RESEARCH on the structure and function of substrate-binding proteins (SBPs) of ATP Binding Cassette (ABC) transporters has been ongoing for at least four decades. Our understanding of these proteins expanded drastically in the 1980s and 1990s when a large number of SBPs, belonging to different families and having distinct substrate specificities, were crystallized. Florente Quiocho and his group stood for a large fraction of these structures (Quiocho et al., 1974; Jacobson and Quiocho, 1988; Luecke and Quiocho, 1990; Quiocho and Ledvina, 1996). For an overview on substrate-binding proteins, see Chapter 1. In short, the majority of these proteins are made up of two domains connected by a hinge-region, with the ligand binding site buried in between the two domains. SBPs can bind a wide variety of substrates (e.g. amino acids, di- and oligopeptides, carbohydrates, metals, oxyanions and vitamins), although individual proteins are generally very specific. They range in size from roughly 20 to 65 kDa. With regard to peptide-binding proteins, a breakthrough came with the elucidation of the structure of the oligopeptide binding protein from Salmonella typhimurium (Tame et al., 1994, 1996; S. Sleigh et al., 1997). This thesis describes the progress in research on SBPs, with specific attention on the lactococcal oligopeptide binding protein A (OppA), the receptor component of the ABC-transporter oligopeptide permease (Opp).
OppA from *L. lactis* (here termed OppA*) is one of the best-characterized SBPs in terms of biochemistry and enzymology (Lanfermeijer et al., 1999; Picon et al., 2000; Detmers et al., 2000; Lanfermeijer et al., 2000; Doeven et al., 2004, 2008), and now with the completion of this thesis, the data on the function of this protein can be placed in the context of high-resolution structural data. There were major discrepancies between the biochemical data and the binding model derived from the structure of homologous proteins, mostly OppA from *Salmonella typhimurium* (Tame et al., 1994). Since OppA from *S. typhimurium* was initially the only peptide binding protein for which structural data was available, the interpretation of the data of the *L. lactis* OppA* was skewed towards this protein. The *S. typhimurium* protein binds tri- to pentapeptides, possibly up to hexapeptides. The model proposed for OppA* had six residues buried inside the core structure, and additional residues of long oligopeptides (peptides up to 35 residues have been shown to bind to OppA* (Doeven et al., 2004)) sticking out of the binding site and possibly interacting with the surface of the protein. This model explained all biochemical data at the time, but when the structure of AppA from *B. subtilis*, with a bound nonapeptide completely buried inside the protein, was elucidated (Levdikov et al., 2005), alternative possibilities arose. AppA is, as OppA*, lipid-anchored and accessible at the surface of the cell, whereas OppA from *S. typhimurium* is located in the periplasm. A protein in the periplasm would not need to bind long peptides, since these molecules cannot pass the outer membrane. A surface-accessible protein on the other hand, may encounter long peptides. Since AppA could bury a bound nonameric peptide inside the protein, the question arose whether OppA* could do so as well. To reveal whether that was indeed the case, or whether the aforementioned model was still valid, effort was put into crystallizing OppA* to understand how this protein interacts with peptides at a atomic level.

**Chapter 2** presents structures of OppA* in its open-liganded and closed-liganded conformations, together with different ligands. Crystal structures in the open-liganded conformation were obtained with four different synthetic peptides, but only the backbone density of (most of) those peptides could be modeled into the structure. The structures of the open-liganded conformations suggested that peptides did not bind in a defined place in the open-liganded state, but were free to ‘slide’ in the binding site. In its closed-liganded conformation, OppA* was co-crystallized together with an endogenous mixture of ligands, which were co-purified with the protein, as well as with the synthetic nonapeptide bradykinin (RPPGFSFPR). The composition of the endogenous ligands was determined by
mass spectrometry and shown to originate from protein breakdown products in the cytosol of *L. lactis*. The structures of the closed-liganded conformation showed that the peptides were buried within a voluminous cavity of almost 5000 Å³. The peptide termini were not fixed with salt-bridges, and the only interactions between OppA* and the peptide were hydrogen bonds formed with the peptide backbone. Only one well-defined side-chain pocket was found, which was very hydrophobic and which seemed to play a significant role in peptide binding. The mixture of endogenous ligands that was co-crystallized had weak electron density at the peptide termini, indicating that peptides of different lengths were bound within OppA*. This information, together with molecular dynamics simulation of the open-liganded conformation, led to the formulation of a model on how OppA* binds its ligands. The new model states that the peptides can bind in different registers inside a very large cavity that on the basis of its volume could accommodate peptides up to a length of 35 residues. In this model, peptide selection is based on amino-acid composition rather than sequence.

Subsequent research, described in **Chapter 3**, supported the ‘binding in different register’ model. Here, a new crystal form of OppA*, with the nonapeptide SLSQSLSQS present, showed that the peptide was bound in a different register when compared to the structure with bound bradykinin, also a nonapeptide. The SLSQSLSQS peptide was designed to have a hydrophobic residue (leucine) at position six instead of position five as in bradykinin (which has a phenylalanine at position 5). The phenylalanine of bradykinin was located in the only well-defined side-chain pocket of the binding cavity, a pocket that is very hydrophobic. The structure of OppA* with SLSQSLSQS bound showed that the leucine occupied the hydrophobic side-chain pocket, indicating that the entire peptide had shifted register relative to bradykinin. The protein was further studied by mutagenesis of the hydrophobic side-chain pocket and a thermodynamic analysis of peptide binding, using isothermal titration calorimetry and intrinsic protein fluorescence measurements. These experiments revealed that the well-defined hydrophobic side-chain pocket plays a significant role in peptide binding as well as in stabilizing the closed conformation of OppA*.

In **Chapter 4**, the focus is on OpuAC, the SBD of OpuA, the osmoregulatory glycine betaine transporter from *L. lactis*. Within OpuA, OpuAC acts as the initial receptor for glycine betaine binding, which is subsequently transferred to the transmembrane domain of OpuA. OpuAC was crystallized in its open and closed-liganded conformation in a complex with glycine betaine. The binding site is made
up of a so-called tryptophan-prism, which coordinates the quaternary ammonium group of glycine betaine; hydrogen bonds are formed between the protein and the carboxyl-group of the ligand. The open conformation structure had electron density in the ligand binding site, although not of sufficient quality to be unambiguously assigned. It is likely however that the electron density represents copurified endogenous glycine betaine. Upon closure of OpuAC, the entire binding site is formed when the third tryptophan of the trp-prism moves into position. The domain closure also puts the ligand in a perfect position to form three hydrogen bonds, one to the side-chain of His392 and one each to the backbone of Gly437 and Val438. OpuAC was further biochemically characterized by analyzing the binding of substrates, such as proline betaine, proline and carnitine. Even though OpuAC can bind proline and carnitine, albeit with very low affinities (high millimolar range), transport of these substrates was not observed in OpuA reconstituted in proteoliposomes. Interestingly, even though the binding site of OpuAC from \textit{L. lactis} is identical to the one of OpuAC from \textit{B. subtilis}, it binds both glycine betaine and proline betaine with a factor four higher affinity. This indicates that the differences in binding affinities between the two proteins must lie outside of the binding pocket.

In the process of solving the first structure of OppA*, it became apparent that having a method to incorporate selenomethionine (SeMet) into proteins expressed in \textit{L. lactis} would be valuable. \textbf{Chapter 5} presents a detailed protocol that makes SeMet incorporation possible. A high incorporation efficiency was proven both by solving the structure of an open-unliganded conformation of OppA*, using single-wavelength anomalous diffraction (SAD) data from the incorporated selenium atoms, as well as via mass spectrometry by comparing the molecular weight of OppA* and OpuA with and without the incorporation of SeMet. It was found that the incorporation efficiency was $>90\%$.

The research presented in this thesis helps understanding how SBPs bind their ligands, in particular the mode of binding of oligopeptides. The work forms part of a greater puzzle, that of our understanding of the complete translocation cycle of ABC-transporters. In recent years, the field has taken several leaps ahead, especially with the elucidation of several crystal structures of complete ABC-transporters. The first structure was that of the vitamin B12 importer BtuCD, which was crystallized in its nucleotide free state (Locher et al., 2002). This provided unprecedented knowledge on how the NBDs interact with the TMD, as well as information on the mechanism of vitamin B12 transport. Later on the structures of
other uptake systems were solved, such as the molybdate transporter ModBC in its inward-facing resting state (Hollenstein et al., 2007) and the maltose transporter MalFGK2 in both its inward-facing resting state as well as in an outward-facing catalytic intermediate, with MBP bound and the maltose delivered to a binding site in the TMD (Oldham et al., 2007; Khare et al., 2009). BtuCD in complex with its SBP BtuF in a catalytic intermediate state and the methionine transporter MetNI in its inward-facing resting state were also elucidated (Hvorup et al., 2007; Kadaba et al., 2008). Information on ABC-type efflux systems has become available with the elucidation of the structures of Sav1866 and P-glycoprotein (Dawson and Locher, 2006; Aller et al., 2009). Sav1866 is a bacterial homologue of P-glycoprotein (P-gp) and is probably involved in drug efflux. Both proteins provide preliminary understanding of the structural basis for drug binding in MDR (multi drug resistance) type efflux systems, which includes hydrophobic and aromatic interactions between protein and ligand. P-gp was crystallized both in its apo state and together with two different cyclic peptide inhibitors, all in the inward facing conformation. The structure represents an initial stage of the translocation cycle.

These structures together with a wealth of biochemical data give a more and more complete picture how these proteins work. Three types of ABC-transporters exist: ABC exporters, Type I ABC importers and Type II ABC importers. ABC exporters are found in all genomes that have so far been sequenced, and have diverse functions. Many, like Sav1866 and P-gp are involved in multi-drug extrusion. ABC importers have so far only been found in prokaryots. Type I ABC importers generally contain a total of 12 transmembrane helices (TM), and transports ions, sugars, amino acids and other substrates (Davidson et al., 2008). Type II ABC importers generally import larger substrates than type I, and are also larger proteins, with a total of 20 TMs. A mechanistic model based on type I ABC-importers (to which ModBC, MalFGK2 as well as Opp and OpuA belong) is presented in figure 6.1.

Perspectives

We now have a good understanding on how peptide binding to OppA* works. Future research should be focused on understanding how the entire Opp system functions. In that area there are numerous open questions. We still do not know how the substrate is transferred from OppA to OppBC, or how the peptides are translocated. Does the N- or the C-terminal of the peptide go first, or are they translocated simultaneously? The peptides transported by Opp can be long enough
Figure 6.1. Schematic overview of a type I ABC-importer transport cycle. The model is based on the structures of ModBC and the different conformations of MalFGK$_2$ (Hollenstein et al., 2007; Oldham et al., 2007; Khare et al., 2009). Some ABC-transporters display a basal ATPase activity (shown within the striped box), although the biological relevance of this futile cycle is not clear. Some ABC-transporters, like OpuA, has a very tightly controlled ATP-hydrolysis, and subsequently does not have this basal ATPase activity (Patzlaff et al., 2003).

Figure 6.1. Schematic overview of a type I ABC-importer transport cycle. The model is based on the structures of ModBC and the different conformations of MalFGK$_2$ (Hollenstein et al., 2007; Oldham et al., 2007; Khare et al., 2009). Some ABC-transporters display a basal ATPase activity (shown within the striped box), although the biological relevance of this futile cycle is not clear. Some ABC-transporters, like OpuA, has a very tightly controlled ATP-hydrolysis, and subsequently does not have this basal ATPase activity (Patzlaff et al., 2003).

Figure 6.1. Schematic overview of a type I ABC-importer transport cycle. The model is based on the structures of ModBC and the different conformations of MalFGK$_2$ (Hollenstein et al., 2007; Oldham et al., 2007; Khare et al., 2009). Some ABC-transporters display a basal ATPase activity (shown within the striped box), although the biological relevance of this futile cycle is not clear. Some ABC-transporters, like OpuA, has a very tightly controlled ATP-hydrolysis, and subsequently does not have this basal ATPase activity (Patzlaff et al., 2003).

Figure 6.1. Schematic overview of a type I ABC-importer transport cycle. The model is based on the structures of ModBC and the different conformations of MalFGK$_2$ (Hollenstein et al., 2007; Oldham et al., 2007; Khare et al., 2009). Some ABC-transporters display a basal ATPase activity (shown within the striped box), although the biological relevance of this futile cycle is not clear. Some ABC-transporters, like OpuA, has a very tightly controlled ATP-hydrolysis, and subsequently does not have this basal ATPase activity (Patzlaff et al., 2003).

Figure 6.1. Schematic overview of a type I ABC-importer transport cycle. The model is based on the structures of ModBC and the different conformations of MalFGK$_2$ (Hollenstein et al., 2007; Oldham et al., 2007; Khare et al., 2009). Some ABC-transporters display a basal ATPase activity (shown within the striped box), although the biological relevance of this futile cycle is not clear. Some ABC-transporters, like OpuA, has a very tightly controlled ATP-hydrolysis, and subsequently does not have this basal ATPase activity (Patzlaff et al., 2003).

to span the membrane, meaning that there is a possibility for one peptide termini to have translocated while the other termini is still attached to OppA. Is this how translocation proceeds or is the entire substrate transferred from OppA to OppBC before the actual translocation proceeds? Does OppBC have a defined binding site for the oligopeptide, or is the peptide just translocated through a ‘non-sticky’ tunnel? Even though the cavity of OppA is large enough to theoretically accommodate peptides up to 35 residues long, does this actually occur, or do long peptides stick out of the cavity and interact with the surface of the protein, as previously proposed for peptides longer than six residues (Detmers et al., 2000)? Does OppA stay attached to the TMD until the translocation of peptide across the
membrane has taken place, or does it dissociate as soon as the initial binding to the TMD has occurred? Previous research has shown that the entire Opp complex can be purified and reconstituted into lipid vesicles (Doeven et al., 2004) and thus many of the heretofore posed questions can in principle be addressed. Also methods have been developed to study protein-protein interactions in proteoliposomes by using fluorescence (cross)-correlation measurements (Doeven et al., 2008).

Other outstanding questions are related to the energetics of the system. One of the more important ones would be to understand how much ATP is required for the transport of peptides. Is the same amount of ATP required for the transport of a nonameric peptide as it is for a 35 residue long peptide? It seems unlikely, and uptake experiments with radiolabeled peptides together with measurements of ATP consumption, as carried out for OpuA (Patzlaff et al., 2003) could in this issue. Further understanding of how Opp works would also be gained from crystallizing the entire complex, although that currently still remains a very laborious process for membrane proteins.