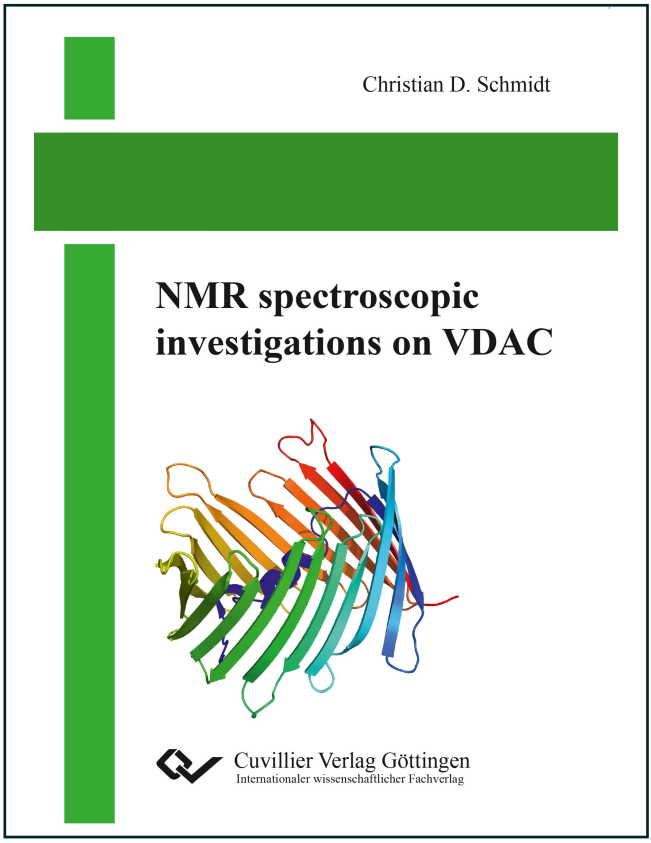

1 Introduction

1.1 Membrane proteins

Membrane proteins are a large subgroup of proteins along with soluble globular proteins, fibrous proteins and disordered proteins. It has been estimated by various methods that about 20 to 30% of all open reading frames (ORFs) over all organisms encode membrane proteins.[1–3] Unlike soluble proteins, that are able to fold with high variability, membrane proteins are only known to occur in two topology categories: α-helical bundles and β-barrels. Membrane proteins give the membranes their characteristical functional properties and allow them to carry out their respective activities. As such, the membrane protein content expressed as mass ratio in a membrane ranges from less than 25% (in the myelin membrane) to up to 75% (in internal membranes of mitochondria and chloroplasts).[4]

α-helical membrane proteins can be found in literally all types of biological membranes: besides occuring in the plasma membrane they are also found in the inner membranes (mitochondria, chloroplasts, endoplasmatic reticulum, peroxisomes and in bacteria). Their function comprises transport processes across membranes, metabolism, cell signaling and regulation processes and as such they are heavily adressed targets for drug discovery.[5]

In contrast to that, β-barrel membrane proteins are only found in the outer membranes of Gram-negative bacteria, scarcely in the cell walls of Gram-positive bacteria and in the outer membranes of mitochondria and chloroplasts. However, in those membranes β-barrel membrane proteins oftentimes make up the majority of the integral membran proteins. The importance of membrane proteins becomes evident with the fact, that they are targeted by more than 50% of small molecule drugs used in the treatment of human diseases: they have been and still are an intensively addressed drug target.[6]

1.1.1 Membrane proteins & NMR

The fact that membrane proteins are such a prominent drug target reasons, why there is also a big interest in the scientific community in determining the 3D structures of membranes proteins. In the Membrane Protein Structures Database of the Stephen White laboratory at University of California Irvine (blanco.biomol.uci.edu/mpstruc/; status of May 2017) 692 unique membrane protein structures are deposited. Of these, only 53 structures were solved with the means of NMR spectroscopy. Until now, X-ray crystallography is still the most commonly applied technique, although structure determination by cryo-EM is becoming more and more the method of choice in recent years.[7] However, the fact that of roughly 130'000 structures in the RCSB Protein Data Bank (www.rcsb.org/pdb; status of April 2017) only about 4130 structures are membrane proteins (which is only 3 %, i.e. far less than 20 to 30 % of all structures, cf. above) shows that all three techniques are experiencing difficulties in structure determination of membrane proteins.

NMR has the unique ability to access not only rigid structures but also dynamics and different states. Furthermore, there is still ongoing advancement in liquid and solid state NMR spectroscopy, e.g. APSY[8] or non-uniform sampling[9] for liquid state NMR or the development of 0.7 mm rotors that allow spinning frequencies of up to 111 kHz and thus the direct detection of protons for solid state NMR. Further increased field strengths of at the moment up to 1 GHz proton resonance frequency raise the overall sensitivity of the spectrometers. Magnets with 1.2 GHz proton resonance frequency are being in development.

That NMR is still a viable tool to determine 3D structures of membrane proteins can be seen by the fact that in 2017 already two structures of membrane proteins or domains of proteins embedded in membranes solved with NMR have been reported: the structure of the C-terminal transmembrane domain of the HDL receptor, SR-BI,[10] and the structure of the epidermal growth factor receptor transmembrane domain dimer.[11] These are faced with 20 structures solved by X-ray diffraction or cryo-EM in 2017 (according to blanco.biomol.uci.edu/mpstruc/; status of May 2017).

1.2 Mitochondria

Mitochondria are organelles found in nearly all eukaryotic organisms (monocercomonoides sp. being the only one known exception[12]) and are the main source of the cells' energy currency by generating ATP through oxidative phosphorylation.[13] They are separated from the cytoplasm of the surrounding cell by two membranes, the inner and the outer mitochondrial membrane (IMM and OMM), which in turn are separated by the inner membrane space (IMS). Following the endosymbiotic hypothesis, the two membranes are leftovers from the mitochondrias' origin as prokaryotic cells.[14] Besides their key function in the energy balance of the cell, mitochondria also play a role in the regulation of cell metabolism, cell-cycle control, development, antiviral responses and cell death.[15]

The IMM contains a series of respiratory enzyme complexes. Electrons passing through these complexes are used to build up a proton gradient across the IMM which is then enabling the ATP synthase to synthesize ATP.[13] Furthermore, the IMM containts the ADP/ATP carrier, which is catalyzing the ADP versus ATP exchange over the IMM and is thus critical for the supply of ATP from the mitochondria to the cytosol.[16,17]

The OMM is populated by four major integral membrane protein families. The translocase of the outer membrane (TOM complex), the sorting and assembly machinery (SAM complex) and the mitochondrial distribution and morphology (Mdm) complex all consist of a core subunit, which is an integral β-barrel membrane protein, and several smaller proteins that regulate the assembly and dynamics of the complexes.[18–22] These three OMM protein families are responsible for the translocation and insertion of nearly all synthesized proteins into the mitochondria.[23]

The last family is mediating the flux of all metabolites (including ADP and ATP) and ions across the OMM.[24] It is the family of voltage dependent anion channels (VDACs). They are named this way due to their anion selectivity and their voltage-dependent conductance.[25,26] In high density regions they cover up to 80 % of the membranes surface[27] and thus are the most abundant protein family in the OMM. They are known to form various oligomers[28], which was determined by EM as early as in the 1960s.[29] The following chapter will describe this family in more detail.

1.3 VDAC – The voltage-dependent anion channel

VDAC, the voltage dependent anion channel, was discovered in two ways independently of each other: by electrophysiological measurements of the outer mitochondrial membrane and by electron microscopic (EM) observations of the same.[29]

Electrophysiological measurements of single-channels in membranes were made possible by the stable formation of planar membranes, which was described by Mueller *et al.* in 1962[30] and by Montal and Mueller in 1972.[31] However, in the first years after the invention of this technique, mainly channel-forming antibiotics that were reconstituted into these membranes were studied, as the patch-clamp technique was not invented before 1976.[32] Indeed, VDAC was the first intrinsic membrane channel that was reconstituted and studied at the single-channel level with planar membranes by Schein, Colombini *et al.* in the beginning of 1975, published in 1976.[25] Also the designation of the channel as "voltage dependent anion channel" is found in this paper first. Interestingly, Schein *et al.* originally wanted to study a voltage-gated calcium channel, but finally found out, that they reconstituted something completely different into the membrane. Different from what they expected, the reconstituted channels showed a conductance that depended on the applied voltage and favored anions like chloride over calcium. Since at that time the presence of channels in the mitochondrial membrane seemed to be unlikely, it was a big surprise that they were stemming from the mitochondrium. Further studies revealed that the VDAC channels were located in the OMM[25] and Marco Colombini concluded that they are the reason for its apparent "leakiness", that had functionally been described by Werkheiser and Bartley more than 20 years before.[33,34]

Studies carried out in parallel by Zalman *et al.* were not published before 1980.[35] They described a protein with an apparent molecular weight of 30 kDa obtained from rat liver mitochondria, but found this protein to be a nonspecific diffusion channel for saccharides of up to 8 kDa. This publication is the origin of the termin "mitochondrial porin". Only later it became clear that the mitochondrial porin and VDAC are the exact same protein.

Already more than ten years earlier in the middle of the 1960s, plant mitochondrial membranes had been studied with the means of negative-stain EM and revealed densely packed stain-filled subunits

with 2–3 nm wide "pits", in a collaboration of Bonner, Parsons *et al.*[36] The identity of these pores was finally clarified by Carmen Mannella in the beginning of the 1980's using high-voltage electron microscopy (HVEM) and in collaboration with Parsons. Working as a PhD candidate in the lab of Bonner, she had previously found with SDS-PAGE experiments that over 50 % of the protein mass of plant outer mitochondrial membranes belong to a protein with a mass of around 30 kDa, labeled "Band I".[37] The obvious hypothesis was, that this band belonged to the subunits that Parsons *et al.* had observed with EM over ten years ago. The HVEM studies in the beginning of the 1980s finally revealed 2D crystals on the fungal outer mitochondrial membrane – which in turn were identified as being composed of VDAC, closing the circle to the observation in 1965.[38]

In the following years a wide variety of functions was described for the VDAC channels and it has been suggested that there is more than one isoform of the protein.[39] This hypothesis turned out to be true, as up to three isoforms were found, depending on the respective organism that is considered. In mammals like mice or the human, three isoforms of VDAC have been found[40,41] in tissue-specific expression-levels, while for example mitochondria from yeast contain only one or two isoforms.[42]

The isoform that was discovered in the 1970s is VDAC1, which is also by far the most abundant (factor 10 over VDAC2 and factor 100 over VDAC3 in HeLa cells) and consequently the best characterized isoform of the protein.[43] All three isoforms induce quite similar permeability when reconstituted into liposomes (molecular weight cutoff between 3.4 and 6.8 kDa based on the permeability of polyethyleneglycol). Electrophysiological studies show, that the prototypical gating properties (described in the following chapter) shown by VDAC1 are highly conserved among different species. However, VDAC2 exists in two forms, one of them with a lower conductance. VDAC3 does not insert into membranes as readily as VDAC1 and VDAC2 and generally does not show distinct gating properties.[44]

1.3.1 Voltage gating of VDAC

As described in the previous chapter, one of the most distinctive features of VDAC, and in particular of VDAC1, is its ability to show a conductance depending on the applied voltage, *i.e.* to switch from a high conducting state to a low conducting state depending on the applied potential. This property

of VDAC1 is highly conserved among the different mammalian VDACs: they all show very similar conductance, selectivity and voltage-gating.[45]

All values presented here have been determined with VDAC reconstituted into planar phospholipid bilayers. The high conducting state of the pore is accessible at low voltages (below ± 20 mV) with a typical conductance of around 4 to 4.5 nS in 1 M KCl (or around 0.5 nS in 0.1 M KCl).[45–52] The conductance of the channels is linear with the KCl activity up to 4 M KCl.[53] In this state, VDAC exhibits a slight selectivity for anions ($P(Cl^-)/P(K^+) \approx 2$)[45], and is permeable to large anions such as ATP.[54] Under special conditions (a 5-fold KCl gradient across the membrane) also a cation-selective open state has once been described.[55]

At potentials above ± 20 mV, the probability of the channel to be in the high conducting ("open") state decreases until it reaches 0 % at ± 40 to 50 mV, meaning that all channels observed are in lower conducting ("closed") states.[45] The channel reaches different closed states at positive or negative potential, respectively.[53] The conductance of these closed states is usually around 50 % compared to the conductance of the open state with KCl as electrolyte in the aqueous phase – around 2 nS in 1 M KCl.[24] Deeper and more refined investigation on mVDAC1 revealed that there are two not completely identical closed substates with conductances of 2.6 nS and 1.9 nS – hence the conductance of the closed conformation of the channel is a composite of these states.[52] Upon closure, the selectivity of the channels switches to favor small cations over anions[24,56] and the pore is furthermore no longer permeable for ADP or ATP.[54,56] (See Figure 1.1.)

1.3.2 3D structure of VDAC

The 3D structure of VDAC was a long standing question that could finally be solved in 2008 – over 30 years after the protein's discovery – by three groups independently. (See below.) Before, models with 12[57–59], 13[60], 16[61–63], 18[64,65] or 19[66] β-strands and an amphipathic N-terminal α-helix had been discussed. Most of the efforts to develop these models were done by Carmen Mannella, Elizabeth Blachly-Dyson, Michael Forte and Marco Colombini over the course of over 20 years.

The number of β-strands proposed by Blachly-Dyson, Forte and Colombini increased over time from 12 to 13, while Mannella always pointed out that there is a large amount of disagreement over

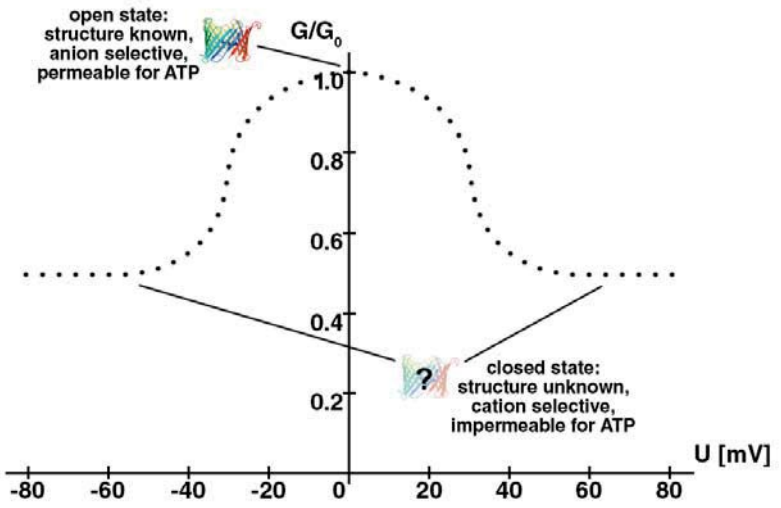

Figure 1.1: Typical voltage gating profile of VDAC. Conductance G of a VDAC channel as function of the applied potential with respect to the zero-voltage conductance G_0. The profile represents typical values yielded from studies in 1 M KCl as electrolyte. Structure shown is the 3D X-ray structure of mVDAC1 (PDB id: 3EMN).

the number of strands[67], as there are models with 16 β-strands published in the 1990s[61,62], that were never really considered by the former. The 12-β-strand model was derived in the late 1980s and early 1990s from a simple integration over the hydropathy values[68] of the amino acid side chains in groups of 10[57] and supported by site-directed mutagenesis that aimed to check for effects on channel selectivity or voltage dependence.[57,59] By biotinylation experiments the folding pattern was revised to include an N-terminal α-helix and 13 β-strands in 1998[60], which is still the model Marco Colombini prefers nowadays for the native state of the protein.[69]

By neural network-based predictors, 16 β-strands were proposed in 2002 again[63], before 18 β-strands were predicted by different authors based on computer analyses[65] and NMR data[64]. The three high resolution structures published in 2008, however, described the channel to be consisting of 19 β-strands – in line with a prediction by the Delphi algorithm from as early as 1987.[66]

In contrast to a structure from this group, that was based on a combination of X-ray crystallography and solution state NMR data[70], the other two structures were solely based on X-ray crystallography[46] or solution state NMR data[48], respectively. All three 3D structures show that the N-terminus is not part of the barrel wall and located in the interior of the pore (unstructured in the NMR structure, as an α-helix in the NMR/X-ray and X-ray structure). Furthermore, the structures showed that there is an unusual glutamate at position 73 in the sequence, which does not point

to the interior of the pore, but to the surrounding hydrophobic environment. The odd number of β-strands and the fact that the N-terminal helix is not part of the barrel wall was somewhat surprising, as this implies the need of a parallel arrangement of β-strands 1 and 19 and all previously solved β-barrel structures of procaryotes had an even number of β-strands.[48] The three high resolution structures from 2008 are shown in Figure 1.2.

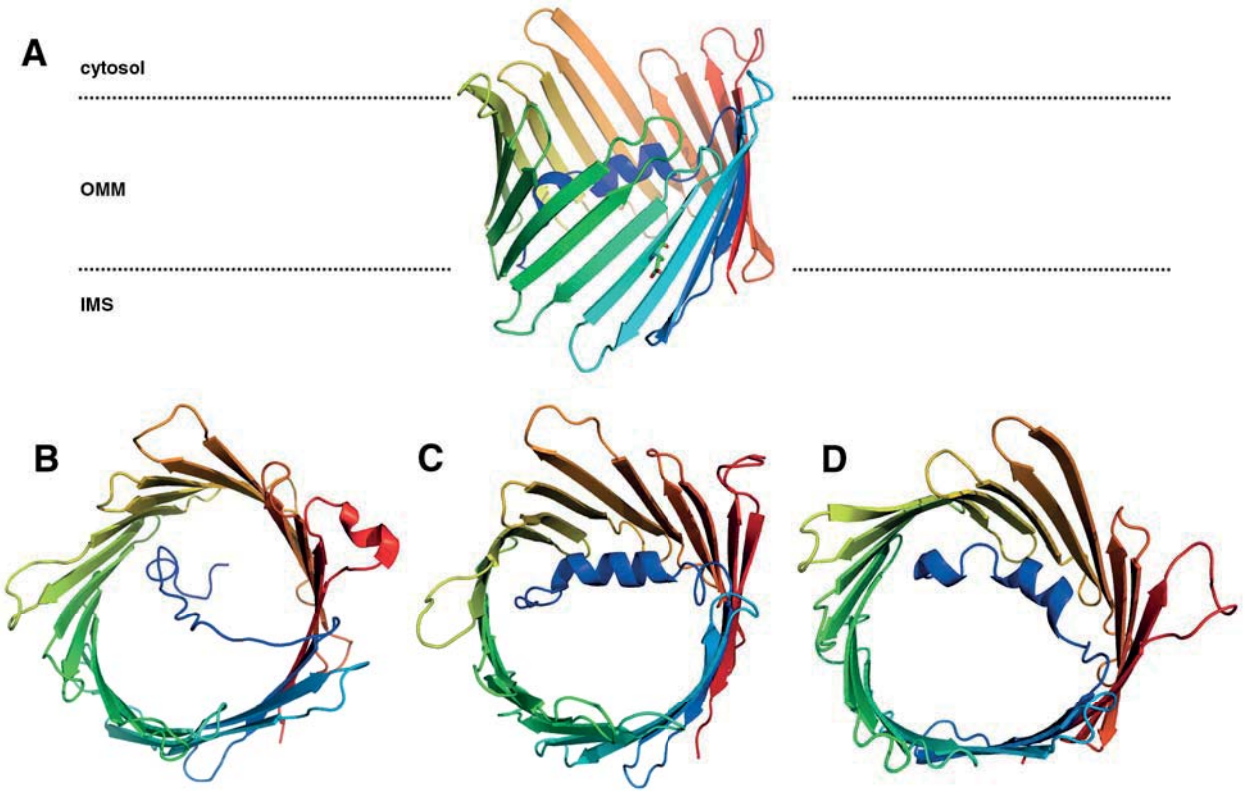

Figure 1.2: 3D structures of VDAC1 from 2008. All structures are shown in a cartoon representation and colored from N-terminus to C-terminus from blue to red. Visualization was done with PyMol.[71] A: X-ray structure of mVDAC1 (PDB id: 3EMN)[46] in a tilted view. The unusual glutamate pointing to the surrounding membrane is shown in stick representation. The dotted lines and the annotations indicate the cytosole, the Outer mitochondrial membrane (OMM) and the inter membrane space (IMS) as suggested by latest studies.[72,73] B: Lowest energy NMR structure of hVDAC1 from an ensemble of 20 structures (PDB id: 2K4T).[48] C: Combined NMR and X-ray structure of hVDAC1 (PDB id: 2JK4).[70] D: X-ray structure of mVDAC1 (PDB id: 3EMN).[46]

While the X-ray structure shows that the N-terminal helix should have a short kink from residues 10 to 12, it does not show this feature in the structure based on combined NMR and X-ray data. In the structure based on NMR data alone, the N-terminus is completely unstructured. Also, in the X-ray structure the N-terminal helix is connected tightly to the barrel, while in the structure based on combined data only the first part of the helix is connected to the barrel and in the NMR structure,

the N-terminus is floating freely in the middle of the barrel. Lastly, the N-terminus is rotated by approximately 90° in the structure based on combined data with respect to the X-ray structure. It was discussed in the PhD thesis of Saskia Villinger[74], whether this is due to differences in the resolution of the structures or due to dynamics of the helix. Based on newer results, the former is most likely the case.[75]

After the publication of these structures the question was raised especially by Marco Colombini, whether these structures represent a native and functional state of the channel or are an artifact of the refolding procedure that was necessary to study the respective samples.[76] However, the diameter of the published structures is in excellent agreement with values obtained from EM and AFM experiments[27,77,78] and the tilt angle of the β-strands fits well to other resolved structures of β-barrel proteins.[79] This would not be the case for the 13-stranded model favored by Marco Colombini.[69] Except of him it is thus now universally accepted that the 3D structures published in 2008 indeed represent a native conformation of the channel.

The structures were once again confirmed by further publications in 2014, when the crystal structure of mVDAC1 in presence of ATP was solved. It reveals a low affinity binding site of ATP in the center and bottleneck of the pore: between the N-terminal helix and the C-terminal end of the sequence.[80]

Also in 2014 the first 3D high resolution structure of the VDAC2 isoform was reported. The structure of zfVDAC2 was solved by X-ray crystallography and reveals a β-barrel with 19 β-strands and an N-terminal α-helix attached to the barrel wall at the usual position and with a kink consisting of residues 10 to 12 (PDB ID: 4BUM).[75] Overall, the X-ray structure of zfVDAC2 is strikingly similar to the X-ray structure of mVDAC1 with an RMSD of 0.98 Å. The only big difference between the structures is the loop from β-strand 1 to 2 with a displacement of 10 Å. The X-ray structure of zfVDAC2 is shown in Figure 1.3.

1.3.3 Gating models & evolution of the models

The development of the gating models of VDAC happened in close connection to the amount of structural information available about the channel, as with more structural information also the mode

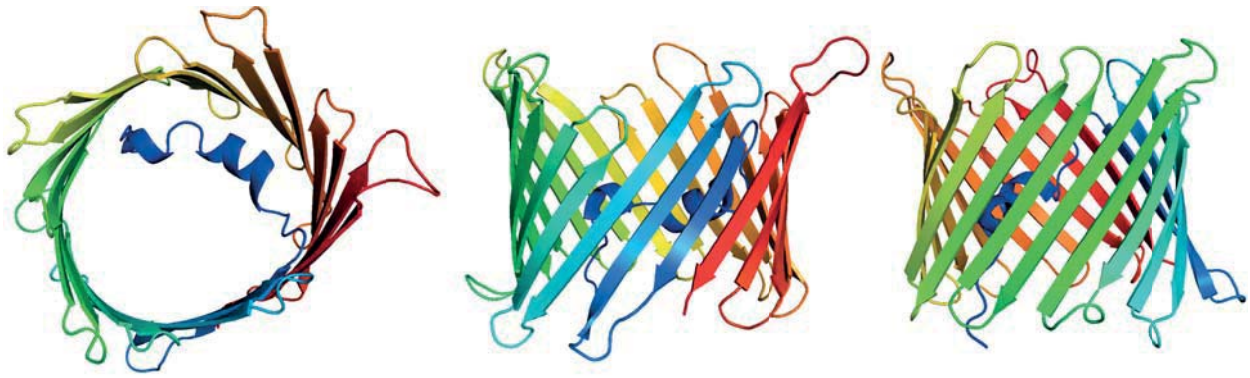

Figure 1.3: 3D structure of zfVDAC2 from 2014 (PDB ID: 4BUM). All views are shown in a cartoon representation and colored from N-terminus to C-terminus from blue to red. Visualization was done with PyMol.[71] zfVDAC2 from above shows the N-terminal helix with the characteristic kink (left) and the usual barrel fold known from the previous 3D high resolution structures (middle and right).

of gating could be proposed with more accuracy and detail. However, even with several high resolution structures available, the mode of gating is still not fully understood.

One of the earliest and most basic insights into the gating of VDAC was, that the inaccessible volume for polyethyleneglycol (PEG, a hydrophilic polymer) of the channel during closure is decreased by roughly 50 % of its initial volume in the open state. This decrease is in the order of $2 \cdot 10^4$ Å^3.[81]

Furthermore, it became evident that there is a positively charged moiety, also called "voltage sensor" region in the protein, which moves during gating and that there are actually two distinct gating processes – one at negative and one at positive potential. These gating processes can be influenced e.g. by altering the voltage sensor with the means of site-directed mutagenesis (to make the voltage dependence asymmetric, stronger or to remove it)[59,82], titration of the voltage sensor (to weaken the gating at higher pH values)[83], phosphorylation (also resulting in asymmetric gating)[84] or the composition of the membrane in which VDAC is embedded[85]. The voltage sensor was determined using a VDAC model with 12 β-strands, but based on the available 3D structures, the voltage sensor would comprise the N-terminal α-helix and parts of β-strands 2–5, 9–10 and 19.[59] (See Figure 1.4.) The residues probed in β-strands 9 and 10 (Glutamate-145 and -152 in scVDAC1, analogous to Glutamate-147 and Glutamine-154 in hVDAC1) were shown to influence only one of the two gating processes.[82] If all of the determined voltage sensor regions would move upon closure of the channel, this would entail a large scale rearrangement of the protein domains.[86]

It was proposed that the voltage sensor of the protein moves in a way, that exposes it to the membrane surface during the gating process. This was probed with site-directed mutagenesis: single