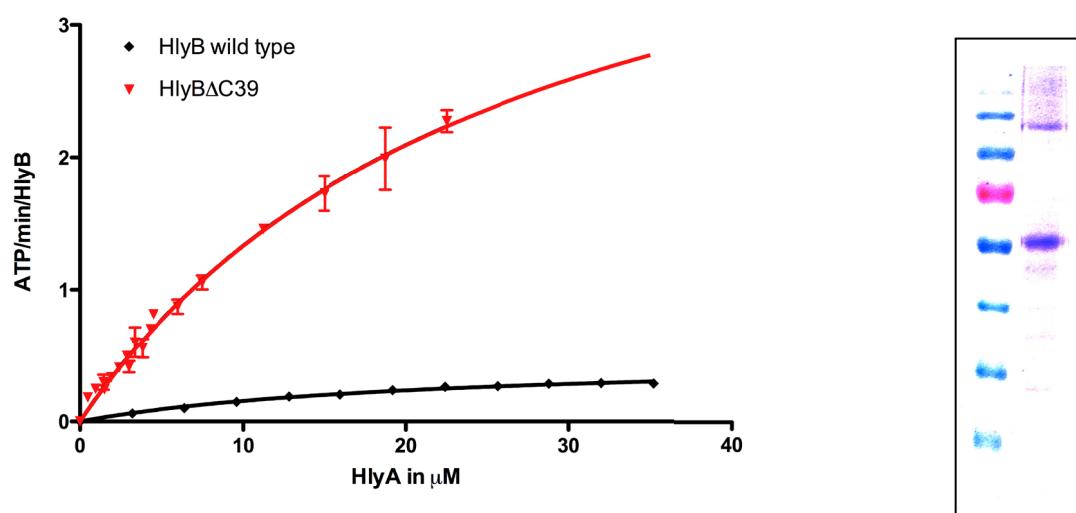

Figure 5: Left panel: HlyA dependent ATPase activity of HlyB full-length and truncation of C39. Increasing amounts of HlyA stimulate ATP hydrolysis rate leading to K_D values of 21 and 26µM, respectively. Right panel: SDS-PAGE showing purified HlyBΔC39.

Interestingly this effect is more than six fold larger for the truncation mutant. Affinity constants are in the same range namely 21 ± 4 and 26 ± 3 µM (Fig 5) for wild type protein and ΔC39, respectively. This is in line with the measured interaction by SPR where a K_D of 4 ± 2 µM was determined for HlyA (25kD, see below) to the isolated HlyB-NBDs (*15*). As C39-domains are known for cleaving their substrates after double glycine motifs we checked if this stimulation effect relies on the signal sequence at the HlyA C-terminus or on GG-repeats in front of it.
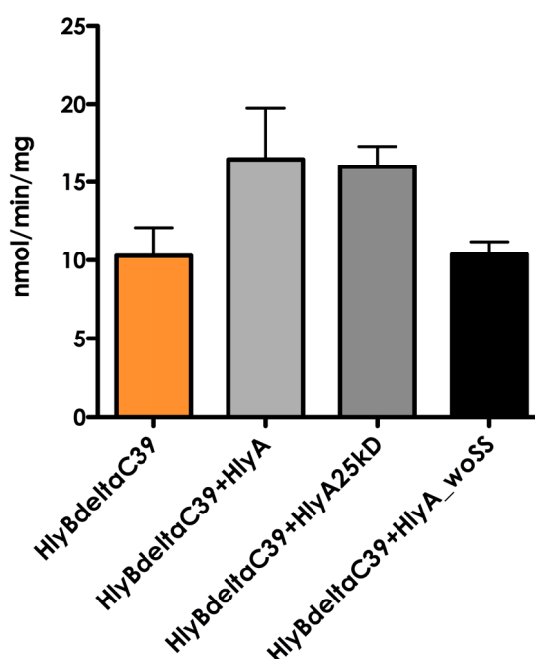

Figure 6: Impact of different variants of HlyA of ATPase activity of HlyBΔC39. HlyA25kD includes the 216 amino acids of the C-terminus of HlyA. HlyA_woSS consists of amino acids 807-966, starting at the same amino acid as the other fragment but lacking the last 57 amino acids at the very C-terminus, i.e. the secretion sequence. HlyA variants were added at a concentration of 3uM, ATP concentration was 1mM.

Figure 6 summarizes the ATP-hydrolysis rates of HlyBΔC39 in the presence of different HlyA proteins. The figure clearly shows, that the secretion sequence is the important determinant of ATP-hydrolysis stimulation. Full-length as well as a 25kDa C-terminal fragment enable HlyBΔC39 to hydrolyze additional 0.4 ATP molecules per minute (0.41 ± 0.064 ATP/min for full length HlyA, 0.36 ± 0.033 ATP/min for 25kDa HlyA). This effect cannot be detected if a degradation fragment of 25kD-HlyA lacking the 57 C-terminal amino acids corresponding to the secretion sequence is added.

Our results suggest a transport mechanism where HlyA has to bind to the C39-like domain via its glycine repeats but the secretion sequence triggers the ABC-transporter to hydrolyze ATP and fuel the transport process with energy.

Discussion

Homologous expression of the Type I secretion system component HlyB and the substrate HlyA in *E. coli* yield sufficient amount of both proteins. We successfully isolated HlyB from the membrane fraction of *E. coli* and prepared membrane vesicles to extract the protein from. However this step was quite challenging as only Fos-choline14 and 16 were able to extract more than 30% of HlyB from the membrane. Moreover one has to handle this detergents with care as several former studies have stated that especially these zwitterioninc harsh detergents tend to unfold proteins (*17*). Our study also resulted in inactive protein in buffers containing FC14 (data not shown). Thus we exchanged the detergent right after extraction and screened for a detergent that stabilizes HlyB in an active conformation. This screening resulted in a protocol using FC14 and $C_{12}E_8$. With this method HlyB could be maintained in its dimeric state, which was proven by size exclusion chromatography with online light scattering analysis. After two purification steps HlyB was more than 95% pure and showed low basal ATPase activity, with a $K_{0.5}$ constant in the range as determined for the isolated NBDs (*3*)

Any uncertainty concerning the detergent can be ruled out if the activity of the protein can be proven and especially if you can detect specific activity. As the light scattering data showed that our preparation is slightly inhomogeneous in terms of protein-detergent ratio, we could dispel all doubts by detecting HlyA dependent ATPase acivity. Furthermore, the H662A mutant of full-length HlyB displayed no activity sustaining our interpretation. Former data on the isolated HlyB-NBDs demonstrated an interaction of the NBDs with the secretion signal at the C-terminus of the substrate HlyA using Biacore measurements (*15*). Here it was possible to confirm these data on full-length protein with ATPase activity measurements as hydrolysis rate increased when adding increasing amounts of substrate and follows a one-site binding kinetic. A modulation of ATPase hydrolysis upon substrate addition has been reported for several ABC-transporters (*18-20*). Inline with these reports one can imagine that the substrate might induce a more efficient dimerisation of the NBDs compared to dimer formation upon binding of the loaded substrate binding protein in import systems.

To rule out the role of the additional N-terminal extension of HlyB, the C39-domain, this domain was removed genetically and the same purification protocol as for wild type protein was applied. When comparing HlyA dependent ATPase activity of wild type and ΔC39 protein ΔC39 shows six fold higher stimulation than wild type, but most importantly no HlyA stimulation was detectable. Thus C39 domain has a regulatory role on ATP hydrolysis during transport and ATP hydrolysis appeared to be uncoupled if this domain is lacking.

Literature

1.   Holland, I. B., Schmitt, L., and Young, J. (2005) Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review), *Mol Membr Biol 22*, 29-39.

2.   Balakrishnan, L., Hughes, C., and Koronakis, V. (2001) Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by Escherichia coli, *J Mol Biol 313*, 501-510.

3.   Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B., and Schmitt, L. (2005) Functional characterization and ATP-induced dimerization of the isolated ABC-domain of the haemolysin B transporter, *Biochemistry 44*, 9680-9690.

4.   Zaitseva, J., Jenewein, S., Oswald, C., Jumpertz, T., Holland, I. B., and Schmitt, L. (2005) A molecular understanding of the catalytic cycle of the nucleotide-binding domain of the ABC transporter HlyB, *Biochem Soc Trans 33*, 990-995.

5.   Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B., and Schmitt, L. (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB, *EMBO J 24*, 1901-1910.

6.   Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B., and Schmitt, L. (2006) A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer, *EMBO J 25*, 3432-3443.

7.   Kenny, B., Taylor, S., and Holland, I. B. (1992) Identification of individual amino acids required for secretion within the haemolysin (HlyA) C-terminal targeting region, *Mol Microbiol 6*, 1477-1489.

8.   Marsischky, G., and LaBaer, J. (2004) Many paths to many clones: a comparative look at high-throughput cloning methods, *Genome Res 14*, 2020-2028.

9.   Bakkes, P. J., Jenewein, S., Smits, S. H., Holland, I. B., and Schmitt, L. (2010) The rate of folding dictates substrate secretion by the Escherichia coli hemolysin type 1 secretion system, *J Biol Chem 285*, 40573-40580.

10.   Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680-685.

11.   Chan, K. M., Delfert, D., and Junger, K. D. (1986) A direct colorimetric assay for Ca2+ - stimulated ATPase activity, *Anal Biochem 157*, 375-380.

12.   Kennedy, K. A., and Traxler, B. (1999) MalK forms a dimer independent of its assembly into the MalFGK2 ATP-binding cassette transporter of Escherichia coli, *J Biol Chem 274*, 6259-6264.

13.   Nikaido, K., and Ames, G. F. (1999) One intact ATP-binding subunit is sufficient to support ATP hydrolysis and translocation in an ABC transporter, the histidine permease, *J Biol Chem 274*, 26727-26735.

14.   Janas, E., Hofacker, M., Chen, M., Gompf, S., van der Does, C., and Tampe, R. (2003) The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p, *J Biol Chem 278*, 26862-26869.

15.   Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B., and Schmitt, L. (2003) A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A, *J Mol Biol 327*, 1169-1179.

16.   Aurade, R. M., Jayalakshmi, S. K., and Sreeramulu, K. (2010) P-glycoprotein ATPase from the resistant pest, Helicoverpa armigera: purification, characterization and effect of various insecticides on its transport function, *Biochim Biophys Acta 1798*, 1135-1143.

17. Kiefer, H. (2003) In vitro folding of alpha-helical membrane proteins, *Biochim Biophys Acta 1610*, 57-62.

18. Sharom, F. J. (1995) Characterization and functional reconstitution of the multidrug transporter, *J Bioenerg Biomembr 27*, 15-22.

19. Davidson, A. L., Shuman, H. A., and Nikaido, H. (1992) Mechanism of maltose transport in Escherichia coli: transmembrane signaling by periplasmic binding proteins, *Proc Natl Acad Sci U S A 89*, 2360-2364.

20. Patzlaff, J. S., van der Heide, T., and Poolman, B. (2003) The ATP/substrate stoichiometry of the ATP-binding cassette (ABC) transporter OpuA, *J Biol Chem 278*, 29546-29551.

Danksagung

An allererster Stelle möchte ich mich hier bei Lutz bedanken. Für die Freiheiten, die Du mir gegeben hast, für die Freude, die Dir ein SDS-Gel bereiten kann, aber auch den Einsatz für mich in jeglicher Hinsicht, egal ob es nur um die Abschlussfeier ging, die Verwaltung oder Messzeiten.

Sander, Du hattest Recht: „Ich hab immer zu Dir gehört!". Wo soll ich da anfangen? Danke für die vielen Stunden Deiner Freizeit, die Du mit mir am Synchrotron verbracht hast, für die Zeit, die Du in den letzten Monaten in meine Arbeit hier gesteckt hast, dafür, dass Du immer nach pragmatischen Lösungen suchst, dass Du ganz im Gegenteil zu mir Deinen Blutdruck im Griff hast, einfach für mittlerweile fast vier lehr-/ erfolgreiche und auch schöne Jahre.

Danke Thorsten! Für jeden *Abstract*, den Du nicht nur in eine allgemein verständliche Sprache gebracht hast, für all die *Sampler*, die ich gebraucht habe, für Deine einfachen Lösungen, dafür, dass Du es geschafft hast eine Laborbank mit mir zu teilen (ich weiß, dass das sicher nicht einfach war und Spuren hinterlassen hat); und nicht zu vergessen: die schönen Cocktailstunden am Abend....

Mir ist bis heute nicht klar, wie ein Mensch so viel Geduld und Ruhe ausstrahlen kann, wenn einem kleine naive Mädels monatelang hinterherlaufen und immer noch nicht den Unterschied zwischen Marker und Laufpuffer kennen? Vielen Dank, Stefan, für Deinen Enthusiasmus und Deine Hingabe für die Arbeit im Labor. Dank Dir habe ich sicher eine der besten Entscheidungen in meinem Leben getroffen.

André Du musst auch mal „Nein" sagen.... Aber ich bin Dir sehr dankbar, dass Dir dieses Wort bisher immer gefehlt hat, wenn ich Dich brauchte.

Vielen Dank, Silke! Immer wenn ich Dich gebraucht habe, warst Du zur Stelle: Korrekturlesen, Kaninchen hüten, mich quer durch das Land fahren, u.v.m., alles kein Problem für Dich gewesen. Ich hoffe, dass das auch so bleiben wird.

Danke, Christian, dass Du wieder ein bisschen Schwung in die Bude gebracht hast. Es macht einen Riesenspaß mit Dir an einem Projekt zu arbeiten!

Nun zum stillen Lexikon in unserem Büro: Nils! Danke, dass Du immer da warst um uns vor groben Unfug zu bewahren. Ich glaube nicht, dass ich mich erinnere mal von Dir keine Antwort bekommen zu haben.

Petra wie Du siehst habe ich doch noch das mit den ATPasen in den Griff bekommen (und einen eigenen Standard hab ich nicht!). Vielen Dank für den ausgleichenden Pol in unserer Ecke. Und das Parkbank Event verlangt noch nach Wiederholung.

Jan und Philipp: Blots und DNA!!!! Beides Dinge, die ich nicht beherrsche! Schön, dass Ihr mir da immer zur Seite standet.

Vielen Dank Frau Blum, dass sie mir den ganzen unwegsamen Verwaltungsaufwand abgenommen haben. Ohne sie wäre ich spätestens nach einem halben Jahr insolvent und ohne Arbeitsvertrag gewesen!

Mein Dank gilt auch allen meinen anderen Kollegen, die vieles erst ermöglicht haben.

Und dann noch die hilfsbereiten, zuverlässigen und allzeit bereiten Partner in Marburg: Danke Marco und Erhard!

Schließlich die Menschen außerhalb der Forschung: Meine Eltern und Geschwistern, sowie mein Freund Kevin: Ohne Eure Unterstützung über die ganzen Jahre hinweg wäre vieles nicht möglich gewesen. DANKE!

Lebenslauf

Persönliche Daten:

Name:	Britta Tschapek
Geburtsdatum, -ort:	16.02.1984, Viersen
Familienstand:	ledig
Staatsangehörigkeit:	deutsch

Hochschulausbildung

11/2007:	Beginn der Promotion am Institut für Biochemie in der Arbeitsgruppe von Prof. Dr. Lutz Schmitt
	Thema: Biophysical characterization of ABC transporters catalyzing import and export function
10/2007:	Abschluss zur Diplom Wirtschaftchemikerin
	Titel der Diplomarbeit:
	Einzelmolekül-Fluoreszenzspektroskopie an der ABC-ATPase OpuAA in Liposomen
	Angefertigt im Institut für molekulare physikalische Chemie der Heinrich Heine Universität Düsseldorf, Prof. Dr. Claus A. M. Seidel
10/2003-10/2007	Studium der Wirtschaftchemie an der Heinrich Heine Universität Düsseldorf.

Schlulausbildung

8/1994-7/2003	Clara-Schumann-Gymnasium Dülken
8/1990-7/1994	Martinschule Süchteln

List of Publications

1. Horn, C., Jenewein, S., Tschapek, B., Bouschen, W., Metzger, S., Bremer, E., and Schmitt, L. (2008) Monitoring conformational changes during the catalytic cycle of OpuAA, the ATPase subunit of the ABC transporter OpuA from Bacillus subtilis, *Biochem J 412*, 233-244.

2. Jumpertz, T., Tschapek, B., Infed, N., Smits, S. H., Ernst, R., and Schmitt, L. (2011) High-throughput evaluation of the critical micelle concentration of detergents, *Anal Biochem 408*, 64-70.

3. Pittelkow, M., Tschapek, B., Smits, S. H., Schmitt, L., and Bremer, E. (2011) Osmoregulatory ABC-transporter OpuB from Bacillus subtilis: crystal structure of the solute receptor OpuBC in complex with choline, *in preparation*.

4. Tschapek, B., Pittelkow, M., Sohn-Bösser, L., Holtmann, G., Smits, S. H., Gohlke, H., Bremer, E., and Schmitt, L. (2011) A single amino acid defines the high affinity state in the glycine betaine binding protein ProX from the hyperthermophile Archaeoglobus fulgidus, *Nature Structural & Molecular Biology under revision*.

5. Tschapek, B., Smits, S. H., and Schmitt, L. (2011) High Throughput analysis of mixed detergent micelles, *J. Am. Chem. Soc under revision*.

6. Schwarz, C. K. W., Tschapek, B., Jumpertz, T., Jenewein, S., Lecher, J., Willbold, D., Panjikar, S., Holland, B., Smits, S. H., and Schmitt, L. (2011) Crystallisation and preliminary X-ray crystallographic studies of an oligomeric species of a refolded C39 peptidase-like domain of the Escherichia coli ABC transporter Haemolysin B., *Acta Cryst. under revison*.

7. Tschapek, B., Schwarz, C. K. W., Jumpertz, T., Jenewein, S., Holland, B., Smits, S. H., and Schmitt, L. (2011) Investigating the ATPase activity of the E. coli transporter HlyB in vitro, *in preparation*.

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 31.01.2011

(Britta Tschapek)