Aus dem
CharitéCentrum 02 für Grundlagenmedizin
Institut für Medizinische Physik und Biophysik
der Medizinischen Fakultät der Charité – Universitätsmedizin Berlin
Direktor: Professor Dr. Christian Spahn

Habilitationsschrift

Structure-function relationships of helical membrane proteins: implications for 3D modeling

zur Erlangung der Lehrbefähigung

für das Fach Biophysik

vorgelegt dem Fakultätsrat der Medizinischen Fakultät

Charité - Universitätsmedizin Berlin

von

Herrn Dr. rer. nat. Peter Hildebrand

geboren am 14.11.1968 in Burglengenfeld

Eingereicht: Dezember 2010

Dekanin: Professor Dr. med. Annette Grüters-Kieslich

1. Gutachter: Prof. Dr. Helmut Grubmüller

2. Gutachter: Prof. Dr. Roger Sunahara

“What is here, is elsewhere.
What is not here, is nowhere”

Vishvasara Tantra

Meiner Familie gewidmet

Contents:

Abbreviations

APP: Amyloid precursor protein

AUC-value: Area under the curve value

$A\beta_{42}$: amyloid beta with 42 amino acid length

β_2AR: β_2 adrenergic receptor

β-CTF: C-terminal fragment of APP

Cα–H•••O: Hydrogen bond between main chain Cα carbon atom and acceptor oxygen

C-cap: C-terminal reverse turn (of the α-subunit of G_t)

Channels: helical membrane proteins with functional pores

ε_r: relative permittivity

EM: Electron Microscopy

FAD: Familiar Alzheimer's disease

Gαβγ: hetero trimeric G protein

G protein: guanine nucleotide-binding protein

GDP, GTP: guanosine diphosphate, guanosine triphosphate

GPCR: G protein coupled receptor

GSM: γ-secretase modulator(s)

G_t: Transducin

LIP: Loops in proteins

LIMP: Loops in membrane proteins

Membrane coils: helical membrane proteins without functional pores

NMR: Nuclear magnetic resonance

PDB: Brookhaven Protein Data Bank

PSSM: Position Specific Scoring Matrix

R*: active form of GPCR

R*·G_t·GDP: active form of GPCR in complex with GDP bound G_t

R*·G_t[empty]: active form of GPCR in complex with GDP free G_t

TMH: transmembrane helix/helices

X-ray: Röntgen radiation

0 Abstract

About one third of the presently mapped gene sequences encode for membrane proteins, which are also major targets for pharmaceutical interventions. In contrast, only a minor fraction (September 2010: 1.8%) of the protein structures deposited in the protein data bank (PDB) belongs to this structural class. Due to difficulties in overexpression and crystallization, their tertiary structure - that is needed i.e. for protein-based virtual screenings of chemical databases - is often assessed using computational methods. Homology modeling may be applied when an appropriate template structure is available. In most cases, however, the sequence similarity between the template (if there is any) and the target structure is low and specialized knowledge based tertiary structure modeling methods such as 'Rhythm' or 'Superlooper' that will be described in detail in this work are becoming important to improve the quality of the models. Such methods are based on the analysis of known high-resolution crystal structures. For the outcome of protein structure analyses varies with the method used, we developed 'MPlot', a framework that integrates various tools for the analysis and visualization in a web based, easy to use workbench that also provides functionality for sharing data, analyses and workflows. Analyses and associated datasets are thereby supplied in a completely transparent way enhancing reproducibility or updating of results.

The analysis of geometrical features reveals that hydrogen bonds in helical membrane proteins - that significantly differ in their structural flexibility as subsumed by their cellular function - have two different effects: Electrostatic interactions amplify or together with internal waters modulate the effects of helix-helix packing. The crossing angle between transmembrane helices (TMH) largely determines, whether the residues at the packing interface intercalate (left-handed packing) or point away (right-handed packing) from the packing core. In the first case, side chain – side chain hydrogen bonds lock the geared and mainly anti-parallel helices close together forming tightly packed and rigid interfaces. In this architecture hydrogen bonds generally seem to have a 'rigidifying effect'. We believe that this special architecture is the main reason why certain protein structures, such as respiratory proteins or photo systems, subsumed here as *membrane coils* are less flexible. In the second case, main chain – main chain hydrogen bonds stabilize the interaction between smooth helical interfaces. The TMH are arranged in loosely packed parallel and anti-parallel arrangements that can sample a significantly wider range of crossing angles. This type of packing is typically found in *channels*, where hydrogen bonding networks open or close in a

coordinated manner to define the solid state motion or sliding of TMH. Accordingly, hydrogen bonds could modulate or even route protein dynamics rather than solely stabilizing a certain protein conformation.

The type of interaction typical for *channels* also seems to be involved in the molecular events taking place in the course of Alzheimer's disease. It is known that dimerisation of the transmembrane domain of the Amyloid precursor protein (APP) enhances the production of cytotoxic amyloid-β ($A\beta_{42}$) products after cleavage by the γ-secretase. The tertiary structure model of APP suggests that a right-handed parallel TMH dimer is stabilized by characteristic electrostatic attractive forces between flat TMH interfaces. In fact, the dynamic equilibrium between monomer and dimer is shifted towards the monomer by γ-secretase modulators (GSM), small molecular compounds perfectly mimicking that interface. Understanding the nature of TMH interactions accordingly also has a strong impact on medical sciences to find new strategies for the development of effective drugs against Alzheimer's disease.

Finally, it is discussed how some basic structural principles eventually underlying membrane protein stabilization and function also play a role during signal transduction from the prominent membrane protein family of G protein coupled receptors to their G protein. Kinetic analysis, structural data from X-ray crystallography and computational modeling support a model where a structural rearrangement of the C-terminal helix of the alpha subunit of transducin at the active receptor interface catalyses GDP release and thus the production of downstream effectors. This transient interaction likely involves a two stage mechanism, where the cognate G protein of rhodopsin first binds to the receptor before it undergoes a drastic structural rearrangement. The underlying 'helix switch' of transducin is very likely at the core of a molecular reaction where the signal that has primarily manifested in the ligand binding pocket at the extracellular site of the receptor, has reached the intracellular site of the cell.

1 Introduction

1.1 Sequence-structure relationships of helical membrane proteins

1.1.1 The (membrane) protein folding problem

The three dimensional fold of a protein is encoded by the amino acid sequence. Since more and more genomes were decoded and high resolution structures of proteins were resolved it appeared that it could just be an issue of computer power until the events underlying protein folding can be revealed *in silico*. However, it turned out that computation power is not the only bottleneck. At the moment, too many parameters necessary to simulate this process remain unknown (1). Consequently, the term 'folding problem' that first emerged around 1960 is still the term used to describe the fact that the process of protein folding is still not well understood (2). Nevertheless, the efforts to solve this problem did not remain without any success. Especially the elucidation of various three dimensional folds of membrane proteins using X-ray, EM or NMR techniques, together with biochemical, biophysical and computational techniques resulted in a first glimpse on the complicated events underlying the folding and stabilization of proteins (3). Accordingly, various forces contribute to the folding and stabilization of a proteins three dimensional fold.

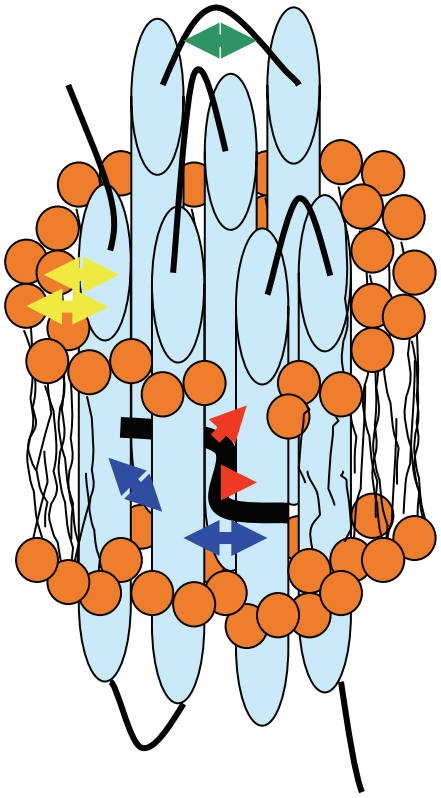

Fig. 1. Stabilizing interactions in a helix bundle protein. Blue arrows: Interactions between transmembrane helices (TMH), or between TMH and lipids, red arrows: between TMH and ligands. Yellow arrows: between TMH caps and the lipid head groups, green arrows: constraints imposed by the loops, or extra- and intracellular domains

These forces include interactions of charged residues with ions, hydrogen bonding, hydrophobic interactions and van der Waals interactions (1). Besides the enthalpy term, the entropy term contributes significantly to the stability and function of proteins. The 'hydrophobic effect', namely the desolvation of (mainly hydrophobic) protein surfaces and the adjunctive gain in entropy, is likely the main driving force of the folding of water soluble globular proteins. However, the hydrophobic effect is toned down within the lipid bilayer and the contributions of the forces stabilizing the tertiary structure of membrane proteins are still not well understood (4).

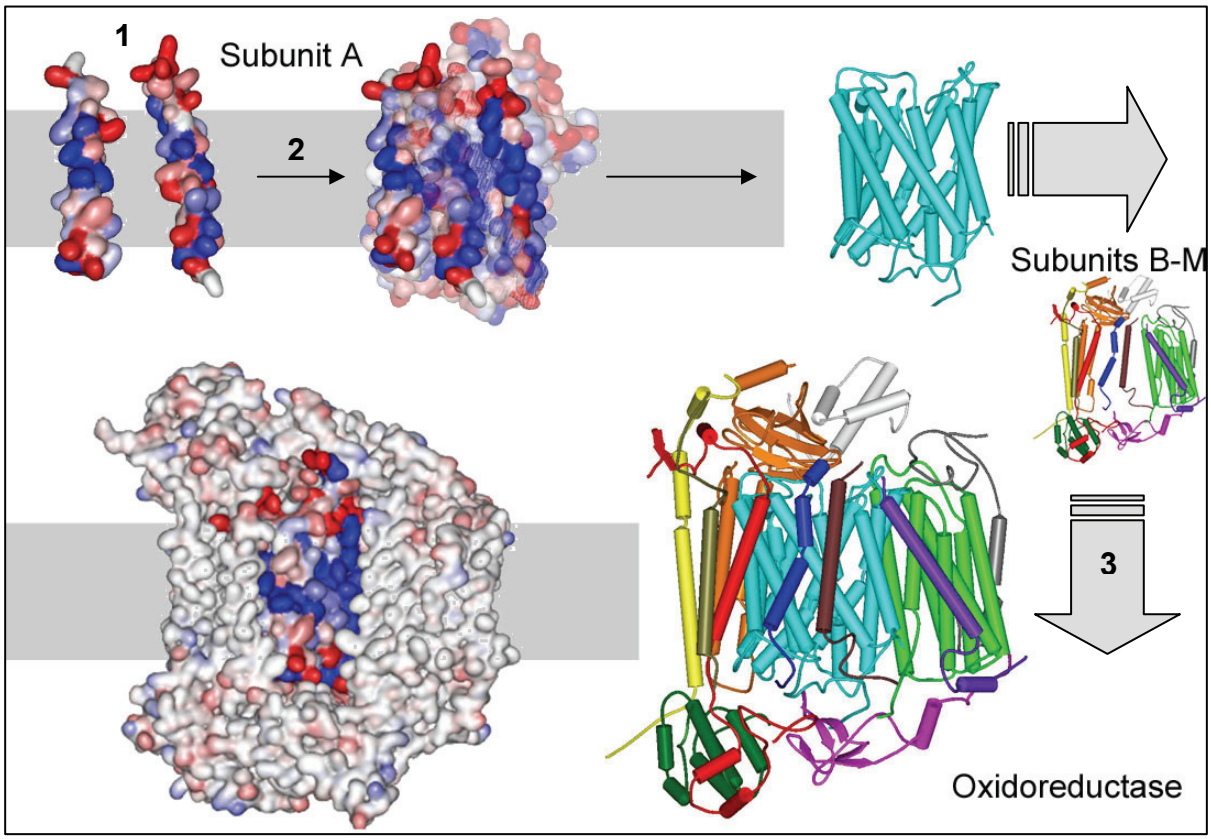

Fig. 2. Schematic model of the two (three) stage model. The TMH fold upon contact with the lipid bilayer (step one). In step two, the TMH assemble. In a third step, large complex structures may arise from the assembly of subunits or entire proteins. In the oxidoreductase, the subunits are even encoded at different loci of the cell. Subunits A-C are encoded by the mitochondrial DNA, subunits D-M are encoded by the nucleus DNA

The folding of helical membrane proteins is best described in the two (or three) stage model (fig. 2), where the TMH are (i) folded upon contact with the lipid bilayer before they (ii) assemble to complete tertiary or (iii) quaternary structures (5). The folding of the secondary structure results in a vast effective shielding of polar backbone atoms from the hydrophobic milieu of the lipid bilayer. This shielding is thus the most likely driving force for

the folding of the TMH that are subsequently fully inserted into the lipid bilayer as single helices or as helical hairpins (6). The forces that drive the assembly of TMH and that promote the folding of helical membrane proteins, however, are still a matter of dispute (4). It is clear that the lipid bilayer causes a weakening of the 'hydrophobic effect', which is supposed to be the main driving force for the folding of water-soluble globular proteins (7). But what are the forces that compensate energetically for the absence of the hydrophobic effect in the membrane? As one possibility, it has been anticipated that optimized van der Waals interactions between TMH could compensate for the lack of the hydrophobic effect (8). This, again, would implicate that membrane proteins are more tightly packed than water-soluble proteins (1). However, as will be demonstrated here, this is actually not the case (9). It therefore remains an important outstanding question how the assembly of TMH is actually accomplished (4).

1.1.2 Prediction of the orientation of TMH

Much effort has been made predicting secondary structure elements and the inside-out topology of helical membrane proteins (10-12). Actually, these tools can be considered to be one of the best working bioinformatics tools, reaching a prediction quality of up to 80%. The topology prediction was further optimized applying consensus predictions also identifying and excluding signal peptides that had often resulted in false-positive predictions (13). Some programs have also improved the prediction of the exact lengths of TMH from the sequence (14) so that the predicted secondary structure information can be directly used as input for tertiary structure modeling. However, tools that perform or assist low resolution tertiary structure modeling of helical membrane proteins are still rare (15-19).

The type of packing of α-helices is fundamental for the stabilization and function of helical membrane proteins (8, 9, 20, 21). Residues involved in helix-helix interactions are therefore regularly more conserved than others and are often arranged in specific sequence motifs that reflect the type of packing (21-24). The higher conservation of residues involved in helix-helix contacts was consequently applied in methods predicting tertiary structure contacts (25-27). This approach was further improved combining conservation criteria with amino acid propensity scales (18, 28-30). The prediction tools ProperTM (18), LIPS (16), RANTS (15) and TMX (19) depend on multiple sequence alignments to produce predictions about TMH orientations or solvent accessibility. However, the quality of prediction of these tools largely depends upon the quality of the multiple sequence alignment provided by the user. Due to the

small size of several transmembrane protein families such alignments are not always at hand. Moreover, most outputs are not presented in a user friendly format and thus again cannot be directly used for modeling purposes. Finally, the combination of statistical potentials with fragment based modeling and energy minimizations were also applied for *de novo* modeling approaches (31-34). However, the structural peculiarities of membrane proteins are so far not adequately respected by these methods.

RHYTHM, a prediction tool that will be described in detail in this work, is the first server that predicts the exposure or burial of transmembrane residues taking into account the structural specificities of membrane proteins with different functions. The prevalent motif of helix-helix interactions in channels and transporters - collectively referred to as '*channels*' - is the right-handed packing motif, whereas this type of contact is clearly underrepresented in other membrane proteins that are expected to be more rigid will collectively be referred to as '*membrane coils*' (9). It will be shown that the quality of prediction (expressed by AUC-values) of helix-helix contacts of *channels* rises by 16% to an average value of 76% when the sequence motifs typical for *channels* are applied, compared to the same approach when a non specific matrix is taken (23).

1.1.3 Prediction of membrane protein loops

Loop prediction is generally one of the most challenging tasks in protein structure determination and modeling (17, 35-46). The preferred conformation of loops often remains unclear even when the rest of the protein is resolved at high resolution, due to the high intrinsic flexibility of loops (47). Loops are again regularly involved in the recognition and binding of modulators or associated proteins. Medically highly relevant interactions, such as the coupling of receptors to G proteins are mediated by membrane protein loops (48) (**see 2.2.3 – 2.2.4**). Therefore, the conformation or the conformational space of a loop must be known in order to reveal details of protein-protein or protein-ligand interactions.

For loop modeling, two different methods, *ab initio* (17, 35, 37, 39, 46, 49, 50) and comparative modeling (40, 42, 45) are applied. *Ab initio* methods calculate possible loop conformations with the help of various energy functions and minimizations. These methods do not depend on large template libraries, but are generally time consuming and are therefore less appropriate for interactive searches. Comparative modeling approaches allow quick searches, but the quality of prediction largely depends on the availability of a suitable

template loop structure. Thus, the potential of comparative modeling methods grows, as the diversity of available templates enlarges (45). It is estimated that, at the moment, the conformation of any loop up to the length of at least 14 residues is already represented very well by protein fragments in the PDB (51, 52). Therefore, the performance of knowledge based methods to find the native loop conformation particularly depends on the size of the loop databank and on the scoring function. The web server SuperLooper that will be described in detail later in this work is a knowledge or fragment based method that allows extensive searches within two regularly updated databanks LIP (Loops In Proteins) and LIMP (Loops In Membrane Proteins). The scoring function was optimized so that the quality of prediction is similar or better even than available commercial tools. Definitely, SuperLooper is one of the most comfortable loop modeling software published to date. Finally, the specificities of membrane protein loops (53) are considered, too (fig. 3). Accordingly, SuperLooper is the first tool to specifically model membrane protein loops.

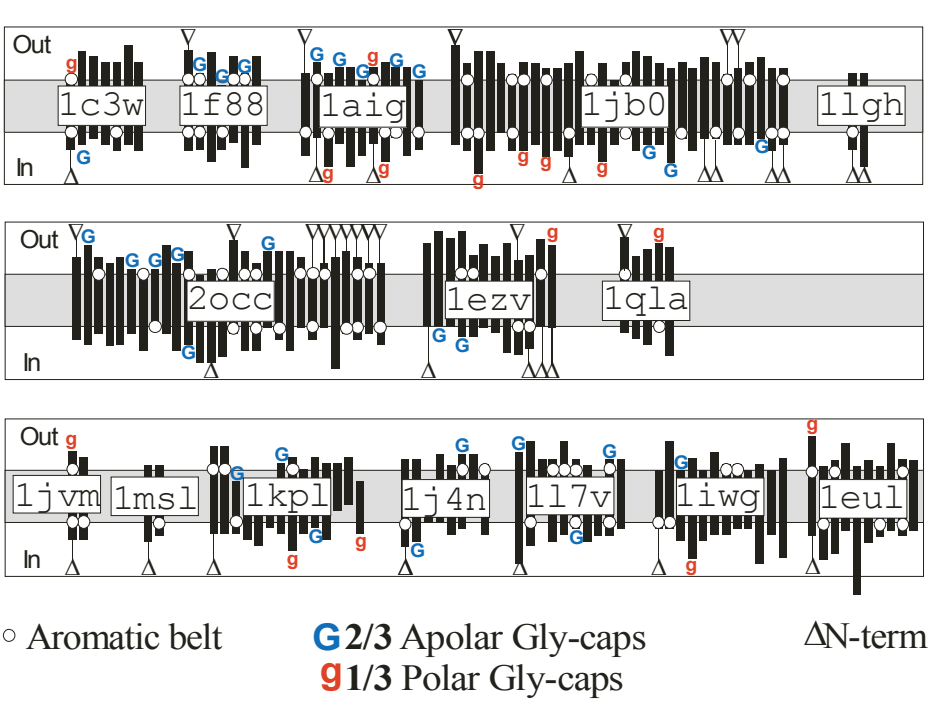

Fig. 3. Arrangement of TMH within the lipid bilayer. The lengths of the helical sections that span the hydrophobic interior of the membrane (gray) vary with their tilt angles (not imaged). Trp and Tyr that label the border to the aqueous milieu are additionally denoted as are the N-termini of the protein subunits. Loop residues capping TMH are regularly found in certain structural motifs that indicate the exact positioning of the helix cap relative to the lipid bilayer. The most common N-terminal motif is the Gly motif (Gly in N position) where the residues in N-2 and N-3 position are exposed to the solvent. If polar residues are found in these positions (g), the according loop will preferentially be exposed to the polar, otherwise (G), to the lipophilic milieu.

1.2 Structure-function relationships of helical membrane proteins

1.2.1 Different types of membrane proteins mediate between the compartments

Membrane proteins are embedded into the hydrophobic environment of lipid bilayers where they modulate the exchange of information and mass between the different partitions of cells and tissues. This protein class includes bio medically and pharmaceutically highly relevant proteins such as G protein coupled receptors, channels, and transporters. Most human membrane proteins are assemblies of hydrophobic TMH that bind coenzymes or ligands or, alternatively, form protein channels. Generally, this type of membrane channel must be distinguished from β-barrels that form rather unselective pores in membranes that originate from bacterial or mitochondrial outer membranes (54). The β-barrels are constructed from β-sheets in which the polar amino acids lining the pores and the hydrophobic residues facing the membrane are arranged in an alternating manner. By contrast, there is no such alteration in helical membrane proteins where the amino acid composition is primarily hydrophobic, with ~90%–95% of the membrane-spanning residues being no polar (53).

1.2.2 Helical membrane proteins are stabilized by hydrogen bonds

The secondary structural elements of helical membrane proteins are supposed to be folded upon contact with the hydrophobic environment of the lipid bilayer (6, 55). Besides several structural peculiarities as π-bulges or Pro induced kinks, TMH are formed as a result of regular patterns of hydrogen bonds between polar main-chain atoms (53). Hydrogen bonds between backbone amide and carbonyl groups are key components of TMH and of all other secondary structures. But, since about 5-10% of transmembrane residues are polar and the helical backbone is not entirely shielded by the side chains (56), electrostatic interactions between polar side-chains and backbone atoms are likely also to play an important role in the interactions between TMH.

Because of the low value of the relative permittivity $\varepsilon_r \sim 1$ within the hydrophobic core of the lipid bilayer compared to $\varepsilon_r \sim 80$ in the aqueous solution the free energy of formation of classical amid hydrogen bonds, is much higher than in the polar milieu: estimates of this value range from -2 to -5 kcal mol−1 per bond (7, 57). Consistent with this observation, it has been shown that a single polar amino acid such as Asn or Glu can drive homomeric association of model transmembrane peptides (58, 59). This implies that membrane proteins may be

vulnerable to loss of function through formation of membrane-buried interhelical hydrogen bonds by partnering of proximal polar side chains (60, 61). In fact, such charged and polar amino acids are rare in helical membrane proteins, probably because of the free energy cost of desolvating polar side chains. Still at least half of all helix pairs in membrane proteins are stabilized by classical hydrogen bonds (21, 62). But, most hydrogen bonds are found in motifs involving medium polar residues such as Ser or Thr. Indeed, such Ser/Thr motifs can also drive the association of model TMH (63). Thus, analysis of helical membrane proteins has suggested and it will be further confirmed in this study, that classical hydrogen bonds are clearly important for the stability of helix-helix interactions of membrane proteins.

1.2.3 Structural flexibility of helical membrane proteins

Packing interfaces containing the small amino acids Gly or Ala allow the formation of Cα–H•••O main chain hydrogen bonds in helical membrane proteins (64) (fig. 4). The residues Gly and Ala are highly abundant in helical membrane proteins (53) and since there is no entropic cost in burying the backbone atoms of small amino acids, such positions may even serve as initial points for the folding of the tertiary structure of helical membrane proteins. For the low value of the relative permittivity within the lipophilic milieu of the membrane, even main chain electrostatic interactions may be strong, depending on the burial state of a residue and on its local environment. It has been estimated that in the membrane Cα–H•••O bonds have a stabilizing effect of up to −1 kcal/mol (65) and accordingly could be nearly as strong as classical amide hydrogen bonds of globular proteins (66), where the strength of hydrogen bonds is diminished by the high dielectric effect in water ($\varepsilon_r \sim 80$) and the competition from water for hydrogen bonds (4).

Nevertheless, the contribution of Cα–H•••O hydrogen bonds to the stability of the tertiary structure of helical membrane proteins has been somewhat controversial (65, 67, 68). In fact, membrane proteins having a functional pore are often highly flexible in order to channel or carry substrates. As carried out in this work, the coordinated opening or closing of hydrogen bonds is a prerequisite of such movements and thus, weak hydrogen bonds are probably preferred over strong hydrogen bonds in order to allow such movements. In fact, a double-mutant cycle analysis applied to bacteriorhodopsin indicates that most hydrogen-bond interactions in membrane proteins are only modestly stabilizing (−0.6 kcal/mol and bond) and that weak hydrogen-bonding should be reflected in considerations of membrane protein folding, dynamics, design, evolution and function (69).

The strength of both classical and Cα–H•••O hydrogen bonds depends on the distance, the chemistry of the donor and acceptor atoms, the relative arrangement of donor and acceptor atoms, and the nature of the surrounding milieu (70, 71). As a consequence, even small conformational changes may result in the breaking of hydrogen bonds (68). Conformational changes such as shifting, rotating, or tilting of helices are believed to occur frequently in some membrane proteins and may be facilitated by specific structural characteristics such as helix kinks, the smoothness of helical surfaces, or local packing defects (9, 72, 73). The analysis of the packing and the geometrical features of helix pairs that are also involved in hydrogen bonds therefore provide clues to understanding the structure-function relationship of membrane proteins (**see 2.2.2**).

Fig. 3 from Ref. (74)

Fig. 4. Parallel right-handed helix–helix interactions with extended networks of Cα—H•••O contacts (GpA-like motifs) (74). (A) Schematic representation of the structure of the glycerol facilitator (GlpF, Protein Database ID 1fx8), the calcium ATPase (1eul), and GpA (model 19 in 1afo). The colour coding corresponds to the interactions shown in B. (B) Apparent networks of Cα–H•••O hydrogen bonds at the interface of four right-handed helix–helix interactions with distance indicated, as well as interhelical distance (a.d.) and packing angle (Ω).

1.2.4 Signal transduction by G protein coupled receptors

G protein-coupled-receptors (GPCRs) are the largest family of membrane-spanning receptor proteins connecting the extracellular environment to the cell interior. Environmental and physiological signals such as hormones, neurotransmitters, odorants, gustatory substances and light, are received by these receptors, which are also the targets for many drugs, including beta blockers and antihistamines. GPCRs use the free energy of agonist binding to transmit physical or chemical signals into the cell. Bound agonists stabilize the seven transmembrane (7TMH) helix bundle of the receptor in an active conformation (R*). R* in turn interacts with intracellular heterotrimeric G proteins (Gαβγ, G) to catalyze the exchange of GDP for GTP in the Gα subunit and thus activate downstream effectors.

Fig. 1 / Box 2 from Ref. (75)

Fig. 5. Crystal structures of inactive rhodopsin and active opsin (Ops*) conformations. Cytoplasmic and lateral views of (a) inactive rhodopsin ground state (shown in green, PDB accession 1U19) and (b) ligand-free opsin (in orange, PDB, 3CAP).Residues of the TM3–TM6 ionic lock (Glu134, Arg135, Glu247), the $Y(x)_7KR$ motif and the TM7–H8 micro domain (Tyr306, Phe313) are shown as stick models. Side chain movements on receptor activation are indicated by black arrows, whereas helix movement is indicated by a yellow arrow. (c) Lateral view and close-up of the cytoplasmic domain of opsin in complex with a C-terminal peptide derived from the last 11 residues of the transducin Gα subunit (shown in blue and magenta, respectively; Gα peptide residues are labelled in italics; PDB accession 3DQB). The peptide binds into the cytoplasmic crevice of opsin opened by movement of TM5 and TM6. The C cap of the peptide helical structure is involved in a hydrogen bonding network with Arg135 in the E(D)R3.5 and Gln312 in the NPxxY(x)5,6F motifs, respectively.

With the X-ray structure of opsin - the ligand-free form of the photoreceptor rhodopsin - which active form is stabilized by a large excess of the G-protein fragment (R*·GαCT) (48), it is now clear that GPCRs respond to the binding of extracellular ligands with a conformational change in the ligand binding site (76) which extends via their 7TMH scaffold into the intracellular domain (fig. 5). It is known that the active cytoplasmic receptor surface enables binding of cognate heterotrimeric G proteins (Gαβγ) and catalysis of GDP→GTP exchange in the Gα subunit. The GTP-bound G protein then decouples from the receptor and dissociates into Gα-GTP and Gβγ subunits that can elicit cell-specific responses via particular effector proteins and regulation of intracellular second messenger levels. The G protein cycle is completed by hydrolysis of GTP to GDP inside Gα and reassociation of Gα–GDP with Gβγ. In concert with the shut-off of the activated receptor by interactions with receptor kinase and arrestin, the second messenger is set back enzymatically to its original level. The catalytic nature of receptor-G protein interaction results in the generation of many copies of the GTP-bound activated G protein, establishing a first step of signal amplification and regulation. The three steps of reception, amplification and negative feedback constitute a signalling module which may be common to signal transduction systems in general (77).

The cycle described explains how in GPCR mediated signal transduction, the signal is eventually established in the GTP-bound form of the G protein, which is the form that activates downstream effectors. The key step in which the signal transits the membrane is represented by formation of the receptor-G protein complex, with the nucleotide binding site empty and ready for uptake of GTP (R*·G_t[empty]) (78, 79). In the absence of GTP, the R*·G_t[empty] complex is stable (80-82). The R*·GαCT crystal structure very likely shows part of the R*·G_t[empty] complex (48), consistent with earlier EPR spectroscopic work (83). Comparison with the G_t crystal structure (84) indicates a rotational and translational movement of the G_tα C-terminal α5 helix in R*·G_t[empty], which is a structural perturbation in the G protein necessary for GDP release (48, 83, 85-87). Kinetic analysis has suggested, that GDP release is triggered by conversion of an intermediate complex (R*·G_t·GDP) into the nucleotide-free R*·G_t[empty] complex (88). Since the intermediate complex does not accumulate under realistic biochemical conditions, crystallization and X-ray analysis cannot be employed. We therefore used a computational modeling approach to gain insight into the dynamic changes of the R*/G_t interface linked to the conversion of R*·G_t·GDP to R*·G_t[empty] (**see 2.2.4**).

2 Results

2.1 Sequence-structure relationships of helical membrane proteins

2.1.1 MPlot, a framework for the analysis of helical membrane proteins

The elucidation of the sequence-structure and the structure-function relationship is one of the main goals in structural biology. Statistical analysis allows highlighting structural features that are important for certain functional classes of proteins, such as *channels* or receptors. The sequence-structure relationship is again relevant for understanding the basic principles underlying protein folding or function. This knowledge can therefore directly be applied by knowledge based methods for tertiary structure predictions as will be shown in one of the following chapters. Despite its general importance, web based tools for the analyses of (membrane) proteins are rare. Here we introduce MPlot, the first online tool allowing the analysis of structural features of helical membrane proteins. The MPlot membrane protein analysis framework integrates tools for the analysis and visualization in a web based, easy to use workbench that also provides functionality for sharing data, analyses and workflows. Analyses and associated datasets can be supplied in a completely transparent way to others enhancing reproducibility or updating of results. The analyses of tertiary structure contacts and geometrical features of helical membrane proteins, however is at the present stage, by no means exhaustive. More tools will be integrated to address other issues dealing with membrane protein structures or to simply broaden the analysis by adding alternative tools for existing analyses. The framework in which MPlot is integrated facilitates such extensions. By reducing the complexity of installing and maintaining programs, MPlot allows us and also other researchers to instantly deal with their tasks at hand and less with the administrative problems around them. The results of MPlot are automatically computed from a membrane protein structure in PDB format. For analysis and statistics, all results can be downloaded as text files that may serve as inputs for or as standard data to validate the output of knowledge based tertiary structure prediction tools. For visualization, the results can be viewed online in the Jmol based protein viewer or in PyMOL on a local computer running a script. For illustration most results can be depicted by a 2D TMH interaction graph, showing a clearly laid out view of the calculated interaction measures while retaining the relative helix positioning in the middle of the membrane.

Ref. (89):

Rose A, Goede A, Hildebrand PW (2010) MPlot--a server to analyze and visualize tertiary structure contacts and geometrical features of helical membrane proteins. *Nucleic Acids Res*.

2.1.2 The prediction of helix-helix contacts of membrane proteins using RHYTHM

Membrane proteins span a large variety of different functions such as cell-surface receptors, redox proteins, ion channels, and transporters. Proteins with functional pores have different characteristics of helix–helix packing compared to other helical membrane proteins. Here we describe the development and performance of a knowledge based method to predict helix-helix vs. helix-membrane interactions, assisting tertiary structure modeling of helical membrane proteins. This method makes use of the characteristic and different motifs of interaction in *channels* and *membrane coils*. Right-handed parallel and anti-parallel interactions are typically found in *channels*. These interactions are mainly accomplished by weakly polar amino acids (G > S > T > F) that preferably create contacts every fourth residue. Left-handed anti-parallel interactions are predominantly found in *membrane coils*. There, large and polar residues (D > S > M > Q) create characteristic contacts every 3.5th residues. These characteristics were stored in Position Specific Scoring Matrices (PSSM) derived from *channels* and *membrane coils*, respectively, with known tertiary structure. A query sequence is then compared to the sequence profile of *channels* or *membrane coils* stored in the corresponding PSSM. The idea behind this approach is that the profile of the matrices allows the back calculation of tertiary structure contacts from the primary structure of a query protein. This is, of course, only valid when the structure-sequence relationship of the proteins used for generating the PSSMs is similar to that of the query protein. In fact, that quality of prediction (expressed by AUC-values) of helix-helix contacts of *channels* rises by 16% to an average value of 76% when the PSSM for *channels* is applied, compared to the same approach when the PSSM of *membrane coils* is taken. We therefore conclude that predictions of tertiary structure contacts are optimized when the structural characteristics of *channels* and *membrane coils*, are respected.

Ref. (90):

Rose A, Lorenzen S, Goede A, Gruening B, Hildebrand PW (2009) RHYTHM--a server to predict the orientation of transmembrane helices in channels and membrane-coils. *Nucleic Acids Res* 37:W575-580.

2.1.3 The prediction of membrane proteins loops using SuperLooper

Loops of membrane proteins are recognition regions for effectors and associated proteins. The probably most prominent example for such an interaction -the coupling of membrane receptors to G proteins - will be discussed below. In fact, knowing the tertiary structure of loops is the prerequisite to understand all the details of these interactions. However, the tertiary structure of loops often remains unknown. Loops are frequently not resolved well in protein structures. This is due to the high flexibility of loops as often also indicated by high B-factor values. Here we describe the development and benchmarking of a fragment based algorithm for the automated prediction of loops in membrane proteins. The possible conformations of loops up to at least fourteen residues length are already represented well in the Brookhaven Protein Databank (PDB). Knowing the coordinates of the N-terminal and C-terminal stem atoms and the primary structure of the missing loop, the tertiary structure can be modeled from fragments of available PDB-structures. We have developed a scoring function that performs very well even compared to several commercial methods. One of the main advantages of our method is that it is fast enough that searches can be done interactively. For user guidance, the candidate loops are visualized by a JMol plug-in and each candidate can be accepted or rejected after visualization. Such user interference significantly enhances the performance of tertiary structure predictions. This is especially important for the modeling of membrane protein loops that should normally not reentrant the membrane. The membrane planes are therefore automatically detected and visualized using the TMDET algorithm. Moreover, loops can be selected taking the restraints resulting from interactions with other proteins into account, too. Finally, the web server provides information on sequence identities or proline and glycine exchanges between the template and the target as well as close distances between a selected loop and the remainder of the protein.

Ref. (91):

Hildebrand PW, Goede A, Bauer RA, Gruening B, Ismer J, Michalsky E, Preissner R (2009) SuperLooper--a prediction server for the modeling of loops in globular and membrane proteins. *Nucleic Acids Res* 37:W571-574.

2.2 Structure-function relationships of helical membrane proteins

2.2.1 Molecular packing and packing defects in helical membrane proteins

The value of the internal packing density is as a measure to estimate the contribution of van der Waals forces to the stability of a protein. Tight packing indicates an optimization of van der Waals forces. In the hydrophobic milieu of the lipid bilayer the hydrophobic effect is diminished by the low dielectric field. The hydrophobic effect, the gain in entropy when hydrophobic residues are desolvated, is the likely driving force for the folding of water soluble proteins. Therefore, unspecific effects of the lipid bilayer, hydrogen bonding or van der Waals forces could energetically compensate for the absence of the hydrophobic effect. We therefore developed and applied the Voronoi cell procedure to accurately calculate the internal atomic packing density of helical membrane proteins. This value is compared to the packing density of all-helical water soluble globular proteins. As a result, we found that helical membrane proteins are packed at an average less tightly than water-soluble helical bundle proteins. Thus van der Waals forces cannot energetically compensate for the absence of the hydrophobic effect in the membrane. In order to understand the functional relevance of this finding, we investigated the packing densities of membrane proteins with different functions: As a result we found that *channels* are packed significantly less tightly than *membrane coils*. The loose packing of *channels* is in turn mainly caused by focal packing defects, i.e. cavities, rather than by global packing defects i.e. steadily increased distances between midpoints of atoms. These cavities cluster at functionally important sites of *channels* along the channels pore or at the proposed hinge regions of transporters. We conclude that cavities contribute to the structural flexibility of *channels*. By contrast, *membrane coils* (photo systems and respiratory proteins) where molecular rearrangements are supposed to occur only on a small scale, have about the same average atomic packing densities as helix bundles of globular proteins. The investigation of the packing density of TMH thus leads to the detection of structural details that promote a better understanding of the relation between stability and function of membrane proteins.

Ref. (9):

Hildebrand PW, Rother K, Goede A, Preissner R, Frommel C (2005) Molecular packing and packing defects in helical membrane proteins. *Biophys J* 88:1970-1977.

2.2.2 Hydrogen bonding and packing features in helical membrane proteins

As stated in the latter chapter, the proper functioning of *channel*s is likely based on the clustering of large cavities at functionally important sites. This structural feature could be the basis of the high flexibility of *channels* that is a prerequisite of their function. Flexibility would not be provided by an architecture, where the helices are packed tightly, a structural feature common to *membrane coils* and most globular proteins. Against this background the question arises how *channels* are actually stabilized. Here we report a comprehensive analysis of the abundance, types, and location of hydrogen bonds in known membrane protein structures. We also inspect the regions around hydrogen bonds for large cavities to identify whether the presence of hydrogen bonds correlates with nearby putative buried waters. Finally, we investigate how the geometry of helix-packing contributes to the formation of hydrogen bonds. We found that hydrogen bonds are clearly overrepresented in *channels* and tend to originate from highly conserved residues. This is mainly due to the significant enrichment of the characteristic structural motif of *channels* - the right handed packing - with Cα–H···O hydrogen bonds. In *channels* helix pairs regularly cross at large right handed angles where the (Cα-Cβ vector of the) side chains point away from the packing core. This geometry results in the formation of smooth helix-helix contact sites and is accompanied by weak electrostatic interactions between backbone atoms. In *membrane coils*, the interdigitation of side chains results in left-handed packing arrangements and a tight packing of helix-helix contacts. The presumed weakness of the hydrogen bonds in *channels*, along with the prevalence of nearby cavities that may contain water molecules, together support a hypothesis that breaking and reformation of different networks of hydrogen bonds may facilitate or even direct conformational changes in *channels*. More broadly, because of the importance of hydrogen bonding in the membrane interior, the detailed description of hydrogen bonds obtained from this analysis may be considered as another step toward a solution to the membrane protein folding problem.

Ref. (24):

Hildebrand PW, Gunther S, Goede A, Forrest L, Frommel C, Preissner R (2008) Hydrogen-bonding and packing features of membrane proteins: functional implications. *Biophys J* 94:1945-1953.

2.2.3 Structure of opsin in its G protein interacting conformation

G protein coupled receptors (GPCR) are the most prominent representatives of membrane spanning receptors. These proteins consist of seven transmembrane spanning helices (7TM), three extra (E1-E3) and intracellular/cytoplasmatic loops (C1-C3) and a short cytoplasmatic helix (H8) at the C-terminus lying parallel to the membrane plane. Distance mapping has shown that after the binding of agonists to the extracellular domain the TMH of these receptors undergo a drastic structural rearrangement. This rearrangement allows the G-protein on the opposite site of the lipid bilayer to couple to the receptor. Though the general mechanism of this coupling has been intensely studied for many decades, the structural details of the coupling reaction remained in the dark. The structure of the inactive conformation of the receptor rhodopsin resolved at high resolution in the year 2000 could give no final clue on the mechanism of receptor activation. Therefore, the elucidation of the active conformation was a long expected breakthrough in the field of membrane receptor research. Here we present the high resolution structure of the ligand free receptor opsin in its G protein interacting conformation. The structure of opsin carrying a high-affinity peptide derived from the main interaction site of the G protein was resolved at 3.2 Å resolution. The prominent outward tilt of TM6 achieves the structural requirements of the activate conformation. The bound high affinity peptide from the C-terminus of the alpha subunit of the G protein reveals the geometry of interaction with the G protein. It shows how C3 and H8 work together to bind this peptide which forms a C-terminal reverse turn (C-cap) upon receptor binding. The receptor binding site involves residues arranged in highly conserved motifs, corroborating that the observed mechanism of G protein binding by rhodopsin is representative also for other GPCRs. Importantly the superposition of the G protein with the high affinity peptide reveals that the G protein also has to undergo a significant structural change upon binding to the receptor. The structural and kinetic details of the coupling reaction were therefore subject of another analysis that will be discussed in the next chapter.

Ref. (48):

Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455:497-502.

2.2.4 Structural and kinetic modeling of rhodopsin-transducin coupling

The high resolution crystal structure of opsin in its G protein interacting conformation reveals several structural details of GPCR activation and of G protein-GPCR coupling. The high affinity peptide (and presumably also the C-terminus of the α-subunit) only become structured upon binding to the active receptor R*. It forms a reverse turn (C-cap) which is again recognized and bound by R*. For hydrogen bonds are formed to main chain atoms of the C-cap much of the specificity of binding that is strengthened by a hydrophobic interaction between the C-cap and R*, originates from the geometry of the C-cap. Comparison to the crystal structure of opsin without peptide reveals that the structural rearrangements upon binding the peptide are nearly negligible. However, the likely structural mechanism of the coupling reaction does only become clear when a combinatorial approach of structural and kinetic modeling is applied. The kinetic analysis articulately shows that the coupling reaction is at least a two stage process. Accordingly, the G protein with bound GDP interacts with R* before in a second and kinetically distinguishable complex GDP is released from its binding pocket. The structural correlates of the two kinetically different states were reproduced using fragment based tertiary structure modeling and flexible molecular docking. Flexible docking strongly indicates that the C-terminus of the α-subunit interacts with R* in two distinct modes. In the first mode, the G protein would not have to undergo any structural rearrangement when it binds to R*. Only the flexible C2 loop was remodeled using SuperLooper. This state was accordingly related to the R* G-[GDP] state. To attain the second mode of interaction the C-terminal helix of the α-subunit has to execute a 90° rotational and a 42° translational movement that we termed ‘helix-switch’. This movement likely catalysis the release of GDP that is located at the N-terminal end of α5. In the resulting complex the conformation of α5 is identical to that observed in the crystal structure and is related here to the R* G-[empty] state.

Ref. (92):

Scheerer P, Heck M, Goede A, Park JH, Choe HW, Ernst OP, Hofmann KP, Hildebrand PW (2009) Structural and kinetic modeling of an activating helix switch in the rhodopsin-transducin interface. *Proc Natl Acad Sci U S A* 106:10660-10665.

2.2.5 $A\beta_{42}$-lowering compounds interfere with APP transmembrane dimerisation

The GxxxGxxxG motif and related motifs are strong determinants for right-handed helix-helix pairing within the lipid bilayer. As described above, these packing arrangements are likely energetically stabilized by polar main chain interactions. In such packing geometries, side chains are not necessarily specifically involved, nor necessarily restricted, for they point away from the packing core. Compared to other kind of interactions the loss of entropy upon binding is therefore expected to be significantly lower likely improving the overall binding affinity. In the amyloid precursor protein (APP), the GxxxGxxxG motif triggers dimerisation of TMH, thereby enhancing the production of toxic amyloid-β ($A\beta_{42}$) cleavage fragments by the γ-secretase. Mutations within the APP GxxxG motif lower $A\beta_{42}$ levels by attenuating transmembrane sequence dimerisation. Here we show using a combination of biochemical, biophysical and computational methods that sulindac sulfide and derived compounds (GSMs) lower the $A\beta_{42}$ production through direct interaction with the APP-TMH. These compounds are generally very flat and rigid, due to their aromatic character and the conjugated π-electron system along the backbone atoms and are quite prone binding to flat protein surfaces. In fact, all compounds under investigation were found to bind to the flat surface provided by the four glycine residues G25, G29, G33 and G37 of the GxxxG motif, exactly at the proposed dimer interface. The polarity of the aromatic substitutes and their hydrogen bonding capacity appear to strongly influence the binding strength. Taken together, the observed ability of certain GSMs to modulate γ-secretase cleavages by shifting the monomer-dimer equilibrium of the substrate via mimicking the dimer interface reveals structural details of this kind of TMH interaction. These structural details of binding facilitate the search for more potent and selective $A\beta_{42}$-lowering GSMs and could therefore lead to a novel drug strategy.

Ref. (93):

Richter L, Munter LM, Ness J, Hildebrand PW, Dasari M, Unterreitmeier S, Bulic B, Beyermann M, Gust R, Reif B, Weggen S, Langosch D, Multhaup G (2010) Amyloid beta 42 peptide (Abeta42)-lowering compounds directly bind to Abeta and interfere with amyloid precursor protein (APP) transmembrane dimerization. *Proc Natl Acad Sci U S A* 107:14597-14602.

3 Discussion

3.1 Sequence-structure relationships of helical membrane proteins

3.1.1 Statistical analysis of helical membrane proteins: the method makes the difference

Computational methods applying neuronal networks, Hidden Markov models, or other machine learning approaches often facilitate the analysis or improve the prediction but clearly complicate the interpretation or comparison of the results, because the way the outputs are generated is often cryptic. We did not apply such kind of methods for the seminal analyses and predictions presented here. This clearly facilitates comparison with other methods and tools.

The Voronoi Cell method developed by Andrean Goede and colleagues (94) that is now also available as a web application (95) calculates packing densities from atomic volumes by sophisticated allocation of the extra volumes remaining when atoms that are subsumed as balls with distinct radii are packed together. By sharp contrast, the occluded surface method that was applied by Eilers and colleagues is not directly calculating packing densities, but instead uses distance criteria to estimate the packing density of proteins (8). As pointed out in detail using the occluded surface method leads to an overrating of the packing densities of atoms in direct neighborhood to protein cavities (**see 2.2.1**). For cavities are common structural features of helical membrane proteins (9) the average packing densities of helical membrane proteins were clearly overestimated by that method.

Comparison of the two methods applied to calculate packing densities in helical membrane proteins highlights that antithetic results and interpretations are obtained. It is therefore important that methods are not only described adequately to facilitate the interpretation of their outputs, but are also available in a user friendly format, to allow reproducibility of the results. The protein structure analysis workbench MPlot facilitates the integration, combination or comparison of various tools, such as the occluded surface method or Voronoia (http://proteinformatics.charite.de/mplot) (89). It provides structural biologists to investigate different aspects of protein structures and facilitates interchange of analysis data with cooperation partners and reviewers. The data can be used for learning or reference data sets, for tertiary structure prediction tools as well as for further statistical analysis. Traceability of protein structure analyses is therefore one important issue addressed by this work.

3.1.2 Tertiary structure prediction of helical membrane proteins: the function matters

The prediction of tertiary structure contacts from the amino acid sequence requires that a correlation between the primary and the tertiary protein structure can be drawn. For membrane proteins which amino acid mixture is very homogeneous, composing of 90-95% unpolar residues (53), it is not sufficient considering a single amino acid to predict its environment, i.e. to decide, whether the amino acid is buried or exposed (96). The performance of such predictions, however, can be improved as soon as specific sequence patterns can be related to specific structural patterns (22). We therefore searched for consensus sequence motifs that relate to cognate tertiary structure motifs. As a condensate two different types of interactions were described: Left-handed anti-parallel interactions that relate to heptad repeat sequence patterns typically for *membrane coils* and right-handed parallel or anti-parallel interactions that relate to octad sequence repeat patterns characteristic for *channels*. As a result, the quality of tertiary structure contact predictions was significantly improved to a level of 76 %, when *channels* and *membrane coils* are predicted using different PSSMs (23, 90).

The quality of prediction will further improve as the data set of non homologous high resolution membrane protein structures grows (3). At the moment the prediction is limited for several reasons: A significant part of buried residues is close to internal cavities (24, 97). Such residues are not judged in our analysis to be part of helix-helix contacts due to insufficient contacts to other residues and are thus often evaluated as false positives in our prediction (90). Moreover, about one quarter of the residues contact both, another helix and the membrane. These residues are frequently not assigned at high specificity thresholds but will be assigned to be buried or exposed at lower thresholds. This ambiguity clearly complicates the prediction, as well as the fact that many *channels* are highly flexible. Residues that are buried in one functional state may become exposed in another (34, 98). Finally, residues that appear to be lipid exposed may become (and may also be predicted to be) buried in quaternary complexes (99). With more structural data of membrane proteins a prediction of residues that are exposed or buried depending on the functional state shall be possible.

3.2 Structure-function relationships of helical membrane proteins

3.2.1 Permanent versus transient interactions between proteins

Protein-protein interactions (PPI) play diverse roles in biology and differ based on the composition, affinity and whether the association is permanent or transient. Since changes in the complex state are regularly coupled with biological function or activity, transient PPI are important biological regulators. In contrast to a permanent interaction that is usually very stable and thus only exists in its complexed form, a transient interaction associates and dissociates in vivo (100). Weak transient interactions that feature a dynamic oligomeric equilibrium in solution, where the interaction is broken and formed continuously can be distinguished from strong transient associations that require a molecular trigger to shift the oligomeric equilibrium (fig. 6). For example, the heterotrimeric G protein dissociates into the Gα and Gβγ subunits upon guanosine triphosphate (GTP) binding, but forms a stable trimer with guanosine diphosphate (GDP) bound. Structurally or functionally obligate interactions are usually permanent, whereas non-obligate interactions may be transient or permanent (100).

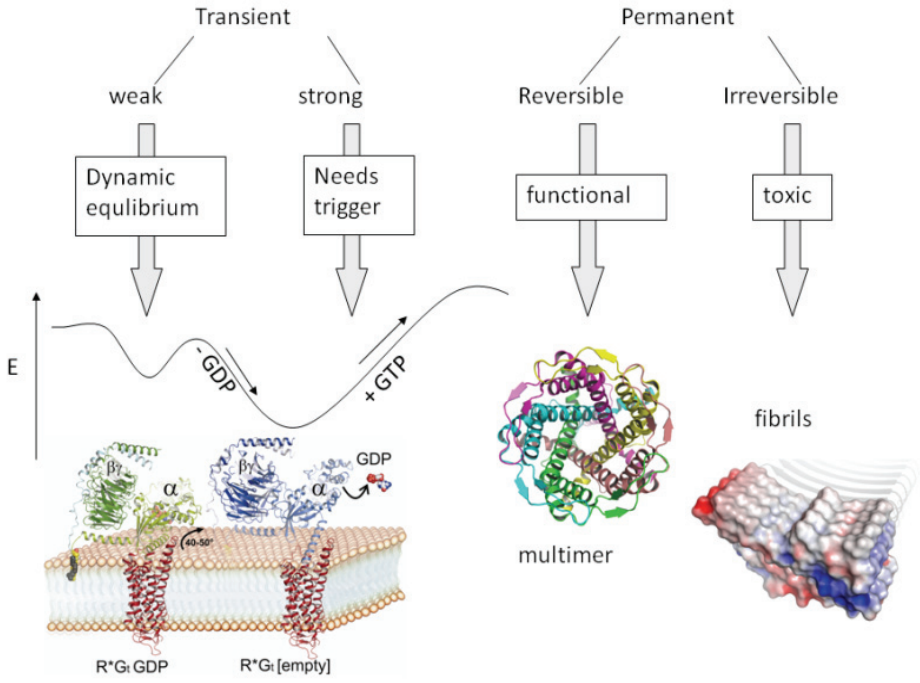

Fig 6. Permanent vs. transient interactions in proteins. (left) The highly metastable transient complex R*Gt-GDP is in a dynamic equilibrium with uncomplexed states. It is shifted by GDP release towards the more stable transient R*Gt-empty complex that needs GTP binding to dissociate. **(right)** Membrane protein complexes form permanent complexes that as in case of channels can be highly flexible. Fibrils are the toxic variants of permanent complexes that persist even after cell death.

However, many PPIs do not fall into distinct types. Rather, a continuum exists between non-obligate and obligate interactions, and the stability of all complexes very much depends on the physiological conditions and environment. An interaction may be mainly transient in vivo but become permanent under certain cellular conditions. Data on the dynamics of the assembly at different physiological conditions or environments are therefore one prerequisite to estimate the in vivo relevance of an interaction (75, 100). Our present work shows that the characteristics of transient PPIs are defined by the structural properties of the interaction partners. Rhodopsin i.e. has to change its conformation significantly upon activation to bind its cognate G protein transducin (48). Thus, an inherent structural flexibility that in case of rhodopsin is lowered by binding of *11-cis*, or elevated by *all-trans* retinal, is needed to allow the cognate G protein to bind (75). Structural flexibility is also an inherent probability of *channels* that form permanent complexes. In order to understand more about the relevance or functionality of PPIs, the structure-function relationship of the different states has to be known.

3.2.2 Structural flexibility versus rigidity in proteins

The view on protein structures has changed dramatically during the last decades. Proteins are not longer considered as 'stiff molecules' trapped in a certain conformation, but rather to be in constant well defined motions. While sampling an ensemble of different conformations, a protein state may selectively be recognized by a certain effector or interaction partner shifting the equilibrium from multiple protein states to a well defined single i.e. active state (fig. 7).

Fig. 1 from Ref. (101)

Fig 7. Remodeling of the energy landscape available to a protein from higher energy (red) to lower energy (blue) (101). (**A**) The large ensemble of non-native states moves down the energy funnel towards the native state that samples distinct energetically accessible conformational states under physiological conditions, given thermal fluctuations (boxed region). (**B**) A signal can remodel the energy landscape by narrowing the size of the ensemble of states in a single energy well, leading to a structural rigidification. (**C**) Alternatively, a protein may exist in equilibrium between two distinct conformational states, and an incoming signal can alter the relative energies of the two states, leading to a redistribution of their occupancies. (**D**) A slight variation on (C) may occur if the sampling of a higher-energy state in the absence of ligand provides a partial pathway toward a signal-induced conformation, as shown by partially overlapping wells of the two states.

In terms of the proteins' energy landscape, the conformational space that a protein can explore is defined by the inter conversion rates between energetically different states, dependent on the energy barriers between states. These dynamic properties allow proteins to transmit signals by acting as switches and transducers (101). The activation of rhodopsin is realized by a dramatic rearrangement of the 7-TMH bundle, the structural prerequisite for G protein coupling and signal transduction but also for regeneration of the ligand. By contrast, the function of other proteins does not require large structural rearrangements. Therefore the question arises: What are the intrinsic structural properties allowing proteins to act as switches and transducers? Our statistical analysis of *channels* and the comparison with helical membrane proteins that are considered to be rather rigid (*membrane-coils*) provides some preliminary answers to that question.

In summary, *channels* are packed less densely than *membrane-coils* and coiled helices of water soluble proteins (9). Focal packing defects at functionally important protein sites very likely induce local structural flexibility, presumably by minimizing van der Waals interactions as revealed by our atomic packing analysis. Reformations of hydrogen bonding networks cross-linking TMH in the switch regions of membrane proteins, guide the coordinated movements of *channels* (24) (**see 3.2.3**) and of receptors (48). The hydrogen bonding networks also include internal waters that are not always resolved in the crystal structures, but seem to be crucial for function (**see 3.2.5**). Our investigations imply that the proximity of waters and the specific local geometry of hydrogen bonded helices in *channels* are such that they likely facilitate their breakage and reformation needed for switching.

Internal water seems not only to be relevant for protein flexibility but also for the folding of proteins. The simulation of helix-helix formation or separation has shown that surfaces of solvated contacts are more easily formed or separated (102). These computational experiments endorse that the entropic cost of forming a tightly packed protein core is probably just too high, to be countervailed by the enthalpic gain of such an interaction (**see 3.2.4**). Actually proteins are not perfectly packed (9, 20, 103). This structural feature may therefore at least partially reflect the energetically requirements of protein folding. The fact that switches and transducers are even packed less densely than other proteins highlights the functional role of a non optimized packing of van der Waals surfaces in *channels*, and receptors.

3.2.3 Role of hydrogen bonds: The sub structural context makes the difference

According to our analysis of geometrical features, hydrogen bonds may have two different effects: Electrostatic interactions amplify or - together with internal waters - modulate the effects of helix-helix packing. The crossing angle largely determines whether the residues at the TMH packing interface intercalate (left-handed packing) or point away (right-handed packing) from the packing core (24, 104). In the first case, side chain – side chain hydrogen bonds lock the geared and mainly anti-parallel helices close together forming tightly packed and rigid interfaces. In this architecture hydrogen bonds generally seem to have a 'rigidifying effect'. They will lead to high interaction affinities and to rigid protein structures, as seen in respiratory proteins and photo systems, subsumed here as *membrane coils*.

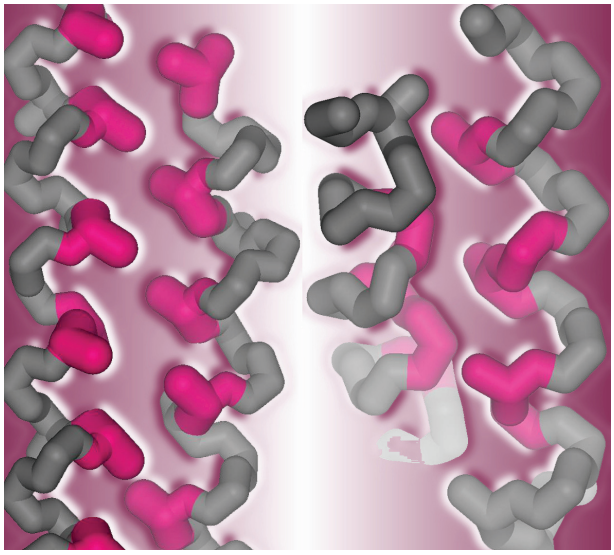

Fig. 8. Schematic representation of the knobs-into-holes packing of an anti-parallel left-handed helix pair of membrane-coils (PDB code: 1c3w) (left) and of a right-handed parallel helix pair of channels (1kpl) (right). The Cα-Cβ vectors of the core residues (in magenta) in right-handed interactions point away from the contact, where small residues are again preferred. The knobs are therefore articulately flattened and the sterically restrictions imposed on the conformation of these helix pairs are limited. For the interdigitations of side-chains the conformation of left-handed helix pairs are sterically constrained and anti-parallel arrangements are strongly preferred.

In the second case, main chain – main chain hydrogen bonds stabilize the interaction of smooth helical interfaces. The TMH are arranged in loosely packed parallel and anti-parallel arrangements that can sample a wide range of crossing angles. The average gain in enthalpy of these helix-helix interactions is presumably low, because they are stabilized by a high portion of long distance hydrogen bonds involving main chain atoms. This type of packing is typically found in *channels*, where hydrogen bonding networks open or close in a coordinated manner to define the solid state motion or sliding of TMH. Accordingly, hydrogen bonds could modulate or even route protein dynamics (24) rather than solely stabilizing a certain protein conformation (105). Actually, there is first evidence that hydrogen bonds in

membrane proteins maybe only modestly stabilizing (69). This is probably due to the fact that within the hydrophobic milieu of the membrane, the strength of hydrogen bonds will be diminished by the high dielectric effect of buried waters and the competition from water for hydrogen bonds (24).

In terms of the underlying energy landscapes, the energetic barriers between structurally different states may be less pronounced in *channels* and receptors facilitating the transition from one state to the other. In *membrane coils* these barriers could be generally higher. Future experiments are needed to understand to what extend and how the energy landscapes of membrane proteins have to be modulated to quantify the energetically contributions of the various stabilizing forces.

3.2.4 Role of entropy in pathological Aβ TMH dimerisation and GSM binding

In the right-handed TMH packing geometry of channels, main chain interactions are strongly preferred (9, 24). Though the electrostatic interactions between polar or polarized main chain atoms are presumably weak in terms of the enthalpy of interhelical bond formation, they still seem to contribute significantly to the interaction affinity for the loss of entropy upon bond formation is only minimal. Investigations of homo dimerisations using model systems like Glycophorin A, or the amyloid-β peptide (Aβ), are an appropriate way to elucidate the role of certain residues or motifs for TMH assembly and to discuss their energetical contributions. The GxxxGxxxG motif that promotes right handed helix packing is present in the Aβ peptide that is contained within the C-terminal fragment (β-CTF) of the amyloid precursor protein (APP). *In vivo*, Aβ is generated by sequential cleavage of β-CTF within the γ-secretase module (106) (fig. 9).

The identification of hereditary familial Alzheimer disease (FAD) mutations in the amyloid precursor protein (APP) and presenilin-1 (PS1) corroborated the causative role of amyloid-β peptides with 42 amino-acid residues ($A\beta_{42}$) in the pathogenesis of AD. While most FAD mutations are known to increase $A\beta_{42}$ levels, mutations within the APP GxxxG motif are known to lower $A\beta_{42}$ levels by attenuating transmembrane sequence dimerisation (107). The combination of the APP-GxxxG mutation G33A with APP-FAD mutations i.e. yielded a constant 60% decrease of $A\beta_{42}$ levels and a concomitant 3-fold increase of $A\beta_{38}$ levels compared to G33-wild-type (106). Since $A\beta_{42}$ is the toxic species, dimerisation of β-CTF and $A\beta_{42}$ and its inhibition was subject of several biochemical and biophysical studies in the last

few years and a model of the pathological mechanisms caused by APP-FAD and presenilin 1 (PS1)-FAD mutations was proposed (fig. 9) (108, 109).

Fig. 6 from Ref. (106)

Fig 9. Model of the pathological mechanisms caused by APP-FAD and presenilin 1 (PS1)-FAD mutations (105). Oxygen atoms are depicted in *red*, nitrogen atoms are depicted in *blue*, and sulphur atoms are colored *dark yellow*. The glycine residues of the interface are highlighted in *yellow*, and APP-FAD mutations are colored *pink*. *Scissors* indicate peptide bonds that are cleaved by the γ-secretase. **(A)** Dimeric wt substrates are degraded by the γ-secretase predominantly by the $A\beta_{40}$ line. *Left*, corresponding peptide bonds are indicated and are located at the dimer interface. APP-FAD mutations cause a general shift between the two product lines so that the $A\beta_{40}$ line is down-regulated, and the $A\beta_{42}$ line becomes a major degradation pathway. *Right*, APP-FAD, peptide bonds of the $A\beta_{42}$ line are indicated. Note that amino acid side chains from FAD mutations do not reach into the dimer interface. PS1-FAD mutants seem to cause substrate flux inhibition leading to a retarded processing within the $A\beta_{42}$ line causing increased $A\beta_{42}$ and decreased $A\beta_{38}$ levels (*vertical arrow*). The dimer crossing point mediated by Gly^{29} (Gly^{700}) and Gly^{33} (Gly^{704}) may cause a steric hindrance and inhibit the consecutive γ-secretase processing leading predominantly to $A\beta_{40}$ ($A\beta_{40}$ line) or to $A\beta_{42}$ ($A\beta_{42}$ line). Mechanistically, the effect of G33A occurs after the effects of APP-FAD mutations explaining why G33A causes a constant reduction by 60% for $A\beta_{42}$ and a 3-fold increase of $A\beta_{38}$ in the presence of all individual APP-FAD mutations analyzed. **(B)** The APP-FAD-TMS dimer in *side view* for better illustration of the FAD-causing amino acid side chains. T714I, V715M, I716V, and V717F stick out of the dimer interface and thus likely affect processing by modulating the substrate-enzyme recognition.

$A\beta_{42}$-lowering compounds directly bind to Aβ and interfere with amyloid precursor protein (APP) transmembrane dimerisation. Besides the pharmacological relevance of this finding, basic knowledge about the nature of the forces promoting dimerisation can be derived investigating the molecular mechanism of the interference of so called γ-secretase modulators (GSMs) with the APP dimer. As described in detail in **2.2.5** these compounds are quite prone binding to flat protein surfaces, because they are generally very flat and rigid. Due to their aromatic character and the conjugated π-electron system along the backbone atoms they seem to perfectly mimic the geometry and the electrostatics of the APP dimerisation interface.

Binding to the flat surface provided by the four glycine residues G25, G29, G33 and G37 of the GxxxG motif, monitored by surface plasmon resonance analysis, NMR spectroscopy and molecular docking therefore provides indirect evidence that weak electrostatic forces together with geometrical fitting of flat interaction surfaces stabilize right-handed transmembrane dimerisations.

Flexible docking of sulindac sulfide to APP G33I indicates that GSMs have the property of adapting themselves geometrically to differently formed binding sites (fig. 10). Rotation by 180° and lateral shifting of the compound even permits the formation of the same hydrogen bonding patterns as observed in the wt structure, where a continuous flat binding site is preserved. Comparison with the wt complex reveals that no additional torsional constraints are imposed on the aromatic ring system of sulindac sulfide. The extensive van der Waals contacts formed between the aromatic ring system and I33 are therefore likely to increase the overall binding enthalpy. However, this effect seems to be countervailed by a dramatic loss in entropy, because the vibrational and translational freedom of the isoleucine side chain will be significantly restricted upon binding of sulindac sulfide (110). In fact, sulindac sulfide was found to bind with a reduced affinity to the G33I mutant compared to $A\beta_{42}$ wt. TMH mediated dimerisation of β-CTF is also attenuated in G33I supporting the hypothesis that entropy also plays an important role in GxxxG mediated TMH dimerisation of $A\beta_{42}$ in general.

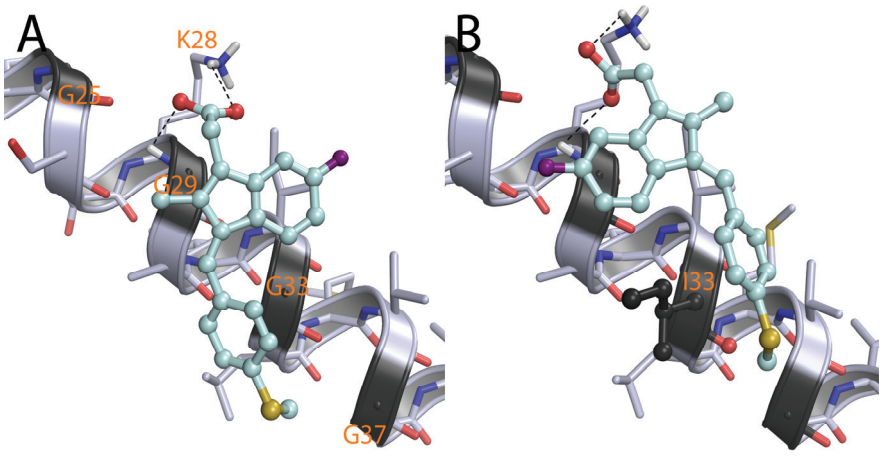

Fig. 10. Model of sulindac sulfide flexibly docked to the APP-TMS of wt **(A)** and G33I **(B)** mutant. The GSM is rotated by 180° around its longitudinal axis in the mutant compared to the wt structure and laterally shifted away from position 33. Extensive van der Waals contacts are formed between I33 and sulindac sulfide, restricting the vibrational and translational freedom of this side chain. Oxygen atoms are depicted in red, nitrogen atoms in blue, sulfur atoms in dark yellow, fluorine atoms in bright blue and chloride atoms in green. Potential hydrogen bonds are indicated as black dashed lines.

3.2.5 Functional relevance of water in signal transduction

Most knowledge of GPCR activation originates from spectroscopic investigations of rhodopsin (111). For rhodopsin several metastable intermediates have been revealed to be important for the activation pathway. These states correlate to some extent to the high- and low-affinity ligand binding states of other class A GPCRs (75, 111). However, in sharp contrast to other GPCRs in rhodopsin there is no constitutive activity. Inactive (dark) rhodopsin is trapped in a deep energy well while the energy landscape of other class A GPCRs seems to be rather shallow, with several conformational states separated by relatively low energy barriers (112) (fig. 11).

Fig. 3 from Ref. (112)

Fig. 11. Energy landscapes of rhodopsin and β2AR activation (112). (A) *Cis-trans* isomerization of retinal provides the energy for the receptor to overcome the high initial activation barrier of the inactive state and proceed along the energy landscape through small conformational changes (Batho and Lumi) to the Meta I state (MI). Subsequently, activation proceeds by formation of the Meta IIa (MIIa), Meta IIb (MIIb; by rigid-body movement of TM6), and Meta IIbH+ intermediates, and establishment of an equilibrium between the Meta forms. **(B)** G protein binding further changes the energy landscape (blue line), displacing the equilibrium to the active ternary complex, capable of catalyzing the GDP-GTP exchange in the G protein. **(C) S**hallow energy landscape of β2AR with several conformational states (R, R', R", that differ in small structural changes in TM5 and TM7) separated by relatively low energy barriers (blue line). This correlates to an inherent flexibility that allows the ligand-free receptor to explore different conformations. Ligand binding to certain intermediates (R" in this example) changes the shape of the energy landscape (green line), and activation proceeds to populate conformations of lower energy (R"'L and R*L). These conformations probably involve a similar set of conformational changes as rhodopsin, i.e., rearrangement of TM6 and neutralization of Asp130 in the DRY motif of TM3. **(D)** G protein binding to these latter states lowers the energy and stabilizes the active ternary complex, changing the energy landscape (blue line).

The main difference between rhodopsin as a photoreceptor and other GPCRs is the photo functional core with its fast light-induced transformations that precede G protein-dependent conformational conversions in equilibrium. Retinal isomerization provides the energy for the receptor to "jump" the high initial activation barrier and proceed through activation along the energy landscape through small structural changes to the Meta I state (112). Subsequently, larger-scale conformational rearrangements and changes in protonation states lead to formation of an equilibrium between the Meta forms. Finally, G protein binding further changes the energy landscape, displacing the equilibrium to the active form of the receptor, capable of catalyzing the GDP-GTP exchange in the G protein (112).

What is the basis for the structural flexibility of rhodopsin? The functional relevance of internal waters has been discussed for rhodopsin, the structurally and functionally so far best characterized GPCR. The activation mechanism of rhodopsin shows that relevant interactions occur not only between residues of the protein, but also between water molecules in strongly hydrogen-bonded water networks, or in protonated water clusters (113), based on the precise and evolutionary conserved arrangement of water molecules in the protein matrix. Thereby, water is likely to have a role not only in the protonation switches between the global 'solid state' motions of proteins. It could also support the sliding motion of helices by solvating helix-helix contacts (75).

Fig. S2 from Ref. (92)

Fig. 12. Hydrogen bonding network and surface potential of $G_t\alpha_{332–350}$. R* interaction. Side chains of the $G_t\alpha_{332–350}$ helix (α5, *blue*) hydrogen bonded (*black dashes*) to side chains of R* (*orange*) and to potential waters (*green*) are labelled and shown as sticks. **(A)** Top view, **(B)** extended hydrogen bonded network between the C-cap and R* and **(C)** extended hydrogen bonded network between α5helix and TM5/TM6. **(D)** Surface potential of R* in the S-interaction. Electrostatic surface potentials were calculated using the program APBS (114) with nonlinear Poisson-Boltzmann equation and contoured at ± 8 kT/e, and negatively and positively charged surface areas in red and blue, respectively. For clarity, the α5 helix is depicted in *orange*.s

The structure of opsin seems only minimally altered when it forms the complex with the G-protein where the GDP is released and the signal transits the membrane barrier (48, 115). Moreover, the specific contacts of the G-protein contributed to this complex are of the main chain type. Thus, there is obviously only little 'induced fit' taking place during this important step of signal transduction on the receptor side. This is an interesting and also unexpected observation, because molecular recognition has been thought to require protein flexibility that facilitates conformational rearrangements upon protein-protein interaction. However, the mutual reorientation and rearrangement of contacting protein surfaces implicated by the term 'induced fit' likely describes a thermodynamically and kinetically extensive process. High activation energy is needed to bury bulky side chains in a newly formed contact mainly because of the loss of entropy (102). This may be less critical to the formation of permanent protein complexes as parts of the large machineries like the ribosome (116, 117) or the proteasome (118), or enzymatic reactions where both partners are present in high concentrations. However, the formation and dissociation of complexes effective for signal transduction describes a highly dynamic process that takes place within only several milliseconds.

This, however, brings up the question when and how the G protein has been spatially oriented to the receptor in three dimensions. In the sequential fit model it has been postulated that an encounter complex between the G protein and R* is formed (119). The attractive forces between G protein and R* likely include long range electrostatic interactions between the positively charged intracellular surface of the receptor and the complimentary charged surface of the G protein (120) (fig. 12). This recognition maybe speeded up by the fact that the C-terminus of the Gα subunit that is only folded upon receptor interaction explores various conformations and is therefore recognized more easily by the receptor. For this process could already have been started when a G protein approaches the positively charged surface of a not fully activated receptor (i.e. in Meta I) and for the orientations of the G protein is limited by the fact that it is bound to the membrane by lipid anchors, the orientation is probably not time limiting at all. The entropic cost of folding the Gα's C-terminus could again contribute to the low interaction affinity of this first transient complex. As soon as both binding partners are arranged properly, the G protein would be ready to undergo the proposed helix-switch leading to GDP release.

4 Conclusion

In order to develop tools that allow the *in silico* folding of helical membrane proteins or assist modeling of their tertiary structure, the mechanisms that drive their folding and the forces that stabilize their three dimensional fold have to be clarified. We applied computational and statistical analyses of known membrane protein structures to estimate and discuss how the various forces energetically contribute to their folding, stabilization and function (4, 121). With help of biophysical measures such as packing densities, distances and geometries of hydrogen bonds, the presence of water and the magnitude of the dielectric constant we described the ability of membrane proteins to adapt their shape in response to a signal or binding of an interaction partner (4, 24). Our investigations highlight that the trade-off between rigidity and flexibility of a protein depends on the architecture and amino acid composition of interacting protein domains and on the local specification of the environment. These local specifications allow defining functional micro domains in helical membrane proteins that cooperate to constitute their biological functions as described here exemplarily for the activation of the G protein coupled receptor rhodopsin (75). Finally, we reveal that the different sequence-structure relationships of helical membrane proteins with different functions and different flexibilities entail that distinct tools are applied to predict their tertiary structures (23, 90).

5 References

1. Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289-316.

2. DeGrado WF, Wasserman ZR, Lear JD (1989) Protein design, a minimalist approach. *Science* 243:622-628.

3. White SH (2009) Biophysical dissection of membrane proteins. *Nature* 459:344-346.

4. Bowie JU (2005) Solving the membrane protein folding problem. *Nature* 438:581-589.

5. Popot JL, Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 29:4031-4037.

6. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, von Heijne G (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433:377-381.

7. White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* 28:319-365.

8. Eilers M, Shekar SC, Shieh T, Smith SO, Fleming PJ (2000) Internal packing of helical membrane proteins. *Proc Natl Acad Sci U S A* 97:5796-5801.

9. Hildebrand PW, Rother K, Goede A, Preissner R, Frommel C (2005) Molecular packing and packing defects in helical membrane proteins. *Biophys J* 88:1970-1977.

10. Moller S, Croning MD, Apweiler R (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17:646-653.

11. Cuthbertson JM, Doyle DA, Sansom MS (2005) Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Eng Des Sel* 18:295-308.

12. Bernsel A, Viklund H, Falk J, Lindahl E, von Heijne G, Elofsson A (2008) Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A* 105:7177-7181.

13. Amico M, Finelli M, Rossi I, Zauli A, Elofsson A, Viklund H, von Heijne G, Jones D, Krogh A, Fariselli P, Luigi Martelli P, Casadio R (2006) PONGO: a web server for multiple predictions of all-alpha transmembrane proteins. *Nucleic Acids Res* 34:W169-172.

14. Granseth E, Viklund H, Elofsson A (2006) ZPRED: predicting the distance to the membrane center for residues in alpha-helical membrane proteins. *Bioinformatics* 22:e191-196.

15. Adamian L, Liang J (2006) Prediction of buried helices in multispan alpha helical membrane proteins. *Proteins* 63:1-5.

16. Adamian L, Liang J (2006) Prediction of transmembrane helix orientation in polytopic membrane proteins. *BMC Struct Biol* 6:13.

17. Barth P, Schonbrun J, Baker D (2007) Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci U S A* 104:15682-15687.

18. Beuming T, Weinstein H (2004) A knowledge-based scale for the analysis and prediction of buried and exposed faces of transmembrane domain proteins. *Bioinformatics* 20:1822-1835.

19. Park Y, Hayat S, Helms V (2007) Prediction of the burial status of transmembrane residues of helical membrane proteins. *BMC Bioinformatics* 8:302.

20. Adamian L, Liang J (2001) Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J Mol Biol* 311:891-907.

21. Gimpelev M, Forrest LR, Murray D, Honig B (2004) Helical packing patterns in membrane and soluble proteins. *Biophys J* 87:4075-4086.

22. Walters RF, DeGrado WF (2006) Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci U S A* 103:13658-13663.

23. Hildebrand PW, Lorenzen S, Goede A, Preissner R (2006) Analysis and prediction of helix-helix interactions in membrane channels and transporters. *Proteins* 64:253-262.

24. Hildebrand PW, Gunther S, Goede A, Forrest L, Frommel C, Preissner R (2008) Hydrogen-bonding and packing features of membrane proteins: functional implications. *Biophys J* 94:1945-1953.

25. Taylor WR, Jones DT, Green NM (1994) A method for alpha-helical integral membrane protein fold prediction. *Proteins* 18:281-294.

26. Fleishman SJ, Ben-Tal N (2002) A novel scoring function for predicting the conformations of tightly packed pairs of transmembrane alpha-helices. *J Mol Biol* 321:363-378.

27. Fuchs A, Martin-Galiano AJ, Kalman M, Fleishman S, Ben-Tal N, Frishman D (2007) Co-evolving residues in membrane proteins. *Bioinformatics* 23:3312-3319.

28. Pilpel Y, Ben-Tal N, Lancet D (1999) kPROT: a knowledge-based scale for the propensity of residue orientation in transmembrane segments. Application to membrane protein structure prediction. *J Mol Biol* 294:921-935.

29. Park Y, Helms V (2006) How strongly do sequence conservation patterns and empirical scales correlate with exposure patterns of transmembrane helices of membrane proteins? *Biopolymers* 83:389-399.

30. Adamian L, Nanda V, DeGrado WF, Liang J (2005) Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. *Proteins* 59:496-509.

31. MacKenzie KR, Engelman DM (1998) Structure-based prediction of the stability of transmembrane helix-helix interactions: the sequence dependence of glycophorin A dimerization. *Proc Natl Acad Sci U S A* 95:3583-3590.

32. Pellegrini-Calace M, Carotti A, Jones DT (2003) Folding in lipid membranes (FILM): A novel method for the prediction of small membrane protein 3D structures. *Proteins* 50:537-545.

33. Park Y, Elsner M, Staritzbichler R, Helms V (2004) Novel scoring function for modeling structures of oligomers of transmembrane alpha-helices. *Proteins* 57:577-585.

34. Yarov-Yarovoy V, Baker D, Catterall WA (2006) Voltage sensor conformations in the open and closed states in ROSETTA structural models of K(+) channels. *Proc Natl Acad Sci U S A* 103:7292-7297.

35. Spassov VZ, Flook PK, Yan L (2008) LOOPER: a molecular mechanics-based algorithm for protein loop prediction. *Protein Eng Des Sel* 21:91-100.

36. Sellers BD, Zhu K, Zhao S, Friesner RA, Jacobson MP (2008) Toward better refinement of comparative models: predicting loops in inexact environments. *Proteins* 72:959-971.

37. Soto CS, Fasnacht M, Zhu J, Forrest L, Honig B (2008) Loop modeling: Sampling, filtering, and scoring. *Proteins* 70:834-843.

38. Olson MA, Feig M, Brooks CL, 3rd (2008) Prediction of protein loop conformations using multiscale modeling methods with physical energy scoring functions. *J Comput Chem* 29:820-831.

39. Rapp CS, Strauss T, Nederveen A, Fuentes G (2007) Prediction of protein loop geometries in solution. *Proteins* 69:69-74.

40. Peng HP, Yang AS (2007) Modeling protein loops with knowledge-based prediction of sequence-structure alignment. *Bioinformatics* 23:2836-2842.

41. Rossi KA, Weigelt CA, Nayeem A, Krystek SR, Jr. (2007) Loopholes and missing links in protein modeling. *Protein Sci* 16:1999-2012.

42. Fernandez-Fuentes N, Zhai J, Fiser A (2006) ArchPRED: a template based loop structure prediction server. *Nucleic Acids Res* 34:W173-176.

43. Lasso G, Antoniw JF, Mullins JG (2006) A combinatorial pattern discovery approach for the prediction of membrane dipping (re-entrant) loops. *Bioinformatics* 22:e290-297.

44. Monnigmann M, Floudas CA (2005) Protein loop structure prediction with flexible stem geometries. *Proteins* 61:748-762.

45. Michalsky E, Goede A, Preissner R (2003) Loops In Proteins (LIP)--a comprehensive loop database for homology modelling. *Protein Eng* 16:979-985.

46. Fiser A, Sali A (2003) ModLoop: automated modeling of loops in protein structures. *Bioinformatics* 19:2500-2501.

47. Lawson Z, Wheatley M (2004) The third extracellular loop of G-protein-coupled receptors: more than just a linker between two important transmembrane helices. *Biochem Soc Trans* 32:1048-1050.

48. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455:497-502.

49. Zhu K, Pincus DL, Zhao S, Friesner RA (2006) Long loop prediction using the protein local optimization program. *Proteins* 65:438-452.

50. Forrest LR, Woolf TB (2003) Discrimination of native loop conformations in membrane proteins: decoy library design and evaluation of effective energy scoring functions. *Proteins* 52:492-509.

51. Fernandez-Fuentes N, Querol E, Aviles FX, Sternberg MJ, Oliva B (2005) Prediction of the conformation and geometry of loops in globular proteins: testing ArchDB, a structural classification of loops. *Proteins* 60:746-757.

52. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235-242.

53. Hildebrand PW, Preissner R, Frömmel C (2004) Structural features of transmembrane helices. *FEBS Lett* 559:145-151.

54. Wimley WC (2003) The versatile beta-barrel membrane protein. *Curr Opin Struct Biol* 13:404-411.

55. Zimmer J, Nam Y, Rapoport TA (2008) Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* 455:936-943.

56. Baldwin RL (2003) In search of the energetic role of peptide hydrogen bonds. *J Biol Chem* 278:17581-17588.

57. Faham S, Yang D, Bare E, Yohannan S, Whitelegge JP, Bowie JU (2004) Side-chain contributions to membrane protein structure and stability. *J Mol Biol* 335:297-305.

58. Zhou FX, Cocco MJ, Russ WP, Brunger AT, Engelman DM (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins [see comments]. *Nat Struct Biol* 7:154-160.

59. Gratkowski H, Lear JD, DeGrado WF (2001) Polar side chains drive the association of model transmembrane peptides. *Proc Natl Acad Sci U S A* 98:880-885.

60. Partridge AW, Therien AG, Deber CM (2002) Polar mutations in membrane proteins as a biophysical basis for disease. *Biopolymers* 66:350-358.

61. Therien AG, Grant FE, Deber CM (2001) Interhelical hydrogen bonds in the CFTR membrane domain. *Nat Struct Biol* 8:597-601.

62. Adamian L, Liang J (2002) Interhelical hydrogen bonds and spatial motifs in membrane proteins: Polar clamps and serine zippers. *Proteins* 47:209-218.

63. Dawson JP, Weinger JS, Engelman DM (2002) Motifs of Serine and Threonine can Drive Association of Transmembrane Helices. *J Mol Biol* 316:799-805.

64. Senes A, Ubarretxena-Belandia I, Engelman DM (2001) The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* 98:9056-9061.

65. Arbely E, Arkin IT (2004) Experimental measurement of the strength of a C alpha-H...O bond in a lipid bilayer. *J Am Chem Soc* 126:5362-5363.

66. Ben-Tal N, Sitkoff D, Topol IA, Yang A, Burt SK, Honig B (1997) Free Energy of Amide Hydrogen Bond Formation in Vacuum, in Water, and in Liquid Alkane Solution. *J Phys Chem* 101:1089-5647.

67. Yohannan S, Faham S, Yang D, Grosfeld D, Chamberlain AK, Bowie JU (2004) A C alpha-H...O hydrogen bond in a membrane protein is not stabilizing. *J Am Chem Soc* 126:2284-2285.

68. Mottamal M, Lazaridis T (2005) The contribution of C alpha-H...O hydrogen bonds to membrane protein stability depends on the position of the amide. *Biochemistry* 44:1607-1613.

69. Joh NH, Min A, Faham S, Whitelegge JP, Yang D, Woods VL, Bowie JU (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453:1266-1270.

70. Baker EN, Hubbard RE (1984) Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol* 44:97-179.

71. Lindauer K, Bendic C, Suhnel J (1996) HBexplore--a new tool for identifying and analysing hydrogen bonding patterns in biological macromolecules. *Comput Appl Biosci* 12:281-289.

72. Perozo E, Cortes DM, Sompornpisut P, Kloda A, Martinac B (2002) Open channel structure of MscL and the gating mechanism of mechanosensitive channels. *Nature* 418:942-948.

73. Cordes FS, Bright JN, Sansom MS (2002) Proline-induced distortions of transmembrane helices. *J Mol Biol* 323:951-960.

74. Senes A, Ubarretxena-Belandia I, Engelman DM (2001) The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* 98:9056-9061.

75. Hofmann KP, Scheerer P, Hildebrand PW, Choe HW, Park JH, Heck M, Ernst OP (2009) A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem Sci* 34:540-552.

76. Swaminath G, Xiang Y, Lee TW, Steenhuis J, Parnot C, Kobilka BK (2004) Sequential binding of agonists to the beta2 adrenoceptor. Kinetic evidence for intermediate conformational states. *J Biol Chem* 279:686-691.

77. Hofmann KP, Spahn CM, Heinrich R, Heinemann U (2006) Building functional modules from molecular interactions. *Trends Biochem Sci* 31:497-508.

78. Kühn H (1980) Light- and GTP-regulated interaction of GTPase and other proteins with bovine photoreceptor membranes. *Nature* 283:587-589.

79. Hofmann KP (1985) Effect of GTP on the rhodopsin-G-protein complex by transient formation of extra metarhodopsin II. *Biochim Biophys Acta* 810:278-281.

80. Bornancin F, Pfister C, Chabre M (1989) The transitory complex between photoexcited rhodopsin and transducin. Reciprocal interaction between the retinal site in rhodopsin and the nucleotide site in transducin. *Eur J Biochem* 184:687-698.

81. Jastrzebska B, Golczak M, Fotiadis D, Engel A, Palczewski K (2009) Isolation and functional characterization of a stable complex between photoactivated rhodopsin and the G protein, transducin. *FASEB J* 23:371-381.

82. Ernst OP, Gramse V, Kolbe M, Hofmann KP, Heck M (2007) Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc Natl Acad Sci U S A* 104:10859-10864.

83. Oldham WM, Van Eps N, Preininger AM, Hubbell WL, Hamm HE (2006) Mechanism of the receptor-catalyzed activation of heterotrimeric G proteins. *Nat Struct Mol Biol* 13:772-777.

84. Lambright DG, Sondek J, Bohm A, Skiba NP, Hamm HE, Sigler PB (1996) The 2.0 A crystal structure of a heterotrimeric G protein. *Nature* 379:311-319.

85. Natochin M, Moussaif M, Artemyev NO (2001) Probing the mechanism of rhodopsin-catalyzed transducin activation. *J Neurochem* 77:202-210.

86. Marin EP, Krishna AG, Sakmar TP (2002) Disruption of the alpha5 helix of transducin impairs rhodopsin-catalyzed nucleotide exchange. *Biochemistry* 41:6988-6994.

87. Ridge KD, Abdulaev NG, Zhang C, Ngo T, Brabazon DM, Marino JP (2006) Conformational changes associated with receptor-stimulated guanine nucleotide exchange in a heterotrimeric G-protein alpha-subunit: NMR analysis of GTPgammaS-bound states. *J Biol Chem* 281:7635-7648.

88. Heck M, Hofmann KP (2001) Maximal rate and nucleotide dependence of rhodopsin-catalyzed transducin activation: initial rate analysis based on a double displacement mechanism. *J Biol Chem* 276:10000-10009.

89. Rose A, Goede A, Hildebrand PW (2010) MPlot--a server to analyze and visualize tertiary structure contacts and geometrical features of helical membrane proteins. *Nucleic Acids Res*.

90. Rose A, Lorenzen S, Goede A, Gruening B, Hildebrand PW (2009) RHYTHM--a server to predict the orientation of transmembrane helices in channels and membrane-coils. *Nucleic Acids Res* 37:W575-580.

91. Hildebrand PW, Goede A, Bauer RA, Gruening B, Ismer J, Michalsky E, Preissner R (2009) SuperLooper--a prediction server for the modeling of loops in globular and membrane proteins. *Nucleic Acids Res* 37:W571-574.

92. Scheerer P, Heck M, Goede A, Park JH, Choe HW, Ernst OP, Hofmann KP, Hildebrand PW (2009) Structural and kinetic modeling of an activating helix switch in the rhodopsin-transducin interface. *Proc Natl Acad Sci U S A* 106:10660-10665.

93. Richter L, Munter LM, Ness J, Hildebrand PW, Dasari M, Unterreitmeier S, Bulic B, Beyermann M, Gust R, Reif B, Weggen S, Langosch D, Multhaup G (2010) Amyloid beta 42 peptide (Abeta42)-lowering compounds directly bind to Abeta and interfere with amyloid precursor protein (APP) transmembrane dimerization. *Proc Natl Acad Sci U S A* 107:14597-14602.

94. Goede A, Preissner R, Frömmel C (1997) Voronoi Cell: New Method for Allocation of Space among Atoms: Elimination of Avoidable Errors in Calculation of Atomic Volume and Density.

95. Rother K, Hildebrand PW, Goede A, Gruening B, Preissner R (2009) Voronoia: analyzing packing in protein structures. *Nucleic Acids Res* 37:D393-395.

96. Punta M, Forrest LR, Bigelow H, Kernytsky A, Liu J, Rost B (2007) Membrane protein prediction methods. *Methods* 41:460-474.

97. Rother K, Preissner R, Goede A, Frömmel C (2003) Inhomogeneous molecular density: reference packing densities and distribution of cavities within proteins. *Bioinformatics* 19:2112-2121.

98. Adamian L, Liang J (2003) Interhelical hydrogen bonds in transmembrane region are important for function and stability of Ca2+-transporting ATPase. *Cell Biochem Biophys* 39:1-12.

99. Stevens TJ, Arkin IT (2001) Substitution rates in alpha-helical transmembrane proteins. *Protein Sci* 10:2507-2517.

100. Nooren IM, Thornton JM (2003) Diversity of protein-protein interactions. *Embo J* 22:3486-3492.

101. Smock RG, Gierasch LM (2009) Sending signals dynamically. *Science* 324:198-203.

102. MacCallum JL, Moghaddam MS, Chan HS, Tieleman DP (2007) Hydrophobic association of alpha-helices, steric dewetting, and enthalpic barriers to protein folding. *Proc Natl Acad Sci U S A* 104:6206-6210.

103. Eilers M, Patel AB, Liu W, Smith SO (2002) Comparison of helix interactions in membrane and soluble alpha-bundle proteins. *Biophys J* 82:2720-2736.

104. Walther D, Eisenhaber F, Argos P (1996) Principles of helix-helix packing in proteins: the helical lattice superposition model. *J Mol Biol* 255:536-553.

105. Schneider D (2004) Rendezvous in a membrane: close packing, hydrogen bonding, and the formation of transmembrane helix oligomers. *FEBS Lett* 577:5-8.

106. Munter LM, Botev A, Richter L, Hildebrand PW, Althoff V, Weise C, Kaden D, Multhaup G (2010) Aberrant amyloid precursor protein (APP) processing in hereditary forms of Alzheimer disease caused by APP familial Alzheimer disease mutations can be rescued by mutations in the APP GxxxG motif. *J Biol Chem* 285:21636-21643.

107. Munter LM, Voigt P, Harmeier A, Kaden D, Gottschalk KE, Weise C, Pipkorn R, Schaefer M, Langosch D, Multhaup G (2007) GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Abeta42. *EMBO J* 26:1702-1712.

108. Harmeier A, Wozny C, Rost BR, Munter LM, Hua H, Georgiev O, Beyermann M, Hildebrand PW, Weise C, Schaffner W, Schmitz D, Multhaup G (2009) Role of amyloid-beta glycine 33 in oligomerization, toxicity, and neuronal plasticity. *J Neurosci* 29:7582-7590.

109. De Strooper B, Vassar R, Golde T The secretases: enzymes with therapeutic potential in Alzheimer disease. *Nat Rev Neurol* 6:99-107.

110. Lee AL, Kinnear SA, Wand AJ (2000) Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nat Struct Biol* 7:72-77.

111. Okada T, Ernst OP, Palczewski K, Hofmann KP (2001) Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem Sci* 26:318-324.

112. Deupi X, Kobilka BK (2010) Energy landscapes as a tool to integrate GPCR structure, dynamics, and function. *Physiology (Bethesda)* 25:293-303.

113. Garczarek F, Gerwert K (2006) Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* 439:109-112.

114. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98:10037-10041.

115. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* 454:183-187.

116. Spahn CM, Penczek PA (2009) Exploring conformational modes of macromolecular assemblies by multiparticle cryo-EM. *Curr Opin Struct Biol* 19:623-631.

117. Ratje AH, Loerke J, Mikolajka A, Brunner M, Hildebrand PW, Starosta AL, Donhofer A, Connell SR, Fucini P, Mielke T, Whitford PC, Onuchic JN, Yu Y, Sanbonmatsu KY, Hartmann RK, Penczek PA, Wilson DN, Spahn CM (2010) Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. *Nature* 468:713-716.

118. Kloetzel PM (2001) Antigen processing by the proteasome. *Nat Rev Mol Cell Biol* 2:179-187.

119. Herrmann R, Heck M, Henklein P, Kleuss C, Hofmann KP, Ernst OP (2004) Sequence of interactions in receptor-G protein coupling. *J Biol Chem* 279:24283-24290.

120. Kosloff M, Alexov E, Arshavsky VY, Honig B (2008) Electrostatic and lipid-anchor contributions to the interaction of transducin with membranes: Mechanistic implications for activation and translocation. *J Biol Chem*.

121. Sachs JN, Engelman DM (2006) Introduction to the Membrane Protein Reviews: The Interplay of Structure, Dynamics, and Environment in Membrane Protein Function. *Annu Rev Biochem*.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zur Erstellung dieser Arbeit beigetragen haben.

An erster Stelle gilt mein Dank meinen wissenschaftlichen Lehrern Cornelius Frömmel, Klaus-Peter Hofmann und Christian Spahn, die mir die Möglichkeit zu dieser Arbeit gaben, sie durch Anregungen und Diskussionen förderten und mir den Raum und die Freiheit gewährten, die ich benötigte, um mein wissenschaftliches Tun zu entfalten. Cornelius Frömmel, der den Grundstein zu meiner wissenschaftlichen Laufbahn gelegt hat, möchte ich darüber hinaus sehr herzlich für sein fortwährendes Interesse an gemeinsamen Forschungsthemen danken. Klaus-Peter Hofmann gilt mein spezieller Dank für die intensiven und sehr persönlichen wissenschaftlichen Gespräche und Diskussionen. Bei Christian Spahn bedanke ich mich im ganz besonderen Maße für die zukunftsweisende Zusammenarbeit bezüglich des Ribosoms.

Namentlich möchte ich mich bei meinen Kooperationspartnern und Kollegen bedanken ohne deren Anstrengungen und Wirken diese Arbeit nicht möglich gewesen wäre (in alphabetischer Reihenfolge): Heiko Bittner, Oliver Ernst, Jan Giesebrecht, Christoph Gille, Andrean Goede, Björn Grünings, Stefan Günther, Anja Haarmeier, Martin Heck, Hergo Holzhütter, Andreas Hoppe, Jochen Ismer, Justus Loerke, Stephan Lorenzen, Lisa Münter, Gerd Multhaup, Robert Preißner, Kristian Rother, Luise Richter, Alexander Rose, Patrick Scheerer, Johanna Tiemann und Dominique Theune. Ohne die intensive Kooperation mit der Arbeitsgruppe um Gerd Multhaup wäre der Einstieg in die Alzheimerforschung nicht möglich gewesen. Ohne die vertrauensvolle Zusammenarbeit mit meinem lieben Kollegen Patrick Scheerer zur strukturellen Aktivierung des Rhodopsins hätten unsere gemeinsamen Arbeiten nicht diese Detailtiefe und Darstellungskraft erhalten.

Der Deutschen Forschungsgemeinschaft und der Charité möchte ich meinen Dank für die langjährige finanzielle Unterstützung meines wissenschaftlichen Wirkens aussprechen.

Zu guter Letzt möchte ich mich ganz herzlich bei meiner Familie für die liebevolle Unterstützung und den starken Rückhalt bedanken, der es mir ermöglicht hat, auch in schwierigen Zeiten stets vorwärts zu schreiten. In diesem Zusammenhang möchte ich mich auch bei meinen Freunden bedanken, die mich z. T. seit meiner Kindheit begleiten und die mir stets gezeigt haben, was in mir steckt.

Erklärung

§ 4 Abs. 3 (k) der HabOMed der Charité

Hiermit erkläre ich, dass

- weder früher noch gleichzeitig ein Habilitationsverfahren durchgeführt oder angemeldet wurde.

- die vorgelegte Habilitationsschrift ohne fremde Hilfe verfasst, die beschriebenen Ergebnisse selbst gewonnen sowie die verwendeten Hilfsmittel, die Zusammenarbeit mit anderen Wissenschaftlern/ Wissenschaftlerinnen und mit technischen Hilfskräften sowie die verwendete Literatur vollständig in der Habilitationsschrift angegeben wurden.

- mir die geltende Habilitationsordnung bekannt ist.

..... ……..

Datum Unterschrift