Chapter 1

Part 1. Introduction to proteins: physiology, structure and mechanism.

1. Proteins and protein structures

The word “protein” is derived from the Greek word prôtos, meaning “primary” or “first rank of importance.” Proteins form the very basis of life. They regulate a variety of activities in all known organisms on all levels, from replication of the genetic code to transporting oxygen, and are generally responsible for regulating the cellular machinery and consequently, the phenotype of an organism. Proteins accomplish their task by an incredible fine-tuning of interactions between various substrates such as DNA and RNA, metabolites and other proteins. Knowing the structure of the protein, we can probe for its function and potentially apply the new knowledge to various genome projects, such as mapping the functions of proteins in metabolic pathways for whole genomes and deducing evolutionary relationships.

Protein structural organization

Proteins have several different levels of organization. They become highly organized and efficient biological machines through many types of ionic and molecular interactions within the protein itself. The first level of protein structure is its primary structure. The primary structure of a protein is the linear sequence of its constituent amino acids. The next level of organization is the secondary structure of the protein.

Secondary structure is a regular structure brought about via hydrogen bonding mainly within the peptide backbone. The most common secondary structure elements in proteins are the alpha helix and the beta sheet. In membrane proteins, long membrane-spanning alpha
Introduction

Helices and beta sheets are the essential components in making the next level of folding, the tertiary structure. Covalent interactions between cysteine groups, noncovalent electrostatic interactions between polar groups, and Van der Waals interactions between non-polar R groups are very common in tertiary structures. A major driving force in determining the tertiary structure of globular proteins is the hydrophobic effect. The polypeptide chain generally folds such that the side chains of the non-polar amino acids are "buried" within the structure and the side chains of the polar residues are exposed on the outer surface.

Quaternary Structure is the combination of two or more chains or subunits, to form a complete unit. The interactions between the chains are not different from those in tertiary structure, but are distinguished only by being interchain rather than intrachain. Many membrane proteins are multisubunit complexes, and all proteins discussed in this theses are examples. In some cases protein subunits can form flexible, multimeric rings of over 10 copies, and these are the perfect examples where EM can make a major contribution in solving structures, as will be presented in this thesis.

1.2 Membrane proteins

Membrane proteins play many crucial roles in the cell and take center stage in processes ranging from basic small-molecule transport to sophisticated signaling pathways. They are also prime contemporary or future drug targets, and it has been estimated that more than half of all drugs currently on the market are directed against membrane proteins (Klabunde et al. 2002). It is still frustratingly hard to obtain high-resolution three-dimensional (3D) structures of membrane proteins, and they represent less than 1% of the structures in the Protein Data Bank (Berman et al. 2000). Even if the number of experimentally known membrane protein structures is on the rise (White 2004; Oberai et al. 2006), methods to predict their topology (i.e., the transmembrane segments and their in-out orientation across the membrane) and fold type from the amino acid sequence will be needed for many years to come.

As mentioned before, α-helices and β-barrels are important in integral membrane proteins and many of them have two basic architectures: an α-helix bundle or a β-barrel sheet. Helix-bundle proteins are found in all cellular membranes, except possibly the bacterial outer membrane, and represent an estimated 20% to 25% of all open reading frames (ORFs) in fully
sequenced genomes (Keseler et al. 2005). The number of β-barrel membrane proteins is more uncertain because they are more difficult to identify by sequence gazing; for bacteria, a rough estimate, based on the fact that all known β-barrel proteins are in the outer membrane and hence are made with an N-terminal signal peptide, suggests that they account for no more than a few percent of all ORFs. The EcoCyc database currently lists 58 outer membrane and 511 inner membrane proteins out of a total of 4332 proteins; the number of inner membrane proteins in *Escherichia coli* has been estimated to be close to 1000 (Krogh et al. 2001). Because all current membrane protein topology and structure prediction schemes first seek to identify the transmembrane segments, they are obviously quite different, and their variations depend on the class of protein for which they are designed. In general long hydrophobic transmembrane helices are easier to recognize in an amino acid sequence than the much shorter and less hydrophobic transmembrane β-strands, and partly for this reason, much more bioinformatics work has been devoted to the helix bundle proteins—another instance of the well-known dictum “always go for the easy problems.”

### 1.3 Structure

Membrane protein structure determination is proceeding and exciting, driven by the hope that structures can connect decades of biochemical and biophysical observations to protein function and mechanism. As science as a whole unravels increasingly complex and finely tuned functionality, can structure determination continue to keep pace, or will new methodologies and perspectives be required? Evolution has exploited three categories of molecular-level organization in order to achieve efficient and diverse membrane protein functions: structure, molecular dynamics, and environmental constraints. As suggested schematically in Figure 2, they are inextricably linked, each influencing the other, collectively dictating membrane protein function. By recasting the structure-function relationship in this way, we suggest that a comprehensive view of membrane protein function may be more readily achieved (Sachs et al. 2006).
Introduction

Any model for predicting function should start with a structure. Static structures can elegantly illuminate the structure-function relationship. Nowadays, articles presenting atomic structures include astonishing details including protein oligomerization and complexation, lipidation, glycosylation, and the presence of structural water and lipid molecules, suggesting that future models may need to account for a high level of complexity and structural variability. Oligomerization is increasingly seen as a common motif in membrane proteins (Engelman et al. 2005). As the authors point out, distinguishing between dimers and higher order oligomers is far from trivial. Although this difficulty may be primarily due to experimental limitations such as dissociation by detergent, it may also reflect true biological variability. In most studied cases, oligomerization serves a functional role, and this function most likely drove its evolution.

1.4 Protein-protein interactions

The final step on the structure prediction ladder is the prediction of quaternary structure, i.e., protein-protein interactions. This is especially pertinent for membrane proteins because membrane-integral protein domains in most cases seem to be encoded by separate polypeptides rather than as multidomain polypeptides as often found in globular proteins (Liu et al. 2004). Large-scale experimental protein-protein interaction studies tend to ignore membrane proteins, although some data are now starting to appear in the literature (Gavin et al. 2002; Stenberg et al. 2005; Lasserre et al. 2006).
The current state of predicting interactions between membrane proteins may be summarized in a few words: much remains to be done. For example, a recent attempt to predict interacting proteins in the *Saccharomyces cerevisiae* membrane proteome by integrating data, such as amino acid sequence, annotated function, subcellular localization, mRNA and protein abundance, transcriptional co-regulation, and gene knock-out phenotype, resulted in a predictor that could only identify 40% of 304 experimentally well-documented gold standard interactions while minimizing the number of false-positive predictions (Xia et al. 2006).
Part 2. Structure investigation of matrix and membranes proteins from different organism by Electron Microscopy

2 The peroxisome

2.1 The peroxisome: structure, function, and peroxisomal biogenesis

The peroxisome is a single-membrane organelle ubiquitous in nearly all eukaryotic cells. It is the most recently discovered subcellular organelle (De Duve and Baudhuin 1966). Peroxisomes, together with glyoxysomes of plants and glycosomes of trypanosomes, belong to the microbody family, which has gained a great functional diversity, from yeast to humans. Unlike chloroplasts and mitochondria, these organelles do not contain endogenous DNA, and hence all their proteins are nuclear encoded, synthesized in the cytosol, and post-translationally imported into the peroxisome (Lazarow and Fujiki 1985). Peroxisomes are very dynamic and display large plasticity in response to changes in environmental conditions. Under normal growth condition, the mature peroxisome presents a globular shape with a diameter of 0.1-1 µm (Fig. 2; van den Bosch et al. 1992).

Figure 2. An electron micrograph of a Saccharomyces cerevisiae cell during restoration of peroxisome biogenesis shows a cluster of peroxisomes (PO) and their distribution relative to the lipid droplets (V), endoplasmic reticulum (ER), mitochondria (M) and nucleus (N).
The peroxisomal protein content is variable among species, and therefore, the metabolic functions can be adjusted according to cellular needs. Peroxisomes in fungi and plants play a central role in the β-oxidation of fatty acids and the detoxification of the hydrogen peroxide thereby produced (Poirier et al. 2006). The importance of the peroxisome is underscored by the existence of numerous human genetic disorders. Peroxisomal disorders are defined as a group of human inherited diseases in which either the peroxisome biogenesis or one or more peroxisomal functions are impaired.

The discovery that peroxisome proliferation can be induced in bakers’ yeast *S. cerevisiae* by oleic acid was the start of a period of extensive research on the biogenesis of the organelle in yeast (Veenhuis et al. 1987). When cells were grown on oleic acid as the sole carbon source, peroxisomes become essential for growth because they represent the exclusive site of fatty acid degradation in yeast. This important discovery allowed the screening of *pex* mutants affected in the biogenesis of peroxisomes (Erdmann et al. 1989; Erdmann et al. 1991) with *PEX* being the acronym for the corresponding gene, the gene products were named peroxins (Distel et al. 1996). Peroxins are involved in ways in the peroxisomal development: 1) formation of the peroxisomal membrane, 2) peroxisome proliferation and 3) compartmentalization of peroxisomal matrix proteins.

Important studies also involved the emerge of peroxisomal membrane, which has been a point of discussion for a long time. First data suggested that peroxisomes generate by budding from the endoplasmic reticulum (ER) (Novikoff and Shin 1964), which were based on ultrastructural studies using electron microscopy. Later, Lazarow and Fujiki (1985) demonstrated that peroxisomal matrix proteins are synthesized on free ribosomes in the cytosol and these proteins are imported posttranslationally in pre-existing peroxisomes. Based on these results, a “growth and division model” was proposed (ER-independent model), which predicted that peroxisomes multiply autonomously like mitochondria or chloroplasts. More recent studies, based mainly on real-time fluorescence microscopy combined with biochemical data, provided evidence for the ER being the source of origin of peroxisomal membranes, at least during de novo formation (Van der Zand et al. 2006; Hoepfner et al. 2005). This process requires the membrane protein Pex3p, which is localized at first in the ER. It concentrates in foci that bud off in a Pex19p-dependent manner which mature to functional peroxisomes. Based on data from plant and mammalian cells, peroxisomes are believed to constitute a semi-autonomous part of the secretory pathway (Hoepfner et al.
Introduction

2005). However, in yeast new peroxisomes are believed to arise primarily by duplication of the pre-existing peroxisomes (Mottley and Hettema, 2007; Nagotu et al. 2007). In conclusion, we can say that peroxisomes contain an elaborate fission and proliferation machinery. Pex11p was among the first components of this device which were discovered, and together with Pex25p and Pex27p play a role in controlling size and number of *S. cerevisiae* peroxisomes. The scission of peroxisomal membranes has recently been connected to the function of the dynamin-related proteins (DLPs) Vps1p and Dnm1p, as well as the Dnm1p-anchoring protein Fis1p, all of which are also involved in fission processes of mitochondria (Van der Zand et al. 2006). The peroxisomes are still clustered after the fission event. These clusters are thought to be dissociated by the membrane proteins Pex28p and Pex29p of *S. cerevisiae*. Pex30p, Pex31p and Pex32p in *S. cerevisiae* further regulate the size of peroxisomes. These peroxins act downstream of Pex28p and Pex29p, but not much is known about their molecular function.

Last but not least Inp1p regulates the movement of peroxisomes during cell division and thus peroxisomal inheritance in yeast. It is supposed to link peroxisomes to a cortical anchor to retain them in the mother cell and bud. Inp2p is a peroxisomal receptor for Myo2p which moves peroxisomes along polarized actin filaments into the bud of the dividing cell (Fagarasanu et al. 2006).

2.2 Peroxisomal protein targeting and sorting

*Cargo recognition and peroxisomal targeting*

The peroxisomal matrix enzymes are translated by polyribosomes and translocated into the peroxisomal matrix by specific targeting signals. Most of the peroxisomal matrix proteins are equipped with a C-terminal targeting motif, which has been classified as the targeting signal 1 (PTS1). The PTS1 motif consists of just three amino acids—SKL—or a conservative variant (Subramani et al. 2000; Gould and Collins 2002). Only some peroxisomal enzymes use a different targeting signal, the amino-terminally located PTS2. Import of PTS1 and PTS2 proteins requires around 20 *PEX* genes and their products; the peroxins (Table 1).
### Table 1. Features of known peroxins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pex1p</strong></td>
<td>A large (100-150 kDa) AAA ATPase in yeast and humans. Interacts with Pex6p and other peroxins. Defects in Pex1p are by far the most common cause of PBDs (peroxisome biogenesis disorders).</td>
</tr>
<tr>
<td><strong>Pex2p</strong></td>
<td>An ~40-kDa integral PMP (peroxisomal membrane protein) with a carboxy-terminal, cytosolically exposed zinc RING domain. Has been identified in yeasts and humans, interacts with Pex10p and is defective in casual genes for CG (complementation group) 10 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex3p</strong></td>
<td>A ~40-kDa protein which traffics through the ER; membrane receptor for PMP in yeast and humans that binds Pex19p and is defective in CG12 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex4p</strong></td>
<td>A small (20–24 kDa) peroxisome-associated ubiquitin-conjugating enzyme that interacts with Pex22p. Has been identified in several yeast species, but so far there is no report of Pex4p in any metazoan.</td>
</tr>
<tr>
<td><strong>Pex5p</strong></td>
<td>An ~70-kDa, predominantly cytosolic/partly peroxisomal protein that is found from yeasts to humans. Contains a PTS1-binding domain in its carboxy-terminal, tetra-tripeptide-repeat-containing half, interacts with several peroxins (Pex8p, Pex10p, Pex12p, Pex13p and Pex14p) and is defective in CG2 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex6p</strong></td>
<td>An ~100-kDa AAA ATPase found in all eukaryotes. Interacts with Pex1p and is defective in CG4 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex7p</strong></td>
<td>An ~40-kDa, WD40-repeat-containing protein that binds the PTS2. Defective in CG11 of the PBDs. Pex7p is conserved from yeast to humans but is absent in the nematode Caenorhabditis elegans.</td>
</tr>
<tr>
<td><strong>Pex8p</strong></td>
<td>A variably sized (60–80 kDa), Pex8p is the sole peroxin that resides in the peroxisome matrix; it interacts with Pex5p. It is found in yeast and other fungi, but has not been identified in higher eukaryotes.</td>
</tr>
<tr>
<td><strong>Pex10p</strong></td>
<td>An ~35-kDa integral PMP, RING finger protein; E3 ligase for Ubc4p dependent ubiquitination of Pex5p. Has been identified in yeasts, plants, fungi and humans, interacts with Pex2p, Pex5p and Pex12p, and is defective in CG7 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex12p</strong></td>
<td>An ~40-kDa integral PMP, RING finger protein. It has been identified in yeasts and humans, interacts with Pex5p and Pex10p, and is defective in CG3 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex13p</strong></td>
<td>An ~44-kDa integral PMP with a carboxy-terminal, cytosol exposed SH3 domain. Has been identified in yeasts and humans, interacts with Pex5p and Pex14p, and is defective in CG13 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex14p</strong></td>
<td>An ~40-kDa PMP, receptor docking complex component. Present in yeasts and humans; binds to Pex5p, Pex13p and Pex17p.</td>
</tr>
<tr>
<td><strong>Pex15p</strong></td>
<td>A 44-kDa integral PMP.</td>
</tr>
<tr>
<td><strong>Pex16p</strong></td>
<td>In humans, Pex16p is a 36-kDa PMP that traffics through the ER; membrane receptor for PMP recruitment in mammalian cells.</td>
</tr>
<tr>
<td><strong>Pex17p</strong></td>
<td>An ~25-kDa integral PMP that interacts with Pex14p. Pex17p is the third component of the proposed receptor docking complex. It has been identified in yeast and other fungi, but so far has not been detected in higher eukaryotes, including humans.</td>
</tr>
<tr>
<td><strong>Pex18p</strong></td>
<td>A 31-kDa soluble protein involved only in PTS2-protein import. It is highly similar to Pex21p, and might act as a Pex7p chaperone. Identified only in S. cerevisiae.</td>
</tr>
<tr>
<td><strong>Pex19p</strong></td>
<td>A 33-kDa farnesylated protein of yeasts and humans. Predominantly cytosolic/partly peroxisomal, binds all known integral PMPs and recognizes many PMP-targeting signals. Is defective in CG14 of the PBDs.</td>
</tr>
<tr>
<td><strong>Pex20p</strong></td>
<td>A 46-kDa soluble protein involved only in PTS2-protein import. Co-receptor required for Pex7p binding cargo; S. cerevisiae has functional homologs Pex21p/Pex18p instead.</td>
</tr>
</tbody>
</table>
PTS1 proteins are recognized by Pex5p, whereas PTS2 proteins are recognized by Pex7p. In mammals, the long isoform of Pex5p (Pex5L) is also required for PTS2-protein import, as it binds to Pex7p and is required for Pex7p transit to peroxisomes. These observations indicate that Pex5p is the direct or indirect import receptor for all newly synthesized and folded peroxisomal matrix enzymes (nsPMEs) in these species. The matrix proteins with PTSs are recognized in the cytosol. Pex5p is a predominantly cytosolic, partly peroxisomal protein that cycles between the cytosol and peroxisomes, which indicates that peroxisomal-protein import involves a complex interplay of cytosolic and peroxisomal events.

PTS2, the second peroxisomal targeting signal is located near the amino-terminus and consists of the consensus sequence RLXXXXX(H/Q)L (Lazarow 2006). In yeasts, only a small number of matrix proteins contain a PTS2. Many more PTS2 proteins contain plants
Like Pex5p, Pex7p shuttles between the cytosol and peroxisome during cargo translocation. Pex5p is known to be an globular protein, probably oligomer. Fluorescent measurement revealed that the pH affects the oligomeric state of Pex5p. Upon shifting the pH from 6.0 to pH 7.2 the Pex5p from monomers forms larger oligomeric forms (Boteva et al. 2003).

**Receptor –cargo docking to the peroxisomal membrane**

After the protein-receptor interaction, the protein complex will interact with docking partners at the peroxisomal membrane (Purdue et al. 2001). At least 12 of the known membrane associated peroxins are directly involved in the import process for peroxisomal matrix proteins. In *S. cerevisiae* the docking complex is comprised of Pex13p and Pex14p, and also Pex17p. Pex17p is a peroxisomal membrane protein (PMP) which associates with Pex14p in a tight core complex. It has been reported that Pex17p possibly interacts through coiled-coil motifs found in several families of proteins (Elgersma et al. 1996). The fact that Pex17p interacts with Pex14p (Huhse et al. 1998; Snyder et al. 1999), led to the conclusion that these three peroxins might form a receptor-docking complex. Pex13p is an integral PMP whose amino- and carboxyl-termini both extend into the cytosol. In *S. cerevisiae* the amino-terminal domain binds the PTS2 receptor Pex7p. The carboxy-terminal region contains a Src-homology-3 (SH3) domain which directly binds the PTS1 receptor Pex5p as well as Pex14p (Pires et al. 2003). Pex14p is a peroxisomal membrane protein involved in the docking of Pex5p. In addition, Pex5p loaded with cargo exhibits a higher binding affinity to Pex14p than to Pex13p (Urquhart et al. 2000). Pex14p is therefore believed to be the docking-site for Pex5p.

**Translation, receptor recycling and cargo release**

Until now, there was much discussion in the field of peroxisomes about the function of the receptor-cargo complex. It can either enter the peroxisome lumen, as proposed in the “extended shuttle hypothesis” (Fig. 3), or remaining embedded in the peroxisome membrane. In the latter case the cargo can be released into the peroxisome matrix with the receptor remaining protease protected, according to the “simple shuttle hypothesis” (Kunau 2001; Smith and Schnell 2001).
Introduction

Figure 3. Overview of peroxisomal matrix protein import. Matrix proteins have either a PTS1 or PTS2 targeting signal that binds to acytosolic receptor, Pex5p and Pex7p respectively (with or without co-receptors) which targets the protein to the docking complex (Pex13p, Pex14p and also Pex17p in yeast). The receptors translocate with their cargo into the peroxisome lumen, release their cargo and are recycled back into the cytosol. The RING-finger complex (Pex2p, Pex1p0 and Pex12p) associates with the docking complex via Pex8p (in yeast) and is required for receptor recycling. Pex10p is linked to the E2 ubiquitin conjugating enzyme, Pex4p, which is anchored to the membrane by Pex22p. Pex5p release from the membrane requires Pex4p-dependent monoubiquitination, which is ATP-dependent. The RING-finger peroxins are the putative E3 ligases for this process. The AAA-proteins Pex1p and Pex6p are also involved in receptor release and are attached to the peroxisomal membrane via Pex15p in yeast and Pex26p in mammals. Little is known about the recycling of Pex7p, although two of its co-receptors, Pex18p and Pex20p, have been shown to be ubiquitinated (Brown and Baker 2008).

However, a fraction of Pex5p is found in the lumen of peroxisomes which gave rise to the “extended shuttle hypothesis” of peroxisomal protein import (Van der Klei and Veenhuis 1996). This model suggests that the PTS receptors do not release the cargo proteins after the docking step but instead enter the peroxisomal matrix together with the cargo proteins. The receptors release their cargo in the peroxisomal lumen and are subsequently recycled to the cytosol. The same hypothesis was found in the mammals (Gouveia et al. 2003; Dammai and
Subramani 2001). Also, in *S. cerevisiae* Pex5p was found to interact with Pex8p, the only known peroxin bound to the trans side of the membrane. This indicates that Pex5p can traverse the membrane or at least be inserted into it to such an extent that it is exposed to the *trans* side.

**Receptor-release to the cytosol**

In addition to the docking complex, multiple Pex proteins participating in the import cycle have been identified. In addition to a role in the docking complex, multiple Pex proteins participating in the import cycle have been identified. The receptor recycling complex involves: (1) the RING finger proteins Pex2p, Pex10p, and Pex12p transiently anchored to the docking complex by Pex8p (in the yeast - Agne et al. 2003), (2) the ubiquitin-conjugating enzyme Pex4p anchored by Pex22p, and (3) a complex of AAA ATPases: Pex1p and Pex6p, anchored to the membrane by Pex15p (in *S. cerevisiae* or Pex26p in mammals).

Ubiquitination of the import receptor (Pex5p) is emerging as an important aspect of import. The ubiquitination of Pex5p or Pex18p has been considered a signal for their export back to the cytosol (Platta et al. 2004; Kiel et al. 2005; Kragt et al. 2005). Monoubiquitination of Pex5p by Pex4p is required for release of Pex5p from peroxisomal membranes (Platta et al. 2007), whereas polyubiquitation of Pex5p by Ubc4p (Kiel et al. 2005; Kragt et al. 2005) and the E3 ligase activity of Pex10p is required for a mechanism to target receptors for degradation when receptor recycling is impaired. However, the detailed machinery for this process has not been elucidated yet. Remarkably, proteins locked in a folded conformation by chemical crosslinking or stabilizing drugs can also be imported into peroxisomes. Moreover, proteins lacking a PTS can be imported into peroxisomes in association with a protein containing a PTS. These facts suggest that proteins may be imported into peroxisomes in a folded or multimeric state, which is different from the mitochondrial and chloroplast protein import process.

Structure determination of peroxisome matrix and membrane complexes is still ongoing research. Several results describing Pex5p and the docking-complex containing Pex14p, Pex17p and Pex13p have been published (Boteva et al. 2003; Agne et al. 2003). Despite the accumulating biochemical knowledge structural studies on Pex protein complexes are lacking almost completely, although some individual proteins, such as Pex5p, have been
Introduction

studied by X-ray diffraction (Gatto et al. 2000). This lack of knowledge stimulated the study of protein complexes by EM and single particle analysis, as we will report in this Thesis. Chapter 2 of this Thesis presents structural data of a complex between Pex5p and Pex20p, involved in peroxisomal matrix protein import from *Hansenula polymorpha* using single particle image analysis. Chapter 3 describes the different biochemical isolation methods together with BN-PAGE (Blue Native Page electrophoresis) and mass spectrometry to obtain one of the most important peroxisomal import machinery multimeric complex: the docking complex containing Pex14p, Pex17p, and a minor fraction of Pex13p.

3. Secondary transporters

3.1 Introduction

All together, archaeal, eukaryotic and bacterial organelles have a large number of integral membrane proteins and protein complexes involved in transport across the membrane (Konings et al. 1996). Most of the transport proteins have recently been classified in different families based on their sequence similarities. This led to the identification of 250 transporter families, of which secondary transporters represent the largest functional category, comprising 85 families (Boush and Saier 2004).

Secondary transporters drive the transport across the membrane by using free energy stored in ion and/or solute gradients. In general, they are classified in three groups based on the mode of energy coupling: 1- uniporters catalyse the translocation of a single solute across the membrane, 2- symporters couple transport of two or more solutes in the same direction, and 3- antiporters couple the movement of solutes in opposite directions. Many of the secondary transporters exchange a substrate at one side of the membrane for another substrate at the other side of the membrane. This exchange mode differs from the antiport mode in that the translocation of the two substrates in the two directions is not obligatorily coupled. Often, exchange is a partial reaction catalyzed by symporters.

3.2 Structure

Secondary transporters are found in the membranes of all biological cells. They come in a great sequence diversity and to date the Transporter Classification system (TC) developed
by Saier and co-workers lists some 85 families in the Porters (uniporters, symporters and antiporters) category (Saier 2000). The 85 families of secondary transporters do not represent as many different global structures. It has been shown by hydropathy profile analysis that many of the families belong to the same structural class (Lolkema et al. 1998) which probably means that they are evolutionary related. The hydropathy profiles report a specific folding pattern of proteins in membranes and are therefore capable of detecting distant evolutionary relationships between protein families.

The folding of secondary transporters is as a bundle of hydrophobic $\alpha$-helices that are more or less oriented perpendicular to the membrane plane axis. At the two sides of the membrane the transmembrane segments are connected by hydrophilic loops of variable lengths. So far we know only four high-resolution structures of secondary transporters. These structures have provided a first glimpse on the structural and mechanistic diversity that may be expected to be present in the many different families of secondary transporters. One of them is the glutamate transporter homologue Glt$_{\text{Ph}}$ from the archaeon *Pyrococcus horikoshii* (Yernool et al. 2004). It has been crystallized and the three-dimensional structure was solved at a resolution of about 2.7 Å (Boudker et al. 2007). The glutamate transporter family is one of the families that forms a separate structural class, also called dicarboxylate/cation symporters. GltT is a secondary transporter which translocates two cations, a proton and a sodium ion, and, therefore, is driven by the proton and sodium ion motive forces (Von Heyne et al. 1991). The DAACS family forms a structural class by itself (ST[4], named Secondary Transporters [4]).

The other two secondary transporters, CitS of *Klebsiella pneumoniae* and GltS of *Escherichia coli*, are both found in the structural class ST[3]. CitS belongs to the 2-hydroxycarboxylate transporter (2HCT) family and it is found in bacteria. CitS of *Klebsiella pneumoniae* is well-studied member of the family which is termed Na$^+$-citrate symporter. GltS instead, from *Escherichia coli* is a secondary transporter that transports glutamate in symport with Na$^+$ ions and belongs to the Glutamate: Na$^+$ (ESS) symporter family (Heyne et al. 1991). 2HCT and ESS family members show highly similar hydropathy profiles which is the base of their classification in one and the same class. Studies of CitS proteins of *Klebsiella pneumoniae* give insight into the structure of class ST[3]. The structural model of the transporters shows a core of two homologous domains consisting of five trans membrane
segments (TMS) each that are connected by a large cytoplasmic loop region. The hydrophathy profile of the protein suggested the presence of 12 putative transmembrane domains (TMSs) that traverse the membrane in a zigzag fashion. However, a direct alignment of the hydropathy profile of the GltS protein family and the profile of the 2-hydroxycarboxylate transporter (2HCT) family revealed a high similarity suggesting a similar folding in the membrane. The CitS protein of *Klebsiella pneumoniae* of the 2HCT family consists of 11 transmembrane segments and comparison of both family profiles suggested that GltS would consist of 10 transmembrane segments with both termini exposed to the periplasm. Accessibility of cysteine residues introduced at the N- and C- termini of GltS and the three C-terminal loops of GltS were in agreement with the model based on the hydropathy profile alignment with 10 transmembrane segments (Sobczak and Lolkema 2003).

### 3.3 Quaternary structure

The structure determination of these proteins is still limited. Some evidence indicates that CitS of *K. pneumoniae* in the 2HCT family and GltS of *E. coli* in the ESS family exist as multimers, probably as a homodimers. The enzymes, solubilized in the detergent Triton X-100, were shown to run as a dimer on blue native PAGE (Kästner et al. 2003; Veenhoff et al. 2002). Moreover, an observation made during the purification of biotin acceptor domain (BAD)-tagged CitS, already strongly suggested the dimeric structure of the protein (Pos et al. 1998).

Electron microscopy (EM) is a proven method to study the structure of (membrane) proteins. EM data indicate that the two transporters from the ESS and 2HCT families, share a similar subunit structure – quaternary structure. There is evidence for differences in size and shape of CitS and GltS (~ 100 kDa) and in size and shape of CitS transporters to which the Biotin Acceptor Domain (BAD) of the oxaloacetate decarboxylase of *Klebsiella pneumoniae* with a mass of 10 kDa was attached.

Chapter 4 of this thesis presents a structural study of the quaternary structure of the secondary transporters GltT, CitS and GltS in which two-dimensional projections maps of several GltT, CitS and GltS particles were analyzed.
4. Secretin of Type IV Pilus system in pathogenic *Neisseria*

4.1 Introduction

*Neisseria gonorrhoeae* and *Neisseria meningitidis* are two pathogenic members of the genus *Neisseria* that cause infections which are associated with significant mortality of their exclusive human hosts. These pathogenic bacteria interact with blood, plasma and exudate fluid. They also adhere to and invade epithelial and endothelial cells (McGee et al. 1983; Van Putten et al. 1998). Several bacterial components are involved in the modulation of pathogen–host cell interactions, including type IV pili, Opa proteins and porins. The expression of pili, emanating from the bacterial surface, appears to be of paramount importance to the pathogenic process. Neisserial pili are ordered arrays of polymerized protein subunits termed pilin.

*Neisseria gonorrhoeae* pilus variation results from homologous recombination between a single complete pilin gene or expression locus and multiple partial pilin gene copies or silent alleles (Haas and Meyer 1986; Swanson et al. 1986; Koomey and Falkow 1987) and an analogous mechanism is thought to operate in *Neisseria meningitidis* pilus variation (Aho et al. 1987; Perry et al. 1988; Blake et al. 1989). Neisserial pilins show a high degree of homology with pilins of other pathogenic Gram-negative bacteria, including *Pseudomonas aeruginosa* (Strom and Lory 1986), *Vibrio cholerae* (Faast et al. 1989), enterotoxigenic *Escherichia coli* (Giron et al. 1994) and enteropathogenic as well as opportunistic pathogens within the genera *Eikenella* (Tønjum et al. 1993) and *Moraxella* (Frøholm and Sletten 1977). Collectively, this family of surface appendages has been termed type-IV pili (Ottow 1975). As in both *Neisseria* species, these related pilus colonization factors appear to be essential for infectivity and disease.

4.2 Type IV pili

Type IV pili are very important in the early stages of infection of human hosts. Type IV pili form filaments of a regularly ordered protein with a length of 1,000-4,000 nm and a diameter of 5 to 6 nm, which protrude from the cell surface of many bacteria of medical, environmental and industrial importance. They are also responsible for attachment to specific
Introduction

epithelial tissue receptors during tissue colonization and mediate a variety of functions, including adhesion, twitching motility, and competence for DNA uptake (Heckels 1989; Virji et al. 1992). The type IV pilus is a helical polymer of pilin protein subunits and is capable of rapid polymerization or depolymerization, generating large motor forces in the process.

Currently, there is a lot of research ongoing concerning the biochemistry, genetics and functions of type IV pili. Collins and co-workers (2005) showed specific interaction between the outer membrane secretin PilQ and the type IV pilus fiber by using transmission electron microscopy. The authors showed that purified type IV pili, to which the purified PilQ oligomer had been added, caused PilQ to be located at one end of the pilus fiber. Determination of the three-dimensional structure of the PilQ-type IV pilus complex at 26 Å resolution showed that the cavity within the protein complex was filled. Craig and co-workers (2006) revealed a detailed Neisseria gonorrhoea Type IV pili structure by quantitative fitting of a 2.3 Å full-length pilin crystal structure into a 12.5 Å resolution 3D reconstruction map solved by cryo-electron microscopy.

Besides that, very little is known about Neisserial pilus biogenesis, and the genes encoding the component proteins are scattered throughout the N. meningitidis and N. gonorrhoeae genomes (Tønjum and Koomey 1997). The only pilus biogenesis proteins that are associated with the outer membrane and that have been characterized to date are PilC, PilP, and PilQ. They are called secretins of Type IV pili.
Identification and individual characterization of pilus biogenesis together with their coordinated functions and interactions in pilus assembly are still poorly understood. One of the integral elements in these processes is the PilQ multimer. PilQ of pathogenic *Neisseria strains* is an abundant outer membrane protein which forms a large multimer composed of 10 to 12 subunits (Drake et al. 1997; Drake and Koomey 1995). The PilQ complex is involved in the surface expression of Tfp, and meningococcal mutants missing or expressing defective PilQ are devoid of pili and pilus-associated phenotypes. A total of 200 to 300 conserved C-terminal residues in the transmembrane region of the PilQ monomer exhibit high sequence

**Figure 4.** Electron micrographs of negatively stained WT type IV pili cells and the PilQ-type IV pilus complex interaction.

(A) A typical Neisseria bacterium with a diameter of about 2 µm with attached pili. The long pili are pointed by black arrows. Scale bar equals 0,5 µm.

(B) Tfp fibers visualized by TEM in uranyl acetate. Several fibers can be seen (pointed by white arrows). Scale bar represents 100 nm.

(C) PilQ-type IV pilus complexes with a distinctive hammer-like side-on view.

4.3 PilQ secretin from *Neisseria species*

Identification and individual characterization of pilus biogenesis together with their coordinated functions and interactions in pilus assembly are still poorly understood. One of the integral elements in these processes is the PilQ multimer. PilQ of pathogenic *Neisseria strains* is an abundant outer membrane protein which forms a large multimer composed of 10 to 12 subunits (Drake et al. 1997; Drake and Koomey 1995). The PilQ complex is involved in the surface expression of Tfp, and meningococcal mutants missing or expressing defective PilQ are devoid of pili and pilus-associated phenotypes. A total of 200 to 300 conserved C-terminal residues in the transmembrane region of the PilQ monomer exhibit high sequence
homology to other secretins and are essential for oligomerization into a functional complex (Guilvout et al. 1999; Nouwen et al. 2000).

Collins and co-workers in 2001 showed that PilQ in *Neisseria meningitidis* forms a large 960-kDa homododecameric complex, which was analyzed by transmission electron microscopy by using negatively stained specimens. The projection map generated from 296 particles had an estimated resolution of 2.6 nm. They indicate that oligomeric PilQ adopts a donut-like structure with an external ring that is 16.5 nm in diameter surrounding a central cavity that is 6.5 nm in diameter. Self-rotation and power spectrum analysis demonstrated the presence of 12-fold rotational symmetry, showing that PilQ is organized as a ring of 12 identical subunits. This symmetry determination is questionable, because the data hardly permit to see consistent features at the subunit level. A few years later, the same authors presented a 3D structure of the PilQ complex visualized by cryo-negative EM staining. The structure shows a large central cavity, which was closed at the poles by ‘plug’ and ‘cap’ features, and four ‘arm’ features at the sides. The results suggested that the PilQ complex is the channel through which the moving pilus fibre (polymerized PilE) is extruded and retracted from the bacterial surface (Collins et al. 2004 and 2005). The dodecameric complex of the PilQ subunits forms an extremely stable structures that is resistant to heat and SDS treatment.

So far, the structure of the outer membrane PilQ secretin complex of *Neisseria meningitidis* is well studied. Despite that, the structure of the PilQ secretins of Type IV pili systems from related species, such as *Neisseria gonorrhoea*, remains unsolved. Therefore, in the last chapter of this Thesis, chapter 5, we have analyzed the PilQ complex from *Neisseria gonorrhoea* by using EM and single particle image analysis. Structural differences of the PilQ secretin system of type IV pili from two pathogenic *Neisseria* strains are also compared and analysed by EM.

5. Application of Electron Microscopy to the Structural Analysis of Proteins

Electron microscopy (EM) is nowadays a wide-spread technique to study protein structure. Over the last two decades the technique of single particle EM has been successfully applied to many types of water-soluble and membrane-bound proteins. This Thesis describes
EM studies performed on matrix, membrane- and membrane-associated complexes from yeasts and bacteria.

5.1 Protein structure determination

Understanding of the processes underlying the passage from a DNA sequence to a protein structure and from a structure to its function and dynamics it is still a demanding process. Among the disciplines in the structural biology field that can contribute to answering these fundamental questions are X-ray crystallography, nuclear magnetic resonance (NMR), electron microscopy (EM) and molecular dynamics (MD). These are the most powerful approaches for macromolecular structure determination. Up till now, X-ray crystallography is one of the most successful tools to solve the structure from small to very large molecules. NMR is the other well-established method for determining atomic structures of relatively small proteins and allows the dynamics of a protein to be measured. Despite progress in understanding crystallization and robotic automation of systematic crystallization approaches, many biological proteins do not arrange themselves in perfectly ordered 3D crystals necessary to perform X-ray diffraction and also many other proteins are simply too large to be approached by NMR spectroscopy.

Electron microscopy has become a progressively more important technique for structural determination of small and large macromolecules. Over the last 3 decades, EM has also evolved into a powerful technique able to deliver three-dimensional (3D) structures of the studied complexes at the higher resolution necessary to visualize structural details of molecules (on the scale of nanometers) rather than of the gross architecture of cells (on the scale of micrometers) (Stahlberg and Walz 2008).

EM structure analysis can be performed on samples from ordered assemblies such as 2D crystals to asymmetric randomly oriented individual particles. The latter objects are rather easy to examine by EM, especially for soluble proteins. Electron crystallography of two-dimensional crystals yields the highest resolution data, since crystals present the largest numbers of proteins in well defined positions and orientations. Until very recently, electron crystallography has been the only EM technique that has reached sufficient resolution to produce atomic models of proteins. The 2D crystals used in electron crystallography are the
reason for the high resolution that can be achieved by this technique, because the crystallization accomplishes the alignment of the molecules, which therefore does not need to be done computationally. Thus, the better the order of the 2D crystals, the higher the resolution that can be achieved. Moreover it has the potential to provide structural information at near-atomic level. Because electron crystallography uses 2D crystals, it has proven particularly useful in structural studies of membrane proteins (Hite et al. 2007), although it was also used to determine the structure of the tubulin dimer (Nogales et al. 1998), light-harvesting complex II (Kühlbrandt et al. 1994), aquaporin (Murata et al. 2000; Ren et al. 2001) and bacteriorhopsin (Henderson et al. 1990). Notably, a recent density map obtained with doublelayered 2D crystals of aquaporin-0 at a resolution of 1.9 Å revealed water molecules in the channel of the protein as well as nine lipid molecules surrounding each monomer (Gonen et al. 2005). Besides electron crystallography, electron tomography is another novel technique to determine the 3D structure of unique objects, which can bring completely new insights into the proteins. Currently, the resolution of 3D reconstructions of biological, beam-sensitive samples rarely exceeds 50 Å because electron tomography still has some limitations. However, it has made stunning progress. In a landmark paper, electron tomography of a vitrified Dictyostelium cell revealed the organization of subcellular structures in the filipodium, including the rough endoplasmatic reticulum and the actin cytoskeleton (Medalia et al. 2002). The tomogram also revealed individual proteasomes, demonstrating that even the structure of macromolecular complexes can be determined in their native environment. Other recent successes of electron tomography include the visualization of the architectures of enveloped viruses (Grünewald et al. 2003), nuclear pore complex (Beck et al. 2007), bacterial cytoskeleton (Kürner et al. 2005), flagellar motor (Murphy et al. 2006), axonemes (Nicastro et al. 2006), magnetosomes (Komeili et al. 2006), and clathrin-coated vesicles (Cheng et al. 2007).

Although lagging behind in resolution with respect to electron crystallography, this technique can reveal structural features and conformational changes that underlie the function of the protein.

**EM specimen preparation**

EM specimen preparation techniques aim to prevent fundamental problems in the radiation sensitivity of the biological material. Biological specimens consist of up to 80%
Chapter 1

water, requiring the samples to be prepared in a way that protects them from structural
damage upon dehydration in the vacuum of the electron microscope. Important issue factor is
that biological specimens also consist mainly of light atoms, and the density of proteins is
very close to that of vitrified ice, making them low-contrast objects.

For individual particles two main specimen preparation methods have been developed
to achieve high contrast. Initially, biological specimens were prepared for EM by negative
staining. Negative staining was introduced by Brenner and Horne at the end of the fifties
(Brenner and Horne 1959). One of the methods of providing high contrast of the image in the
electron microscope is to dry and embed the specimen in a layer of electron-dense heavy
metal salts. The biological material surrounded by a heavy-metal (‘negative stain’) layer, is
considerably increasing the scattering contrast lacking in the absence of stain. This method is
not only very fast and straightforward, but also highly reproducible.

Amongst the most effective heavy metal salts that have been found to be suitable for
this method are uranyl acetate, sodium- and potassium phosphotungstate and ammonium
molybdate. Uranyl oxalate can replace unbuffered uranyl acetate, which has a very low pH, if
proteins turn out to be unstable at low pH, as proposed by the former group leader E.F.J. Van
Bruggen (Mellema et al. 1967).

One of the important aspects, in addition to negative staining is a glow-discharge
device. It is often employed to make the surface of the support film more hydrophilic. It
facilitates the absorption of the protein to the film and gives a better stain distribution. Using
this method, the signal-to-noise ratio is much improved. Negative staining is effective at
revealing the outer surface and therefore the quaternary structure of a protein molecule.
However, there are a few dissadvantages of this method. While negative staining is fast and
easy to use, there are possible distortions of the molecules resulting from the staining/drying
procedure, causing the low attainable resolution of around 20 Å. Because of the limited
penetrating ability of the stain, and the fact that the image is formed mainly by the stain-
protein boundary, the structural information is restricted to topographical features of the
molecule surface.
However, distortions of the molecules can be reduced by cryo-negative staining methods (Adrian et al. 1998; Ohi et al. 2004). This is the second and advanced fast preparation technique for a thin layer of hydrated specimen at liquid nitrogen temperature. Preparation methods to obtain ice-embedded protein specimens were developed in the eighties by Dubochet and colleagues (Adrian et al. 1984). Vitrification of the unstained specimen in its native buffer solution is the method of choice. In the preparation of cryosamples, the specimen is rapidly frozen (by plunging the EM grid into liquid ethane) and becomes thus embedded in a layer of vitreous (amorphous) ice (Dubochet et al. 1988). Then the grid is transferred to the electron microscope in special nitrogen-cooled holders. Although this technique preserves the specimen virtually artifact-free in a near-native environment, images of vitrified specimens are very noisy and have a contrast that is an order of magnitude lower than that of stained specimens.

To prevent proteins and supermolecular assemblies from denaturation and dehydration one can use air-drying in the presence of sustaining media such as glucose (Unwin and Henderson 1975) or trehalose (Hebert et al. 1997). Embedding in sugar in this way generates a non-volatile water-like environment for protein molecules, thus sustaining the specimen and taking advantage of the decrease of radiation damage at low temperature. Unfortunately the contrast using this method is even lower than in ice, since sugar almost matches the density of the protein.

One of the other methods invented by Dubochet and colleagues in 1998, combine the two techniques of vitreous-ice embedding and negative staining. The specimen preparation is very similar to the thin-film vitrification technique used for preparing cryo-EM specimens, except for the addition of a heavy metal salt at a concentration of several percent. The advantages of this method are two-fold. Firstly, the high contrast of cryo-negatively stained sample can be imaged close to focus and thus better preserves the high-resolution information. Secondly, the signal-to-noise ratio is higher, so improving the visualisation of the sample with a high amount of detail.

5.2 Electron Microscopy

The first electron microscope prototype was built in 1931 by the German engineers Ernst Ruska and Max Knoll in Berlin. Although this initial instrument was capable of
magnifying objects by only four hundred times and had a modest resolution of about 200 Å it demonstrated the principles and potency of an entirely different type of microscope. The Siemens company already produced commercial TEM microscopes in 1939, but actually the first practical electron microscope was built at the University of Toronto in 1938, by Eli Franklin Burton and students Cecil Hall, James Hillier, and Albert Prebus. A steadfast search for improvements has perfected the electron microscope of the last 70 years towards a powerful research tool to investigate matter at a resolution beyond 1.0 Å. Currently researchers use various types of microscopes to examine biological materials (such as microorganisms and cells), a variety of large molecules, medical biopsy samples, metals and crystalline structures and the characteristics of various surfaces. The electron microscope is also used extensively for inspection, quality assurance and failure analysis applications in industry, including, in particular, semiconductor device fabrication.

**Image formation**

In a conventional transmission electron microscope, a thin specimen is irradiated with a coherent electron beam of uniform current density. A schematic representation of the image formation in the electron microscope is depicted in Figure 5. Electrons are emitted from the electron gun and the electron beam is deflected through a two or three stage condenser lens system, which ensures that the incident electron beam falling on the specimen is parallel (for simplicity this is not depicted on the picture) with a tunable size. The incident electron beam passes through the specimen by which the electrons are scattered or non-scattered by the atoms of the specimen. Scattered electrons further interact with the magnetic field of the objective lens. The objective lens provides the formation of either an image or diffraction pattern of the specimen. The electron intensity distribution behind the specimen is magnified with a three or four stage lens system and viewed on a fluorescent screen or recorded by cameras. If the electrons are considered to interact with the specimen atoms as elementary particles the contrast in an electron micrograph arises from elastic and inelastic scattering. When the beam of electrons is deflected by the positively charged nucleus without loss of energy, the deviation angle is mostly large and scattered electrons can be removed by the objective aperture. This part of the contrast is called scattering contrast. For thick specimens or for heavy-metal stained specimens it is the major contribution in the final contrast. When the electrons undergo inelastic scattering with loss of energy, the interaction with the lenses is different due to chromatic aberration effects. The result is a blurring of the image and the
consequence of the energy loss is an increase in background noise and radiation damage as a side-effect. The contribution of inelastically scattered electrons can, however, be removed from the final image by filtering, as will be discussed in the next section. Filtering is useful for unstained specimens, but for negatively stained membrane proteins, as studied in this thesis the removal is not crucial for obtaining higher quality data.

When the electron beam is considered as waves, the contrast arises from interference between non-scattered waves without a phase shift and scattered waves with additional phase shift. This is called phase contrast and it can be optimized by a right amount of defocusing of the objective lens and properties of the objective lens such as its spherical aberration constant. Thin specimens or unstained biological specimens are typically weakly scattering and such specimens are considered as phase objects. In practice, the final image always results from a combination of phase contrast and scattering contrast at a ratio depending on the type of specimen.

The final image can be recorded by direct exposure of a photographic emulsion or an image plate or digitally by a CCD camera. The acceleration voltage of up to date routine instruments is 120 to 200 kV. Medium-voltage instruments work at 200-500 kV to provide a better transmission and resolution, and in high voltage electron microscopy (HVEM) the acceleration voltage is in the range 500 kV to 3 MV. The acceleration voltage determines the velocity, wavelength and hence the resolution (ability to distinguish the neighbouring microstructural features) of the microscope.

Electron microscopes can also be used to generate and view diffraction patterns of samples. These are usually from samples that contain a repeated motif, such as 2D crystals (sheets one layer thick) of proteins. In the back focal plane of the objective lens, a diffraction image is formed. Depending then on the positioning of the intermediate lens, the diffraction pattern or the image will be magnified and displayed on the view screen.
Chapter 1

Figure 5. Schematic diagram showing the principle of image formation in an EM. An arrow shows the part of the image which is magnified by the sets of the objective and projection lenses and reprojected on the image plane.

5.3 Image analysis as a method to increase resolution and signal-to-noise ratio

Generally speaking, an electron microscopy image of biological material suffers from a very low signal-to-noise ratio (SNR). The introduction of digital image processing made it possible to determine the 2D and 3D structures of biological molecules from very noisy EM images. The idea of increasing the SNR in electron images started by Aaron Klug, a Nobel prize laureate in 1982, and his colleagues from Cambridge in the sixties. Application to unstained specimens was first discussed by Glaeser in 1971 (Glaeser 1971) and applied by Nigel Unwin and Richard Henderson to bacteriorhodopsin crystals in two classical papers, the most cited piece of work in protein EM (Unwin and Henderson 1975; Henderson and Unwin 1975).

The first attempts to average single particle projections came just at the time of the crystallographic approach in the seventies (Frank 1975), although most of the basic tools were developed in the eighties, such as the use of statistical analysis to separate particle projections...
from different angles (van Heel and Frank 1981) and schemes to obtain 3D reconstructions (Frank et al. 1988). In the single particle averaging approach computer programs are used to average images of equivalent molecules to increase the SNR, but improvement of the SNR depends not only on the number of images that are averaged but also on the correct alignment of the images such that the averaging actually improves the signal. In principle, the attained resolution (the minimum size of the structural detail that can be resolved) depends on the number of images averaged, their homogeneity (they should represent the same protein in the same orientation), and the accuracy of alignment; however, in practice the problem is difficult and remains the subject of vigorous research.

In addition, the computer is used to combine projection images of the molecule in different orientations to calculate a 3D reconstruction, thus overcoming the problem that electron microscopes can only record 2D images. Underlying 3D reconstruction are two implicit notions, namely, that electron microscope images are true projections of the imaged molecule and that all images represent identical molecules.

5.4 Novel developments in electron microscopy

Much of the current and future progress in molecular EM relies on developments in instrumentation, from specimen preparation to data acquisition. There is a lot of recent progress, for instance, the FEI Vitrobot has made it possible for the non-experienced user to obtain, within a short time, usable cryo-EM grids for single-particle EM and electron tomography. In general, automation of procedures of data acquisition and particle picking has tremendously shortened processing time. Most of the progress has been achieved in the instrumentation of transmission electron microscopes driven by the materials sciences. The introduction of highly coherent and bright field emission gun (FEG) electron sources made it possible to record highly defocused images with only a limited loss in resolution. Because recording pictures at higher defocus values boosts the image contrast tremendously, FEG instruments have been crucial for obtaining higher resolution structures of biological specimens with their weak-phase-object characteristics. Stable EM electronics and specimen stages allowed for longer exposure times, which is beneficial for the coherence of the illumination and may result in lower beamdamage per dose (Chen et al. 2008).
Charge-coupled-device (CCD) cameras revolutionized imaging because they allow images to be captured in digital format, avoiding the time-consuming processes of developing and scanning photographic film. Initial CCD cameras had small imaging areas (typically 1K x 1K chips); recently the first camera with an 8K x 8K chip was introduced. Improvement in better data collection strategies are still under development. Together with higher acceleration voltages, energy filters made it possible to image thicker specimens and thus became essential for electron tomography. Operated in the so-called zero-loss mode, the energy filter removes inelastically scattered electrons, which are responsible for much of the noise in electron micrographs, thus improving the SNR of images, and also of electron diffraction patterns (Yonekura et al. 2002). Unstained biological samples are weak phase objects and thus create little image contrast.

The current way to enhance contrast is to take images at high defocus. Similar to the effect of a Zernike phase plate for light microscopy (Zernike 1935; Zernike 1955), the phase-shifted electrons then produce strong phase contrast when they interfere with the non-phase-shifted electrons in the image-recording medium. Such phase plates are currently under development, but the various designs still need to overcome significant technical challenges to become of practical use (Schroeder et al 2008). Another part of progress was done on aberration-corrected electron lenses. It has allowed material scientists to record much higher resolution images. Instruments free of spherical aberration have a clean transfer of features to a very high resolution, but the concomitant loss of contrast for the low-resolution frequencies makes this innovation less useful for biological samples with their low inherent contrast (Evans et al. 2008). The combination of an aberration-corrected imaging system with a phase-shifting device could, however, deliver an instrument with strong contrast and high-resolution transfer characteristics. However, even when combined with a phase-shifting device, the low tolerance of focus variations for an aberration-corrected lens system may prevent its use for imaging large, tilted biological samples. This problem could potentially be addressed by using “spot scan” imaging with dynamic focus adjustment, in which a series of small image frames is made with a focussed electron beam, the actual “spot”. At least for untitled biological samples, an electron optical imaging system that combines lenses corrected for spherical and also chromatic aberration (Haider et al. 2007) with a phase-shifting device, may deliver images of biological specimens of unprecedented quality, potentially revolutionizing molecular EM as we know it today.
Introduction

Outline of this Thesis

The work described in this thesis is mainly performed within the research group Electron Microscopy. The research in this group is concentrating on the structural investigation and characterisation of protein complexes by using single particle electron microscopy. This thesis is focused on investigation of matrix and membrane proteins from different organisms belonging to yeasts or bacteria. Transmission electron microscopy in combination with single particle analysis and proteomics techniques have been the main tools in this study.

Chapter 2 of this thesis deals with a structural characterisation of Pex5p and Pex20p, peroxins involved in peroxisomal matrix protein import in the yeast *Hansenula polymorpha*. EM analysis showed that Pex5p is a tetramer. Upon addition of multimeric Pex20p, the conformation of tetrameric Pex5p changed from a closed conformation with a diameter of 115 Å into an open conformation of 134 Å. The results also indicated that the Pex5p-Pex20p complex was capable to bind native, folded catalase, a peroxisomal PTS1 protein. This suggests that the Pex5p-Pex20p complex may be functional as receptor complex.

Chapter 3 discusses Pex proteins involved in matrix transport over the peroxisomal membrane. We have performed a survey of the major peroxisomal membrane protein complex Pex14p, Pex13p and Pex17p by using Blue-Native polycrylamide gel electrophores (BN-PAGE) in combination with biochemical characterization and mass spectrometry. It appeared that purification of the peroxisomal membranes is crucial, but difficult to achieve by classical methods such as differential and sucrose density centrifugation since cells have a large excess of mitochondrial membranes. The mitochondrial protein complexes obscure the peroxisomal ones in 1D and 2D PAGE analysis and a structural characterisation of the major peroxisomal membrane protein complexes is difficult to achieve.

In chapter 4 the results of an analysis of the secondary transporters GltT, CitS and GltS using single particle electron microscopy is presented. Since the CitS and GltS particles are expected to be at the lower size limit of what is feasible for the technique, GltT protein with known structure was included as a reference particle. Evidence has been presented that the GltT protein assembles into a trimeric structure. EM analysis indicated that the other two
transporters from the ESS and 2HCT families, share a similar subunit structure. Both CitS and GltS (~ 100 kDa) showed to be dimers. Unfortunatelly, because of the small dimensions of proteins (~100 kDa) together with the low signal-to-noise ratio of EM images, and the unstructured detergent shell around membrane proteins, we were unable to reveal much detail in these proteins. Additional data were obtained by applying single particle averaging on proteins of mutants. We could show differences in size and shape of CitS transporters to which the Biotin Acceptor Domain (BAD) of the oxaloacetate decarboxylase of *Klebsiella pneumoniae* with a mass of 10 kDa was attached.

**Chapter 5** focuses on a structural characterization of the Type IV secretion system from prokaryotes. For this investigation, intact membrane structures from *Neisseria species* were studied by electron microscopy. We have concentrated on PilQ, which is a member of the secretin family of outer membrane proteins and is specifically involved in secretion of type IV pili in *Neisseria meningitides* and *Neisseria gonorrhoeae*. We have investigated the quaternary structure of PilQ from *N. gonorrhoeae* analyzed by transmission electron microscopy. Oligomeric PilQ adopts a structure with two rings: an undefined inner and anh outer that is composed of 14 protein densities in *Neisseria gonorrhoeae*. At the periphery, seven external spikes bind to the second ring. In *Neisseria meningitides* the second ring is absent in many of the particles, but when it is present it is composed on 19 copies of an protein. The structure data have been extended by analysing the structures from specific knock-outs of Pili proteins in *Neisseria gonorrhoeae*. In one of them, a change in the second ring plus absence of the spikes was found.
REFERENCES


Glaeser, R.M. (1971) Limitations to significant information in biological electron microscopy as a result of radiation damage. J Ultrastruct.Res. 36, 466-482.


Introduction


Introduction


Introduction


