# Insights into the structural and functional properties of the eukaryotic porin Tom40

Von der Fakultät Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

**Dennis Benjamin Gessmann**

aus Ostfildern, Deutschland

| | |
|---|---|
| Hauptberichter: | Prof. Dr. Stephan Nussberger |
| Mitberichter: | Prof. Dr. Andreas Kuhn |

Tag der mündlichen Prüfung: 3. April 2012

Biologisches Institut der Universität Stuttgart

2012

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Ostfildern, den 08.01.2012

*This work is dedicated to Karl and Gertrud Schiffert*

*I did not become a rocket scientist as you always wished for.*
*But I am sure that you can forgive me,*

*for "falling out of line"…*

# Contents

# Abstract

Tom40 forms the preprotein conducting channel in the outer membrane of mitochondria enabling transport of up to 1500 different preproteins through an optimized pore environment. Moreover, Tom40 exhibits a voltage-dependent gating mechanism in terms of an 'electrical switch' making this eukaryotic β-barrel a promising target for nanopore based applications. In this work, new bioinformatics methods were developed and verified through practical approaches to shed light on the structural elements of Tom40 facilitating its particular function in mitochondria. Based on these results, Tom40 proteins were designed with modified and optimized structural properties.

TmSIP, a physical interaction model developed for TM β-barrel proteins, was used to identify weakly stable regions in the TM domain of Tom40 from mammals and fungi. Three unfavorable β-strands were determined for human Tom40A. Via CD and Trp-fluorescence spectroscopy it was shown that substitution of key amino acid residues in theses strands resulted in an improved resistance of the protein to chemical and thermal perturbations. Further, the mutated form of hTom40A was strictly found in its monomeric state. Equal improvements were gained for the apparent stability of Tom40 from *Aspergillus fumigatus*.

Tom40 was isolated and purified in its native state from *Neurospora crassa* mitochondria. Time-limited proteolysis of native *Nc*Tom40 coupled to mass spectrometry revealed comparable protease-accessibility to VDAC isoform 1 from mammals suggesting a similar fold. Thus, a homology model of *Nc*Tom40 was developed on the basis of the solved mouse VDAC-1 crystal structure. It was found that Tom40 forms a 19-stranded β-barrel with an N-terminal α-helix inside the pore. Further, a conserved 'polar slide' in the pore interior is possibly involved in preprotein translocation and a second conserved domain, termed 'helix anchor region', in arresting the helix inside the Tom40 pore.

Based on the homology model of *Nc*Tom40, the structure and function of the N-terminal domain of Tom40 was addressed. Examination of the model structure

revealed two different domains for the N-terminus, the inner-barrel and outer-barrel N-terminus. *In vivo* investigations showed that both parts prevent a heat-induced dysfunction of Tom40 in *N. crassa* mitochondria independently. By applying CD spectroscopy the predicted N-terminal α-helix could be allocated to the inner-barrel N-terminus. Further, in combination with Trp-fluorescence spectroscopy it was found that the N-terminal α-helix unfolds independently from the Tom40 β-barrel, but is not necessary for pore stability or integrity. However, a conserved amino acid residue, Ile47 of *Nc*Tom40, in the inner-barrel N-terminus is essential for the structural integrity of the N-terminal α-helix.

In conclusion, these results may offer a basis for future works on TM β-barrel proteins with the aim to alter the structural properties in the absence of a high atomic resolution structure or an established knowledge of the biochemical and biophysical properties.

# Zusammenfassung

Tom40 bildet den Präprotein leitenden Kanal in der Außenmembran von Mitochondrien aus und ermöglicht den Transport von bis zu 1500 verschiedenen Präproteinen durch ein optimiertes Porenmilieu. Weiterhin besitzt Tom40 einen spannungsabhängigen Schaltmechanismus in der Form eines elektrischen Schalters. Diese Eigenschaften machen das eukaryotische Porin Tom40 zu einem vielversprechenden Objekt für Nanoporen-basierte Anwendungen. In dieser Arbeit wurden neue bioinformatische Anwendungen entwickelt und über empirische Ansätze verifiziert, um die strukturellen Elemente von Tom40, welche dessen spezielle Funktion in Mitochondrien ermöglichen, aufzudecken. Basierend auf diesen Ergebnissen wurden neuartige Tom40 Proteine entwickelt, welche modifizierte und optimierte strukturelle Eigenschaften aufwiesen.

TmSIP ist ein physikalisches Interaktionsmodell, welches für die Identifizierung schwach-stabiler Transmembran (TM) Regionen in TM β-Barrel Proteinen entwickelt wurde. In dieser Arbeit wurde TmSIP verwendet, um schwach-stabile Regionen in der TM Domäne von Tom40 aus Säugetieren und Pilzen zu identifizieren. Drei energetisch ungünstige β-Stränge wurden in der TM Region von humanem Tom40A aufgedeckt. Unter Anwendung von Circulardichroismus (CD) und Tryptophan Fluoreszenzspektroskopie wurde gezeigt, dass der Austausch von bestimmten Aminosäureresten in diesen drei Strängen zu einer verbesserten Resistenz gegen chemische und thermische Perturbation führte. Weiterhin wurde aufgedeckt, dass die mutierte Form von humanem Tom40A grundsätzlich als Monomer vorlag. Es wurden gleichwertige Verbesserungen in der ‘scheinbaren’ Stabilität von Tom40 aus *Aspergillus fumigatus* erzielt.

Tom40 wurde in nativer Form aus *Neurospora crassa* Mitochondrien isoliert und gereinigt. Unter Kombination von zeitlich begrenzter Proteolyse und Massenspektrometrie von nativem *Nc*Tom40 wurde entdeckt, dass *Nc*Tom40 die gleiche Proteasezugänglichkeit wie VDAC Isoform 1 aus Säugetieren aufweist. Es kann daher eine ähnliche Faltung der Proteine angenommen werden und basierend auf der Kristallstruktur von Maus VDAC Isoform 1 wurde

ein Homologiemodell von *Nc*Tom40 entwickelt. Tom40 bildet ein 19-strängiges β-Barrel Protein mit einer N-terminalen Extension aus, welche als α-Helix in das Poreninnere gefaltet ist. Weiterhin ist eine konservierte 'Polare Gleitfläche' im Innern der Pore möglicherweise an der Translokation von Präproteinen beteiligt. Eine zweite konservierte Region auf der Innenseite der Pore, welche als 'Helix Anker Region' definiert wurde, hat dem Anschein nach die Funktion, die N-terminale α-Helix zu arretieren.

Basierend auf dem Homologiemodell von *Nc*Tom40 wurde die Struktur und Funktion der N-terminalen α-Helix untersucht. Die N-terminale Extension konnte in zwei unterschiedliche Domänen unterteilt werden, in den inneren und äußeren N-Terminus. Unter Anwendung von *in vivo* Analysemethoden wurde gezeigt, dass beide N-terminale Domänen eine hitzeinduzierte Fehlfunktion von Tom40 in *N. crassa* Mitochondrien unabhängig voneinander verhindern. Durch CD Spektroskopie konnte die im Modell vorhergesagte α-Helix dem inneren N-Terminus zugewiesen werden. In Kombination von CD mit Tryptophan Fluoreszenzspektroskopie wurde weiterhin aufgedeckt, dass die N-terminale α-Helix sich unabhängig vom Tom40 β-Barrel entfalten lässt und nicht für die Stabilität oder Integrität des Barrels erforderlich ist. Jedoch ist ein konservierter Aminosäurerest, Ile47 in *Nc*Tom40, für die Aufrechterhaltung der Struktur der N-terminalen α-Helix notwendig.

Diese Ergebnisse können als Grundlage für zukünftige Arbeiten an TM β-Barrel Proteinen dienen, welche das Ziel verfolgen, die strukturellen Eigenschaften des Proteins zu verändern. Weiterhin können die hier beschriebenen Methoden und empirischen Ansätze zur Untersuchung von TM β-Barrel Proteinen angewendet werden, falls ein fundiertes biochemisches und biophysikalisches Wissen, sowie eine atomar aufgelöste Proteinstruktur noch nicht vorhanden sind.

# Introduction

## Towards the understanding of biological membranes and membrane proteins

Biological membranes are the structures that define life in general as they form natural hydrophobic permeation barriers around the hydrophilic interior and exterior of cells and cellular organelles (Singer, 1992). This structural arrangement explains the need for peripheral and integral membrane proteins (Fig. 1) enabling such essential cellular processes as the oxidative phosphorylation, amongst many others. First described by Singer and Nicolson in 1972 as 'fluid mosaic model', the textbook description of biological membranes underwent refinements to a more mosaic than fluidic portrayal of natural membranes. It was shown that a cellular membrane varies in its diversity of membrane thickness, its lipid components as well as the distribution of membrane proteins. Therefore, newly developed terms, for example lipid rafts, gave rise to the understanding of a highly organized and structured membrane environment rather than a stochastic arrangement of its modules (Singer and Nicolson, 1972; Vereb, Szollosi *et al.*, 2003; Engelman, 2005).

In general, the structure of a biological membrane consists of lipids which are built of a nonpolar stretch with a short polar head group. These lipids form a planar lipid bilayer with a thickness of around 45 Å (Fig. 1). In a simplistic model, a membrane can be separated into three domains: the hydrophobic bilayer center, the head group layer and the aqueous region (Korn, 1966; Popot and Engelman, 2000). Naturally, this defines the energetic requirements of an integral membrane protein to insert into the membrane. The amino acid residues of a transmembrane protein face not only these three regions, but also the amino acids of the protein itself and possibly a water-filled environment inside of the membrane protein (White, Ladokhin *et al.*, 2001). Therefore, integral membrane proteins fold upon providing amino acid residues, which are energetically tolerated in each domain of the bilayer (Fig. 1) (White and Wimley, 1999).

In 1982, a landmark in the field of membrane biology was achieved by Hartmut Michel and colleagues when they successfully crystallized the first

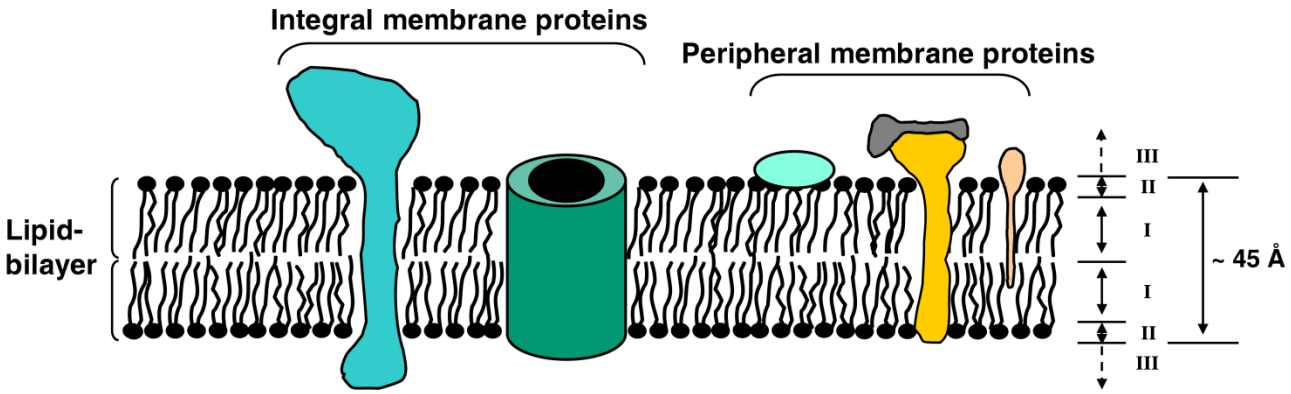



**Fig. 1** Schematic representation of a biological membrane and the types of membrane proteins. A cellular membrane is composed of lipid molecules that form a planar bilayer with a thickness of ~ 45 Å. Three domains can be distinguished: I, the hydrophobic bilayer center (total of ~ 30 Å); II, the head group layer (total of ~ 15 Å); and III, the aqueous region. Integral membrane (or transmembrane) proteins reach into both aqueous domains. Peripheral membrane proteins are attached to the membrane either at its hydrophilic surface (domain II) or through protein-protein interactions/lipid-anchors at one side of the membrane (domain II/III) (Popot and Engelman, 2000).

integral membrane protein, whose structure could subsequently be solved to atomic resolution. Michel succeeded by using detergent molecules to reconstitute the membrane-embedded photosynthetic reaction center (PRC) of *Rhodopseudomonas viridis* into water-soluble detergent micelles (Michel, 1982). In doing so, three dimensional protein crystals could be produced for X-ray diffraction, sufficient to determine the protein structure at high atomic resolution of 3 Å (Deisenhofer, Epp *et al.*, 1985). Since then, the number of solved membrane protein structures has increased exponentially (White, 2004) and new techniques have become equally important tools in the field of molecular membrane protein biology, for instance soluble NMR spectroscopy (Hiller and Wagner, 2009).

Due to an increased knowledge about the structure of membrane proteins, integral membrane proteins can be separated into two main classes, with each class being defined by its characteristic secondary structure build-up. The first type of an integral membrane protein is composed of α-helices, so called transmembrane α-helices (Fig. 2A). The second type is mainly composed of β-strands forming three-dimensional β-sheets that result in a barrel-like structure (Fig. 2B) (Schulz, 2000).

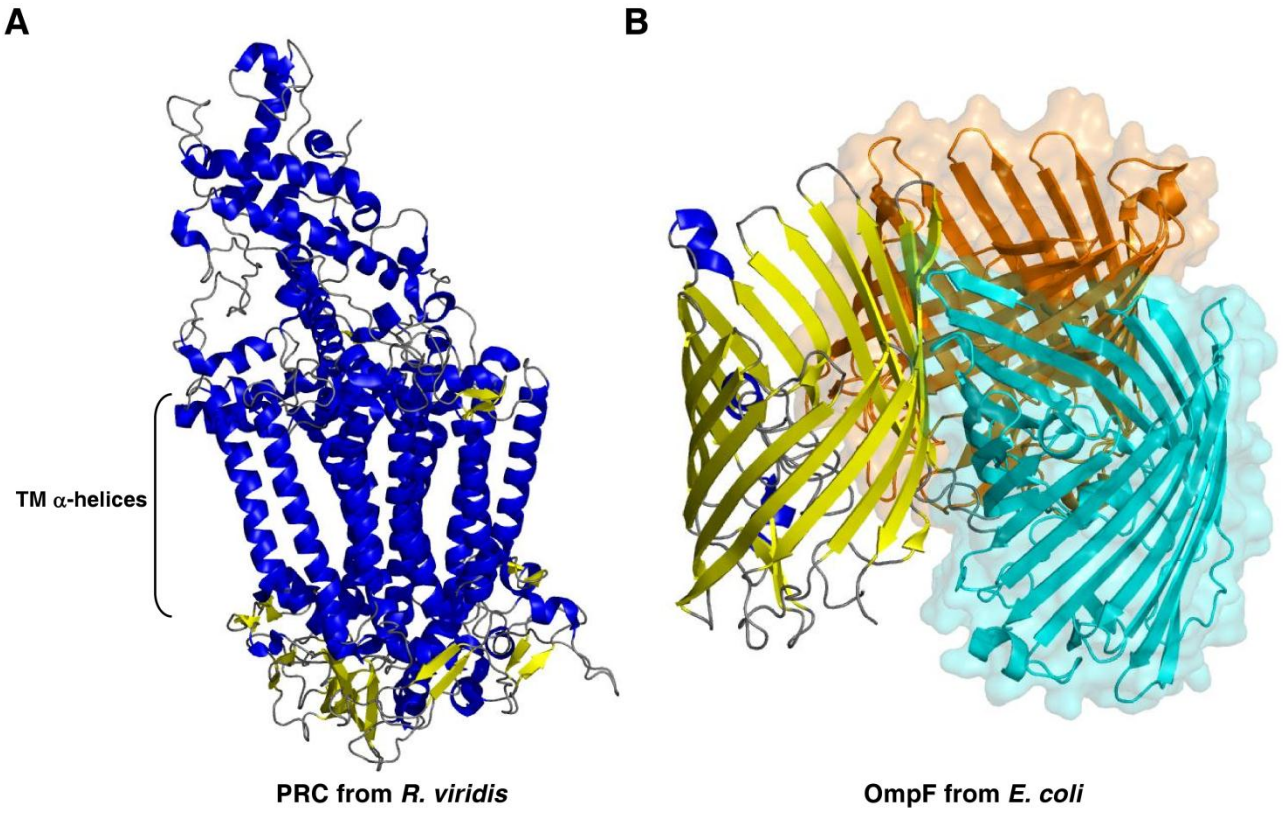



**Fig. 2** The types of integral membrane proteins. (A) Helical membrane proteins. Crystal structure of the α-helical photosynthetic reaction center (PRC) from *Rhodopseudomonas viridis* at 2.3 Å (PDB ID 1PRC) (Deisenhofer, Epp *et al.*, 1995). (B) TM β-barrel proteins. Crystal structure of the homotrimeric β-barrel OmpF from *Escherichia coli* at 1.6 Å (PDB ID 2ZFG) (Yamashita, Zhalnina *et al.*, 2008). α-helices are colored in blue, β-strands are presented in yellow and loops are shown in grey; except for 2nd and 3rd subunit of OmpF, which are displayed in cyan and orange with their 'Connolly' surfaces, respectively. All structure figures were created with PyMOL (DeLano, 2006). TM – transmembrane.

## Helical membrane proteins

Analyses of complete genome sequences revealed that 20-30 % of all open reading frames code for membrane proteins with helix bundle motifs (Arkin, Brunger *et al.*, 1997; Wallin and von Heijne, 1998; White and Wimley, 1999). It is therefore not surprising that helical membrane proteins serve various functions in cell membranes, for instance in metabolic pathways (Picot, Loll *et al.*, 1994), transduction of cellular signals (Gordeliy, Labahn *et al.*, 2002) and transport of molecules across or into a membrane (Sirrenberg, Bauer *et al.*, 1996; Long, Tao *et al.*, 2007). Helix proteins are often found as subunits in higher order protein complexes, for example the C-subunit of the ATP synthase (Girvin, Rastogi *et al.*,

1998), and their build-up varies from single hydrophobic TM helices to complex helix bundles (Fig. 3).

A typical single membrane spanning α-helix is composed of around 15-30 largely nonpolar amino acid residues, adequate to pass the hydrophobic core of the lipid bilayer (Fig. 1 and 3A) (Popot and Engelman, 2000; White and von Heijne, 2005). Helix formation is driven by the hydrophobic effect and H-bonds are generally formed by the polypeptide backbone in between residue i and i+4 (Popot and Engelman, 2000; Ubarretxena-Belandia and Engelman, 2001). The most common TM α-helix displays a regular right-handed fold (Fig. 3), however, π- and $3_{10}$-helices have also been reported in helical membrane proteins (Heinz, Baase *et al.*, 1993; Keefe, Sondek *et al.*, 1993; Ostermeier, Harrenga *et al.*, 1997).

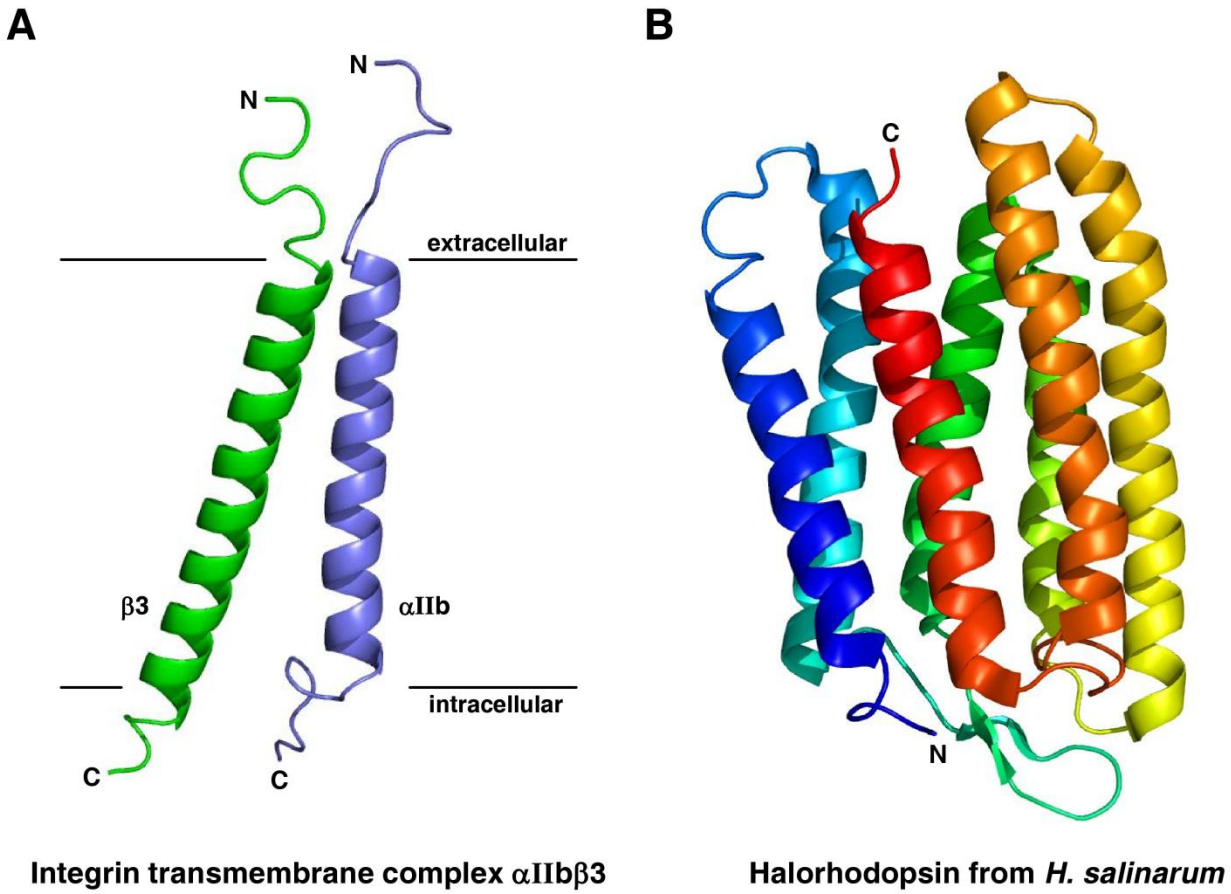



**Fig. 3** Membrane proteins of the helix bundle class. (A) Single-pass transmembrane helices. NMR structure of the transmembrane subunits αIIb and β3 of the integrin complex (PDB ID 2K9J) (Lau, Kim *et al.*, 2009). TM helices αIIb (slate) and β3 (green) cross the membrane from the extracellular to the intracellular side with 24 and 29 amino acid residues, respectively. (B) Helix bundle protein. Structure of halorhodopsin from *Halobacterium salinarum* determined by electron crystallography (PDB ID 1FBB) (Subramaniam and Henderson, 2000). The atomic resolution is 3.2 Å in-plane and 3.6 Å vertical. The protein is rainbow colored from its N- (N) to its C-terminus (C).

In short terms, π- and $3_{10}$ -helices exhibit a bonding pattern of 'i to i+5' and 'i to i+3', respectively. Thus, in contrast to an α-helix, a π-helix is more loosely packed and a $3_{10}$ -helix more tightly (Pauling and Corey, 1951; Pauling, Corey *et al.*, 1951; Low and Baybutt, 1952; Donohue, 1953; Barlow and Thornton, 1988).

Helix bundles are found to be packed in a right- or left-handed manner (Crick, 1953; Chothia, Levitt *et al.*, 1977; Alberti, 1998), yet, helices with left-handed crossing angles are more tightly packed, e.g. halorhodopsin (Fig. 3B). At large, helices in membrane proteins exhibit a broader distribution of interhelical contacts than helices in soluble proteins (Eilers, Patel *et al.*, 2002). It was shown for the helix dimer of the transmembrane domain of glycophorin A (GpA) that van der Waals interactions alone can mediate stable and specific associations between transmembrane helices (MacKenzie, Prestegard *et al.*, 1997). Nevertheless, hydrogen bonds, ion pairs and polar groups have been identified in TM helices and intrahelical interactions, where they serve a functional reason (Kuhlbrandt, Wang *et al.*, 1994; Stock, Leslie *et al.*, 1999; Popot and Engelman, 2000). Both helical membrane and soluble proteins make use of a general motif for helix interactions that relies mainly on four residues (Leu, Ala, Ile, Val) to mediate helix interactions in a fashion characteristic of left-handed helical coiled coils. A second motif for mediating helix interactions is revealed by the high occurrence and high average packing values of small and polar residues (Ala, Gly, Ser, Thr) in the helix interfaces of membrane proteins (Eilers, Patel *et al.*, 2002). In addition, the interactions of TM helices and lipids can be general or specific, providing an environment for TM protein folding but not necessarily to maintain the protein fold (Popot and Engelman, 2000). Further, for instance the helices of bacteriorhodopsin are tightly packed similar to their soluble counterparts, leaving no space for lipid molecules (Subramaniam and Henderson, 2000). Other TM bundles, e.g. the $bc_1$ complex, are more loosely packed with helix-lipid-helix contacts (Xia, Yu *et al.*, 1997).

## Beta-barrel proteins in prokaryotes and eukaryotes

### Diversity and function of β-barrels and porins

Beta-barrel proteins have been identified in the outer membranes of gram-negative bacteria (Kleffel, Garavito *et al.*, 1985), mitochondria (Forte, Adelsberger-Mangan *et al.*, 1987; Song, Midson *et al.*, 1998) and chloroplasts

(Schleiff, Eichacker *et al.*, 2003). Their name results from their tertiary fold that resembles the form of a barrel (Fig. 2B). Although different methods for the prediction of β-barrels from proteomes and genomes have been developed (Wimley, 2002; Bigelow, Petrey *et al.*, 2004), it remains imprecise how many open reading frames of a specific genome encode for β-barrel proteins. One main factor influencing their detection is that the alternate amino acid pattern of β-strands of integral β-barrels is flexible with regard to the amino acid residues facing into the barrel (Schulz, 2002). Nevertheless, multiple independent analyses indicate that 2-3 % of the genes in all Gram-negative bacteria encode β-barrel OMPs (Wimley, 2003). Further, 3.1 % of all genome-derived protein sequences from *Arabidopsis thaliana* are possibly forming β-barrels in the outer membranes of chloroplasts (Schleiff, Eichacker *et al.*, 2003). Furthermore, analyses of the yeast genome identified that 1.5 % of the encoded proteins of *Saccharomyces cerevisiae* are potentially mitochondrial β-barrels (Wimley, 2003).

The functions of TM β-barrels are as diverse as active metabolite transport, passive nutrient intake, protein translocation, membrane anchors, membrane-bound enzymes, membrane biogenesis and defense against attack proteins (Schulz, 2000; Nishimura, Tajima *et al.*, 2010; Fairman, Noinaj *et al.*, 2011). If a β-barrel forms a protein pore and therefore functions as solute filter in a cellular membrane, it is also categorized as porin (pore-protein) (Schulz, 1996). For instance, the bacterial β-barrel maltoporin facilitates the specific diffusion of maltodextrins across the outer membrane of Gram-negative bacteria (Schirmer, Keller *et al.*, 1995), whereas the eukaryotic β-barrel VDAC (voltage-dependent anion channel, defined as 'mitochondrial porin') serves as unspecific hydrophilic solute filter in the outer membrane of all mitochondria (Bayrhuber, Meins *et al.*, 2008). In addition, VDAC was shown to be involved in apoptotic related events (Abu-Hamad, Arbel *et al.*, 2009) and the energy-dependent metabolism of the cell (Lemasters and Holmuhamedov, 2006; De Pinto, Reina *et al.*, 2008).

## The structure of prokaryotic and eukaryotic β-barrels

The β-pleated sheet was first proposed by Pauling and colleagues in 1951, and later confirmed by X-ray analysis, as the dominant conformation in the fibrous proteins silk and β-keratin (Pauling and Corey, 1951; Marsh, Corey *et al.*,

1955; Chothia, 1973). It was shown that β-sheets exhibit a right-handed twist, when viewed along the polypeptide chain direction, due to a lower free energy than straight sheets or those with a left-handed twist (Chothia, 1973). The main-chains of β-sheets form two dimensional surfaces which are rarely even approximately planar (Murzin, Lesk *et al.*, 1994). In other words, amino acid residues are localized either slightly underneath or above the modular β-sheet plane resulting in a "zig-zag" line along the "arrow" (Fig. 2B). In proteins, β-sheets are able to pair in a parallel or antiparallel manner (Ashida, Tanaka *et al.*, 1981), whereas antiparallel pairing exhibits an increased stability over parallel-bonded strands (Fig. 4) (Chou, Pottle *et al.*, 1982). In transmembrane β-barrel proteins, around 9 to 11 amino acid residues are necessary to span the membrane with a strand slope angle α of 20-45° (Fig. 5) (Wimley, 2003).

The first quantitative analyses to determine and define the characteristics of a β-barrel protein were published by McLachlan in 1979 on the soluble enzyme chymotrypsin, a member of the serine protease family. He classified the properties of a basic feature, common to many proteins: a large β-sheet that twists and coils in antiparallel fashion to form a closed structure in which

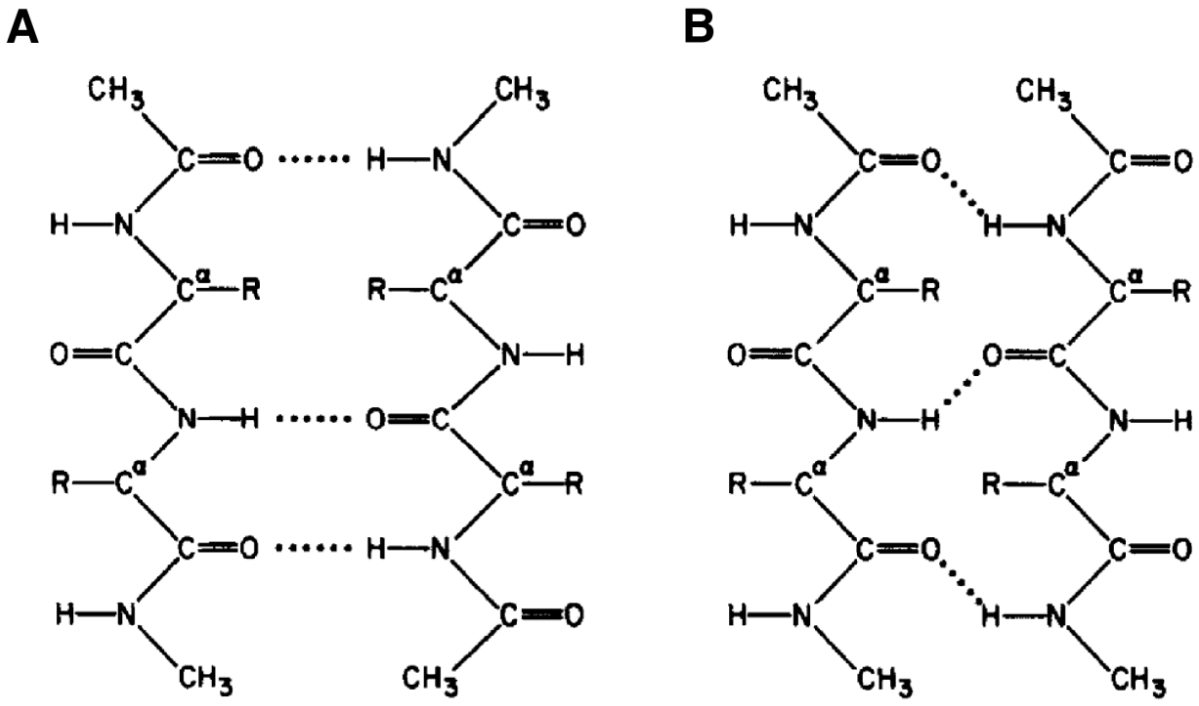



**Fig. 4** Hydrogen bonding pattern between two neighboring β-strands. (A) Antiparallel and (B) parallel strand pairing. R – amino acid residue. ((Chou, Pottle *et al.*, 1982), see reference list for permission details)

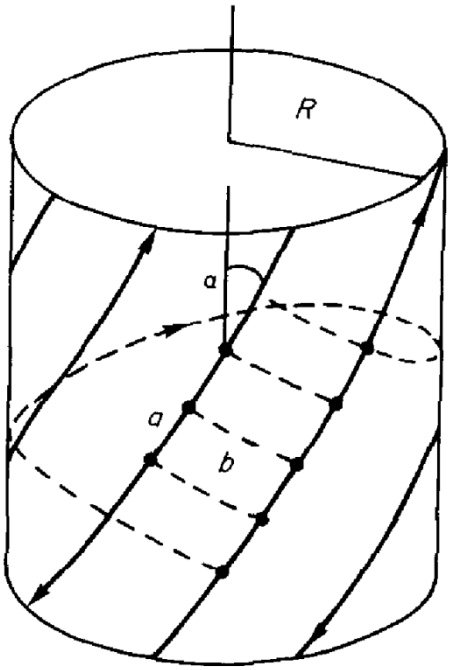



**Fig. 5** Shear number of a twisted β-sheet cylinder. Strands twist to right from bottom to top. Arrows follow left-handed helical path of hydrogen bond connections once round cylinder, emerging *S* residues higher on the first strand (here *S* = 4). *R* – β-barrel radius; *a* – distance along a strand between too adjacent residues; *b* – distance from one hydrogen bonded strand to the next; α – angle of strand slope to cylinder axis. ((McLachlan, 1979), see reference list for permission details)

the first strand is hydrogen bonded to the last, with each strand being connected to the next one by a loop region (Fig. 5 and 6A) (McLachlan, 1979; Murzin, Lesk *et al.*, 1994). In a modular cylindrical β-barrel with a right-handed twist along the strand direction, the pairs of α-carbon atoms, which are hydrogen bonded to one another on the same side of the sheet tie on a left-handed helical trace across the surface of the barrel (Fig. 5). Following this path once round the cylinder one arrives back at the first strand a certain number of residues further on, defined as shear number *S* (McLachlan, 1979). In summary, McLachlan categorized β-barrels in terms of two integral parameters: the number of strands in the β-sheet, *n*, and the "shear number", *S*, a measure of the stagger of the strands in the β-sheet. Further, he showed that the mean radius of a β-barrel, α, and the extent to which strands are tilted relative to its axis are determined by the values of *n* and *S* (McLachlan, 1979; Murzin, Lesk *et al.*, 1994). Murzin and colleagues further established these results and validated them in an

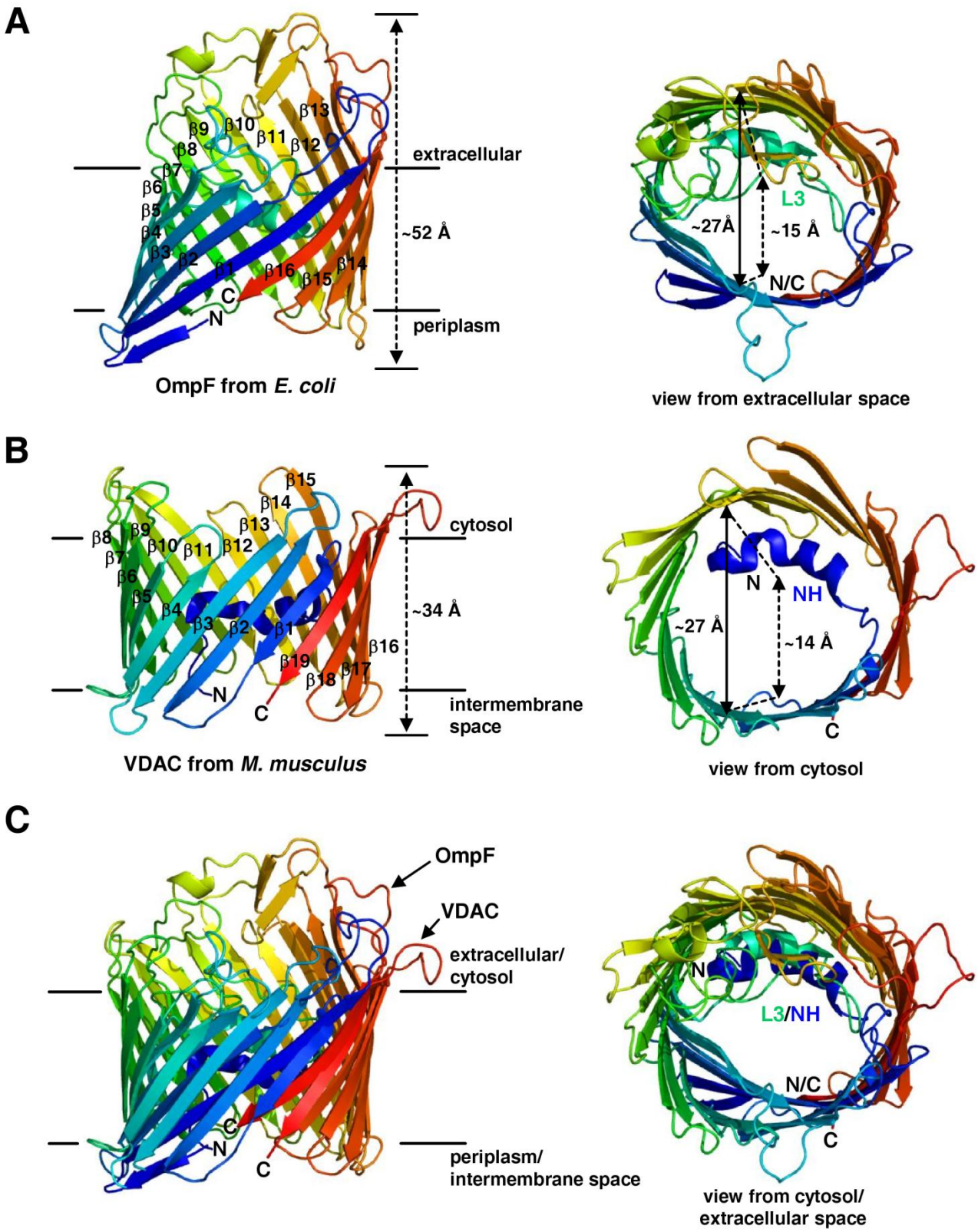



**Fig. 6** The architecture of transmembrane β-barrels. (A) Example of a prokaryotic β-barrel. Refined crystal structure of the porin OmpF from *E. coli* at 2.4 Å (PDB ID 2OMF) (Cowan, Schirmer *et al.*, 1992). (B) Example of a eukaryotic β-barrel. Crystal structure of VDAC from *M. musculus* at 2.4 Å (PDB ID 3EMN) (Ujwal, Cascio *et al.*, 2008). (C) Structural alignment of OmpF and VDAC. Computations were performed

accompanying paper with the observed structures of soluble and integral β-barrel proteins (Murzin, Lesk *et al.*, 1994; Murzin, Lesk *et al.*, 1994).

Until the year 2008, common belief was that the nature of strand pairing in a membrane bound β-barrel protein is strictly antiparallel, resulting in an even number of β-strands. This stands true with regard to prokaryotic β-barrels (Fig. 6A) (Schulz, 2000; Arnold, Poynor *et al.*, 2007). Therefore, a β-hairpin was assumed and further shown to be the minimal module of prokaryotic TM β-barrels (Wimley, 2003; Arnold, Poynor *et al.*, 2007). In contrast, the solution structure of the eukaryotic porin VDAC revealed an uneven number of β-strands with one parallel strand pairing in between the first and last β-strand (Fig. 6B) (Bayrhuber, Meins *et al.*, 2008; Hiller, Garces *et al.*, 2008; Ujwal, Cascio *et al.*, 2008). Pusnik and colleagues later suggested that related mitochondrial porins, e.g. Tom40, share the same exceptional fold (Pusnik, Charriere *et al.*, 2009). Hence, β-barrels of the outer membrane of mitochondria were classified as 'mitochondrial porins' distinct from their bacterial counterparts due to their unique tertiary structure build-up.

To this date, bacterial β-barrels have been observed with a minimum number of 8 (Pautsch and Schulz, 1998) and a maximum number of 24 β-strands (Remaut, Tang *et al.*, 2008). As described above, the eukaryotic porin VDAC consists of an uneven number of β-strands, namely 19, with one parallel pairing of the β-strands 1 and 19 (Fig. 6B). Prokaryotic and eukaryotic β-barrels were observed as monomers (Yildiz, Vinothkumar *et al.*, 2006), dimers (Snijder, Ubarretxena-Belandia *et al.*, 1999), trimers (Cowan, Schirmer *et al.*, 1992) and in the form of complex oligomers (Zalk, Israelson *et al.*, 2005; Robert, Volokhina *et al.*, 2006). Usually, each subunit of a β-barrel complex is formed by a single

---

**(*Fig. 6 continued*)** with the CCP4 program suite 6.2.0 (Winn, Ballard *et al.*, 2011) using the 'superpose' utility according to (Krissinel and Henrick, 2004). Prokaryotic β-barrels exhibit an even number of β-strands (β1-β16) with strict antiparallel pairing (A). Eukaryotic porin VDAC was shown to feature an uneven number of β-strands (β1-β19) with one parallel strand pairing between β1 and β19 (B). OmpF and VDAC have a constricting element in common, the extracellular loop L3 and the N-terminal helix (NH), respectively (C) (Bayrhuber, Meins *et al.*, 2008). In both cases, the pore is narrowed to around 14 Å. N – N-terminus; C – C-terminus.

polypeptide chain (Fig. 2B). In addition, trimeric bacterial OMPs are typically built from 16/18 β-strands (Fig. 6A). However, β-barrels exist where a single pore is formed by multiples of one subunit. Two important examples are the protein toxin α-haemolysin from *Staphylococcus aureus* (Song, Hobaugh *et al.*, 1996) and the bacterial efflux pump TolC from *E. coli* (Koronakis, Sharff *et al.*, 2000).

The domains of a β-barrel which are in close contact with the hydrophobic core domain of the lipid-bilayer (Fig. 1) are cryptically encoded in their amino acid sequences. Beta-strands in this region can be identified by a dyad repeat pattern of alternating hydrophobic/hydrophilic residues, either oriented towards the hydrophobic domain of the bilayer or the aqueous protein pore, respectively. Similar to helical membrane proteins, the composition of amino acid residues located in the lipid-exposed surfaces of β-barrels is characterized by an abundance of phenylalanine, tyrosine, tryptophan, valine and leucine and the strong exclusion of polar and charged residues (Seshadri, Garemyr *et al.*, 1998; Ulmschneider and Sansom, 2001; Wimley, 2002; Wimley, 2003). Importantly, no cysteine residue has been identified in the TM domain of bacterial β-barrels to this date (Freeman and Wimley, 2010). In contrast, eukaryotic porin VDAC1 from human mitochondria revealed one cysteine residue facing into the lipid bilayer (Aram, Geula *et al.*, 2010). The abundance of aromatic residues is especially striking at the bilayer interfaces, where they constitute about 40 % of the lipid-exposed amino acids. Despite the fact that many β-barrels contain a water-filled channel, the abundant residues in the pore interior are predominantly the small or polar amino acids glycine, threonine, serine, asparagine and glutamine. Tyrosine is also relatively abundant, whereas the other hydrophobes are rare (Wimley, 2002; Wimley, 2003).

## Interplay between functional and structural elements of β-barrels with weakly stable TM regions

Along with the structural features of TM β-barrels described in the previous chapter, additional structural elements or deviations from the 'model' β-barrel (Fig. 5) exist, namely the amino acid composition of TM β-strands, a mobile element or plug domain (Schulz, 2002; Wimley, 2002). In 2009, Naveed and colleagues used a computational approach to identify weakly stable regions,

oligomerization state and protein-protein interfaces in the TM domain of prokaryotic β-barrel proteins (Naveed, Jackups *et al.*, 2009). Their *in silico* method is based on a physical interaction model, a simplified conformational space for efficient enumeration, and an empirical potential function from a detailed combinatorial analysis. They found that the instability in a single β-strand or regions of β-strands results from the presence of amino acids disrupting the regular dyad repeat pattern of residues in a module β-strand with the consequence of energetically unfavorable molecular interactions. Moreover, they discovered the existence of three general mechanisms for the stabilization of weakly stable regions: out-clamps, in-plugs and oligomerization. These structural characteristics of integral β-barrels are very often directly linked to the structural organization and/or the function of a TM β-barrel protein in a cellular membrane. For each mechanism, fundamental examples are given in the following.

In the outer membrane β-barrel enzyme PagP from *E. coli*, amino acid R59 in β-strand 2 (also termed β-strand B) was identified as the most 'unstable' or 'destabilizing' residue (Naveed, Jackups *et al.*, 2009). The putative unfavorable interaction of R59 with lipid-molecules (Ahn, Lo *et al.*, 2004) is counterbalanced by the N-terminal amphipathic α-helix of PagP through binding to strand 2 and its neighboring strands on the outside of the β-barrel (Naveed, Jackups *et al.*, 2009). It was further shown, that the helix of PagP is not required for folding and assembly of the protein, but instead for the stabilization of PagP after folding and insertion into the membrane are complete. In that way, the N-terminal helix acts as a clamp (Huysmans, Radford *et al.*, 2007) and this mechanism of TM β-barrel stabilization was defined as 'out-clamp' (Naveed, Jackups *et al.*, 2009). Another example for the mechanism of 'out-clamping' is the 14 stranded β-barrel α-haemolysin. Among the top four most unstable residues in the hairpin, two (K110 and Y148) are stabilized by a barrel out-clamp through extensive hydrogen bond interactions with the N-terminal of the soluble domain (K110 with Q150B and N173B; Y148 with N178B) (Naveed, Jackups *et al.*, 2009).

The second mechanism is similar to an 'out-clamp' with the exception of the location of the 'stabilizer', which binds to weakly stable regions inside of the barrel pore. These so called 'in-plugs' are often formed by interstrand loops and the N-terminus (Fig. 6C) (Naveed, Jackups *et al.*, 2009). Naveed *et al.*

investigated the stability of several in-plug containing β-barrel proteins, e.g. the ferric hydroxamate uptake receptor protein FhuA (Ferguson, Hofmann *et al.*, 1998), by using a statistical mechanics model to calculate the relative melting temperature, $T_m$, which is an indicator for the overall protein stability. They determined that the in-plug domains in these proteins are important in stabilizing the weakly stable strands (Naveed, Jackups *et al.*, 2009), for instance the in-plug of FhuA, which consists of two α-helices and a 4-strand β-sheet (Ferguson, Hofmann *et al.*, 1998). Naturally, an in-plug shapes the interior of a β-barrel, thus the barrier characteristics of the membrane protein, often through constriction of the pore (Fig. 6C). Therefore, the in-plugs of many TM β-barrels are important for controlling the transport of their target molecule(s) (Cowan, Schirmer *et al.*, 1992; van den Berg, 2005; Ujwal, Cascio *et al.*, 2008).

NMR measurements on human VDAC1 revealed that the N-terminal α-helical region is stabilizing the protein pore by forming hydrogen bonds to amino acid residues in strands 8-19 within the protein pore (Fig. 6B) (Ujwal, Cascio *et al.*, 2008; Schneider, Etzkorn *et al.*, 2010). This suggests the function of an in-plug domain as defined by (Naveed, Jackups *et al.*, 2009). Based on the high atomic resolution structure of mammalian VDAC1, several theoretical models for the gating mechanism and voltage sensitivity of VDAC were proposed (Ujwal, Cascio *et al.*, 2008; Choudhary, Ujwal *et al.*, 2010). At large, the channel characteristics of VDAC might involve movement of the helical domain, but does not fully explain the observed channel properties (Choudhary, Ujwal *et al.*, 2010). For bacterial porins such as OmpC or OmpF it has been long established that the extracelluar loop L3 is involved in the permeability properties of these porins (Fig. 6A) (Cowan, Schirmer *et al.*, 1992; Phale, Schirmer *et al.*, 1997; Basle, Rummel *et al.*, 2006). Further, the constriction loop L3 seems to be contributing to the stability of the pore, but compared to the latching loop L2, L3 appears to play a less significant role for the structural integrity of bacterial porin (Lakey, Lea *et al.*, 1991; Phale, Philippsen *et al.*, 1998). It should be noted, that the position and the influence of the helix onto the overall VDAC structure might be reminiscent of the L3 loop in bacterial porins in general (Fig. 6C) (Bayrhuber, Meins *et al.*, 2008). Another structural similarity of mitochondrial porins to their bacterial counterparts is the fact that both N- and C-termini are facing into the intermembrane space/periplasm (Fig. 6). Thus, mitochondrial

porins retain the sidedness found in all bacterial OMPs (Schulz, 2002; Bayrhuber, Meins *et al.*, 2008). In summary, the interplay of in-plug domains and their regarding weakly stable regions inside of the β-barrel pore suggests a direct contribution to the protein function (Naveed, Jackups *et al.*, 2009).

The third mechanism applies to TM β-barrels featuring extended and adjacent weakly stable regions. Naveed *et al.* determined β-strands 1 to 6, 15 and 16 of OmpF from *E. coli* to be weakly stable (Fig. 6A). It turns out that the strands involved in oligomerization of OmpF (Fig. 2B) are strands 1–5 and strand 16 (Cowan, Schirmer *et al.*, 1992; Naveed, Jackups *et al.*, 2009). Oligomerization of β-barrel membrane proteins was revealed as the third possibility to stabilize weakly stable regions in bacterial TM β-barrels. Naveed and colleagues were able to detect interfaces for protein-protein interaction in TM β-barrels and to calculate the oligomerization state of the complex. In addition, Naveed *et al.* developed an index, so called oligomerization index ϱ, which summarizes the energy deviation of unstable β-strands from the overall expected energy value (Naveed, Jackups *et al.*, 2009). Index ϱ essentially describes the average deviation of energy values of unstable strands from the mean energy of all strands. If ϱ of a certain TM β-barrel protein is greater than 2.5, the protein is predicted to form oligomers.

## Working models for the folding and insertion of TM β-barrel proteins

Formation of native TM β-barrels in the outer membranes of bacteria and endosymbiotically derived organelles such as mitochondria is accomplished by an evolutionary conserved assembly machinery (Dolezal, Likic *et al.*, 2006). The β-barrel assembly machinery (Bam) enables OMP maturation in the outer membrane of bacteria (Hagan, Silhavy *et al.*, 2011). Its central component is a membrane protein termed Omp85 (or BamA), which is built of one or more N-terminal polypeptide transport-associated (POTRA) domains and a single C-terminal 16-stranded TM β-barrel (Sanchez-Pulido, Devos *et al.*, 2003; Clantin, Delattre *et al.*, 2007). Omp85 and its eukaryotic homologues constitute the evolutionary link between the prokaryotic and eukaryotic TM β-barrel assembly machineries (Reumann, Davila-Aponte *et al.*, 1999; Paschen, Waizenegger *et al.*, 2003).

In mitochondria, the sorting and assembly machinery (SAM), also termed topogenesis of mitochondrial outer membrane β-barrel proteins (TOB), promotes folding and insertion of TM β-barrel proteins such as VDAC and Tom40. The functional homologue of BamA in mitochondria is a protein termed Tob55 or Sam50 (Kozjak, Wiedemann *et al.*, 2003; Paschen, Waizenegger *et al.*, 2003). Obviously, several other subunits of the β-barrel assembly machinery and numerous other transfer/transport complexes are necessary to enable TM β-barrel maturation in bacteria/eukarya (Tommassen, 2010; Hagan, Silhavy *et al.*, 2011).

The work of many years has shed light on this important cellular process, still, the last step of β-barrel assembly and membrane insertion is the least understood. For instance, it is not clear whether insertion could occur through the β-barrel of a BamA monomer or within a hydrophobic cavity created either by the lipoproteins (BamB, C, D and E), or by the association of multiple Bam complexes (Hagan, Silhavy *et al.*, 2011). The following model has been suggested for OMP folding and insertion: binding of an unfolded OMP to the polypeptide transport-associated (POTRA) domains of BamA by β-strand augmentation initiates β-structure formation. The POTRA domain thereby satisfies the hydrogen bonds on one edge of the β-sheet as folding proceeds. Ultimately, the two edges of the sheet are brought together, and the barrel is inserted into the membrane (Gatzeva-Topalova, Walton *et al.*, 2008; Hagan, Silhavy *et al.*, 2011).

*In vitro* analyses through TDFQ (time-resolved distance determination by fluorescence quenching) in the absence of Bam showed that the four β-hairpins of the eight-stranded prokaryotic β-barrel OmpA insert into the membrane in a concerted fashion (Pautsch and Schulz, 1998; Kleinschmidt, den Blaauwen *et al.*, 1999). Thus, β-barrel folding and insertion likely occur in a combined process. In addition, recent works reported on the assembly of the staphylococcal pore-forming toxin α-hemolysin. Via single-molecule fluorescence imaging, it was found that the assembly from the monomer to the complete heptamer occurs in < 5 ms, thus, extremely rapid (Thompson, Cronin *et al.*, 2011). Consequently, it is incredibly challenging to observe and interpret the structures of substrates at different stages *in vivo* and *in vitro* (Hagan, Silhavy *et al.*, 2011; Thompson, Cronin *et al.*, 2011).

Several *in vitro* approaches have yielded insights into the molecular mechanisms of β-barrel folding in the absence of the assembly machinery (Wimley, 2003). For example, TDFQ analyses by Kleinschmidt and colleagues determined, in addition to above described results, three structurally distinct membrane-bound folding intermediates (Kleinschmidt, den Blaauwen *et al.*, 1999). Further, *in vitro* investigations of OmpA secondary and tertiary structure formation via circular dichroism (CD) and tryptophan fluorescence spectroscopy revealed that the process of barrel formation is cooperative and spontaneous (Kleinschmidt and Tamm, 2002). In detail, the formation of β-sheet secondary structure and closure of the β-barrel of OmpA were correlated with the same rate constant and coupled to the insertion of the protein into the lipid bilayer. This all-or-none insertion process makes sense for β-barrels with regard to their network of nonlocal hydrogen bonds (Wimley, 2003).

The formation of TM β-barrel structure is clearly distinct from the proposed two-stage folding model of TM helix bundles (Popot and Engelman, 1990). Different folding mechanisms of integral membrane proteins are likely a consequence of the very different intramolecular hydrogen bonding and hydrophobicity patterns of TM helix bundles and TM β-barrels (Kleinschmidt and Tamm, 2002). Likewise cooperative assembly of β-barrels has been observed for other bacterial β-barrels such as OmpG (Conlan and Bayley, 2003), hydrophobic model peptides (Wimley, Hristova *et al.*, 1998) and pore-forming toxins (Bayley, 1997; Ramachandran, Heuck *et al.*, 2002; Wimley, 2003).

However, even though most OMPs (bacterial outer membrane protein) fold spontaneously into lipid bilayers *in vitro* (Hagan, Silhavy *et al.*, 2011), bacterial β-barrels do not have the same ability to fold into any single bilayer environment (Burgess, Dao *et al.*, 2008). It was further discovered that prokaryotic β-barrel folding efficiency is dependent on pH, temperature and the lipid type *in vitro*. This suggests that although environmental factors influence folding, bacterial TM β-barrels also have intrinsic qualities that profoundly modulate their folding. Beta-barrel assembly complexes such as Bam must therefore overcome these intrinsic differences (Burgess, Dao *et al.*, 2008; Hagan, Silhavy *et al.*, 2011).

Similar to prokaryotic porins, the eukaryotic β-barrels VDAC and Tom40 spontaneously insert and fold into lipid bilayers in the absence of folding

assistants or energy sources like ATP *in vitro* (Hill, Model *et al.*, 1998; Becker, Bannwarth *et al.*, 2005; Shanmugavadivu, Apell *et al.*, 2007). However, in stark contrast to β-barrel membrane proteins, human VDAC1 exhibited varying secondary structure content in detergent micelles and phospholipid bilayers, with higher content of β-sheet and lower content of α-helix when inserted and folded into lipid bilayers (Shanmugavadivu, Apell *et al.*, 2007). Comparable observations were made for fungal and mammalian Tom40 reconstituted in detergent micelles and liposomes (Becker, Bannwarth *et al.*, 2005; Stutz, 2009).

Nevertheless, membrane insertion of detergent solubilized β-barrel proteins allows the assignment of three formations for bacterial and mitochondrial porins: a water-soluble/aqueous form with low β-sheet percentage, a detergent solubilized form with enriched β-sheet content and the membrane-embedded native state (Conlan and Bayley, 2003; Shanmugavadivu, Apell *et al.*, 2007). Therefore, it can be suggested that the native structure of prokaryotic and eukaryotic TM β-barrels requires the environment of a lipid bilayer.

Importantly, recent investigations based on single-molecule force spectroscopy contradict the concerted folding model for TM β-barrels. Sapra *et al.* mechanically unfolded fully folded bacterial OmpG, which was reconstituted in its native *E. coli* lipid bilayer (Sapra, Damaghi *et al.*, 2009), showing that each of the seven β-hairpins of OmpG can be unfolded individually (Yildiz, Vinothkumar *et al.*, 2006; Sapra, Damaghi *et al.*, 2009). Similar results were gained for the refolding of partly unfolded OmpG (Damaghi, Koster *et al.*, 2011). Conventional unfolding experiments with thermal or chemical denaturants induce very different unfolding scenarios for, in most cases, solubilized OmpG. To this date, reversible folding of TM β-barrels was only achieved at acidic pH, for instance the bacterial outer membrane phospholipase OmpLA (Moon and Fleming, 2011). With this in mind, the question is raised whether the observed folding/unfolding scenarios for TM β-barrels are the result of the applied measurement conditions or not. The hydrophobic environment of the lipid bilayer might energetically 'trap' the fully folded or partly folded β-barrel as the anisotropic environment of the membrane contributes to the structural stability of the embedded protein. Forces might have to be repetitively applied to a sequence of unfolding intermediates until the entire protein has been unfolded

(Sapra, Damaghi *et al.*, 2009; Damaghi, Koster *et al.*, 2011). Thus, one simple explanation for the unfolding pathway reported by Sapra *et al.* could be the effect of the lipid environment on the native β-barrel structure.

## The eukaryotic β-barrel protein Tom40

### Function of Tom40 in the outer mitochondrial membrane

Mitochondria are essential membrane-enclosed subcellular organelles found in almost all eukaryotic cells (Henze and Martin, 2003). They have the ability of governing either life or death of the cell. This description arose from the fact that mitochondria produce ATP as the energy source of the cell on the one hand, and play a crucial role in apoptotic events on the other hand (Mokranjac and Neupert, 2009). Furthermore, mitochondria host other crucial cellular processes, such as signaling and metabolic pathways involving lipids, amino acids and iron (Chacinska, Koehler *et al.*, 2009). It is generally accepted that mitochondria evolved from endosymbiotic purple non-sulphur bacteria, the α-proteobacteria (Cavalier-Smith, 2006).

Mitochondria are surrounded by two distinct membranes, defined as outer membrane and inner membrane (Fig. 7A). This structural arrangement results in three separate mitochondrial spaces: the outer membrane walls the cytosol from the intermembrane space, whereas the inner membrane fences the intermembrane space from the innermost mitochondrial space, termed matrix (Palade, 1953). Mitochondria contain about 1000 (yeast) to 1500 (human) different proteins and only ~1 % of all mitochondrial proteins are synthesized on ribosomes in the matrix, whereas the remainder are synthesized on ribosomes in the cytosol (Chacinska, Koehler *et al.*, 2009). This is thought to be an indirect consequence of a process known as ‘‘Müller’s ratchet’’ where asexually reproducing endosymbiotes tend to accumulate deleterious mutations, therefore favouring the transfer of genes to the host nucleus where sexual recombination can counterbalance such mutation (Adams and Palmer, 2003; Perry, Rimmer *et al.*, 2008). As a result, mitochondrial genomes are severely reduced in size when compared with their free-living α-proteobacterial ancestors (Burger, Gray *et al.*, 2003; Perry, Rimmer *et al.*, 2008). Hence, mitochondria host a complex machinery enabling the transport of preproteins to

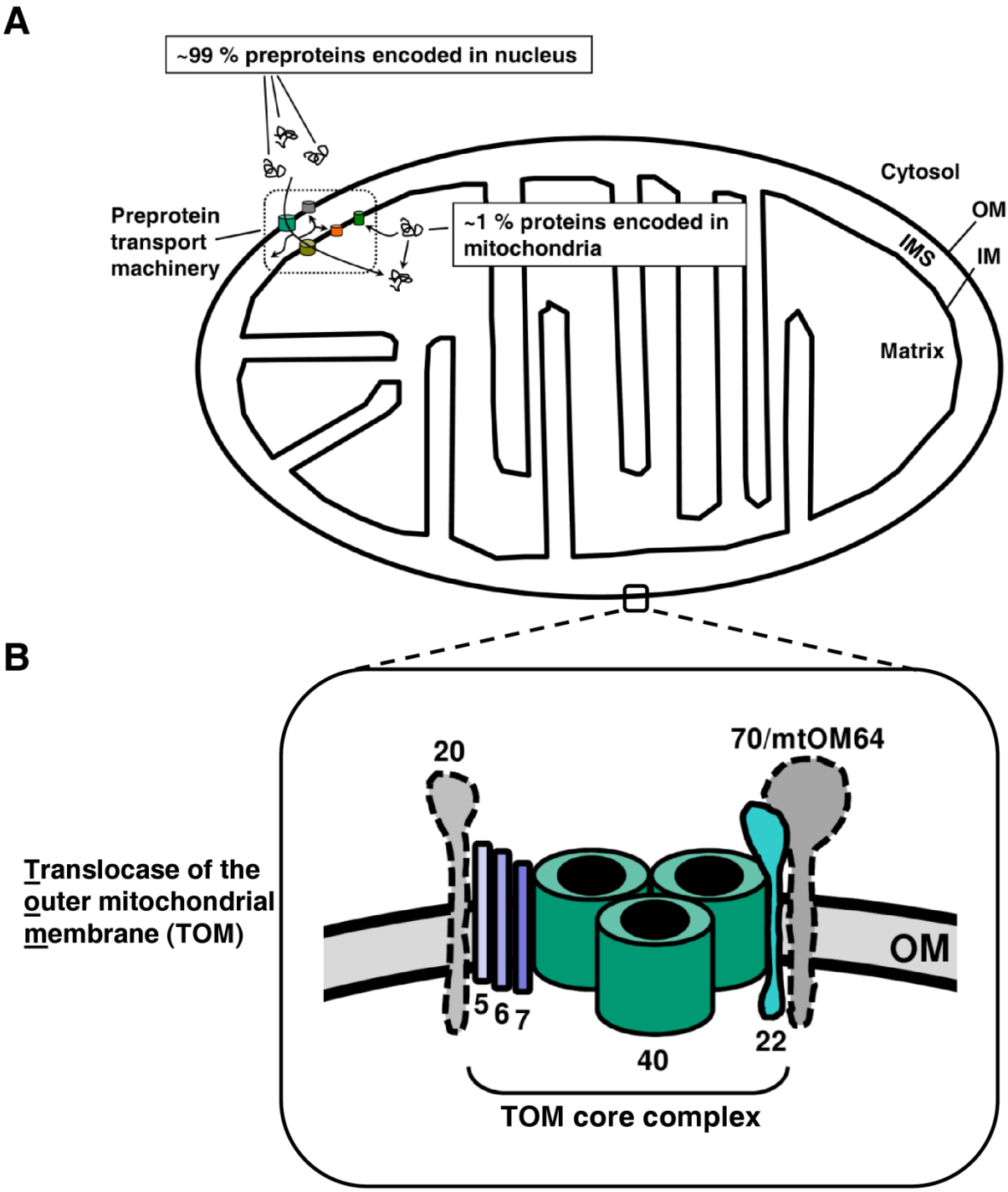



**Fig. 7** Mitochondria and the TOM complex. (A) Structure of mitochondria. The outer membrane (OM) separates the cytosol from the intermembrane space (IMS). The inner membrane (IM) walls the IMS from the matrix. About 99 % of all mitochondrial proteins are encoded in the nucleus and synthesized on cytosolic ribosomes. Remaining proteins are encoded in the matrix. The preprotein transport machinery enables transport of mitochondrial proteins across and into the outer or inner membrane, respectively. (B) Translocase of the outer mitochondrial membrane (TOM). The TOM core complex consists of the pore-forming subunit Tom40, Tom22 and the small Tom proteins 5, 6 and 7. The full TOM complex from animals/fungi compromises the receptors Tom20 and Tom70; in plants/algae, mtOM64 substitutes Tom70.

their specific mitochondrial location (Fig. 7A). Two major preprotein transport systems, defined as TOM (translocase of the outer mitochondrial membrane) and TIM23/TIM22 (translocase of the inner mitochondrial membrane) have been identified in the outer and inner membrane, respectively. However, several other mitochondrial translocases and protein complexes are required for protein transport into mitochondria (Chacinska, Koehler *et al.*, 2009; Mokranjac and Neupert, 2009).

Tom40 was the first component identified as part of the transmembrane machinery which imports preproteins into mitochondria (Vestweber, Brunner *et al.*, 1989). Vestweber and colleagues showed that an engineered precursor protein that sticks in the import site of isolated mitochondria can be specifically photo-crosslinked to a mitochondrial protein. IgG antibodies directed against this ~42 kDa protein blocked the import of preproteins into mitochondria. This protein was therefore revealed as part of the precursor import machinery and it was termed 'import site protein 42' or ISP 42. Under conditions that preserved the integrity of the outer membrane, ISP 42 was accessible to antibody labeling and protease digestion. Thus, ISP 42 is located in the outer membrane as integral membrane protein and exposed to the cytosolic surface. Further studies showed ISP 42 to be indispensable for cell viability and protein import since it constitutes the preprotein conducting, pore-forming subunit of the TOM complex (Baker, Schaniel *et al.*, 1990; Hill, Model *et al.*, 1998; Ahting, Thieffry *et al.*, 2001). Hence, ISP 42 was named Tom40 (Juin, Thieffry *et al.*, 1997).

## The translocase of the outer mitochondrial membrane

*Composition and diversity*

The core element of the translocase of the outer mitochondrial membrane, termed TOM core or general import pore (GIP) complex, consists of the main subunit Tom40 and the small Tom proteins 5,6,7 and Tom22 (Fig. 7B) (Dekker, Ryan *et al.*, 1998; Ahting, Thun *et al.*, 1999). In animals and fungi, the TOM core complex is loosely associated with the receptor subunits Tom20 and Tom70 (Dekker, Ryan *et al.*, 1998), whereas in plants a protein termed mtOM64 might be the functional replacement for absent Tom70 (Chew, Lister *et al.*, 2004; Chan, Likic *et al.*, 2006; Lister, Carrie *et al.*, 2007). All Tom subunits except for Tom40 traverse the outer mitochondrial membrane as single TM helices. TOM core

complex associated with Tom20 and Tom70/mtOM64 is defined as TOM (holo) complex (Fig. 7B).

The TOM core complex components have been found in all eukaryotic lineages, suggesting their presence in a common mitochondrial ancestor. However, the phylogenetic distribution of the additional components that optimize the function of the TOM complex is lineage-specific (Dolezal, Likic *et al.*, 2006; Rada, Dolezal *et al.*, 2011). For instance, the outer membrane of hydrogenosomas or mitosomes of certain parastitic protists revealed a highly reduced TOM machinery such as *Trichomonas vaginalis* which lacks Tom20 and Tom70 (Dagley, Dolezal *et al.*, 2009; Dolezal, Dagley *et al.*, 2010; Rada, Dolezal *et al.*, 2011). The molecular weight of the TOM complex has been reported in between 170-590 kD depending on the organism and, interestingly, also on the method used for molecular weight determination (Dekker, Ryan *et al.*, 1998; Ahting, Thun *et al.*, 1999; Dagley, Dolezal *et al.*, 2009). The exact stoichiometry of the TOM complex from a specific organism is therefore still a matter of controversy (Mager, Sokolova *et al.*, 2010).

*Function of the TOM subunits and preprotein translocation*

The small Tom proteins are found to be important for maintaining the stability and dissociation as well as the biogenesis of the TOM complex (Dembowski, Künkele *et al.*, 2001; Schmitt, Ahting *et al.*, 2005; Becker, Wenz *et al.*, 2011). Tom6 and Tom7 play antagonistic roles, with Tom6 stabilizing the large TOM complex, whereas Tom7 favors its dissociation. This supports the view of a dynamic organization of the TOM complex. Tom5 promotes assembly of the TOM complex and participates in preprotein transfer from Tom22 to the Tom40 channel (Dietmeier, Hönlinger *et al.*, 1997; Model, Meisinger *et al.*, 2001; Wiedemann, Kozjak *et al.*, 2003; Chacinska, Koehler *et al.*, 2009). Tom20 is the initial recognition site for preproteins with presequences and transfers the preproteins to the central receptor Tom22 (Saitoh, Igura *et al.*, 2007; Chacinska, Koehler *et al.*, 2009). Tom70 forms the primary recognition site for precursors of inner membrane metabolite carrier proteins, which carry multiple internal targeting signals. Tom70 further transfers these precursor proteins to Tom22 (Kiebler, Keil *et al.*, 1993; van Wilpe, Ryan *et al.*, 1999; Chacinska, Koehler *et al.*, 2009).

Recent works using NMR spectroscopy revealed competitive binding of the cytosolic domain of plant Tom22 with the presequence segment of preproteins to the cytosolic domain of Tom20 (Rimmer, Foo *et al.*, 2010). The precursor protein is thus placed in the immediate vicinity of the Tom40 translocation pore. It is suggested that displacement of the presequence by the cytosolic domain of Tom22 could facilitate the interaction of the preprotein with Tom40, thus, promoting the translocation of the precursor protein across the outer membrane. *In vivo* and *in organello* site-specific photocross-linking techniques provided snapshots of fungal Tom22 at work (Shiota, Mabuchi *et al.*, 2011). Their findings indicate that the acidic region of Tom22, which is part of the cytosolic domain, competes with the presequence for the interaction with Tom20 in line with (Rimmer, Foo *et al.*, 2010). Shiota *et al.* also hypothesized that Tom20 and Tom22 recognize the hydrophobic side and the positively charged side, respectively, of the amphiphilic helix of the presequence simultaneously (Shiota, Mabuchi *et al.*, 2011). Further, it was shown that the presequence and the IMS domain of Tom22 compete with each other to bind to Tim50, a component of the TIM23 machinery. Shiota and colleagues therefore suggested the following scenario: the transient interaction between Tim50 and Tom22 brings the TIM23 complex closer to the TOM complex, the presequence of the incoming precursor protein binds to the IMS domain of Tom22 when the Tom22 IMS domain is not occupied with Tim50. Then binding of Tim50 to Tom22 clears the presequence from Tom22, and Tim50 receives the presequence, thereby facilitating efficient presequence transfer from the TOM complex to the downstream TIM23 complex (Shiota, Mabuchi *et al.*, 2011). In line with these results, it was suggested early on that Tom40 and Tom22 form the functional core unit which promotes transfer of preproteins across the outer mitochondrial membrane (Meisinger, Ryan *et al.*, 2001).

In summary, preprotein translocation through the TOM complex might function in terms of a ‘relay system’ where competitive binding in between presequences and regarding TOM domains enables transport of preproteins (i) from the cytosolic site to the translocation channel and (ii) further to the machineries in the IMS. It is not understood as to how preproteins are translocated through the Tom40 pore. Further, an import motor involved in TOM promoted transport has not been identified as of this date. Nevertheless, this proposed mechanism can be effectually correlated with the previously

proposed 'binding chain hypothesis': the Tom proteins provide a chain of binding surfaces, including various types of non-covalent interaction, to guide presequence containing preproteins into mitochondria (Pfanner and Geissler, 2001).

**Diversity and structure of Tom40**

With a molecular weight ranging from around 34 to 42 kD, Tom40 has been identified in numerous eukaryotic species including fungi, animals, plants, microsporidians, trimastix, apicomplexa (Macasev, Whelan *et al.*, 2004) and several other parasitic protists, namely *Giardia intestinalis* (Dagley, Dolezal *et al.*, 2009), *Trichomonas vaginalis* (Rada, Dolezal *et al.*, 2011) and *Entamoeba histolytica* (Dolezal, Dagley *et al.*, 2010). A high atomic resolution structure of Tom40 has not been solved as of this date. Nevertheless, secondary structure investigations of Tom40 from various organisms via CD and FTIR (fourier transform infrared spectroscopy) yielded high β-sheet content and low α-helical portion comparable to mitochondrial, but not to bacterial porin (Hill, Model *et al.*, 1998; Ahting, Thieffry *et al.*, 2001; Suzuki, Kadowaki *et al.*, 2004; Becker, Bannwarth *et al.*, 2005; Malia and Wagner, 2007; Stutz, 2009). In addition, a common origin for VDAC and Tom40 was suggested (Pusnik, Charriere *et al.*, 2009) and it is therefore generally accepted that Tom40 constitutes a β-barrel structure similar to VDAC (Fig. 6B).

Examinations of functionally refolded or native Tom40s from different organisms using electron microscope and electrophysiological techniques as well as pore size-probing with gold-modified precursors determined an effective pore diameter of 22-25 Å sufficient for the transport of preproteins (Hill, Model *et al.*, 1998; Schwartz, Huang *et al.*, 1999; Schwartz and Matouschek, 1999; Ahting, Thieffry *et al.*, 2001). Isolation of TOM core complex from *Neurospora crassa* mitochondria allowed three-dimensional reconstruction by electron tomography. The structure exhibits two open spaces traversing the complex with a diameter of ~21 Å and a height of ~70 Å (Ahting, Thun *et al.*, 1999). In an accompanying paper, it was shown that Tom40 is the key structural element of the TOM core complex as it forms the pore in the outer mitochondrial membrane (Ahting, Thieffry *et al.*, 2001). Cryo-electron microscopy on a TOM core-Tom20 complex isolated from mitochondria of *Saccharomyces cerevisiae*

showed a triangular shape with three elliptical pores of ~15 Å × 25 Å. The TOM core-Tom20 complex openings are also thought to be formed by yeast Tom40 (Fig. 7) (Model, Meisinger *et al.*, 2008).

The orientation of the N- and C-terminus of Tom40 in the outer mitochondrial membrane was assessed through protease protection assays. Tom40 from *Neurospora crassa* and Tom40 isoform B from *Rattus norvegicus* revealed an IMS localization for both termini in line with the 'bacterial sidedness' of mitochondrial porin VDAC (Künkele, Juin *et al.*, 1998; Kinoshita, Mihara *et al.*, 2007; Bayrhuber, Meins *et al.*, 2008). However, the N-terminal segment of rat Tom40 isoform A faces into the cytosol with the C-terminus being exposed to the IMS (Suzuki, Kadowaki *et al.*, 2004). Yeast Tom40 N-terminal segment also points into the cytosol, whereas the topology of the C-terminus is not known (Hill, Model *et al.*, 1998).

**Functional properties of Tom40**

The interior of the Tom40 pore is mostly hydrophilic and cation-selective with the ability to bind and prevent the aggregation of precursor proteins (Hill, Model *et al.*, 1998; Künkele, Juin *et al.*, 1998; Esaki, Kanamori *et al.*, 2003). Moreover, Esaki *et al.* concluded that the Tom40 channel (or pore) offers an optimized environment to translocate non-native precursor proteins. It is therefore assumed that Tom40 plays an active role in preprotein transport, in contrast to the nonstick tunnel of the ribosome for polypeptide exit (Esaki, Kanamori *et al.*, 2003; Gabriel, Egan *et al.*, 2003). This hypothesis can be further established by other works revealing an interaction of Tom40 with precursor proteins on the *cis* and *trans* side of the outer membrane (Rapaport, Neupert *et al.*, 1997; Kanamori, Nishikawa *et al.*, 1999). Furthermore, Tom40 has the capability of binding the cytosolic domain of Tom20 during preprotein translocation (Shiota, Mabuchi *et al.*, 2011).

Various electrophysiological investigations have been performed on fungal Tom40 isolated from mitochondria with the addition of multiple studies on recombinant Tom40 that was functionally renatured in detergent micelles or liposomes (Hill, Model *et al.*, 1998; Ahting, Thieffry *et al.*, 2001; Becker, Bannwarth *et al.*, 2005). Although the results seem to diverge to some extent, an overall Tom40 channel property can be assigned. At large, upon voltage excess

Tom40 switches between different subconductance states in terms of channel gating. Addition of preprotein peptides such as Su9-DHFR resulted in obstruction of the recorded ion current (Hill, Model *et al.*, 1998; Ahting, Thieffry *et al.*, 2001; Becker, Bannwarth *et al.*, 2005). This so called 'pore blockage' can either be interpreted as precursor insertion into the pore or binding to Tom40 in a way that the barrel opening is constricted by the preprotein.

Similar blockage experiments with TOM complex isolated from yeast mitochondria displayed 40-60 fold increased sensitivity to mitochondrial presequences in contrast to Tom40 alone. Therefore, a role of the non-channel Tom proteins in regulating the channel activity of Tom40 is assumed (Becker, Bannwarth *et al.*, 2005). This hypothesis was confirmed by Poynor and colleagues in 2008. They showed that Tom40 isolated from *Neuropora crassa* mitochondria is 'frozen' in five different conductance states. Equal examinations of the TOM core complex revealed flickering in between these states when a voltage of higher than ±60 mV was applied. For the lower conductance states inverted *S*-shaped nonlinear current-voltage curves were observed. These curves are reminiscent of much narrower protein pores where the permeating ions have to overcome an electrostatic energy barrier. It was further determined that Tom22, Tom7, Tom6 and Tom5 act as modulators of the Tom40 pore dynamics, explicitly the flickering between the five conductance states. These TOM subunits are able to reduce the energy barrier between different conformational states. When Tom40 alone was incubated at higher than ±100 mV, the pore started gating similar to the TOM core complex, which was reversible when the applied voltage was decreased to ±70 mV. Thus, Tom40 does not only form the preprotein conducting pore but also contains the 'gating machinery' of the TOM complex (Poynor, Eckert *et al.*, 2008). In conclusion, the (sub)conductance states of Tom40 are assigned as structural rearrangements in the β-barrel which are influenced by the non-pore forming Tom subunits.

Additionally, recent works provided first quantitative data regarding the kinetics of polypeptide interaction with the mitochondrial TOM machinery. It was found that structural differences between substrates were strongly reflected in characteristic blockage frequencies as well as the duration of blockage (Romero-Ruiz, Mahendran *et al.*, 2010). Further, the rates for association $k_{on}$ and dissociation $k_{off}$ strongly depended on the applied

transmembrane voltage (Mahendran, Romero-Ruiz *et al.*, 2012). Both kinetic constants increased with an increase in the applied voltage providing strong evidence for peptide translocation. Thus, it is now possible to distinguish between substrate blocking and permeation (Romero-Ruiz, Mahendran *et al.*, 2010; Mahendran, Romero-Ruiz *et al.*, 2012).

# Aims of this study

The main subunit of the translocase of the outer mitochondrial membrane, Tom40, forms the preprotein conducting channel in the outer mitochondrial membrane (Hill, Model *et al.*, 1998; Ahting, Thieffry *et al.*, 2001). The interior of the Tom40 pore offers an optimized environment for the transport of up to 1500 different preproteins in humans (Esaki, Kanamori *et al.*, 2003; Chacinska, Koehler *et al.*, 2009). Moreover, Tom40 contains the 'gating machinery' of the TOM complex, which can function in terms of an 'electrical switch' (Poynor, Eckert *et al.*, 2008). In summary, these properties make Tom40 an interesting target for nanopore based applications (Bayley and Cremer, 2001).

The aims of this study were the identification of the structural elements of Tom40 enabling its specific function in mitochondria, and further, the modification of Tom40 to retrieve a rationally designed nanopore. Since there exists a lack of high atomic resolution structural information of Tom40 from any organism, theoretical approaches were employed and developed as basis of this study. In doing so, I used the following approaches:

(I) A previously developed physical interaction model (Naveed, Jackups *et al.*, 2009) was utilized to identify weakly stable regions on the outside of the TM region of Tom40 and investigated through methods of molecular biology, biochemistry and biophysics. By this means, I set about to engineer Tom40 proteins with optimized structural properties.

(II) Bioinformatics and mass spectrometric methods were carried out to generate a homology model of *Nc*Tom40 enabling development of rational Tom40 mutants with altered inner-pore properties.

(III and IV) Recombinant Tom40 mutants of mammals and fungi were designed based on the developed homology model of *Nc*Tom40 to investigate the structural elements of the Tom40 pore. *In vivo* and *in vitro* methods were then applied to allocate the pore properties of Tom40 and its gating mechanism.

# Manuscript I

## Improving the resistance of a eukaryotic β-barrel protein to thermal and chemical perturbations

**Dennis Gessmann[1], Frauke Mager[1,2], Hammad Naveed[3], Thomas Arnold[2], Sara Weirich[1], Dirk Linke[2], Jie Liang[3] and Stephan Nussberger[1]**

[1]Biophysics Department, Institute of Biology, University of Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany

[2]Department of Protein Evolution, Max Planck Institute for Developmental Biology, Spemannstr. 35-39, 72076 Tübingen, Germany

[3]Department of Bioengineering, University of Illinois, 835 South Wolcott, Chicago, IL 60612, USA

*Running head:* Apparent stability of human Tom40A

*Keywords:* β-barrel; Tom40; apparent protein stability; protein engineering; TOM complex

*Abbreviations*: TM, transmembrane; hTom40A, human Tom40A; AfTom40, Tom40 from *Aspergillus fumigatus*; LDAO, lauryldimethylamine-oxide; GnHCl, guanidine hydrochloride; wt, wild type; IB, inclusion bodies

## Abstract

β-Barrel membrane proteins have regular structures with extensive hydrogen-bond networks between their transmembrane (TM) β-strands, which stabilize their protein fold. Nevertheless, weakly stable TM regions, which are important for the protein function and interaction with other proteins, exist. Here, we report on the apparent stability of human Tom40A, a member of the "mitochondrial porin family" and main constituent of the mitochondrial protein-conducting channel TOM (translocase of the outer membrane). Using a physical interaction model, TmSIP, for β-barrel membrane proteins, we have identified three unfavorable β-strands in the TM domain of the protein. Substitution of key residues inside these strands with hydrophobic amino acids results in a decreased sensitivity of the protein to chemical and/or thermal denaturation. The apparent melting temperature observed when denatured at a rate of 1 °C per minute is shifted from 73 to 84 °C. Moreover, the sensitivity of the protein to denaturant agents is significantly lowered. Further, we find a reduced tendency for the mutated protein to form dimers. We propose that the identified weakly stable β-strands 1, 2 and 9 of human Tom40A play an important role in quaternary protein–protein interactions within the mammalian TOM machinery. Our results show that the use of empirical energy functions to model the apparent stability of β-barrel membrane proteins may be a useful tool in the field of nanopore bioengineering.

## Introduction

The transmembrane (TM) domains of β-barrel membrane proteins have regular structures with an extensive hydrogen-bond network between the individual β-strands.[1,2] Despite this strong network, unfavorable or weakly stable regions exist in several β-barrel membrane proteins.[3,4] They are often important for their function, such as voltage sensing,[5] flux control of metabolites and ion sensing (see Refs. 6 and 7 for detailed reviews).

In general, β-barrel membrane proteins are stabilized through binding of α-helices to weakly stable regions inside or outside of the pore, so-called in-plugs or out-clamps, through formation of oligomers via protein–protein interfaces or interactions with lipids.[3,4,6,8]

In this work, we explore whether the prediction of weakly stable regions in bacterial β-barrels[3,4] can be adapted to human Tom40A (hTom40A) and reveal further insights on the structural organization of this protein.

In eukaryotes, Tom40 proteins represent an essential class of pore proteins that facilitate the translocation of unfolded proteins from the cytosol into mitochondria. They comprise the main subunit of the TOM (*t*ranslocase of the *o*uter *m*embrane) import machinery in mitochondrial outer membranes.[9-14]

It is generally predicted that all eukaryotic Tom40 proteins belong to the "mitochondrial porin" superfamily. Thus, hTom40A most likely constitutes a β-barrel architecture similar to that of VDAC-1 with 19 β-strands and a short α-helix located inside the pore.[15-17] Consistent with this model, circular dichroism (CD) and Fourier transform infrared secondary structure analyses of Tom40 and VDAC proteins from different organisms revealed a dominant β-sheet structure with a small α-helical part.[9,18,19]

In the present study, we calculated the energy of each amino acid in the predicted native conformation of hTom40A using a recently introduced empirical potential function that was developed based on extensive combinatorial analysis of known bacterial β-barrel membrane protein structures.[3] We have identified three β-strands (1, 2 and 9) in the TM domain of hTom40A that contribute to the overall sensitivity of the protein to denaturation. We show that mutagenesis of the predicted specific destabilizing amino acid residues within these strands leads to a higher resistance to thermal or chemical perturbation of the barrel. Similar results were obtained with Tom40 from *Aspergillus fumigatus* (AfTom40). We propose that the unstable β-strands 1, 2 and 9 of the TM domain of hTom40A interact with other Tom40 molecules or subunits of the TOM complex.

## Results

### Weakly stable regions in hTom40A and oligomerization index

The stability of β-barrel membrane proteins is determined by the balance between favorable hydrogen-bond networks, van der Waals and hydrophobic interactions and unfavorable conformational entropy. To identify weakly stable regions in the TM domain of wild-type (wt) hTom40A (Fig. 1a and b), we estimated the energetic contribution of all amino acid residues to the β-strand

stability of the protein by using a computational approach[3] that has recently been applied to model the conformational stability of 25 nonhomologous β-barrel membrane proteins of known structure. Briefly, we calculated the energetics of embedding specific residue types at different regions of the TM domain, the stabilizing effect due to interactions between residues on neighboring strands through strong H-bond, weak H-bond and side chain–side chain interactions using the updated TmSIP empirical energy parameters.[3,20]

We calculated the contribution of residues to the empirical energy for each β-strand (Fig. 1c). Strands 1, 2 and 9 have significantly higher empirical energies and are thus less favorable than the rest of the protein. Then, the oligomerization index $\varrho_{wt}$, which summarizes the energy deviation of unstable strands from the expected energy value for all the strands in the protein, was calculated to be $\varrho_{wt} = 2.48$. This is in good agreement with the observation that highly unfavorable strands are often associated with protein–protein interfaces in the TM region.

In data sets using sequence information of bacterial β-barrel membrane proteins, a protein can be predicted to be monomeric if the oligomerization index is < 2.25 and oligomeric if it is > 2.75.[3] The oligomerization index of hTom40A is between these prediction thresholds. Theoretically, wt hTom40A may thus exist as stable monomers but may also form higher-order complexes through distinct protein–protein interaction interfaces.

We further examined the contribution of all amino acids facing the membrane lipids within the predicted three unfavorable strands 1, 2 and 9 (Fig. 1d). In this analysis, residues K107 in strand 1, H117 in strand 2 and H220 in strand 9 were found to contribute the most to the instability of these strands. Based on our calculations, we predict strands 1, 2 and 9 of hTom40A to be unfavorable/weakly stable. Each strand can have two orientations: the side chain of the first residue can face either the lipid environment or the internal of the barrel. The orientations of strands are predicted using the energy scale in Ref. 3 such that the number of costly burial of ionizable/polar residues facing the lipid environment is overall minimized.

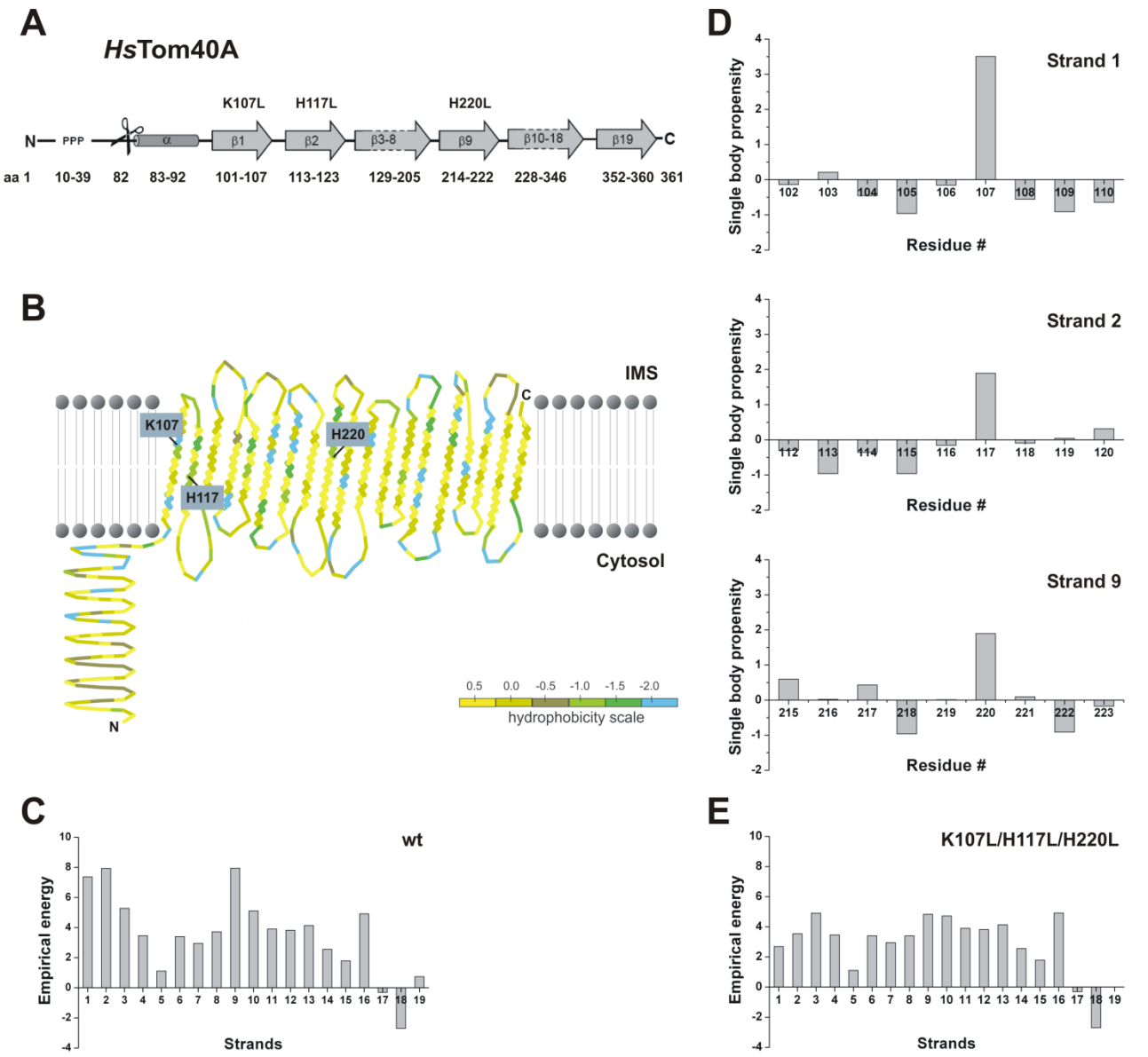



**Fig. 1.** Expected energy levels of strands and residues in the TM domain of hTom40A. (a) The secondary structure prediction of hTom40A is based on PRED-TMBB[30] and TMBETAPRED-RBF.[31] (b) The protein topology of hTom40A was generated using TMRPres2D.[48] (c) Empirical energy of β-strands 1–19 of wt protein. β-Strands 1, 2 and 9 are predicted as the weakly stable strands. (d) Single-body propensities of amino acid residues of the TM domains of β-strands 1, 2 and 9. Amino acids K107, H117 and H220 show the highest values, indicating a destabilizing effect on the regarding strand. (e) Empirical energy of β-strands 1–19 of mutant (K107L, H117L and H220L) hTom40A.

**Secondary and tertiary structures of hTom40A**

In order to account for the differences between bacterial and mitochondrial β-barrel membrane proteins and keeping in mind that hTom40A forms complexes also with other components of the TOM machinery, we wanted to construct a mutant protein to have an oligomerization index ϱ below 1.5 so that the resulting mutant hTom40A would form more stable monomers, possibly without protein–protein interaction interfaces.

To test to what extent do β-strands 1, 2 and 9 determine the overall resistance to denaturation and oligomerization state of hTom40A, we designed five mutants, termed K107L, H117L, H220L, K107L/H117L and K107L/H117L/H220L, where residues K107, H117 and H220 of hTom40A are replaced by leucines. Leucine is predicted to be the most stabilizing amino acid when facing the lipid within the core region of a TM β-strand.[3]

The empirical energy profile of the triple mutant hTom40A$^{mut}$ (K107L/H117L/H220L) is shown in Fig. 1e. The oligomerization index of this mutant ($\varrho_{K107L/H117L/H220L}$ = 1.37) was significantly lower than that of the wt protein ($\varrho_{wt}$ = 2.48), predicting a very robust monomeric β-barrel membrane protein. Energy calculations of single and double mutants revealed higher indices ($\varrho_{K107L}$ = 2.08, $\varrho_{H117L}$ = 2.01, $\varrho_{H220L}$ = 2.03, $\varrho_{K107L/H117L}$ = 1.91, $\varrho_{K107L/H220L}$ = 1.71 and $\varrho_{H117L/H220L}$ = 1.68), predicting less stable proteins.

Mammalian Tom40A isoforms show a very high sequence identity among each other (> 91%, data not shown). All include a remarkable N-terminal poly-proline region, which is only present in mammalian Tom40A. So far, thermal stability analyses of the TM β-barrels of mammalian full-length Tom40A proteins using UV CD spectroscopy have not been successful in our hands (data not shown). The β-barrel UV CD signals appeared to be superimposed by strong CD signals caused by the poly-proline-rich region of the protein. As a consequence of this interference, temperature-induced transitions of the β-barrel itself could not be monitored accurately. Since the N-terminal poly-proline-rich domain in mammalian Tom40A proteins has no effect on the channel formation,[21] we deleted the poly-proline region of all hTom40A proteins to improve the CD signal in heat-induced unfolding measurements. wt and mutant Tom40A proteins were expressed in *Escherichia coli* and nickel affinity

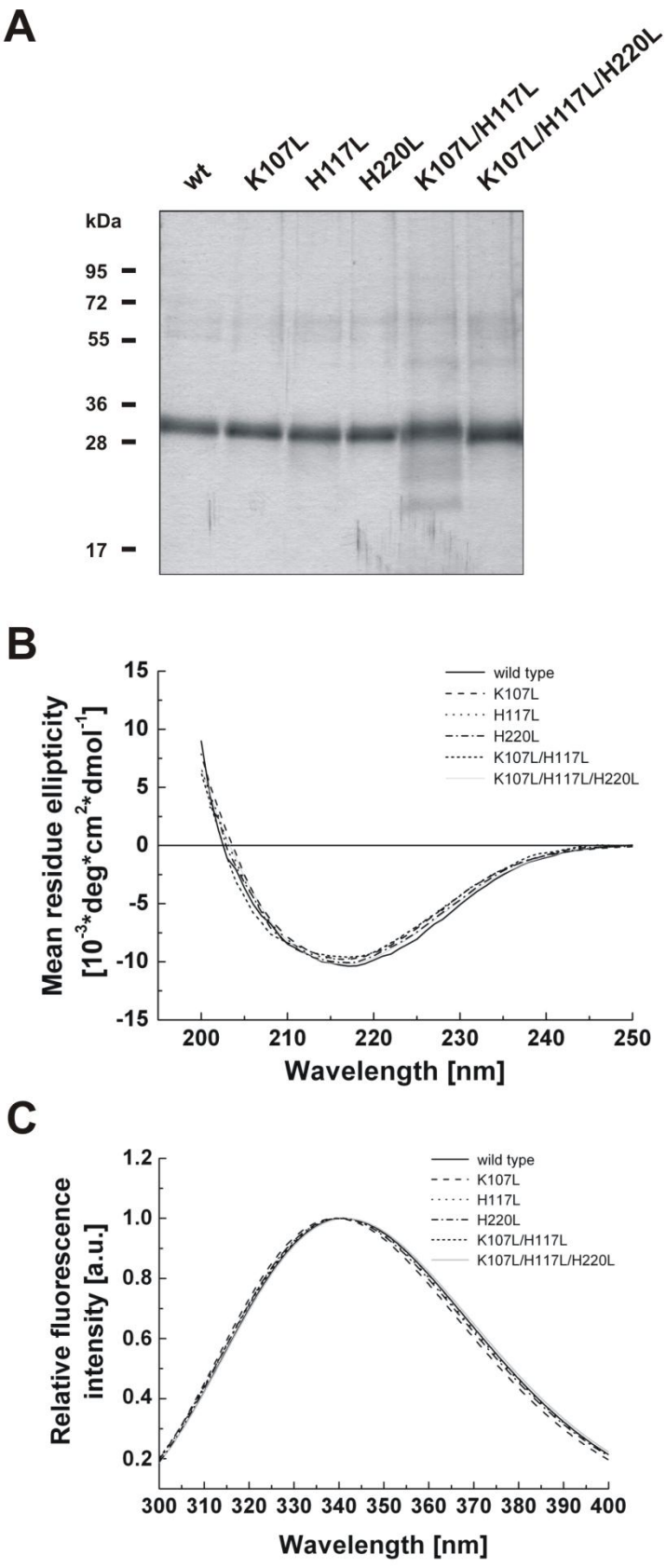



**Fig. 2.** Far-UV CD spectra and tryptophan fluorescence emission spectra of wt and mutant hTom40A. (a) SDS-PAGE showing purified wt and mutant (K107L, H117L, H220L, K107L/H117L and K107L/H117L/H220L) hTom40A. Proteins were visualized by silver staining. (b) Comparison of far-UV CD spectra between wt and mutant

purified from inclusion bodies (IB) under denaturing conditions. The proteins were refolded by rapid dilution of denaturant into lauryldimethylamine-oxide (LDAO)-containing detergent buffer and further purified via size-exclusion chromatography. Analysis of refolded proteins by SDS-PAGE and Coomassie and silver staining (Fig. 2a) indicated that all isoforms were virtually pure. Far-UV CD spectroscopy showed typical β-barrel spectra for wt and mutant proteins with similar curves for all hTom40A isoforms[9,10,18,21] (Fig. 2b). The spectral characteristics of the wt hTom40A and the triple mutant are summarized in Supplementary Table 2. At wavelengths > 250 nm, the CD spectra approached ellipticity values close to 0, indicating that the protein preparations were virtually free of higher-order aggregates, which would cause light-scattering effects and interfere with the interpretation of the data.

Further, tertiary structure was analyzed via tryptophan fluorescence spectroscopy. The emission spectra of wt and all mutant Tom40 proteins (Fig. 2c) were exactly the same, with intensity maxima at approximately 340 nm and an unchanged width of the emission spectra. This can be interpreted as an unaltered environment of the widely spread tryptophan residues, W188, W259 and W322, which are conserved in the amino acid sequence of all isoforms and remain unchanged in all our designed mutant proteins. In summary, we suggest that secondary and tertiary structures match in all hTom40A isoforms and, consequently, differences in resistance to chemical and thermal perturbations and oligomerization state are not due to an altered protein structure.

---

**(*Fig. 2. continued*)** hTom40A. Measurements were carried out at a protein concentration of ~ 0.2 mg/ml in 1% (w/v) LDAO at 25 °C for wt and mutant proteins, respectively. For each experiment, five scans were accumulated at the indicated temperatures. Noisy data below 200 nm due to optical density have been removed. (c) Comparison of tryptophan fluorescence emission spectra of wt and mutant hTom40A. Emission spectra were conducted in 20 mM Tris–HCl (pH 8), 1% (w/v) LDAO, 1 mM β-mercaptoethanol and 350 mM GnHCl at a protein concentration of ~ 0.15 mg/ml for all isoforms. Data points were fitted to the log-normal distribution as described in Materials and Methods. Relative fluorescence intensity is in arbitrary units (a.u.). All isoforms exhibit the exact same far-UV CD and tryptophan fluorescence emission spectra characteristics. We suggest analogue β-barrel formation for wt and mutant hTom40A.

## Thermal and solvent stability of hTom40A

To compare the thermal stability of wt and mutant hTom40A, we measured CD signals at different temperatures at constant wavelength. wt hTom40A unfolded at an apparent melting temperature of about 73 °C when denatured at a rate of 1 °C per minute (Fig. 3). In line with our energy calculations described above, hTom40A with substitutions at positions K107, H117 and H220 revealed an apparent midpoint of resistance to thermal denaturation of approximately 84 °C (Fig. 3). The single and double mutants appeared to be slighty more stable than wt protein; however, differences proved not to be significant (data not shown).

To provide further evidence that substitution of unfavorable amino acids in the TM domain of hTom40A results in conformational stabilization of the protein, we compared tryptophan fluorescence spectra of wt and mutated

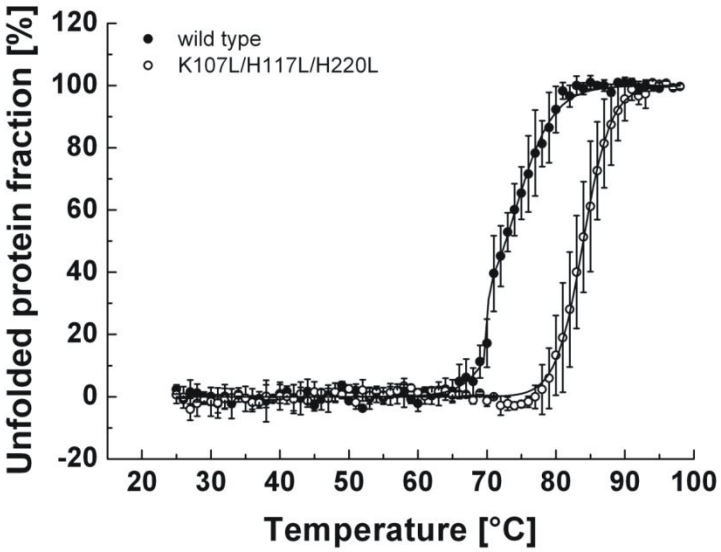



**Fig. 3.** Sensitivity of wt and mutant hTom40A to thermal denaturation. Thermal denaturation of both hTom40A isoforms was monitored under the same conditions as described in Fig. 2b by change in ellipticity at 216 nm. wt and mutant proteins were subjected to temperature increases of 1 °C/min from 25 to 98 °C, respectively. Data points of melting curves were normalized to the minimum and maximum percentages of unfolded protein fraction ($n$=3, error bars=SD)[49] and then fitted to sigmoid functions (black line). Apparent melting temperatures were retrieved at the midpoint of the transition curves. wt Tom40A indicated a three-state unfolding mechanism, whereas the triple mutant hTom40A showed a two-state unfolding mechanism with an approximately 11 °C higher apparent melting temperature.

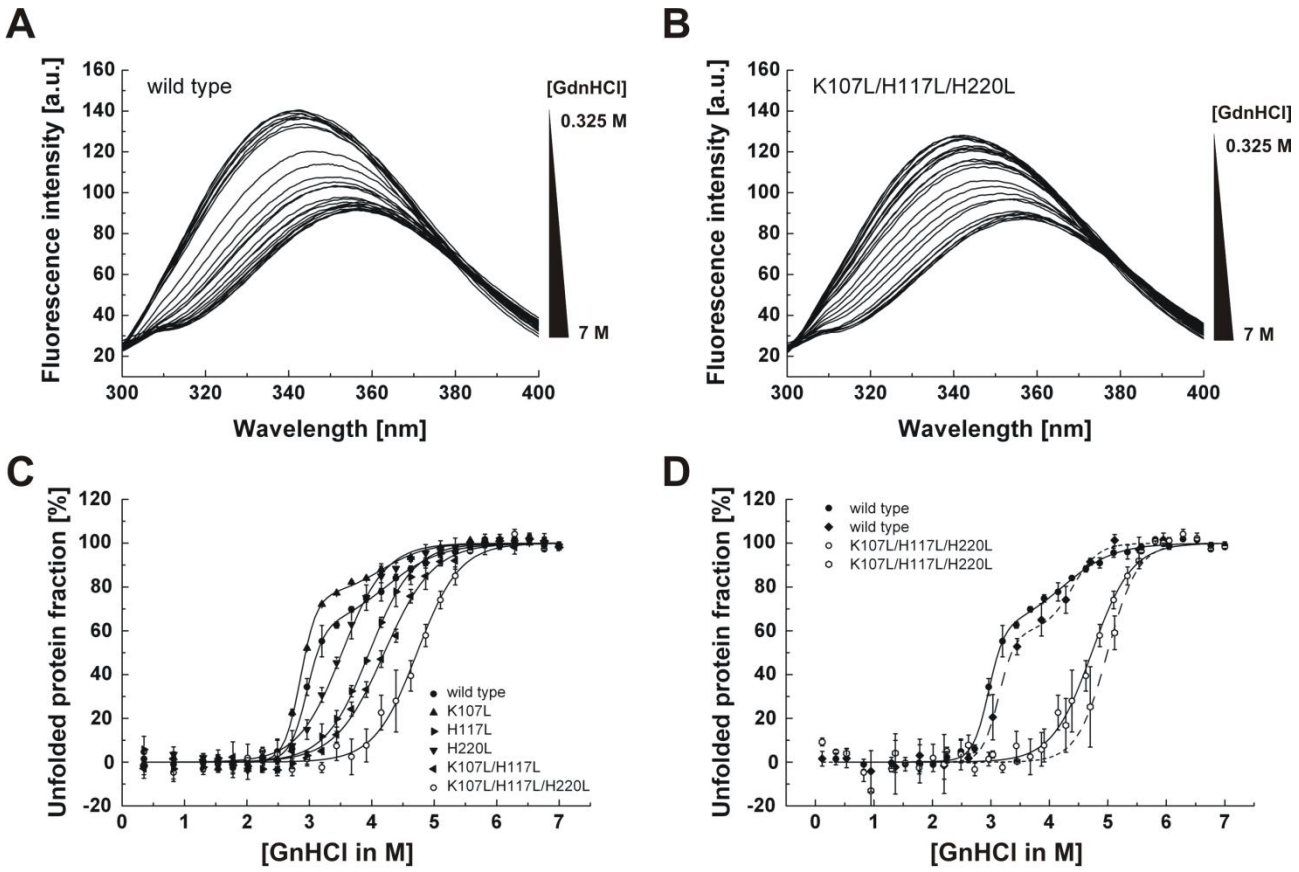



**Fig. 4.** Sensitivity of wt and mutant hTom40A to chemical denaturation. (a and b) Tryptophan fluorescence emission spectra of wt and mutant hTom40A (~ 0.15 mg/ml) were recorded in Tris–HCl buffer containing 1% LDAO after 24 h of incubation with GnHCl at 25 °C. GnHCl concentrations varied from 0.35 to 7 M, respectively. (c) The fractions of unfolded wt and mutant Tom40 were determined at different GnHCl concentrations from fluorescence intensities recorded at 330 nm relative to a denatured protein solution. (d) To show the reversibility of the chemical unfolding, we repeated the experiment by subjecting folded wt and mutant K107L/H117L/H220L hTom40A to increasing GnHCl concentrations (dashed lines). For comparison, we included the data from the refolding reaction in panel (c) (solid lines). The figures in (c) and (d) show an average of three independent experiments (error bars=SD).

hTom40A in the presence of chemical denaturants. The change in tryptophan fluorescence of wt and mutated hTom40A, respectively, was monitored at different guanidine hydrochloride (GnHCl) concentrations (Fig. 4a and b). wt Tom40 was completely denatured in ~ 5 M GnHCl. The apparent midpoint of unfolding occurred at a concentration of ~ 3.1 M GnHCl. On the other hand, the resistance of mutant Tom40 K107L/H117L/H220L to chemical denaturation was greatly enhanced (Fig. 4c). The apparent midpoint of unfolding of hTom40A

was shifted from 3.1 to approximately 4.8 M GnHCl. Mutant hTom40A K107L/H117L/H220L completely unfolded at ~ 6.3 M GnHCl. These results were obtained whether unfolded or refolded Tom40 was incubated with GnHCl (Fig. 4c and d), showing the reversibility of the chemical denaturation.

Chemical unfolding of hTom40A K107L, H220L, H117L and K107L/H117L occurred at around 2.9, 3.5, 3.9 and 4.2 M GnHCl (Fig. 4c). These mutants were thus more resistant to chemical denaturation than wt hTom40A but less resistant than the triple mutant K107L/H117L/H220L. Our data indicate that the resistance of the mutants to chemical denaturation increased with decreasing ϱ-values. They further suggest that wt hTom40A unfolds via a multistate mechanism, whereas unfolding of "stabilized" protein follows a two-state conformational transition.

**Thermal and solvent stability of AfTom40**

To further test the prediction of weakly stable β-strands in eukaryotic porins, we applied stability calculations to full-length AfTom40. Comparison of full-length AfTom40 and wt hTom40AΔ1–82 revealed a moderate identity at the amino acid level (~ 25% identical residues in the conserved TM part, including the N-terminal helix) but a similar β-barrel secondary structure (Supplementary Fig. 1). We identified five unstable regions in AfTom40 to compromise the stability of β-strands 1, 5–7, 9, 13 and 17 (Fig. 5a). The ϱ-value of the wt protein was calculated to be 2.81. To see to what extent do these β-strands determine the stability of the protein, we made an AfTom40 variant with mutations K69H, N150H, S180L, S236A and K302H. These mutated residues were predicted to be the most favorable residues in the respective positions of the β-strand by the TmSIP potential function. The ϱ-index of this mutant was calculated as 1.66 (Fig. 5a).

wt and mutant AfTom40 were over-expressed in *E. coli* cells, purified under denaturing conditions and refolded into polyoxyethylene monolauryl ether (Brij35)-detergent containing buffer in a similar way as hTom40A (Fig. 5b). Far-UV CD spectra of both proteins (Fig. 5c) showed a β-barrel fold as for hTom40A. Tryptophan fluorescence emission spectra were also almost identical, supporting a similar structure for both AfTom40 proteins (Fig. 5d). Thermal

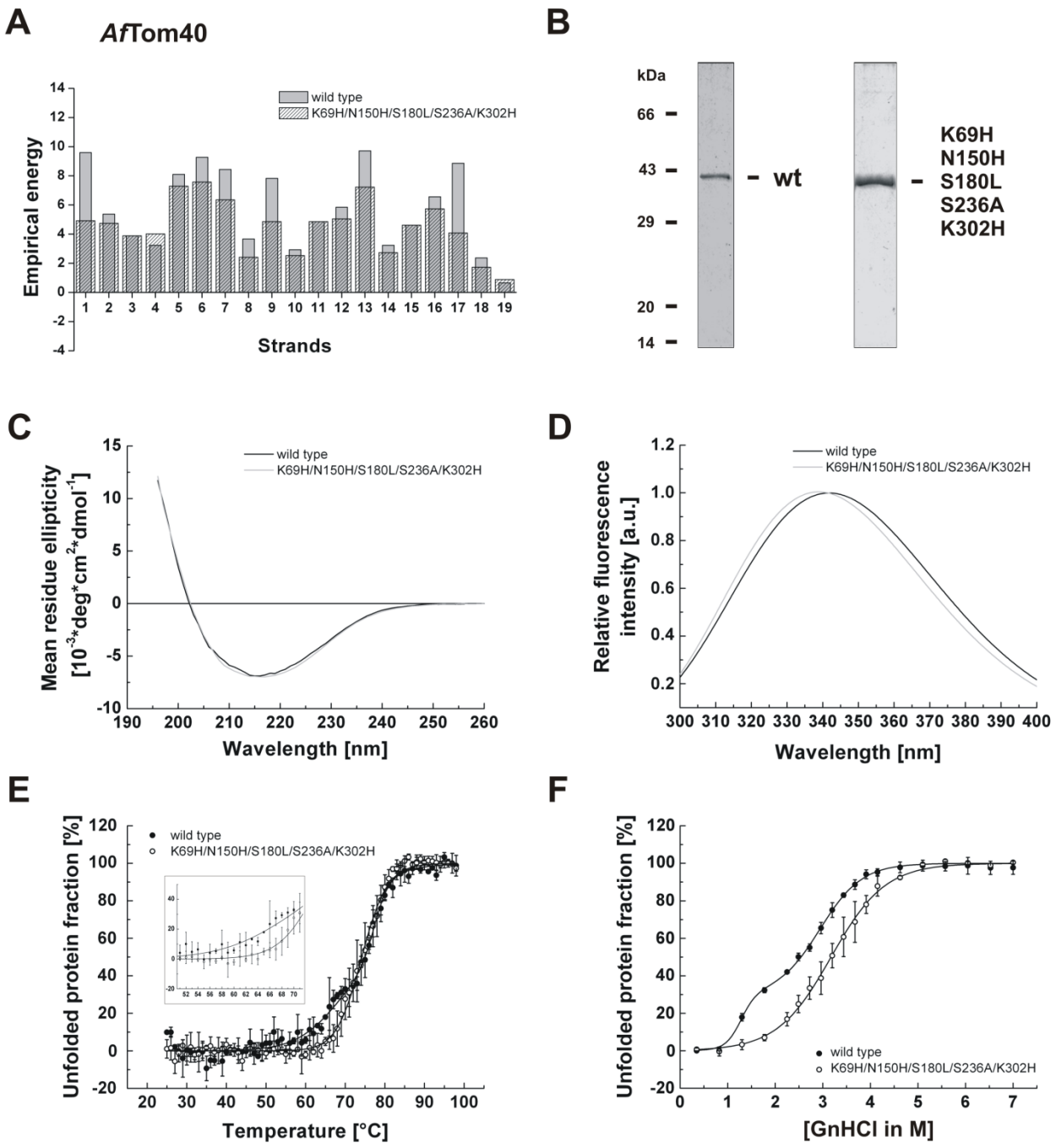



**Fig. 5.** Stability analysis of AfTom40. (a) Empirical energy of β-strands 1–19 of wt and mutant proteins. (b) SDS-PAGE of purified and refolded wt and mutant (K69H, N150H, S180L, S236A and K302H) AfTom40 throughout Coomassie Blue staining. Far-UV CD spectra (c) and tryptophan fluorescence emission spectra (d) of AfTom40 wt/mutant were recorded and deconvoluted as described in Fig. 2, respectively. Both AfTom40 isoforms adapt the same secondary structure (c), but mutant AfTom40 showed an approximately 1.5-nm left-shifted tryptophan fluorescence emission spectrum. Thermal (e) and chemical (f) sensitivities of AfTom40 isoforms were investigated according to Figs. 3 and 4c. Thermal and GnHCl-induced unfolding of proteins revealed a three-state unfolding mechanism for the wt and a two-state unfolding

denaturation of both AfTom40 proteins followed by CD spectroscopy revealed no difference in the apparent melting temperature but a clear change in the cooperativity in unfolding (Fig. 5e). In contrast to thermally induced unfolding, chemically induced unfolding of wt and mutant AfTom40 proteins revealed an apparent stabilization of the mutant protein (Fig. 5f). The apparent midpoint of unfolding of wt protein was observed around 2.5 M and was shifted to about 3.2 M GnHCl in the mutant protein. Both patterns of unfolding were in line with those observed for thermal unfolding (Fig. 5e) and those observed for hTom40A (Fig. 4c and d).

**Oligomerization states of hTom40A**

Previous studies revealed that weakly unfavorable regions in bacterial β-barrel proteins can be stabilized through oligomerization of monomers.[3] Our energy calculations described above predicted a very stable monomeric form for the mutant hTom40A K107L/H117L/H220L. To verify this hypothesis, we conducted chemical cross-linking experiments on both wt and triple mutant Tom40 solubilized in LDAO. Both protein samples were incubated with glutaraldehyde, which forms covalent bonds between primary amine groups. The degree of intermolecular cross-link formation was evaluated by SDS-PAGE and Western blotting (Fig. 6). wt hTom40A was present mostly in its monomeric form but also formed dimers. In contrast, mutant protein was predominantly monomeric and revealed virtually no dimers. Similar results were obtained with proteins in β-dodecylmaltoside-containing buffer (data not shown). It should be noted, however, that long exposures of both proteins to glutaraldehyde resulted in protein aggregates that did not enter the SDS gel. In summary, we conclude successful stabilization of the hTom40A β-barrel. Hence, for mutant hTom40A K107L/H117L/H220L, oligomerization is not required to compensate for unfavorable regions.

**(*Fig. 5. continued*)** mechanism for the mutant protein with a higher stability for the mutated form of AfTom40.

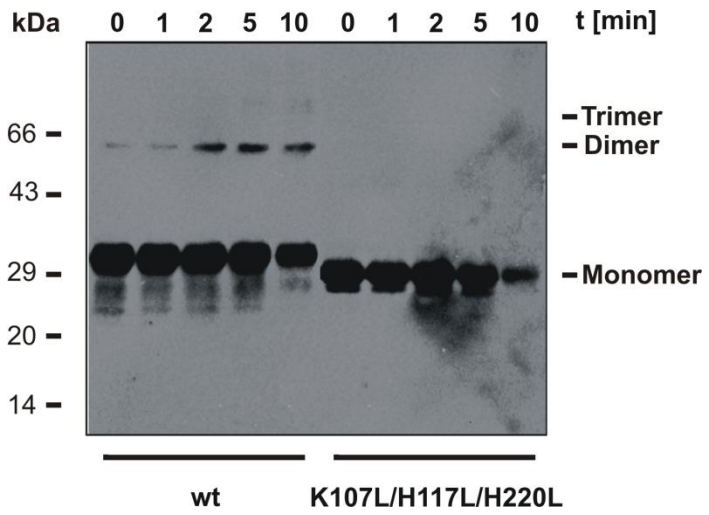



**Fig. 6.** Chemical cross-linking of refolded wt and mutant (K107L, H117L and H220L) hTom40A. wt and mutant proteins (40 μM) were incubated with glutaraldehyde (125 μM) at 37 °C in 20 mM $NaH_2PO_4$, pH 8. Aliquots were taken after the addition of the cross-linking reagent at the time indicated. In these, cross-linking reactions were stopped by addition of 50 mM Tris, pH 8. Proteins and cross-linking products were visualized by SDS-PAGE and Western blotting.

## Discussion

In this study, we applied a recently developed method for the prediction of unfavorable regions in bacterial β-barrel membrane proteins[3] to eukaryotic hTom40A. We identified and experimentally confirmed two regions in the TM domain of the protein (β-strands 1 and 2 and β-strand 9) that have strong impact on its resistance to thermal and chemical perturbations. Oligomerization of hTom40A monomers might provide the mechanism to counterbalance the overall β-barrel instability, indicating possible protein–protein interfaces in mammalian Tom40 proteins. Indeed, the existence of dimeric and trimeric Tom40 forms was shown for fungal Tom40.[22,23] However, for mammalian Tom40 proteins, experimental evidence has yet to be shown. Nevertheless, our results suggest that β-strands 1, 2 and 9 in the TM domain of mammalian Tom40 might be important for the association of single Tom40 molecules. In addition, other subunits of the mitochondrial protein import machinery TOM, for example, the small Tom proteins, might also interact with β-strands 1, 2 and 9 in the TOM complex from mammals.[24]

Interestingly, the key amino acids K107, H117 and H220 in the two unfavorable regions of hTom40A are not strongly conserved between different

species (e.g., fungi and plants). This and the results of our stability calculations for AfTom40 indicate that weakly stable regions in mammalian Tom40 proteins differ from evolutionarily more distant Tom40 β-barrels.

Using the described method above for the prediction of unfavorable regions on fungal AfTom40, we identified five unstable domains comprising strands 1, 5–7, 9, 13 and 17. As in the case of mammalian TOM complex, these five unstable domains in fungal Tom40 proteins might play an important role for the interaction with other subunits of the TOM complex in fungi.[25] It is tempting to suggest that the assembly and dissociation of the TOM complex are caused by the compensation of weakly stable regions in the TM domain of Tom40 by the association of other TOM subunits and Tom40 itself, not only in detergent buffer but also in the *in vivo* membrane environment of mitochondria.

Empirical energy calculations on AfTom40 predicted a significant stabilization of the pore protein upon mutations in β-strands 1, 7, 9, 13 and 17. However, experimental analysis revealed only a moderate resistance of the protein to chemical and thermal denaturations. A possible explanation for this is that the full-length AfTom40 construct used in this study not only comprises a β-barrel domain but also has additional large N- and C-terminal random-coil and α-helical tails. They unfold at low temperatures and low GnHCl concentrations compromising the CD and fluorescence signals of β-barrel unfolding and complicate the empirical energy calculations of the β-barrel. Another reason could be the imprecise nature of the identified locations of the 19 β-strands of AfTom40, which is used for the empirical energy calculations.

We conclude that computational prediction of unfavorable regions in β-barrel membrane proteins displays a useful tool to improve the resistance of eukaryotic β-barrel proteins to thermal and chemical denaturations. With bioinformatics-derived empirical energy parameters and calculation of thermodynamic properties based on a firm statistical mechanical model, we can now begin to combine both experimental and computational approaches to engineer β-barrel membrane proteins with rationally designed biophysical properties. We propose that our method can be easily adapted to the stabilization of other β-barrel membrane proteins, which is of high interest in, for example, nanotechnology[26] and structural biology.[27-29]

## Materials and Methods

### Secondary structure and weakly stable amino acid residue prediction

The software packages PRED-TMBB[30] and TMBETAPRED-RBF[31] were used for the prediction of the secondary structure of hTom40A and AfTom40. Sequence conservation and far-UV CD spectra data were incorporated in the secondary structure prediction. The oligomerization index $\varrho^3$ for the presumably 19-stranded β-barrel indicated the presence of higher-order protein complexes. We examined the predicted lipid-facing residues, and those with relatively higher energy values are predicted to be weakly stable. The calculation of energy values of individual β-strand residues was performed as described in Ref. 3. Briefly, we calculated the energy of each residue in the predicted native conformation using an empirical potential function, TmSIP, derived from bioinformatics analysis of β-barrel membrane proteins,[32] first developed in Refs. 20, 33 and 34 and further refined with additional structural data in Refs. 3 and 4. The energy for each residue consists of two terms. First, each residue is assigned an energy value of burying this residue type at a particular depth in the lipid bilayer and with the orientation of its side chain. There are two possible orientations, namely, side chains facing the lipid environment or facing inside the barrel. This is termed the "single-body term." Second, each residue interacts with two residues on separate neighboring strands through strong backbone H-bond interaction, side-chain interactions and weak H-bond interactions, which collectively make up the two-body energy term. Strand energy is the summation of both single-body and two-body energy terms over all residues in the strand.

### Cloning and strains

For recombinant expression of wt hTom40A in *E. coli*, a truncated version of the protein lacking amino acid residues 1–82 was used (Fig. 1a). The corresponding gene (*hTom40AΔ1–82*) was PCR amplified and cloned into a pET24d vector (Novagen) introducing a C-terminal His6-tag into the protein for purification.[35] The triple mutant *hTom40AΔ1–82$^{mut}$* (K107L, H117L and H220L) was prepared by site-directed mutagenesis using *hTom40AΔ1–82* as a template and Pfu Ultra II DNA Polymerase (Agilent Technologies Inc.). First, replacement of amino acids K107L and H117L was accomplished by using the primers for-K107L-H117L and rev-K107L-H117L (Supplementary Table 1). Second, to

generate the mutation H220L, we further mutagenized the plasmid containing mutations K107L and H117L with primers for-H220L and rev-H220L (Supplementary Table 1). Single hTom40A mutants K107L and H117L were made by site-directed mutagenesis of *hTom40AΔ1–82* by using the primers for-H117L/rev-H117L and for-K107L/rev-K107L, respectively (Supplementary Table 1). In all cases, methylated parental plasmids were specifically digested with DpnI (New England Biolabs). Final plasmids were verified by DNA sequencing. Eventually, all vectors were transformed into *E. coli* BL21-CodonPlus (DE3) (Stratagene, Agilent Technologies) cells for protein expression.

The protein sequence of AfTom40 was retrieved from the *Aspergillus* Genome Database. The corresponding gene and an *AfTom40* gene coding for a protein with mutations K69H, N150H, S180L, S236A and K302H were synthesized and cloned into a pET24d vector (GeneArt/Invitrogen). Both genes encoded for proteins with a carboxy-terminal hexahistidine tag. To produce these AfTom40 proteins, we transformed vectors containing wt and mutant *AfTom40* into *E. coli* C41(DE3) (Lucigen) and BL21(DE3) (Stratagene, Agilent Technolgies) cells, respectively.

**Protein expression and isolation of IB**

wt hTom40A (hTom40AΔ1–82) and mutated hTom40A (hTom40AΔ1–82$^{mut}$) were expressed forming IB in *E. coli* BL21-CodonPlus (DE3) under tight control of expression from the T7 promoter. To obtain wt protein, we transformed BL21-CodonPlus (DE3) cells with pET24d-*hTom40AΔ1–82* and grew them in LB medium (20 ml) containing kanamycin (50 μg/ml) at 37 °C in shaking flasks. This culture was used to inoculate LB medium (2 l) containing kanamycin (50 μg/ml) for high-level protein expression. Protein expression was induced with 1 mM IPTG at a cell density corresponding to an $OD_{600}$ of 0.6. After 6 h of growth, cells were harvested by centrifugation, washed with phosphate-buffered saline and stored at − 20 °C until further use. Mutant hToma40A protein (hTom40AΔ1–82$^{mut}$) was expressed by the same procedure. For purification of wt and mutant Tom40, cells were thawed, and 3 ml lysis buffer [50 mM Tris–HCl (pH 8), 1 mM ethylenediaminetetraacetic acid, 100 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.26 mg/ml lysozyme] was added per gram of cells. After lysis of membranes with 4 mg deoxycholate per

gram of cells, DNaseI (12.5 units/g cells; Sigma-Aldrich) from bovine pancreas was added.[36,37] After incubation on ice for 20 min, IB were separated from cell debris by centrifugation at 20,000*g* for 30 min at 4 °C. IB pellets were washed with buffer containing 50 mM Tris–HCl and 100 mM NaCl, pH 8, and subsequently solubilized in 6 M GnHCl, 20 mM Tris–HCl, 150 mM NaCl and 1 mM β-mercaptoethanol, pH 8, using a glass–glass homogenizer. For removal of insoluble material, the homogenate was centrifuged at 30,000*g* for 30 min at 4 °C, and supernatants were stored at 4 °C. One liter of cell culture yielded 0.5–1 g of unprocessed protein. Protein concentrations were determined by UV absorbance using an extinction coefficient $\varepsilon_{280}$ of 29,900 $M^{-1}$ $cm^{-1}$ for hTom40AΔ1–82 and hTom40AΔ1–82mut.[38]

For the over-expression of wt and mutant AfTom40, protein production was performed as for hTom40A with the difference that, after induction, the cells were further grown at 37 °C for 19 h and cells were harvested by centrifugation. Cell pellet was stored at − 20 °C until further use (~ 4 g/l of cell culture). The isolation of IB was conducted under the same conditions for both AfTom40 wt and mutant and was based on the deoxycholic acid method according to Refs. 36 and 37. Cells were thawed on ice and resuspended in 30 ml of lysis buffer [50 mM Tris–HCl (pH 8.5), 1 mM ethylenediaminetetraacetic acid and 100 mM NaCl] for 10 g of cells. After the addition of PMSF and lysozyme to final concentrations of 0.13 mM and 0.26 mg/ml, respectively, solution was incubated on ice for 20 min and stirred occasionally. Then, solution was transferred into a water bath at a temperature of 37 °C. In the following, 40 mg of deoxycholic acid was added, and solution was stirred until it became viscous. Two hundred fifty units of Benzonase® (Novagen) was added at room temperature, and suspension was stirred until it was no longer highly viscous. Solution was further spun down at 19,600*g* for 30 min at 4 °C. Supernatant was discarded, and pellet was resuspended in 100 ml of TNBP buffer [50 mM Tris–HCl (pH 8.5), 100 mM NaCl, 1 mM β-mercaptoethanol and 1 mM PMSF]. After homogenization using a glass homogenizer, a clarifying spin was applied at 20,000*g* for 30 min at 4 °C. Retrieved IB pellet was resuspended in 95 ml of TNTBP buffer [TNBP buffer containing 0.1% (v/v) Triton X-100] and homogenized on ice. A second clarifying spin was applied under the same conditions. Eventually, the pellet was washed with 100 ml of TNBP buffer, and

IB pellet was retrieved after centrifugation under the named conditions above. Washed IB pellet was then stored at − 20 °C until further use. Determination of protein concentration was conducted as for human Tom40 by using extinction coefficients[38] $\varepsilon_{280}$ of 37,025 $M^{-1}$ $cm^{-1}$ for wt and mutant AfTom40, respectively.

**Protein purification and folding**

IB containing wt and mutated hTom40A (hTom40AΔ1–82 and hTom40AΔ1–82$^{mut}$) were loaded onto a Ni-Sepharose HiTrap column (1–20 ml; GE Healthcare) equilibrated with 6 M GnHCl, 20 mM Tris–HCl, 150 mM NaCl and 1 mM β-mercaptoethanol, pH 8. After washing the column with 2 column volumes of equilibration buffer, we removed unspecifically bound proteins with 20 mM imidazole. Then, hTom40A proteins were eluted with 300 mM imidazole, and fractions containing Tom40 were merged. Protein concentrations were adjusted to 5 mg/ml, and samples were stored at 4 °C.

For refolding of wt and mutant hTom40A, purified protein in GnHCl was diluted 10-fold into 20 mM Tris–HCl, pH 8, containing 0.5% LDAO and 1 mM β-mercaptoethanol. After removal of aggregates by centrifugation at 100,000*g*, samples containing refolded protein were concentrated to ~ 5 mg/ml by Ni-NTA affinity chromatography. Samples were loaded onto Ni-Sepharose HiTrap columns (1–20 ml) previously equilibrated with 20 mM Tris–HCl (pH 8), 0.1% LDAO and 1 mM β-mercaptoethanol. Bound protein was eluted in the same buffer containing 300 mM imidazole. Final purification of protein was achieved by size-exclusion chromatography using a Superose 12 column (GE Healthcare) that had been pre-equilibrated with 20 mM Tris–HCl (pH 8), 0.1% LDAO and 1 mM β-mercaptoethanol. The purity of isolated protein was assessed by SDS-PAGE followed by Coomassie Brilliant Blue staining, silver staining or Western blotting using antibodies against human Tom40 (Santa Cruz Biotechnology, Heidelberg).

For refolding of wt and mutant AfTom40A, the IB pellet of the respective protein was resuspended and homogenized in 6 ml of binding buffer [50 mM Tris–HCl (pH 7.5), 8 M urea and 1 mM β-mercaptoethanol] per gram of IB pellet. Solution was centrifuged at 19,600*g*, and supernatant was loaded onto a Ni-Sepharose HiTrap column (1–5 ml; GE Healthcare) at room temperature. After column was washed with 10–15 column volumes of binding buffer, a gradient of

0–1 M imidazole was applied. Fractions containing AfTom40 were adjusted to a protein concentration of 6 mg/ml, and 6 mg of protein was refolded in 10 ml of refolding solution [20 mM $NaH_2PO_4$ (pH 7), 1 mM β-mercaptoethanol and 0.8% (w/v) polyoxyethylene monolauryl ether (Brij35)]. The refolded protein solution was dialyzed at 4 °C against 50- to 100-fold volumetric excess of refolding buffer using a filter with a 6- to 8-kDa molecular mass cutoff (Spectra/Por®). Purity was then determined throughout SDS-PAGE and Coomassie Brilliant Blue. If necessary, a second purification step was conducted using the same conditions with the variation of exchanging 8 M urea against 0.8% (w/v) Brij35 in the binding and elution buffers. Purified protein was then further dialyzed against 50- to 100-fold volumetric excess of refolding buffer, and protein concentration was then adjusted to 0.2–0.4 mg/ml as needed.

**Chemical cross-linking**

For cross-linking experiments, 40 μg of refolded and purified hTom40AΔ1–82 and hTom40AΔ1–82$^{mut}$ was suspended in 100 μl 20 mM $NaH_2PO_4$, pH 8, and incubated with 125 μM freshly prepared glutaraldehyde at 37 °C for 0–45 min. Aliquots were removed, and cross-linking reactions were stopped by adding Tris, pH 8, to a final concentration of 50 mM. Finally, cross-linking products were analyzed by SDS-PAGE and Western blotting. Polyclonal antibodies against hTom40A were obtained from Santa Cruz Biotechnology.

**Circular dichroism**

CD (far-UV CD) spectroscopy measurements of refolded and purified wt and mutant hTom40A (~ 0.2 mg/ml) in 20 mM Tris–HCl (pH 8), 1% (w/v) LDAO and 1 mM β-mercaptoethanol were performed using a Jasco J-715/815 spectrometer (Tokyo, Japan) in quartz cuvettes of 0.1-cm path length out of three independent protein preparations. Spectra were recorded at 25 °C from 185 to 260 nm with a resolution of 1.0 nm and an acquisition time of 20–100 nm/min. Final CD spectrum was obtained by averaging five consecutive scans and corrected for background by subtraction of spectrum of protein-free samples recorded under the same conditions. Melting curves were recorded at constant wavelength at 216 nm for hTom40-A(Δ1–82) wt and mutant hTom40A from 25 to a maximum of 98 °C by applying a temperature ramp of 1 °C/min. All CD samples were

filtered (Rotilabo® filter; pore size, 0.22 μm) and spun down at full speed with a bench top centrifuge for 5 min at room temperature (Biofuge Fresco; Heraeus, Newport Pagnell) before measurements were conducted. Temperature readings displayed an error of ~ 1 °C, which was added to the experimental error. The mean residue ellipticity Θ($T$) was calculated based on the molar protein concentration and the number of amino acid residues of regarding Tom40 proteins. Secondary structure content was determined using the CDpro package, namely, CDSSTR, CONTIN/LL and SELCON 3.[39-41]

AfTom40 proteins were measured under the same conditions at 0.1–0.4 mg/ml 20 mM $NaH_2PO_4$ (pH 7), 1 mM β-mercaptoethanol and 0.8% (w/v) Brij35.

To fit the unfolding curves of mutant Tom40, we assumed a two-state unfolding mechanism. The fraction $f_U(T)$ of unfolded protein was calculated according to $f_U(T)=[\Theta(T)-\Theta_N(T)]/[\Theta_U(T)-\Theta_N(T)]$ and fitted to the sigmoid function $f_U(T)=1/\left[1+e^{(A/T-B)}\right]$. $\Theta_N(T)$ and $\Theta_U(T)$ represent the ellipticities of the native and unfolded molecules, respectively. At low and high temperatures, $\Theta_N(T)$ and $\Theta_U(T)$ increase linearly with temperature according to $\Theta_N(T)=aT+b$ and $\Theta_U(T)=cT+d$. A and B are fitting parameters. If unfolding is fully reversible, they correlate with enthalpic and entropic changes in the unfolding reaction, respectively.[42-44] wt Tom40 curves were fitted by superposition of two sigmoid functions assuming a three-state unfolding mechanism.[45]

## Tryptophan fluorescence measurements

Tryptophan fluorescence spectra of wt and mutant hTom40A (~ 0.15 mg/ml) were recorded by a FP-6500 spectrofluorimeter (Jasco Inc.) after 24 h of incubation in 20 mM Tris–HCl, (pH 8), 1% (w/v) LDAO and 1 mM β-mercaptoethanol and 0.35 to 7.0 M GnHCl at 25 °C. Tryptophans were excited at 280 nm. Emission spectra were recorded between 300 and 400 nm with an integration time of 1 s. The band width for excitation and emission was set to 3 nm, respectively. Fluorescence spectra were evaluated by fitting the background-corrected spectra $I(\lambda)$ to the log-normal distribution $I(\lambda)=I_0\,e^{-\left[\ln 2/\ln^2\rho\right]\ln^2\left[1+(\lambda-\lambda_{max})(\rho^2-1)/\rho\Gamma\right]}$,[46,47] where $I_0$ is the fluorescence intensity observed at the wavelength of maximum intensity $\lambda_{max}$, $\rho$ is the line shape

asymmetry parameter and $\Gamma$ is the spectral width at half-maximum fluorescence intensity $I_0/2$. wt and mutant AfTom40 (~ 0.1 mg/ml) were investigated using the same parameters in 20 mM $NaH_2PO_4$ (pH 7), 1 mM β-mercaptoethanol and 0.8% (w/v) Brij35.

To characterize denaturant-induced folding and unfolding of mutant hTom40A and AfTom40, we fitted the fraction of unfolded protein $f_U(D)$ by the sigmoid function $f_U([D]) = 1/\left[1+e^{(A-B[D])/T}\right]$. If unfolding and folding are fully reversible, A correlates with the free energy of the protein that describes its stability at zero denaturant concentration. In this case, B would be a measure of the dependence of free energy on denaturant concentration [D].[50-52] Chemically induced unfolding of wt Tom40 was fitted to a three-state unfolding reaction as described in the circular dichroism section.

## Acknowledgements

The authors thank Drs. Uwe Gerken and Robin Ghosh for stimulating discussions and Dr. Kornelius Zeth for continuous support to F.M. We are also indebted to Drs. Andreas Kuhn and Andrei Lupas for access to the UV CD and fluorescence spectrometer and Dr. Markus Bohnsack for advice. Financial support was provided by the Competence Network on "Functional Nanostructures" of the Baden-Württemberg Foundation to S.N. and D.L. (TP/A08), by the Landesgraduiertenkolleg Baden-Württemberg to F.M., by National Institutes of Health grants GM079804 and GM086145 and National Science Foundation grant BMS0800257 to J.L., and by the Fulbright Commission and the Higher Education Commission of Pakistan to H.N.

## Contribution to this manuscript

DG proposed the use of the eukaryotic β-barrel protein Tom40 to adapt and experimentally verify previously developed computational methods by JL, HN and colleagues[3]. JL and HN carried out computational analysis. FM generated the hTom40A clones. DG, FM, TA and SW produced the recombinant human Tom40 proteins, whereas DG produced recombinant AfTom40 proteins. FM made cross-links experiments with human Tom40A proteins. SW, TA and DG made the biochemical and spectroscopic measurements. DG analyzed and fitted

spectroscopic data sets. HN, FM and DG created the figures. DG and SN wrote the manuscript, with the contribution of DL, TA, FM, HN and JL. HN, DG, SN and JL outlined the project. JL and SN oversaw the project.

## References


1. Haltia, T. & Freire, E. (1995). Forces and factors that contribute to the structural stability of membrane proteins. *Biochim. Biophys. Acta*, **1228**, 1–27.
2. Stanley, A. & Fleming, K. (2007). The role of a hydrogen bonding network in the transmembrane beta-barrel OMPLA. *J. Mol. Biol.* **370**, 912–924.
3. Naveed, H., Jackups, R., Jr. & Liang, J. (2009). Predicting weakly stable regions, oligomerization state, and protein–protein interfaces in transmembrane domains of outer membrane proteins. *Proc. Natl Acad. Sci. USA*, **106**, 12735–12740.
4. Adamian, L., Naveed, H. & Liang, J. (2011). Evolutionary conservation of lipid-binding sites in membrane proteins. *Biochim. Biophys. Acta, Biomembr.* **1808**, 1092–1102.
5. Van Gelder, P., Saint, N., Phale, P., Eppens, E. F., Prilipov, A., van Boxtel, R. *et al.* (1997). Voltage sensing in the PhoE and OmpF outer membrane porins of *Escherichia coli*: role of charged residues. *J. Mol. Biol*. **269**, 468–472.
6. Phale, P. S., Philippsen, A., Kiefhaber, T., Koebnik, R., Phale, V. P., Schirmer, T. & Rosenbusch, J. P. (1998). Stability of trimeric OmpF porin: the contributions of the latching loop L2. *Biochemistry*, **37**, 15663–15670.
7. Delcour, A. (2002). Structure and function of poreforming beta-barrels from bacteria. *J. Mol. Micro. Biotechnol.* **4**, 1–10.
8. Evanics, F., Hwang, P. M., Cheng, Y., Kay, L. E. & Prosser, R. S. (2006). Topology of an outer-membrane enzyme: measuring oxygen and water contacts in solution NMR studies of PagP. *J. Am. Chem. Soc.* **128**, 8256–8264.
9. Ahting, U., Thieffry, M., Engelhardt, H., Hegerl, R., Neupert, W. & Nussberger, S. (2001). Tom40, the poreforming component of the protein-conducting TOM channel in the outer membrane of mitochondria. *J. Cell Biol.* **153**, 1151–1160.
10. Hill, K., Model, K., Ryan, M. T., Dietmeier, K., Martin, F., Wagner, R. & Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature*, **395**, 516–521.
11. Kinoshita, J. Y., Mihara, K. & Oka, T. (2007). Identification and characterization of a new tom40 isoform, a central component of

mitochondrial outer membrane translocase. *J. Biochem. (Tokyo)*, **141**, 897–906.

12. Künkele, K. P., Heins, S., Dembowski, M., Nargang, F. E., Benz, R., Thieffry, M. *et al.* (1998). The preprotein translocation channel of the outer membrane of mitochondria. *Cell*, **93**, 1009–1019; issn: 0092–8674.
13. Suzuki, H., Okazawa, Y., Komiya, T., Saeki, K., Mekada, E., Kitada, S. *et al.* (2000). Characterization of rat TOM40, a central component of the preprotein translocase of the mitochondrial outer membrane. *J. Biol. Chem.* **275**, 37930–37936.
14. Werhahn, W., Jänsch, L. & Braun, H. P. (2003). Identification of novel subunits of the TOM complex of *Arabidopsis thaliana*. *Plant Physiol. Biochem.* **41**, 407–416.
15. Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S. *et al.* (2008). Structure of the human voltage-dependent anion channel. *Proc. Natl Acad. Sci. USA*, **105**, 15370–15375.
16. Pusnik, M., Charriere, F., Maser, P., Waller, R. F., Dagley, M. J., Lithgow, T. & Schneider, A. (2009). The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane. *Mol. Biol. Evol.* **26**, 671–680.
17. Zeth, K. (2010). Structure and evolution of mitochondrial outer membrane proteins of beta-barrel topology. *Biochim. Biophys. Acta*, **1797**, 1292–1299.
18. Becker, L., Bannwarth, M., Meisinger, C., Hill, K., Model, K., Krimmer, T. *et al.* (2005). Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore. *J. Mol. Biol.* **353**, 1011–1020.
19. Malia, T. J. & Wagner, G. (2007). NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-xL. *Biochemistry*, **46**, 514–525.
20. Jackups, R., Jr. & Liang, J. (2005). Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. *J. Mol. Biol.* **354**, 979–993.
21. Suzuki, H., Kadowaki, T., Maeda, M., Sasaki, H., Nabekura, J., Sakaguchi, M. & Mihara, K. (2004). Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched beta-structure. *J. Biol. Chem.* **279**, 50619–50629.

22. Ahting, U., Thun, C., Hegerl, R., Typke, D., Nargang, F. E., Neupert, W. & Nussberger, S. (1999). The TOM core complex: the general protein import pore of the outer membrane of mitochondria. *J. Cell Biol.* **147**, 959–968.
23. Model, K., Meisinger, C. & Kuhlbrandt, W. (2008). Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase. *J. Mol. Biol.* **383**, 1049–1057.
24. Kato, H. & Mihara, K. (2008). Identification of Tom5 and Tom6 in the preprotein translocase complex of human mitochondrial outer membrane. *Biochem. Biophys. Res. Commun.* **369**, 958–963.
25. Sherman, E. L., Go, N. E. & Nargang, F. E. (2005). Functions of the small proteins in the TOM complex of *Neurospora crasssa*. *Mol. Biol. Cell*, **16**, 4172–4182.
26. Bayley, H. & Cremer, P. S. (2001). Stochastic sensors inspired by biology. *Nature*, **413**, 226–230.
27. Seitz, T., Bocola, M., Claren, J. & Sterner, R. (2007). Stabilisation of a (betaalpha)8-barrel protein designed from identical half barrels. *J. Mol. Biol.* **372**, 114–129.
28. Hocker, B., Lochner, A., Seitz, T., Claren, J. & Sterner, R. (2009). High-resolution crystal structure of an artificial (betaalpha)(8)-barrel protein designed from identical half-barrels. *Biochemistry*, **48**, 1145–1147.
29. Seitz, T., Thoma, R., Schoch, G. A., Stihle, M., Benz, J., D'Arcy, B. *et al.* (2010). Enhancing the stability and solubility of the glucocorticoid receptor ligandbinding domain by high-throughput library screening. *J. Mol. Biol.* **403**, 562–577.
30. Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C. & Hamodrakas, S. J. (2004). PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. *Nucleic Acids Res.* **32**,W400–W404.
31. Ou, Y. Y., Chen, S. A. & Gromiha, M. M. (2010). Prediction of membrane spanning segments and topology in beta-barrel membrane proteins at better accuracy. *J. Comput. Chem.* **31**, 217–223.
32. Liang, J. (2002). Experimental and computational studies of determinants of membrane-protein folding. *Curr. Opin. Chem. Biol.* **6**, 878–884.
33. Jackups, R., Jr & Liang, J. (2010). Combinatorial analysis for sequence and spatial motif discovery in short sequence fragments. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **7**, 524–536.

34. Jackups, R., Jr, Cheng, S. & Liang, J. (2006). Sequence motifs and antimotifs in beta-barrel membrane proteins from a genome-wide analysis: the Ala–Tyr dichotomy and chaperone binding motifs. *J. Mol. Biol.* **363**, 611–623.
35. Mager, F., Gessmann, D., Nussberger, S. & Zeth, K. (2011). Functional refolding and characterization of two Tom40 isoforms from human mitochondria. *J. Membr. Biol.* **242**, 11–21.
36. Harris, T. J., Patel, T., Marston, F. A., Little, S., Emtage, J. S., Opdenakker, G. *et al.* (1986). Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in *Escherichia coli*. *Mol. Biol. Med.* **3**, 279–292.
37. Engel, J. (1987). DNA cloning: a practical approach. *Acta Biotechnol.* **9**, 254.
38. Wilkins, M., Gasteiger, E., Bairoch, A., Sanchez, J., Williams, K., Appel, R. & Hochstrasser, D. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol. (Clifton, N. J.)*, **112**, 531–552.
39. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260.
40. Sreerama, N. & Woody, R. W. (2003). Structural composition of betaI- and betaII-proteins. *Protein Sci.* **12**, 384–388.
41. Sreerama, N. & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods Enzymol.* **383**, 318–351.
42. Becktel, W. J. & Schellman, J. A. (1987). Protein stability curves. *Biopolymers*, **26**, 1859–1877.
43. Pace, C. N., Hebert, E. J., Shaw, K. L., Schell, D., Both, V., Krajcikova, D. *et al.* (1998). Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *J. Mol. Biol.* **279**, 271–286.
44. Fersht, A. (1999). *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. pp. 1–631, W. H. Freeman and Company, New York.
45. Zheng, X. Y. & Yang, B. S. (2010). An improved method for measuring the stability of a three-state unfolding protein. *Chinese Sci. Bull.* **55**, 4120–4124.
46. Ladokhin, A. S., Jayasinghe, S. & White, S. H. (2000). How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? *Anal. Biochem.* **285**, 235–245.

47. Winterfeld, S., Imhof, N., Roos, T., Bar, G., Kuhn, A. & Gerken, U. (2009). Substrate-induced conformational change of the *Escherichia coli* membrane insertase YidC. *Biochemistry*, **48**, 6684–6691.
48. Spyropoulos, I. C., Liakopoulos, T. D., Bagos, P. G. & Hamodrakas, S. J. (2004). TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics*, **20**, 3258–3260.
49. Pace, C. & Scholtz, J. (1997). Measuring the conformational stability of a protein. In *Protein Structure. A Practical Approach* (Creighton, T. E., ed.), pp. 299–321, Oxford University Press, USA.
50. Huyghues-Despointes, B. M., Pace, C. N., Englander, S. W. & Scholtz, J. M. (2001). Measuring the conformational stability of a protein by hydrogen exchange. *Methods Mol. Biol.* **168**, 69–92.
51. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
52. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.7

# Supplementary Material

**Supplementary Table1** Primers used for site-directed mutagenesis of hTom40A

| Mutation | Forward primer (5'-3') | Reverse primer (5'-3') |
|---|---|---|
| K107L/H117L | TCAATCTAGGGTTGAGTAACCATTTTCAGGT CAACCTCACAGTAGCCCTCAGCACAATCGG | CTGTGAGGTTGACCTGAAAATGGTTACTCAA CCCTAGATTGACTGTGAGCTTGACACCCTCCA TC |
| H220L | GGGTTCAGGAATCCTCGTAGCCCTCTACCTCC AGAGCATCACGCCTTGCCTGGC | GGCAAGGCGTGATGCTCTGGAGGTAGAGGGC TACGAGGATTCCTGAACCCACGAGG |
| H117L | GAGTAACCATTTTCAGGTCAACCTCACAGTA GCCCTCAGCACAATCGGG | GTGCTGAGGGCTACTGTGAGGTTGACCTGAA AATGGTTACTCAACCC |
| K107L | GGGTGTCAAGCTCACAGTCAATCTAGGGTTG AGTAACCATTTTCAGGTCAAC | CCTGAAAATGGTTACTCAACCCTAGATTGAC TGTGAGCTTGACACCCTCCATC |

**Supplementary Table2** Secondary structure content of wild type and K107L/H117L/H220L mutant hTom40A

| Protein | $\lambda_{crossover}$ | $\lambda_{min}$ | $\Theta_{R\,min}$ | α-Helix | β-Sheet | Turn | Random coil |
|---|---|---|---|---|---|---|---|
| Wild-type | 205 ± 0.6 | 217 ±0.3 | -11 ± 0.9 | 0.2 ± 0.02 | 0.35 ± 0.03 | 0.2 ± 0.01 | 0.25 ± 0.02 |
| Mutant | 204 ± 0.3 | 216 ± 0.3 | -10 ± 0.9 | 0.2 ± 0.02 | 0.35 ± 0.03 | 0.2 ± 0.01 | 0.26 ± 0.02 |

CD spectra characteristics and secondary structure percentages (SS) were determined by averaging three curves of independent experiments. One example is displayed in Fig. 2 for wild type and the triple mutant of hTom40A, respectively. Using the CDpro package [1,2,3], SS content was determined by averaging data from CDSSTR, CONTIN/LL and SELCON[3]. $\lambda_{crossover}$ and $\lambda_{min}$ in [nm]; $\Theta_{R\,min}$ in [$10^{-3}$ deg cm$^2$ dmol$^{-1}$].

```
hTom40AdN1-82      1 ---------------------------------------MDFEECHRKCK     11
                                                           ...:...|:.:
AfTom40            1 MADFLALPSFLVENSAAVALKDAYNSFSERRAALGLSNPGTVDNIAREVQ     50

hTom40AdN1-82     12 E---LFPIQMEGVKLTVNK--GLSNHFQVNHTVALSTIGESN---YHFGV     53
                     :   |......|::..:.|  |:|..|:|:|  |.|..|..|   |.|..
AfTom40           51 KEVLLSNFMFSGLRADLTKVFGMSPLFRVSH--AFSMGGSGNLPPYAFSA     98

hTom40AdN1-82     54 TYVGTKQLSPTEAFPVLVGDMDNSGSLNAQVIHQLGPGLRSKMAIQTQQS    103
                     .| ||.::       .:.|:..:.|:|.|...::..|.|.:|...|....
AfTom40           99 MY-GTPKV-------FMQGNFGSDGALAAVFNYRWNPKLVTKTNTQIMAG    140

hTom40AdN1-82    104 KFVN-WQVDGEYRGSDFTAAVTLGNPDVLVG--SGILVAHYLQSITPCLA    150
                     .... .|:|.:|.|.||:|::...||..|.|  :||.|..|||||||.||
AfTom40          141 ASQGLLQIDNDYTGDDFSASIKAFNPSCLDGGLTGIFVGSYLQSITPSLA    190

hTom40AdN1-82    151 LGGELVYHR-----RPGEEGTVMSLAGKYTLNNWLATVTLGQAGMH-ATY    194
                     ||.|.::.|     ||   .|.:|...:|..|:|:|:..|...|:. |:|
AfTom40          191 LGFEAIWQRQAMNTRP---ETALSYCARYKANDWIASAQLQAQGVFTASY    237

hTom40AdN1-82    195 YHKASDQLQVGVEFE-------ASTRM----QDTSVSFGYQLDLPKANLL    233
                     :.|.|::::.||:..       |:..|    :|.:.|.|.:.|...:.
AfTom40          238 WRKLSERVEAGVDMNLQFAPNAAAALMGGPSRDGTTSIGAKYDFRAST--    285

hTom40AdN1-82    234 FKGSVDSNWIVGATLEKKLPPLPLTLALGAFLNHRKNKFQCGFGLTIGLE    283
                     |:..|||...|...|||:: .:|::|.....:..|...:  .||.:.||
AfTom40          286 FRAQVDSAGKVSCLLEKRI-AMPISLTFAGEIDQAKQSAK--LGLAVSLE    332

hTom40AdN1-82    284 ----------------------HHHHHH    289
                                            ||||||
AfTom40          333 IAGEELMEQQEKIEAQGMIPPPFHHHHHH    361
```

**Supplementary Figure 1** Amino acid sequence alignment of hTom40AΔN1-82 and full length AfTom40. Alignments were conducted by using a Needleman-Wunsch global alignment algorithm[4] via the EMBL-EBI toolkit site. Amino acid identities and similarities are ~25 and ~40%, respectively. Predicted α-helical and β-barrel secondary structure protein segments (see Materials and Methods) are coloured in red and green, respectively. Mutated amino acid residues are underlined.


1. Sreerama, N. & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods Enzymol* **383**, 318-51.
2. Sreerama, N. & Woody, R. W. (2003). Structural composition of betaI- and betaII-proteins. *Protein Sci* **12**, 384-8.
3. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252-60.
4. Needleman, S. B. & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**, 443-53.

# Manuscript II

## Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry

**Dennis Gessmann[1,#], Nadine Flinner[2,4,#], Jens Pfannstiel[5], Andrea Schlösinger[1], Enrico Schleiff[2,3,4], Stephan Nussberger[1] and Oliver Mirus[4]**

[1]Biophysics Department, Institute of Biology, University of Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany

[2]Centre of Membrane Proteomics, Germany

[3]Cluster of Excellence Macromolecular Complexes, Germany

[4]Department of Biosciences, Molecular Cell Biology of Plants, University of Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt, Germany

[5]Life Science Center Core Facility, University of Hohenheim, August-von-Hartmann Str. 3 Hohenheim, Germanyepartment for Protein Evolution, Max Planck Institute for Developmental Biology, Spemannstr. 35-39, 72076 Tübingen, Germany

*Running head:* D. Gessmann *et al.* / Biochimica et Biophysica Acta 1807 (2011) 1647-1657

*Keywords:* Tom40; VDAC; β-barrel; eukaryotic porin; mitochondria; protein translocation

[#]DG and NF contributed equally to this work.

## Abstract

Most mitochondrial proteins are imported into mitochondria from the cytosolic compartment. Proteins destined for the outer or inner membrane, the inter-membrane space, or the matrix are recognized and translocated by the TOM machinery containing the specialized protein import channel Tom40. The latter is a protein with β-barrel shape, which is suggested to have evolved from a porin-type protein. To obtain structural insights in the absence of a crystal structure the membrane topology of Tom40 from *Neurospora crassa* was determined by limited proteolysis combined with mass spectrometry. The results were interpreted on the basis of a structural model that has been generated for NcTom40 by using the structure of mouse VDAC-1 as a template and amino acid sequence information of ~270 different Tom40 and ~480 VDAC amino acid sequences for refinement. The model largely explains the observed accessible cleavage sites and serves as a structural basis for the investigation of physicochemical properties of the ensemble of our Tom40 sequence data set. By this means we discovered two conserved polar slides in the pore interior. One is possibly involved in the positioning of a pore-inserted helix; the other one might be important for mitochondrial pre-sequence peptide binding as it is only present in Tom40 but not in VDAC proteins. The outer surface of the Tom40 barrel reveals two conserved amino acid clusters. They may be involved in binding other components of the TOM complex or bridging components of the TIM machinery of the mitochondrial inner membrane.

## Introduction

Channeling of proteins across biological membranes through nanometer-scale pores is a process common to all living cells [1,2]. In all identified pathways the transfer of proteins across the lipid bilayer requires an assembly of proteins, which recognize the polypeptide to be translocated, mediate its insertion into a pore, facilitate the transfer through the pore, and drive the movement so that the transport becomes vectorial. In outer membranes of mitochondria, the translocon at the outer membrane of mitochondria (TOM) is such a molecular machine evolved for this function [3–6]. Within the TOM complex, Tom40 represents the pore for transport of virtually all cytosolic proteins destined for mitochondria. It forms a cation-selective, hydrophilic channel [7–11], which

binds to mitochondrial pre-sequences [12,13] and facilitates translocation of unfolded or partially folded proteins from the cytosol to the inter-membrane space of mitochondria [14,15]. Tom40 is an essential component of the TOM complex and already moderate reduction of the Tom40 protein content shows a drastic reduction of precursor protein import manifesting the central role of this protein in translocation [16]. In addition, specific mutations of the Tom40 amino acid sequence resulted in an ineffective transfer of preproteins to the mitochondrial inner membrane protein translocase TIM23 [17]. All results together indicate that the Tom40 channel is not a passive pore, but plays an active role in protein sorting for all sub-mitochondrial locations.

The secondary structure of Tom40 from several organisms has been investigated by CD and FTIR spectroscopy. These studies revealed an enriched β-sheet secondary structure, similar to the mitochondrial voltage-dependent anion channel VDAC [10,18–20]. Consistently, a β-barrel conformation was postulated early on [21,22]. Recent modeling of the Tom40 structure based on the mouse VDAC-1 [23] is consistent with the existence of 19 β-strands and one N-terminal α-helical segment located inside the pore [24–26].

It was demonstrated that precursor proteins can be chemically cross-linked to the *cis* and *trans* site of Tom40 [12]. This interaction is speculated to be essential to prevent precursor protein aggregation [13]. A subsequent mapping of the interaction sites within the pre-sequence revealed that Tom40 is in close contact with the central region of the pre-sequence while the N-terminus is bound to the inter-membrane space domain of Tom22 [27]. In Tom22 mutants interactions between the N-terminus of the pre-sequence and Tom40 are observed [27]. Remarkably, by a similar strategy using full length protein fused to DHFR, cross-links between Tom40 and the mature domain but not with the pre-sequence were observed [27]. Thus, a specific binding motif within pre-sequences for the interaction with Tom40, which can only be transient, has not yet been identified. Based on this finding it is tempting to assume that the mode of interaction between Tom40 and precursor proteins parallels the mode of chaperone-substrate interaction by exhibiting a specific surface for the recognition of hydrophobic or polar stretches [28].

The analysis of purified Tom40 protein and of the TOM core complex after reconstitution into planar lipid bilayers indicated that Tom40 forms a cation

selective membrane pore with an estimated inner pore diameter of 20 to 25 Å [7,10,19,29,30]. At high voltages, the channel revealed a complex gating behavior indicating fluctuation between different conformational states [7,19,31]. As expected, mitochondrial pre-sequence peptides were significantly more potent in blocking the pore than the synthetic non-mitochondrial model peptides [32].

In summary, Tom40 engages multiple interactions with precursor proteins and TOM complex components. Inspired by the solved structure of VDAC, several homology models were generated [24–26]. Remarkably, however, they could not yet explain the multitude of properties observed by biochemical and biophysical means. Here, we aimed at a structural description of the protein topology and its support by biochemical studies. We built a homology model using a consensus alignment approach to construct a multiple sequence alignment from a non-redundant set of ~270 Tom40 and ~480 VDAC amino acid sequences. Our alignment approach allowed detecting homologous secondary structure elements of these two divergent but phylogenetically related protein families. The structural architecture is displayed for Tom40 from *Neurospora crassa* (NcTom40) and the model is supported by peptide-bond accessibility probed by enzymatic hydrolysis. Mapping the properties of the ensemble of Tom40 sequences in our data set onto the model reveals conserved amino acid residues out- and inside the pore, which may play key roles in the assembly of the protein translocation machinery TOM and the transport of precursors, respectively. Based on the model we can reconcile many of the established results.

## Materials and Methods

### Purification and preparation of NcTom40 for proteolysis

Tom40 protein was purified from mitochondria of *N. crassa* strain GR-107 that comprises a hexahistidinyl-tagged form of Tom22. Purification was conducted accordingly as previously described [10,31,33] with minor modifications regarding the buffers. Mitochondria were solubilized in 1% (w/v) n-dodecyl-β-maltoside (DDM), 20 mM Tris-HCl pH 8.5, 20% (v/v) glycerol, 300 mM NaCl, 20 mM imidazole and 1 mM PMSF for 30 min at 4 °C. By applying an ultracentrifugation step (181,000×g, 40 min) extract was clarified and further

bound to a nickel-nitriloacetic acid column (GE Healthcare). The column was washed with 0.1% (w/v) DDM, 20 mM Tris–HCl pH 8.5, 10% (v/v) glycerol, 300 mM NaCl, 1 mM PMSF, before Tom40 was eluted with 3% (w/v) n-octyl-β-D-glucopyranoside (OG, Glycon), 20 mM Tris–HCl pH 8.5, 2% (v/v) DMSO or 10% (v/v) glycerol. Peak fractions of OG elution step were pooled and concentrated with a 50 kDa MWCO filter (Sartorius) to a final protein concentration of 2 to 10 mg/ml. Concentrated protein sample was further dialysed (6–8 kDa MWCO; Spectrum Laboratories, Inc.) against 100× the volume of OG elution buffer at 4 °C overnight by stirring. Final sample was then diluted with OG elution buffer to a final protein concentration of 1 mg/ml for proteolysis and stored at 4 °C. Protein purity was examined by sodium dodecyl sulfate gel electrophoresis and staining with Coomassie Brilliant Blue. Determination of protein concentration was carried out via UV absorbance spectroscopy at 280 nm with a NanoDrop ND-1000 spectrophotometer (Thermo Fischer Scientific). The protein concentration was then calculated using an extinction coefficient of 35410 $M^{-1}$ $cm^{-1}$ estimated by primary sequence analyses of NcTom40 via ProtParam on the EXPASY.org proteomics server [34,35] with the tyrosine and tryptophan extinction coefficients of Pace et al. [36].

**Proteolysis**

Proteolysis of NcTom40 was performed *in vitro* on the basis of previous methods [37]. 2 μg of NcTom40 in 50 μl of 3% (w/v) n-octyl-β-D-glucopyranoside, 20 mM Tris–HCl pH 8.5 and 2% (v/v) DMSO were incubated with either 0.2 μg Asp-N, Lys-C, Glu-C or Trypsin at a protein to protease ratio of 1:10 (w/w) at 25 °C. 50 μl aliquots for SDS-PAGE analysis were taken after 10 s, 1 min, 5 min, 15min and 30 min. 10 μl aliquots for mass spectrometry analysis were taken after 1, 10, 60 and 240 min. Digests were stopped by acidifying the aliquots with 1 μl of 10% TFA and putting them on ice. For ESI-MS/MS experiments samples were passed through 10 kDa cut off Nanosep filtration units (Pall, Ann Arbor, United States) to remove undigested protein and part of the detergent.

**Mass spectrometry**

Peptide mass fingerprints and MALDI-TOF-TOF spectra of NcTom40 digests were acquired on an Autoflex III MALDI-TOF-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). The instrument was operated in the positive ion mode and externally calibrated using peptide calibration standards (Bruker Daltonics, Bremen, Germany). NcTom40 protein digests were desalted and concentrated using micro-$C_{18}$ ZipTips (Millipore, Schwalbach, Germany) according to the manufacturer's protocols. Peptides were eluted directly onto a stainless steel target using a HCCA ($\alpha$-Cyano-4-hydroxycinnamic acid) matrix solution (5 mg/ml in 50% ACN/ 50% 0.1% TFA, v/v). Peptide mass fingerprint data and MS/MS spectra were recorded in reflector mode using an accelerating voltage of 21 kV. Peptide mass fingerprint spectra were analyzed in the MS mode with 2000 laser shots per sample to ensure good S/N quality. MS/MS analysis was done with varying number of laser shots. Flex Analysis 3.0 and Bio-Tools 3.0 software (Bruker Daltonics, Bremen, Germany) were used for data processing.

Nano-LC-ESI-MS/MS experiments were performed on an ACQUITY nano-UPLC system (Waters, Milford, USA) directly coupled to a LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Fisher, Bremen, Germany). Digests of NcTom40 were concentrated and desalted on a precolumn (2 cm×180 μm, Symmetry $C_{18}$, 5 μm particle size, Waters, Milford, USA) and separated on a 20 cm×75 μm BEH 130 $C_{18}$ reversed phase column (1.7 μm particle size, Waters, Milford, USA). Gradient elution was performed from 1% ACN to 50% ACN in 0.1% FA within 1 h. The LTQ-Orbitrap XL was operated under the control of XCalibur 2.0.7. software. Survey spectra (m/z=250–1800) were detected in the Orbitrap at a resolution of 60.000 at m/z=400. Data-dependent tandem mass spectra were generated for the five most abundant peptide precursors in the linear ion trap. For all measurements using the Orbitrap Detector internal calibration was performed using lock-mass ions from ambient air.

Mascot™ 2.2 software (Matrix Science, London, UK) was used for protein identification from MALDI-TOF and ESI-MS/MS experiments. Spectra were searched against the NCBI protein sequence database downloaded as FASTA-formatted sequences from ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz. Search parameters for MALDI-TOF experiments specified fungi as taxonomy, Asp-N,

Glu-C, Lys-C or Trypsin as cleaving enzyme allowing two missed cleavages, a 50 ppm mass tolerance for peptide precursors and 0.6 Da tolerance for fragment ions. Due to the higher mass accuracy of the Orbitrap analyzer mass tolerance for precursor ions in ESI-MS/MS experiments was set to 3 ppm. Carbamidomethylation of cysteine residues was set as fixed modification, methionine oxidation and N-terminal acetylation of proteins were allowed as variable modifications.

**Tom40 and VDAC sequences**

Amino acid sequences of Tom40 and VDAC proteins were collected from GenBank [38] and genome projects (Supplementary Table S2) not included in GenBank by PSI-BLAST [39] and HMMER (http://hmmer.org/). They were manually curated and a non-redundant sequence set was produced by cd-hit [40] leaving 266 Tom40 and 485 VDAC sequences in the full dataset.

**Multiple sequence alignment**

Multiple sequence alignments (MSAs) were created with MAFFT v6.847b [41,42] and its iterative, progressive alignment method L-INSi (localpair, max. 1000 iterations). We tested BLOSUM 30, 45, 62, 80 [43] and JTT 50, 100, 150, 200, 250 matrices [44]. Gap open/extension penalties (MAFFT parameters *op* and *ep*) were tested in the ranges 2.5–3.0 and 0.06–0.175, respectively, for their ability to produce MSAs of VDAC and Tom40, which contain as few gaps as possible in regions of Tom40 aligned to the core regions of transmembrane β-strands of VDAC. For each matrix-parameter combination 50 MSAs were calculated with the input sequences in different orders. The JTT200 matrix and 2.65 gap-open, 0.15 gap-extension worked best and with these parameters 1000 MSAs were calculated with random succession of the input sequences in the FASTA file supplied to MAFFT. But we required a unique order of the first 10 sequences for each FASTA file to avoid a bias in the guide trees. From all 1000 MSAs a statistic was computed on how often which region from NcTom40 was aligned to which β-strand core region of mVDAC-1. Among the 1000 MSAs only 49 showed a gap-free alignment in all 19 β-strand core regions. Out of these 49 MSAs only 27 contained the secondary structure elements aligned in the most frequently occurring fashion. One of these very similar 27 MSAs was chosen for

reconstructing the phylogeny of Tom40/ VDAC and for homology modeling of NcTom40.

### Homology modeling

Homology modeling of Tom40 from *N. crassa* was performed with Modeller 9v7 [45,46] based on the template structure of VDAC from *Mus musculus* solved at a resolution of 2.3 Å with an $R_{free}$ of 0.277 [23]. The alignment of target to template was taken from the MSA of all VDAC and Tom40 sequences. Loop modeling and model optimization were performed with YASARA (www.yasara.org).

### Analysis of sequence properties of Tom40

A plugin was written for YASARA to color a given structure based on various amino acid properties. The plugin takes as input a MSA and the type of analysis to perform. In here we use an MSA filtered with cd-hit [40] to a maximal sequence identity of 90% in order to reduce the bias, which would otherwise be introduced by an overrepresentation of very similar sequences of closely related species. Each residue in the selected structure is colored according to the average of all residues in the respective column of the MSA either by hydrophobicity (Supplementary Table S3) [47], side chain size (Supplementary Table S3), degree of amino acid conservation, or degree of conservation of groups of amino acids of similar properties (Supplementary Table S4).

### Molecular graphics

Molecular graphics were created with Yasara (www.yasara.org) and PovRay (www.povray.org).

## Results and Discussion

### Homology modeling of NcTom40 required a specific alignment process

We built a database of all available Tom40 and VDAC amino acid sequences from GenBank, PFAM and several genome projects (Supplementary Table S2), which could be identified by PSI-BLAST or profile HMM searches. The manually curated, non-redundant database contained ~270 Tom40 and ~480 VDAC

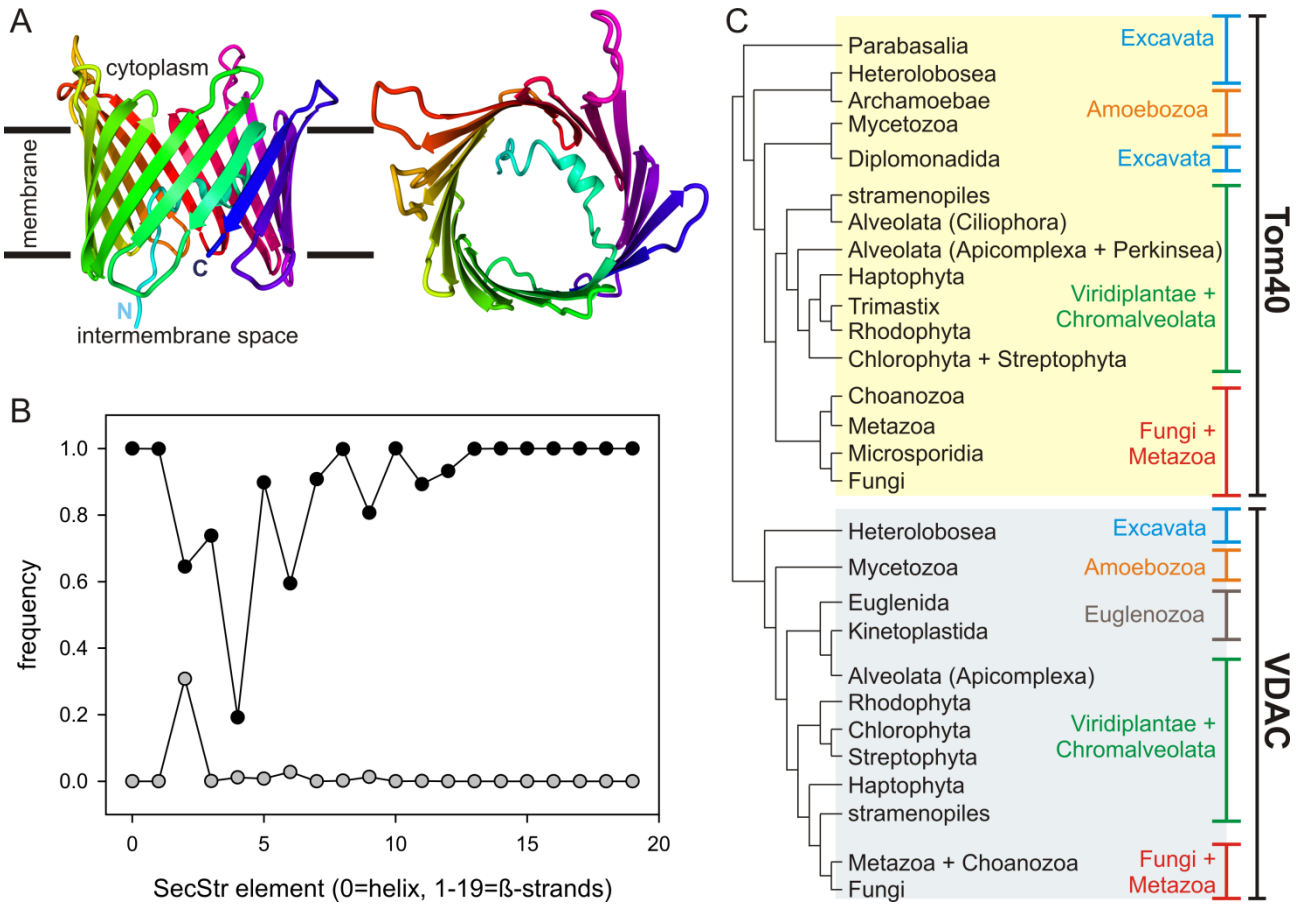



**Fig. 1.** *Homology model of Tom40*. (A) The homology model of Tom40 from *Neurospora crassa* is colored from N- to C-terminus in a gradient from cyan, over green, yellow, red, and magenta to blue. The membrane-inserted region of NcTom40 is approximately indicated on the left panel, and on the right panel the model is rotated by 90° around the x axis, resulting in a view through the pore from the cytoplasm to the intermembrane space. (B) Our multiple sequence alignment strategy was applied to the full, non-redundant set (black lines) of Tom40 and VDAC sequences. The frequency of occurrence of equal alignments of regions of NcTom40 to the helix and β-strand core regions of mVDAC-1 is shown for the most frequently (black circles) and for the 2nd most frequently occurring aligned regions (gray circles) of NcTom40. (C) A schematic representation of the maximum likelihood (ML) phylogeny of Tom40 and VDAC. The branch lengths in this schematic tree do not correspond to the frequency of amino acid substitutions; they only indicate the topology of the ML tree.

proteins. This database provided the base to develop a structural model of Tom40 using the sequence from *N. crassa* as an example (Fig. 1A). The latter was chosen, because this protein is well characterized [10,16,48–50] and was used in the subsequent biochemical analysis. The protein shares ~60% sequence

similarity with the Tom40 protein from *Saccharomyces cerevisiae*, for which numerous data exist as well [7,11,12,17,19].

We employed an iterative, progressive, local alignment method (L-INS-i) implemented in MAFFT [41,42] to align all VDAC and Tom40 contained in our database. This algorithm uses a guide tree, whose topology depends on the order in which sequences are supplied to the alignment program. Prompted by the diversity of the collected sequences of Tom40 and VDAC we tested several parameter and substitution matrix combinations for their ability to align regions from NcTom40 to the core regions of the known transmembrane β-strands of mouse VDAC [23] with a low gap frequency. We reasoned that insertions and deletions show a tendency to occur more frequently between different secondary structure elements, that is between α-helices and β-strands, than within them [51]. Based on this constraint we observed that the JTT200 matrix [44] in combination with gap-open and gap-extension penalty of 2.65 and 0.15, respectively, yielded the best alignments.

From 1000 multiple sequence alignments (MSAs) computed with these settings and different successions of the input sequences we calculated the frequency of occurrence of equal alignments of regions of NcTom40 to the core regions of each β-strand of mVDAC-1 (Fig. 1B). In total 27 MSAs contained all most frequently aligned regions of NcTom40 aligned to the β-strand core regions of mVDAC-1. Analyzing the positioning within the alignments shows that the alignment to all structural segments is robust with the exception of β-strands 2 and 4 (Fig. 1B). Thus, aligned regions of NcTom40 were assigned to the N-terminal helix and the 19 transmembrane β-strands of VDAC (Supplementary Figs. S1).

The highest noise level in the multiple sequence alignments is observed for the β-strands β2, β4, and β6 (Fig. 1B). To assess the robustness of the proposed β-strand positioning, we repeated our MSA strategy for sequence sets reduced to a maximal sequence identity of down to 70% in steps of 10% (Supplementary Fig. S2). With the exception of β2 the results presented in Fig. 1B were reproduced. For β2 we observed an increase of the frequency of an alternative alignment while lowering the maximal sequence identity. The alternative alignment suggests a two-residue shift of β2 in N-terminal direction. Thus, for β2 we cannot completely rule out a slightly different alignment. For the other two

strands the original alignment remained the most frequent one (Supplementary Fig. S2).

The β-strand β4, however, exposes a negatively charged residue (E73) in mVDAC-1 to the lipid bilayer, which forms a binding site for the antiapoptotic Hexokinase [52]. This moderately conserved, negatively charged residue interrupts the typical alternating hydrophilic–hydrophobic pattern of transmembrane β-strands, which increases the noise for β4 in the MSA. An alternative alignment, whose frequency rises with decreasing maximal sequence identity of the Tom40/VDAC sequence set (Supplementary Fig. S2), can be excluded, because it is incompatible with the robust alignment for the neighboring β-strand β3. Interestingly, the support for β-strands β7- β19 is solid (N80%), especially for the extreme C-terminal β-strands β13- β19 (Fig. 1B, Supplementary Fig. S2). The lower support in the N-terminal region, however, reflects a higher degree of sequence variability in this region compared to the C-terminal part. A similar situation, where the N-terminal part of a transmembrane β-barrel is less conserved than the C-terminal part, was reported for Omp85 [53].

To further test the validity of our alignment approach we reconstructed a maximum likelihood phylogeny from one of those very similar 27 MSAs representing the most frequent alignment. Our phylogenetic tree shows the known succession of Viridiplantae, Fungi and Metazoa not only in the VDAC but also in the Tom40 clade (Fig. 1C; Supplementary Fig. S3). The latter was not obtained in a previous analysis [54]. However, sequences from protists, which show strongly varying amino acid compositions, inserted between Viridiplantae/Chromalveolates and Fungi/Metazoa in the VDAC clade. This insertion is only weakly supported by bootstrap analysis (Supplementary Fig. S3). Such a positioning of protist sequences is often a problem, which can only be resolved when more sequenced protist genomes become available.

Based on the analysis of the alignment frequency and the result of phylogenetic tree reconstruction the alignment obtained was considered suitable to build a homology model of Tom40 from *N. crassa* (Fig. 1A) using the crystal structure of murine VDAC [23] as a template. Remarkably, comparing our model with the model by Zeth [25,26] differences in assignment of β-strands β1 to β5, β9 and β10 are revealed. In Zeth's model of human Tom40 β1- β3 and β10

are shifted in C-terminal direction relative to our assignment of β-strands (Supplementary Fig. S1C) by 2, 4, 3, and 1 residues, respectively. In contrast, β4, β5, and β9 are shifted in N-terminal direction by 3, 2, and 1 residues, respectively.

As previously noticed, Tom40 shows the characteristic properties known from prokaryotic β-barrel proteins [55,56]. Mapping the average hydrophobicity derived for each sequence position from the ensemble of Tom40 sequences onto the model reveals the typical alternating pattern of hydrophobic and hydrophilic residues in the transmembrane β-strands (Supplementary Fig. S4). Also, the aromatic girdle is more pronounced on the side of the intermembrane space than the cytoplasm, which is also the case for VDAC. Furthermore, the conserved G323, which is assigned as a central amino acid of the β-signal responsible for targeting of Tom40 to outer mitochondrial membranes [57], is indeed positioned in β-strand β19.

The diameter of the pore of our homology model of NcTom40 is of course similar to that of the structure of the template VDAC. However, these pore dimensions match with prior studies on fungal Tom40s that determined a diameter of 22–25 Å and a size exclusion for particles with a diameter of maximal 25 Å [7,10,14]. Furthermore, the membrane topology of Tom40 from *N. crassa* has been assessed by antibody labeling, revealing both the N- and C-terminus in the inter-membrane space [58]. This strongly suggests that the N-terminus should thread through the pore as is the case in VDAC [23] and as shown (Fig. 1A).

**Protease-accessibility of isolated NcTom40**

To support the homology model of NcTom40 (Fig. 1A) we performed limited protease treatment coupled to mass spectrometry of Tom40 purified from *N. crassa* mitochondria (Fig. 2A). The purified NcTom40 adapted a functional fold after refolding as CD spectroscopic analysis is consistent with a β-barrel-like structure (Fig. 2B, Supplementary Table S1) [7,10,19]. The properties of the protein after reconstitution into black lipid bilayers were in line with previous experiments [31,32].

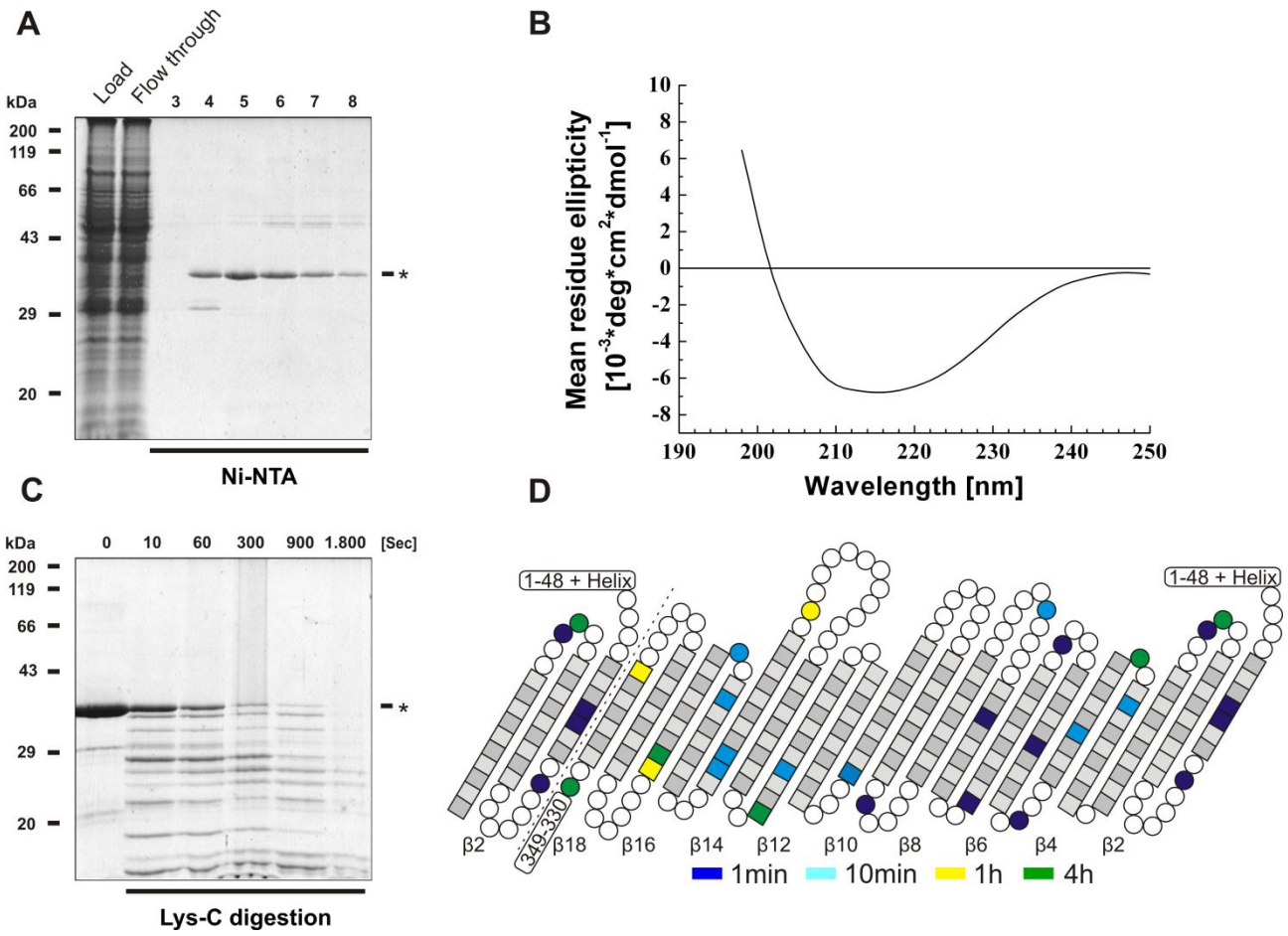



**Fig. 2.** *Assessment of the topology of Tom40 from Neurospora crassa by limited proteolysis and mass spectrometry*. (A) Coomassie-stained SDS-PAGE gel of NcTom40 purified from *N. crassa* mitochondria. (B) CD spectra of Tom40 were measured from 190 to 250 nm. Noisy data below 200 nm due to optical density has been removed. For each CD spectra, 3 scans were accumulated and background corrected at 20 °C. (C) Limited proteolysis of NcTom40 with Lys-C. NcTom40 in 3% (w/v) n-octyl-β-D-glucopyranoside, 20 mM Tris–HCl pH 8.5 and 2% (v/v) DMSO was incubated with Lys-C at a protein to protease ratio of 1 to 10 (w/w) at 25 °C. Samples were taken at depicted times and analyzed by SDS-PAGE. (D) The homology model of NcTom40 is shown in a schematic representation. Light gray boxes indicate residues exposed to the membrane, medium gray boxes indicate residues facing the pore interior. Loop residues are indicated by white circles. The residues at the beginning and end of an indicated β-strand (boxes) are as follows: β1 S63-V69, β2 L78-A86, β3 P93-T102, β4 Q104-D112, β5 G115-Y123, β6 R128-I137, β7 M143-G152, β8 F155-I162, β9 G173-Q181, β10 G188-Q196, β11 D205-Y214, β12 D218-A227, β13 A230-T240, β14 V243-L251, β15 Q266-D276, β16 S280-D287, β17 G290-K298, β18 T306-D314, β19 K321-S327. The cutting sites determined by mass spectrometry from the proteolysis experiments are marked by color for all proteolytic enzymes, i.e. Asp-N, Glu-C, Lys-C

Purified protein was then incubated with proteases Asp-N, Glu-C, Lys-C and trypsin for 1, 10, 60 and 240 min each (one example is shown in Fig. 2 C). After stopping proteolysis by acidification samples were directly subjected to MALDI-TOF and ESI-MS/MS analysis. Treatment of protein with Asp-N (Supplementary Fig. S5A), Glu-C (Supplementary Fig. S5B), Lys-C (Supplementary Fig. S5C) and trypsin (Supplementary Fig. S5D) revealed a characteristic peptide pattern for each protease over the period of incubation. Most peptides could be assigned to NcTom40 and are summarized in Table 1.

Two-thirds of all cleavage sites of NcTom40 cut after 1 min of digestion were located in the predicted loop regions of the protein (8 out of 12) (Fig. 2D and Supplementary Fig. S6). A number of putative cleavage sites located in the core region of the protein (6) and to some extent at the edge of the barrel (16) were not found to be accessible to protease even after 4 h of digestion. Three peptide bonds in the putative transmembrane strands β1 and β6 were readily cleaved by Asp-N (A67), Lys-C (K132) and trypsin (R66) (Fig. 2D and Supplementary Fig. S6). In comparable experiments using human VDAC-1 specific protease cleavage in the middle of β-strands 1 and 6 at corresponding positions K31, D33, F102 and D103 were observed as well (Supplementary Fig. S1C) [37,59]. This supports the model of NcTom40 although it cannot completely rule out a false assignment of β-strands 1 and 6. Alternative modeling taking into account the proteolysis results was not continued, because it was clearly not supported by the MSA. In summary, our results support the hypothesis of a VDAC-like fold for Tom40 proteins.

---

**(*Fig. 2. continued*)** and trypsin. The color code is as follows: cutting site detected after (i) 1 min (dark blue), (ii) 10 min (cyan), (iii) 1 h (yellow), (iv) 4 h (green). Amino acids cut C-terminally are shown and colored according to the time of occurrence during the coupled digestion/mass spectrometry experiments. This means that in case of AspN, which cuts aspartate at its N-terminus, not aspartate itself, but the residue N-terminal to it is marked. In Supplementary Fig. S6 the cutting sites are shown for each enzyme used and also residue numbers of cutting sites are given.

**Table 1** Mass identity of time-dependent protease digestion of NcTom40.

| Lys-C | | | | | Asp-N | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Exp. determined mass [M/z] | | | | | Exp. determined mass [M/z] | | | | |
| 1′ | 10′ | 60′ | 240′ | Peptide | 1′ | 10′ | 60′ | 240′ | Peptide |
| 3764.9 | 3764.8 | 3764.7 | 3764.9 | 2 - 34 | 1451.8 | 1451.8 | 1451.8 | 1451.7 | 2 - 14 |
| 4142.2 | 4142.1 | 4142.0 | 4142.2 | 35 - 71 | 1467.8 | 1467.8 | 1467.8 | 1467.7 | 2 - 14 |
| 3052.3 | 3052.3 | 3052.4 | 3052.3 | 133 - 160 | 2402.3 | - | - | - | 2 - 23 |
| - | 2349.3 | 2349.1 | 2349.1 | 216 - 237 | 968.5 | - | - | - | 15 - 23* |
| - | - | 2921.5 | 2921.4 | 238 - 265 | 3478.9 | 3479.0 | - | - | 24 - 53 |
| - | 1865.0 | 1864.9 | 1864.9 | 275 - 289 | 1601.9 | 1601.9 | 1601.9 | 1601.8 | 54 - 67 |
| - | - | 2409.2 | 2409.3 | 299 - 321 | - | 1617.9 | 1617.9 | 1617.8 | 54 - 67 |
| - | - | - | *#3004.4 | 322 - 349*# | - | 1785.9 | 1785.9 | 1785.8 | 112 - 126 |
| | | | | | - | - | 4846.4 | 4846.3 | 68 - 111 |
| | | | | | - | 1608.9 | - | - | 127 - 141 |
| | | | | | 2903.5 | 2903.6 | - | - | 127 - 152 |
| | | | | | - | 1313.6 | 1313.7 | 1313.5 | 142 - 152 |
| | | | | | - | 1624.9 | - | - | 153 - 167 |
| | | | | | - | - | - | 2563.3 | 218 - 240 |
| | | | | | - | - | - | 3646.7 | 241 - 275 |
| | | | | | - | 1371.8 | 1371.7 | 1371.7 | 276 - 286 |
| | | | | | - | 1387.8 | 1387.7 | 1387.7 | 276 - 286 |
| | | | | | *2114.0 | - | - | - | 312 - 331* |
| | | | | | *158.7 | - | - | - | 332 - 345* |

| Trypsin | | | | | Glu-C | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Exp. determined mass [M/z] | | | | | Exp. determined mass [M/z] | | | | |
| 1′ | 10′ | 60′ | 240′ | Peptide | 1′ | 10′ | 60′ | 240′ | Peptide |
| 1451.7 | 1451.8 | 1451.7 | 1451.7 | 2 - 14 | - | - | - | 2660.3 | 8 - 31 |
| 3480.6 | 3480.7 | - | - | 2 - 32 | 1601.9* | - | - | *1601.9 | 32 - 45* |
| 2047.9 | 2048.0 | 2047.9 | 2047.9 | 15 - 32 | - | - | - | 2800.4 | 90 - 114 |
| 1603.8 | 1603.9 | 1603.8 | 1603.9 | 35 - 49 | 1071.5* | - | - | 1071.5 | 267 - 276* |
| 2043.0 | 2043.1 | - | - | 50 - 66 | - | - | - | 3353.8 | 298 - 329 |
| - | 1530.8 | 1530.8 | 1530.8 | 54 - 66 | 1260.6* | - | - | *1260.6 | 330 - 340* |
| 2136.0 | 2136.1 | 2136.0 | 2136.1 | 72 - 90 | 945.5* | - | - | *945.5 | 341 - 349* |
| - | - | - | 2152.1 | 72 - 90 | | | | | |
| - | 3229.7 | 3229.6 | 3229.6 | 91 - 120 | | | | | |
| - | - | 3052.4 | 3052.4 | 133 - 160 | | | | | |
| 2721.4 | 2721.5 | 2721.4 | 2721.5 | 161 - 186 | | | | | |
| 1240.7 | 1240.7 | 1240.7 | 1240.7 | 187 - 197 | | | | | |
| - | 1724.9 | 1724.8 | 1724.9 | 198 - 213 | | | | | |
| - | 2640.4 | 2640.3 | - | 214 - 237 | | | | | |
| - | - | 2349.1 | 2349.2 | 216 - 237 | | | | | |
| - | - | - | *#3004.4 | 322 - 349*# | | | | | |

All masses were determined by MALDI-TOF except for '*'-marked masses/peptides, which were detected by ESI-MS after overnight incubation#. Note: Time displayed in minutes ′; all peptides starting with amino acid '2' were found to be acetylated; 'ox' corresponds to methionine oxidation; M/z values for ESI experiments (data not shown) are deconvoluted and correspond to the uncharged molecular masses of the peptides; values of peptide masses are truncated to first decimal place from data shown in Supplementary Fig. S5.

**Properties of the pore-inserted helix**

Our model of NcTom40 contains a putative helix located at the N-terminal end of the β-barrel domain inside the NcTom40 pore. Two highly conserved, charged amino acid residues (E45 and R49) are located at the N-terminal end of the helix. In VDAC four conserved charged residues (mVDAC-1: D9, K12, D16, K20) are located at similar positions (Fig. 3A and B). Interestingly, Tom40 and VDAC share only a single overall conserved residue in the helix, namely the conserved positive charge K12 in hVDAC-1, which aligns to R49 in NcTom40. This conservation might reflect an importance of this position for the general pore properties.

Channel conductance measurements of full length and N-terminally truncated versions of VDAC reconstituted into planar lipid bilayers suggested that the N-terminal helix of this protein is involved in channel gating [23,60–62]. Similar truncations of the region N-terminal of the α-helix in human Tom40 have been shown to result in virtually non-gating pores [20]. Thus, it can be speculated that the conserved charged amino acid residues E45 and R49 of NcTom40 are involved in channel gating.

Interestingly, in *N. crassa* and other *Ascomycota*, Tom40 sequences contain a single amino acid insertion in the middle of the N-terminal helix (73 of 102 sequences; in 47 cases a glutamine is inserted). Such an insertion is neither present in Tom40 sequences from other taxonomic groups nor in any VDAC sequence. The presence of this insertion might have been caused by evolutionary adaptation probably resulting in a transition of an α-helix to a π-helix as was recently shown by an analysis of all protein structures deposited in the PDB to be a common evolutionary event [63]. In line with such hypothesis, the two positions which correspond to V51 and T55 in NcTom40 surrounding the insertion (Q52) show a conserved hydrophobic character (Fig. 3C). Therefore, a π-helical conformation would be required in the *Ascomycota* Tom40 to register with residues G246 and G272 located in β-strands β14 and β15 of the NcTom40 barrel. In contrast, an α-helical conformation for the N-terminal helix in *Ascomycota* would lead to unfavorable interactions of the helix with the barrel wall.

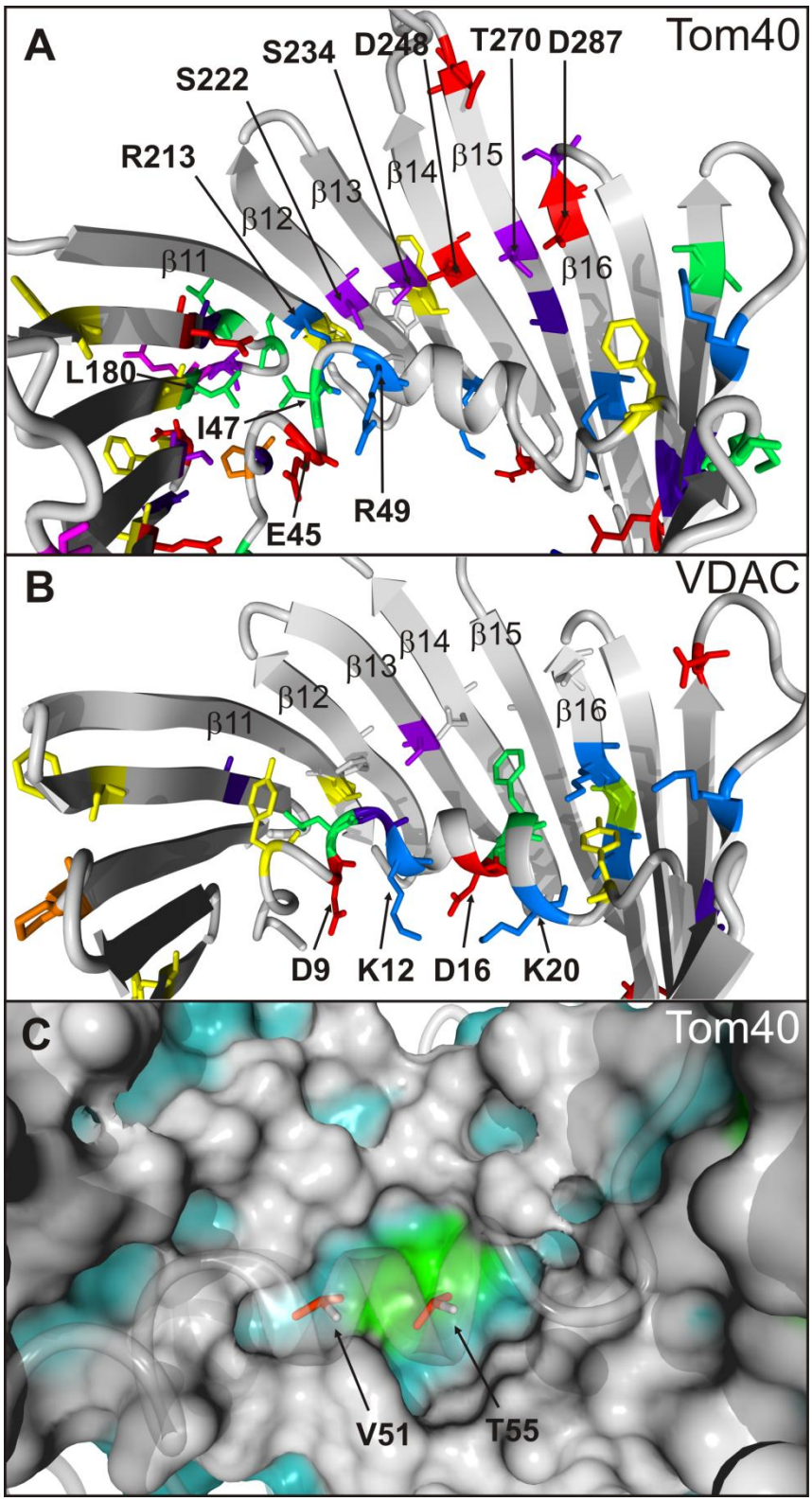



**Fig. 3.** *Positions with conserved residue properties in Tom40*. A close-up view of helix α1 and adjacent β-strands with the helix anchor region is shown. The sequences in the MSA of Tom40 and VDAC were reduced to a maximal sequence identity of 90%. The degree of conservation of amino acid groups (Supplementary Table S4) was calculated for each column in this MSA for either (A) Tom40 or (B) VDAC and mapped onto the surface of the homology model of (A) NcTom40 or (B) the crystal structure of

Six conserved amino acid residues in β-strands 11 to 16 (in NcTom40: R213, S222, S234, D248, T270 and D287) define a clear feature of Tom40. These residues align on a straight line, which crosses the pore from the cytoplasmic to the intermembrane space face of the protein (Fig. 3A). The line is capped by positively (R213) and negatively (D287) charged residues on both ends. It contains three small polar residues (S222, S234 and T270) and a single negatively charged residue (D248) in between. Since all these six residues are highly conserved among Tom40 proteins (Fig. 3A, Supplementary Fig. S7A) we hypothesize that they form a conserved helix anchor region inside Tom40 channels.

Nearly all residues of this helix anchor region are in contact with helix α1, with the exception of D287. Position S234 in NcTom40 is also strongly conserved in VDAC (Fig. 3B; Supplementary Fig. S7B). This serine (position 193) of hVDAC-1 is involved in hydrogen bonding with the helix α1 [59]. In addition, H183 is in contact with the helix in hVDAC-1 [59]. The corresponding amino acid in NcTom40 is serine 222, which is also capable of hydrogen bonding with the helix. Thus, this region most likely forms an interaction surface for helix α1. As the amino acid composition of the helix anchor region is more restricted in Tom40 than in VDAC (Fig. 3A and B; Supplementary Fig. S7) and helix α1 is proposed to move within the pore during gating events [23,60–62], the helix might be more tightly controlled in Tom40 than in VDAC. If the helix moves indeed during gating events, also residue D287 might be able to interact with the helix.

---

**(*Fig. 3. continued*)** mVDAC-1, respectively. Only positions with at least 75% conserved residue properties are shown. Residues are colored according to Supplementary Table S4. (C) The average side chain size of Tom40 sequences was inferred from the MSA prepared as described above and mapped onto the homology model of NcTom40. Only residues with average values below − 1 (Supplementary Table S3) are colored green, the others in gray. Backbone atoms from residues pointing into the membrane are colored in cyan. Green marks G246 and G272 in the binding pocket for helix α1, occupied in all Tom40 by residues with an average sidechain length of − 0.414 and − 0.253, respectively. The average hydrophobicity of V51 and T55, which dock in the groove, is − 2.11 and − 2.58, respectively, according to Kyte and Doolittle (Supplementary Table S3).

An involvement of the helix anchor region in binding helix α1 is supported by our finding that two positions (NcTom40: G246 in β14, G272 in β15) form a binding pocket for helix α1 and are occupied by residues with on average short side chains not only in VDAC, but also in Tom40 (Fig. 3C). Further, the hydrophobic character of two residues corresponding to the interacting hVDAC-1 residues L10 (α-helix) and V143 (β-strand β9) is conserved to at least 75% in Tom40 (NcTom40: I47 and L180, Fig. 3A). The according two residues in VDAC were shown to be important for the stability of the β-barrel of human VDAC-1 [64]. These results are in line with former studies on a temperature-sensitive yeast strain, where Leu66, which corresponds to Ile47 in *N. crassa* Tom40, was mutated to proline [49,65]. We postulate that the N-terminal helix stabilizes the overall β-barrel structure, and that at least Ile47 and Leu180 are involved in arresting the N-terminal helical region to the interior barrel wall.

**The polar slide in the pore interior**

Besides the polar helix anchor region (Fig. 3) we observed a 'polar slide' on the other side of the channel wall (Fig. 4A), which does not exist in VDAC (Fig. 4B, Supplementary Fig. S8). We mapped the average hydrophobicity for each residue derived from the MSA of all Tom40 sequences onto the surface of the homology model of NcTom40. This reveals a conserved 'polar slide' running from the cytoplasm to the intermembrane space spanning β-strands β3 to β8 (Fig. 4C). The slide is framed by two hydrophobic patches (labeled 1 and 2 in Fig. 4A and C). Opposed to the 'polar slide' two hydrophilic patches are located (labeled 3 and 4 in Fig. 4A, Supplementary Fig. S8), which are separated by a binding region for the linker of helix α1 and β-strand β1. The 'polar slide' is not so much characterized by conserved groups of amino acids (Fig. 4D, Supplementary Table S4), but instead by a conservation of hydrophilic property (Fig. 4A and C).

In maltoporin of gram negative bacteria, however, a similar situation exists with a 'greasy slide' formed by aromatic residues [66]. This greasy slide is important for substrate binding [67]. The substrates of Tom40, the mitochondrial targeting signals, have a capacity to form amphiphilic helices [68–70]. Thus, the polar slide might be involved in the recognition of the targeting signal similar to the polar 'glutamine face', which was discovered for the human

Tom20 [71–74] to be involved in precursor protein binding [75]. Additionally, a chaperone-like function of Tom40 has been demonstrated [13], which might be fulfilled by the polar slide. Alternatively, the polar slide might provide a surface, which thermodynamically favors movement of the translocating polypeptide chain relative to the pore.

Remarkably and at the first glance contradicting our model, a truncation mutant of Tom40 from *Rattus norvegicus* lacking the first 165 residues (rTom40-Δ(1–165)) formed a sucrose-permeable pore, assembled into 170 kDa oligomeric complexes *in vitro*, and retained the capability to bind and partially translocate preproteins although to a lesser extent than the wild-type protein [18]. According to our model the first five β-strands are missing in this mutant. However, although half of the β-strands forming the polar slide are absent in the mutant, this particular mutant would result in the formation of an alternative slide assuming a comparable β-strand positioning as in the wild-type protein (Supplementary Fig. S9) as β-strand β19 pairs in a parallel fashion with β6. Consequently, β-strands β18 and β19 would contribute their hydrophilic (labeled '3' in Fig. 4A and C and Supplementary Fig. S8A) and hydrophobic positions to the formation of this alternative slide substituting for some of the lost positions of the polar slide in β2-β5 (Supplementary Fig. S9A). This explains that the protein is still functional, although the first five β-strands are not present. Considering the model by Zeth [25,26] one could not explain the behavior of this mutant, as the polar slide could not be detected in the full-length protein in the first place. The latter is the result of the different β-strand assignments (see Section `*Homology modeling of NcTom40 required a specific alignment process´* ); especially the shifts of the sequence forming β3, β4, and β9 by an uneven number of residues destroyed the structural motif. Another factor is the diameter of the mutant's pore. VDAC has an elliptic minimal cross-section of 27×14 Å with the helix present, but is almost circular without it [23]. Estimating the diameter and cross-section [55] of the mutant – assuming a circular shape of the pore – shows that the diameter decreases to ~20 Å, but the area of the cross-section increases slightly, because of the missing helix. All things considered, the suggested formation of an alternative slide and a sufficiently wide pore in this truncated protein would support the assumption

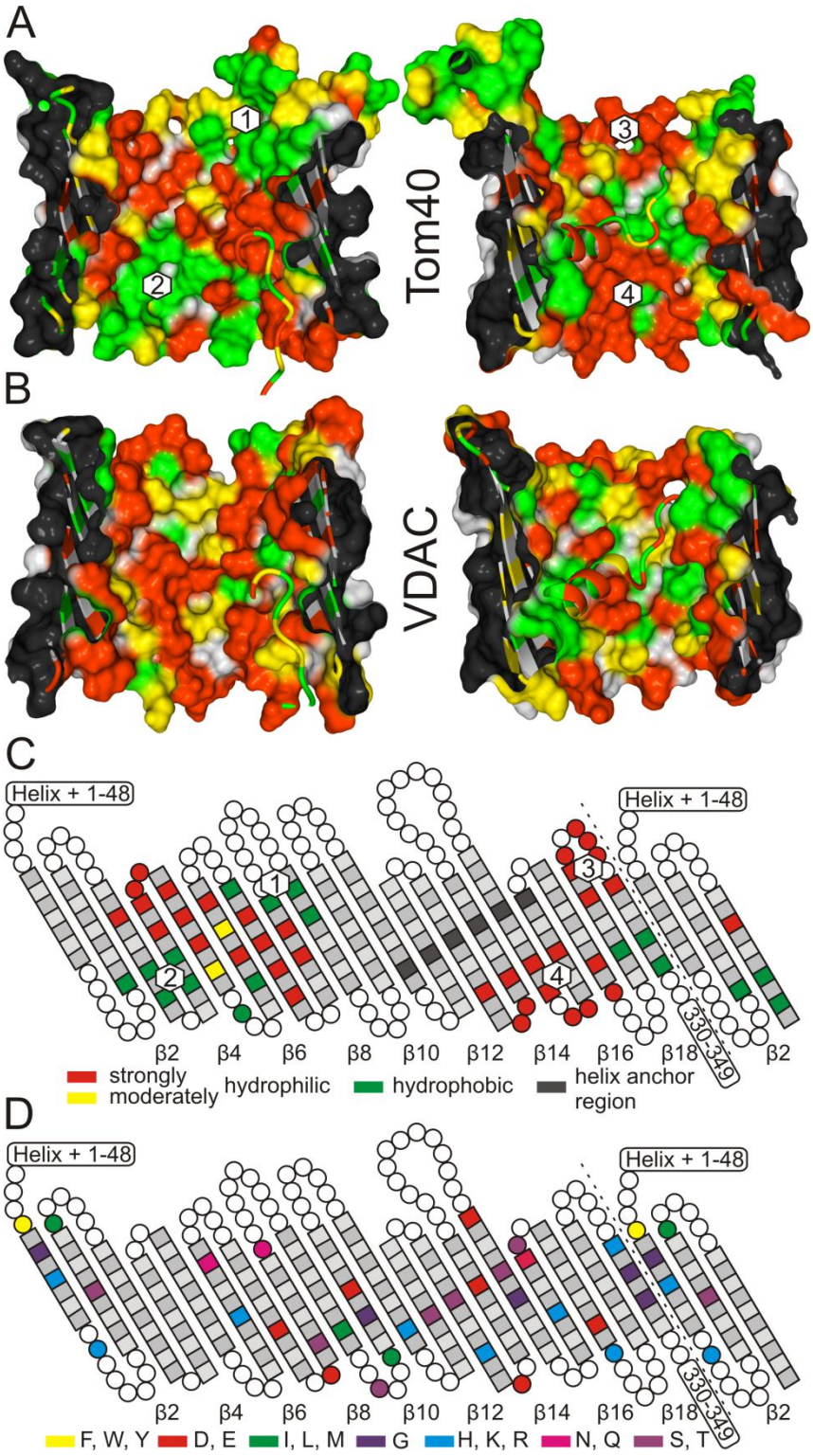



**Fig. 4.** *The polar slide in the pore of Tom40*. (A) The β-barrel of Tom40 was cut open to show in a surface representation (left panel) the polar slide and two hydrophobic patches (labels 1, 2) and the hydrophilic patches (right panel; labels 3, 4). The sequences in the MSA of Tom40 and VDAC were reduced to a maximal sequence identity of 90%. Only Tom40 sequences were used to calculate the average hydrophobicity for each column in this MSA and mapped onto the surface of the homology model of NcTom40. Red marks on average strongly hydrophilic

that the polar slide is involved in precursor protein recognition or translocation, and explains that the affinity for precursor proteins is only reduced but not lost in the mutant.

In agreement with these results, former studies revealed that the inner wall of the Tom40 pore has the ability to interact with an unfolded or loosely folded segment exposed in the translocating protein through hydrophobic interactions [13]. Further, except for the two hydrophobic patches lining the polar slide (Fig. 4A and C), the pore interior, including the α-helix, is mostly hydrophilic. This is in good agreement with single channel analysis on Tom40, revealing it to expose a hydrophilic interior [7]. We suggest that the interaction or binding site for preproteins [7,12,18,24] inside the Tom40 pore is formed by the polar slide (Fig. 4).

---

**(*Fig. 4. continued*)** positions (− 4.5 to − 1.6), yellow an average medium hydrophilicity (− 1.6 to − 0.4), and green marks on average hydrophobic positions (− 0.4 to 4.5). In (B) the crystal structure of murine VDAC is shown in the same orientation as NcTom40 in (A), and is also colored by the average hydrophobicity. In (A) and (B) the structure on the right is rotated clockwise by 120° around the Y axis with respect to the structure on the left side. In (C) and (D) a schematic representation of the NcTom40 homology model is shown. Medium gray boxes indicate residues exposed to the membrane, light gray boxes indicate residues facing the pore interior. Loop residues are indicated by white circles. Residues of the helix anchor region are colored in dark gray. The residues at the beginning and end of an indicated β-strand (boxes) are as follows: β1 S63-V69, β2 L78-A86, β3 P93-T102, β4 Q104-D112, β5 G115-Y123, β6 R128-I137, β7 M143-G152, β8 F155-I162, β9 G173-Q181, β10 G188-Q196, β11 D205-Y214, β12 D218-A227, β13 A230-T240, β14 V243-L251, β15 Q266-D276, β16 S280-D287, β17 G290-K298, β18 T306-D314, β19 K321-S327. In (C) residues assembling the polar slide between the two hydrophobic patches (labels 1, 2), and residues of the two hydrophilic patches (labels 3, 4) are colored by the average hydrophobicity, which was calculated as described in (A). Red indicates strongly and yellow moderately hydrophilic positions, whereas green marks hydrophobic positions; dark gray indicates the positions belonging to the helix anchor region shown in (D). In (D) residues are colored by conservation of properties/amino acid groups as in Fig. 3A; the color code is given in the figure and in Supplementary Table S4.

**Membrane interface of the β-sheet**

Tom40 is a component of the TOM complex. Electron microscopy revealed that large portions of the channel exterior are covered by interactions with other proteins in the complex [10,33,76]. Thus, binding regions on the outer surface of the pore should be discernible by a stronger conservation of residues in comparison to lipid-exposed regions due to the existing evolutionary pressure enforced by protein–protein interactions [77]. Indeed, two regions on the channel surface attracted our attention with respect to conserved patches.

The first region, termed region I, is located in β-strands β1 to β3, where seven positions show conserved amino acid properties (Fig. 5). Four of them are aromatic (F79, Y94, F96, Y100), two are positively charged (K71, H83) and one a glycine (G115). G115 seems only to be conserved to form the sharp β-turn between β-strands β4 and β5. The conserved patch spans the whole membrane. The conserved, charged residue (H83) is present in Tom40 (Fig. 5) whereas in VDAC this position is not conserved. Replacement of amino acids S82, H83 and Q84 of NcTom40 with alanine, respectively, resulted in a severely destabilized TOM complex in mild detergent solution in comparison to wild type TOM [48]. In addition, this mutant of NcTom40 accumulated into a 250 kDa intermediate rather than into the wild type TOM complex with a molecular mass of 400 kDa. Thus, region I most likely represents either a Tom40 dimerization surface or an interaction site for the helical region of other Tom subunits. In line with this notion, the corresponding region in VDAC is directly adjacent to the proposed homo-dimerization interface of hVDAC-1 [59], which is highlighted in Fig. 5.

The second region, termed region II, was identified in β-strands β8 to β11 (Fig. 5). These conserved positions do not span the whole membrane but concentrate in a region reaching from roughly the bilayer midplane to the intermembrane space. This region is characterized by a polar/charged patch, which is formed by residues D154, Q181 and T184. In *N. crassa*, deletion of amino acid T184 and its neighboring residues V183 and P185 has been described to result in a significant TOM assembly defect. Tom40 accumulated in an intermediate below 100 kDa [48]. The same result was observed when these three residues were replaced by alanine, respectively. It can be suggested that not only region I but also region II plays a role in the TOM complex assembly.

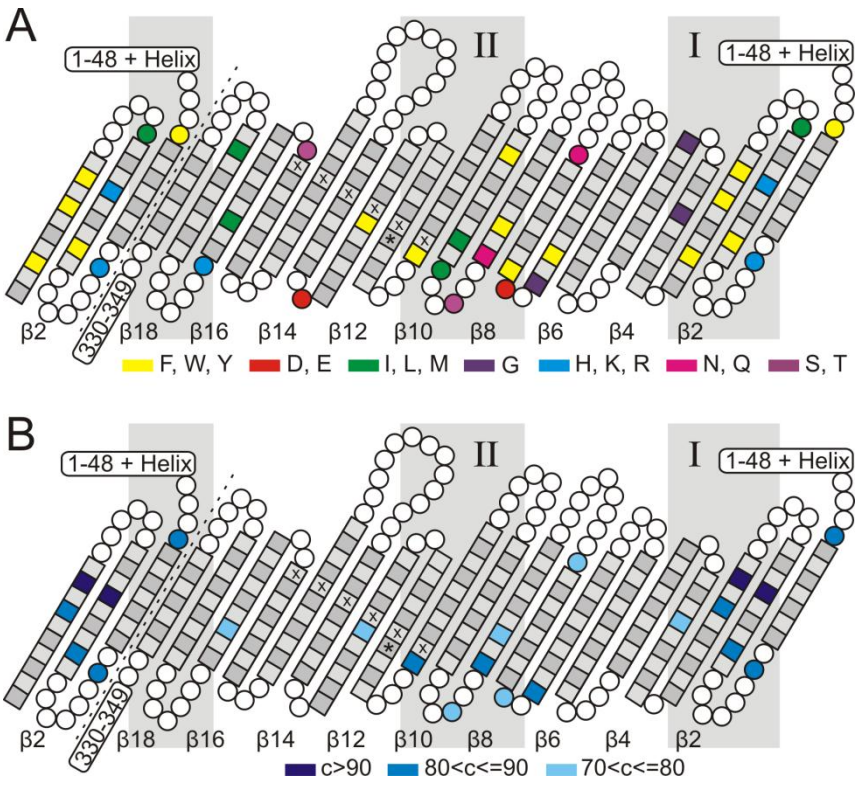



**Fig. 5.** *Putative binding sites on the β-barrel surface of Tom40*. Shown are conserved regions on the membrane-exposed surface of a schematic representation of the β-barrel of Tom40. Medium gray boxes indicate residues exposed to the membrane, light gray boxes indicate residues facing the pore interior. Loop residues are indicated by white circles. Residues of the helix anchor region are colored in dark gray. The residues at the beginning and end of an indicated β-strand (boxes) are as follows: β1 S63-V69, β2 L78-A86, β3 P93-T102, β4 Q104-D112, β5 G115-Y123, β6 R128-I137, β7 M143-G152, β8 F155-I162, β9 G173-Q181, β10 G188-Q196, β11 D205-Y214, β12 D218-A227, β13 A230-T240, β14 V243-L251, β15 Q266-D276, β16 S280-D287, β17 G290-K298, β18 T306-D314, β19 K321-S327. Residues of the helix anchor region are marked by crosses and position W219 by an asterisk. Regions of the β-barrel highlighted in gray mark potential interactions surfaces. The unlabeled gray region on the left in (A) and (B) corresponds to the proposed homodimerisation interface of VDAC. The sequences in the MSA of Tom40 and VDAC were reduced to a maximal sequence identity of 90%. In (A) residues are colored according to their conserved properties. A residue is colored only, if the level of conservation is at least 75%. The amino acid groups and the assigned color are indicated. In (B) residues are colored according to the degree of amino acid conservation. The darker the blue, the more conserved the respective amino acid.

Next to the polar/charged patch in region II is a small patch of hydrophobic residues with long side chains (L187, L189). Both patches are bordered by aromatic positions in β-strands β8, β9 and β11 (Fig. 5). C-terminal of region II between the strongly conserved Y214 in β11 and F235 in β13 (Fig. 5) is the moderately conserved W219, for which a mutant was generated in Tom40 of *Saccharomyces cerevisiae* (W243R [17]). This mutation results in an inefficient transfer of precursor proteins to the Tim23 complex. A similar phenotype was observed for the deletion of the intermembrane space domain of Tom22 [78–81]. However, the transfer of the precursor proteins to the Tim22 complex is not affected by this mutation of Tom40, which means that the interaction of the TOM complex with the tiny Tims is not affected. Thus, the effect of the W243R mutant of ScTom40 might be interpreted as a disturbance of the interaction of Tim21 with the TOM complex [80,81]. Therefore it is tempting to suggest that NcW219 and the conserved amino acids in the neighboring β-strands β11 (NcY214) and β13 (NcF235, Fig. 5) are required to coordinate the interaction of Tom22 and Tim21.

Conserved positions in β-strands β8 to β13 of Tom40 occur only in the region of the intermembrane space leaflet of the membrane (Fig. 5). The highest concentration of conserved positions on the outside of the β-barrel is located in β-strands β8 to β10, which is very intriguing as this conservation occurs where the helix anchor region and the "polar slide" meet on the inside of the barrel (Fig. 4). A logical assumption is that with this region probably component(s) interact, which provide the *trans*-binding region of the TOM complex, guiding the precursor further along its intended import route.

## Conclusion

We have presented a detailed analysis of the general properties of the protein-conducting channel of the TOM complex based on a homology model of Tom40 from *Neurospora crassa*. The model is distinct from earlier proposals and well supported by bioinformatic evaluation and mass spectrometry-coupled limited proteolysis. Transmembrane β-strands in previously published models are often shifted by a few residues relative to our assignment. Especially shifts of an uneven number of residues are problematic, because they hinder the detection of structural motifs. Based on the multiple sequence alignment and

our model we have identified several conserved regions on the in- and outside of the pore, which are probably involved in protein translocation and/or assembly of the TOM complex. The most peculiar features of Tom40 are its HELIX ANCHOR REGION of six conserved polar/charged residues and the POLAR SLIDE, which clearly discriminates Tom40 from VDAC. The helix anchor region is most likely involved in the positioning of helix α1, whereas the polar slide might play a key role in recognition of a presequence in particular or in precursor translocation in general. All in all, we provide a structural basis for Tom40 and present feasible explanations for various biochemical and biophysical data on the protein translocation channel of mitochondria. Thus, our results open the possibility of a structure-guided investigation of Tom40.

## Acknowledgements

The authors are grateful to Beate Nitschke for expert technical assistance, Dr. Frank Nargang for providing the Tom40 *N. crassa* expression strain GR-107 and Dr. Arndt von Haeseler for critical discussions regarding our multiple sequence alignment procedure. Nadine Flinner was funded by a fellowship from the Center of Membrane Proteomics. The work was funded in part by grants of the Competence Network on Functional Nanostructures of the Baden-Württemberg Stiftung to SN and by the Deutsche Forschungsgemeinschaft in the frame of the Excellence Macromolecular Complexes and of the SFB807 project 17 to ES.

## Contribution to this manuscript

DG suggested modeling of Tom40 based on the solved crystal/NMR structure of mammalian VDAC-1 and validation of the NcTom40 structure model by limited proteolysis combined with mass spectrometry of detergent-embedded NcTom40. AS and DG conducted growth of *Neurospora crassa* strain GR-107, isolation of mitochondria from *N. crassa* strain GR-107 and the purification of Tom40 from respective isolated mitochondria. DG carried out circular dichroism (CD) and electrophysiological measurements of NcTom40 and analysis of the regarding data sets, respectively. DG screened the biochemical conditions for the proteolysis of NcTom40. JP performed proteolysis experiments on NcTom40 samples to be subjected to mass

spectrometry immediately. JP executed mass spectrometric analysis and generated the MALDI-TOF spectral graphs. OM and NF performed bioinformatics analyses. ES, SN, DG and OM interpreted the structural model of NcTom40 and the identified conserved regions in the ensemble of Tom40 sequences. JP, DG, OM and NF generated the figures. JP, NF, ES, SN, DG and OM wrote the manuscript. ES supervised bioinformatics and SN biochemical aspects of the project. ES, SN and OM oversaw and outlined the project.

## References


[1] G. Schatz, B. Dobberstein, Common principles of protein translocation across membranes, Science 271 (1996) 1519-1526.

[2] W. Wickner, R. Schekman, Protein translocation across biological membranes, Science 310 (2005) 1452-1456.

[3] W. Neupert, J.M. Herrmann, Translocation of proteins into mitochondria, Annual Review of Biochemistry 76 (2007) 6.1-6.27.

[4] E. Schleiff, T. Becker, Common ground for protein translocation: access control for mitochondria and chloroplasts, Nat. Rev. Mol. Cell Biol. 12 (2011) 48-59.

[5] A. Chacinska, C.M. Koehler, D. Milenkovic, T. Lithgow, N. Pfanner, Importing mitochondrial proteins: machineries and mechanisms, Cell 138 (2009) 628-644.

[6] T. Endo, K. Yamano, S. Kawano, Structural insight into the mitochondrial protein import system, Biochim. Biophys. Acta 1808 (2011) 955-970.

[7] K. Hill, K. Model, M.T. Ryan, K. Dietmeier, F. Martin, R. Wagner, N. Pfanner, Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins, Nature 395 (1998) 516-521.

[8] K.P. Künkele, S. Heins, M. Dembowski, F.E. Nargang, R. Benz, M. Thieffry, J. Walz, R. Lill, S. Nussberger, W. Neupert, The preprotein translocation channel of the outer membrane of mitochondria, Cell 93 (1998) 1009-1019.

[9] H. Suzuki, Y. Okazawa, T. Komiya, K. Saeki, E. Mekada, S. Kitada, A. Ito, K. Mihara, Characterization of Rat TOM40, a Central Component of the Preprotein Translocase of the Mitochondrial Outer Membrane, J. Biol. Chem. 275 (2000) 37930-37936.

[10] U. Ahting, M. Thieffry, H. Engelhardt, R. Hegerl, W. Neupert, S. Nussberger, Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria, J. Cell Biol. 153 (2001) 1151-1160.

[11] C. Meisinger, M.T. Ryan, K. Hill, K. Model, J.H. Lim, A. Sickmann, H. Muller, H.E. Meyer, R. Wagner, N. Pfanner, Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors, Mol. Cell Biol. 21 (2001) 2337-2348.

[12] D. Rapaport, W. Neupert, R. Lill, Mitochondrial protein import. Tom40 plays a major role in targeting and translocation of preproteins by forming a specific binding site for the presequence, J. Biol. Chem. 272 (1997) 18725-18731.

[13] M. Esaki, T. Kanamori, S. Nishikawa, I. Shin, P.G. Schultz, T. Endo, Tom40 protein import channel binds to non-native proteins and prevents their aggregation, Nat. Struct. Biol. 10 (2003) 988-994.

[14] M. Schwartz, A. Matouschek, The dimensions of the protein import channels in the outer and inner mitochondrial membranes, Proc Natl Acad Sci U S A 9 (1999) 13086-13090.

[15] M. Schwartz, S. Huang, A. Matouschek, The structure of precursor proteins during import into mitochondria, J. Biol. Chem. 274 (1999) 12759-12764.

[16] R.D. Taylor, B.J. McHale, F.E. Nargang, Characterization of *Neurospora crassa* Tom40-deficient mutants and effect of specific mutations on Tom40 assembly, J. Biol. Chem. 278 (2003) 765-775.

[17] K. Gabriel, B. Egan, T. Lithgow, Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins, Embo J. 22 (2003) 2380-2386.

[18] H. Suzuki, T. Kadowaki, M. Maeda, H. Sasaki, J. Nabekura, M. Sakaguchi, K. Mihara, Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched beta-structure, J. Biol. Chem. 279 (2004) 50619-50629.

[19] L. Becker, M. Bannwarth, C. Meisinger, K. Hill, K. Model, T. Krimmer, R. Casadio, K.N. Truscott, G.E. Schulz, N. Pfanner, R. Wagner, Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore, J. Mol. Biol. 353 (2005) 1011-1020.

[20] F. Mager, D. Gessmann, S. Nussberger, K. Zeth, Functional refolding and characterization of two Tom40 isoforms from human mitochondria, J. Membr. Biol. 242 (2011) 11-21.

[21] D.A. Court, R. Lill, W. Neupert, The protein import apparatus of the mitochondrial outer membrane, Can. J. Bot. 73 (1995) 193-197.

[22] I.L. Rivera, G.C. Shore, E. Schleiff, Cloning and characterization of a 35-kDa mouse mitochondrial outer membrane protein MOM35 with high homology to Tom40, J. Bioenerg. Biomembr. 32 (2000) 111-121.

[23] R. Ujwal, D. Cascio, J.P. Colletier, S. Faham, J. Zhang, L. Toro, P. Ping, J. Abramson, The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating, Proc. Natl. Acad. Sci. U S A 105 (2008) 17742-17747.

[24] A. Harsman, V. Kruger, P. Bartsch, A. Honigmann, O. Schmidt, S. Rao, C. Meisinger, R. Wagner, Protein conducting nanopores, J. Phys. Condens. Matter 22 (2010) 454102.

[25] K. Zeth, Structure and evolution of mitochondrial outer membrane proteins of beta-barrel topology, Biochim. Biophys. Acta 1797 (2010) 1292-1299.

[26] K. Zeth, M. Thein, Porins in prokaryotes and eukaryotes: common themes and variations, Biochem. J. 431 (2010) 13-22.

[27] T. Kanamori, S. Nishikawa, I. Shin, P. Schultz, E. T, Probing the environment along the protein import pathways in yeast mitochondria by site-specific photocrosslinking, Proc. Natl. Acad. Sci. U S A (1997) 485-490.

[28] S. Walter, J. Buchner, Molecular chaperones - cellular machines for protein folding, Angew. Chem. Int. Ed. 41 (2002) 1098-1113.

[29] P. Juin, M. Thieffry, J.P. Henry, F.M. Vallette, Relationship between the peptide-sensitive channel and the mitochondrial outer membrane protein translocation machinery, J. Biol. Chem. 272 (1997) 6044-6050.

[30] C. Muro, S.M. Grigoriev, D. Pietkiewicz, K.W. Kinnally, M.L. Campo, Comparison of the TIM and TOM channel activities of the mitochondrial protein import complexes, Biophys. J. 84 (2003) 2981-2989.

[31] M. Poynor, R. Eckert, S. Nussberger, Dynamics of the Preprotein Translocation Channel of the Outer Membrane of Mitochondria, Biophys. J. 95 (2008) 1511-1522.

[32] M. Romero-Ruiz, K.R. Mahendran, R. Eckert, M. Winterhalter, S. Nussberger, Interactions of mitochondrial presequence peptides with the mitochondrial outer membrane preprotein translocase TOM, Biophys. J. 99 (2010) 774-781.

[33] U. Ahting, C. Thun, R. Hegerl, D. Typke, F.E. Nargang, W. Neupert, S. Nussberger, The TOM core complex: The general protein import pore of the outer membrane of mitochondria, J. Cell Biol. 147 (1999) 959-968.

[34] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExPASy server, Methods in molecular biology (Clifton, NJ) 112 (1999) 531-552.

[35] H.C. Gasteiger E., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., Protein Identification and Analysis Tools on the ExPASy Server, The Proteomics Protocols Handbook (2005) 571-607

[36] C.N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, Protein Sci. 4 (1995) 2411-2423.

[37] H. Engelhardt, T. Meins, M. Poynor, V. Adams, S. Nussberger, W. Welte, K. Zeth, High level expression, refolding and probing the natural fold of the human voltage-dependent anion channel isoforms I and II, J. Membrane Biol. 216 (2007) 93-105.

[38] D.A. Benson, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank, Nucleic Acids Res. 39 (2011) D32-37.

[39] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389-3402.

[40] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22 (2006) 1658-1659.

[41] K. Katoh, K. Kuma, H. Toh, T. Miyata, MAFFT version 5: improvement in accuracy of multiple sequence alignment, Nucleic Acids Res. 33 (2005) 511-518.

[42] K. Katoh, H. Toh, Parallelization of the MAFFT multiple sequence alignment program, Bioinformatics 26 (2010) 1899-1900.

[43] S. Henikoff, J. Henikoff, Amino acid substitution matrices from protein blocks, Proc. Natl. Acad. Sci. U S A 89 (1992) 10915-10919.

[44] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, Comput. Appl. Biosci. 8 (1992) 275-282.

[45] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779-815.

[46] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using Modeller, Curr. Protoc. Bioinformatics Chapter 5 (2006) Unit 5 6.

[47] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105-132.

[48] E.L. Sherman, R.D. Taylor, N.E. Go, F.E. Nargang, Effect of mutations in Tom40 on stability of the translocase of the outer mitochondrial membrane (TOM) complex, assembly of Tom40, and import of mitochondrial preproteins, J. Biol. Chem. 281 (2006) 22554-22565.

[49] D. Rapaport, R.D. Taylor, M. Kaser, T. Langer, W. Neupert, F.E. Nargang, Structural requirements of Tom40 for assembly into preexisting TOM complexes of mitochondria, Mol. Biol. Cell 12 (2001) 1189-1198.

[50] F. Mager, L. Sokolova, J. Lintzel, B. Brutschy, S. Nussberger, LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM, J. Phys. Condens. Matter 22 (2010) 454132.

[51] M. Gribskov, A.D. McLachlan, D. Eisenberg, Profile analysis: detection of distantly related proteins, Proc. Natl. Acad. Sci. U S A 84 (1987) 4355-4358.

[52] H. Zaid, S. Abu-Hamad, A. Israelson, I. Nathan, V. Shoshan-Barmatz, The voltage-dependent anion channel-1 modulates apoptotic cell death, Cell Death Differ. 12 (2005) 751-760.

[53] R. Bredemeier, T. Schlegel, F. Ertel, A. Vojta, L. Borissenko, M. Bohnsack, M. Groll, A. von Haeseler, S. E, Functional and phylogenetic properties of the pore-forming beta-barrel transporters of the Omp85 family, J. Biol. Chem. 282 (2007) 1882-1890.

[54] M. Pusnik, F. Charriere, P. Maser, R.F. Waller, M.J. Dagley, T. Lithgow, A. Schneider, The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane, Mol. Biol. Evol. 26 (2009) 671-680.

[55] G.E. Schulz, The structure of bacterial outer membrane proteins, Biochim. Biophys. Acta 1565 (2002) 308-317.

[56] W.C. Wimley, The versatile beta-barrel membrane protein, Curr. Opin. Struct. Biol. 13 (2003) 404-411.

[57] S. Kutik, D. Stojanovski, L. Becker, T. Becker, M. Meinecke, V. Kruger, C. Prinz, C. Meisinger, B. Guiard, R. Wagner, N. Pfanner, N. Wiedemann, Dissecting membrane insertion of mitochondrial beta-barrel proteins, Cell 132 (2008) 1011-1024.

[58] K.P. Künkele, P. Juin, C. Pompa, F.E. Nargang, J.P. Henry, W. Neupert, R. Lill, M. Thieffry, The isolated complex of the translocase of the outer membrane of mitochondria. Characterization of the cation-selective and voltage-gated preprotein-conducting pore, J. Biol. Chem. 273 (1998) 31032-31039.

[59] M. Bayrhuber, T. Meins, M. Habeck, S. Becker, K. Giller, S. Villinger, C. Vonrhein, C. Griesinger, M. Zweckstetter, K. Zeth, Structure of the human voltage-dependent anion channel, Proc. Natl. Acad. Sci. U S A 105 (2008) 15370-15375.

[60] B. Popp, D.A. Court, R. Benz, W. Neupert, R. Lill, The role of the N and C termini of recombinant Neurospora mitochondrial porin in channel formation and voltage-dependent gating, J. Biol. Chem. 271 (1996) 13593-13599.

[61] O.P. Choudhary, R. Ujwal, W. Kowallis, R. Coalson, J. Abramson, M. Grabe, The electrostatics of VDAC: implications for selectivity and gating, J. Mol. Biol. 396 (2010) 580-592.

[62] S. Abu-Hamad, N. Arbel, D. Calo, L. Arzoine, A. Israelson, N. Keinan, R. Ben-Romano, O. Friedman, V. Shoshan-Barmatz, The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins, J. Cell Sci. 122 (2009) 1906-1916.

[63] R.B. Cooley, D.J. Arp, P.A. Karplus, Evolutionary origin of a secondary structure: pi-helices as cryptic but widespread insertional variations of alpha-helices that enhance protein functionality, J. Mol. Biol. 404 (2010) 232-246.

[64] R. Schneider, M. Etzkorn, K. Giller, V. Daebel, J. Eisfeld, M. Zweckstetter, C. Griesinger, S. Becker, A. Lange, The native conformation of the human VDAC1 N terminus, Angew. Chem. Int. Ed. Engl. 49 (2010) 1882-1885.

[65] C.K. Kassenbrock, W. Cao, M.G. Douglas, Genetic and biochemical characterization of ISP6, a small mitochondrial outer membrane protein associated with the protein translocation complex, EMBO J. 12 (1993) 3023-3034.

[66] D.I. Meyer, Mimics – or gimmicks?, Nature 347 (1990) 424-425.

[67] P. Van Gelder, F. Dumas, I. Bartoldus, N. Saint, A. Prilipov, M. Winterhalter, Y. Wang, A. Philippsen, J.P. Rosenbusch, T. Schirmer, Sugar transport through maltoporin of *Escherichia coli*: role of the greasy slide, J. Bacteriol. 184 (2002) 2994-2999.

[68] G. von Heijne, Mitochondrial targeting sequences may form amphiphilic helices, EMBO J. 5 (1986) 1335-1342.

[69] Y. Gavel, L. Nilsson, G. von Heijne, Mitochondrial targeting sequences. Why 'non-amphiphilic' peptides may still be amphiphilic, FEBS Lett. 235 (1988) 173-177.

[70] K.A. Rimmer, J.H. Foo, A. Ng, E.J. Petrie, P.J. Shilling, A.J. Perry, H.D. Mertens, T. Lithgow, T.D. Mulhern, P.R. Gooley, Recognition of mitochondrial targeting sequences by the import receptors Tom20 and Tom22, J. Mol. Biol. 405 (2011) 804-818.

[71] E. Schleiff, G. Shore, I. Goping, Interactions of the human mitochondrial protein import receptor, hTom20, with precursor proteins *in vitro* reveal pleiotropic specificities and different receptor domain requirements, J. Biol. Chem. 272 (1997) 17784-17789.

[72] E. Schleiff, J.L. Turnbull, Functional and structural properties of the mitochondrial outer membrane receptor Tom20, Biochemistry 37 (1998) 13043-13051.

[73] E. Schleiff, J.L. Turnbull, Characterization of the N-terminal targeting signal binding domain of the mitochondrial outer membrane receptor, Tom20, Biochemistry 37 (1998) 13052-13058.

[74] Y. Abe, T. Shodai, T. Muto, K. Mihara, H. Torii, S. Nishikawa, T. Endo, D. Kohda, Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20, Cell 100 (2000) 551-560.

[75] E. Schleiff, G.C. Shore, I.S. Goping, Human mitochondrial import receptor, Tom20p. Use of glutathione to reveal specific interactions between Tom20-glutathione S-transferase and mitochondrial precursor proteins, FEBS Lett. 404 (1997) 314-318.

[76] K. Model, C. Meisinger, W. Kuhlbrandt, Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase, J. Mol. Biol. 383 (2008) 1049-1057.

[77] T. Schlegel, O. Mirus, A. von Haeseler, E. Schleiff, The tetratricopeptide repeats of receptors involved in protein translocation across membranes, Mol. Biol. Evol. 24 (2007) 2763-2774.

[78] D.A. Court, F.E. Nargang, H. Steiner, R.S. Hodges, W. Neupert, R. Lill, Role of the intermembrane-space domain of the preprotein receptor Tom22 in protein import into mitochondria, Mol. Cell Biol. 16 (1996) 4035-4042.

[79] A. Chacinska, P. Rehling, B. Guiard, A. Frazier, A. Schulze-Specking, N. Pfanner, W. Voos, M. C., Mitochondrial translocation contact

sites:separation of dynamic and stabilizing elements in formation of a TOM-TIM-preprotein supercomplex, EMBO J. 22 (2003) 5370-5381.

[80] A. Chacinska, M. Lind, A.E. Frazier, J. Dudek, C. Meisinger, A. Geissler, A. Sickmann, H.E. Meyer, K.N. Truscott, B. Guiard, N. Pfanner, P. Rehling, Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17, Cell 120 (2005) 817-829.

[81] R. Albrecht, P. Rehling, A. Chacinska, J. Brix, S. Cadamuro, R. Volkmer, B. Guiard, N. Pfanner, K. Zeth, The Tim21 binding domain connects the preprotein translocases of both mitochondrial membranes, EMBO Rep. 7 (2006) 1233-1238.

# Supplementary Material

**Supplementary Table S1** Secondary structure content of native Tom40 from *N. crassa* mitochondria

| Method | α-Helix | β-Sheet | Turn | Random coil |
|---|---|---|---|---|
| I | 0.163 | 0.312 | 0.22 | 0.298 |
| II | 0.163 | 0.309 | 0.219 | 0.309 |
| III | 0.144 | 0.281 | 0.263 | 0.271 |

Secondary structure percentages were determined with the CDpro package CDSSTR (I), CONTIN/LL (II) and SELCON3 (III), using the CD data shown in Fig. 2b. The procedure is described in:

Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287, 252-60.

Sreerama, N. & Woody, R. W. (2003). Structural composition of betaI- and betaII-proteins. *Protein Sci* 12, 384-8.

Sreerama, N. & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods Enzymol* 383, 318-51.

**Supplementary Table S2** Genome Projects searched for Tom40 and VDAC sequences

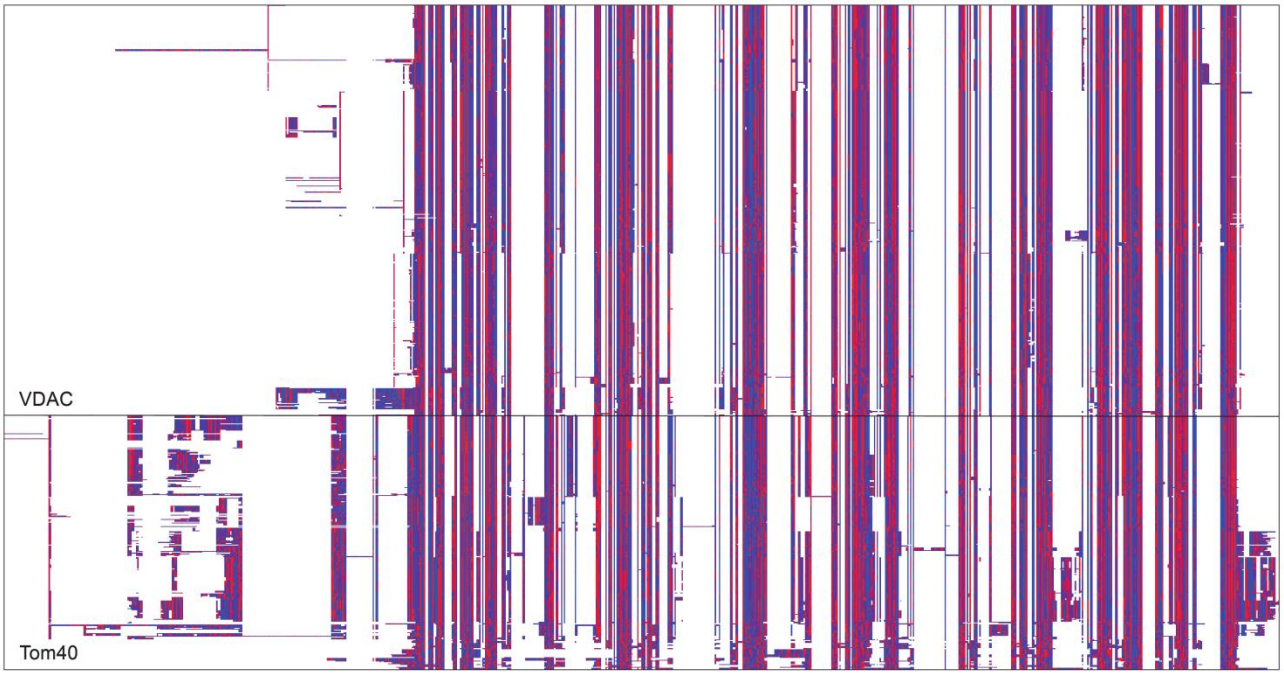
| Organismus | Gruppe | Quelle |
|---|---|---|
| Dictyostelium purpureum | Amoebozoa; Mycetozoa | http://genome.jgi-psf.org/Dicpu1/Dicpu1.home.html |
| Monosiga brevicollis | Choanoflagellida; Codonosigidae | http:// genome.jgi-psf.org/Monbr1/Monbr1.home.html |
| Mucor circinelloides | Fungi; Mucoromycotina | http:// genome.jgi-psf.org/Mucci1/Mucci1.home.html |
| Batrachochytrium dendrobatidis | Fungi; Chytridiomycota | http://genome.jgi-psf.org/Batde5/Batde5.home.html |
| Cochliobolus heterostrophus | Fungi; Dikarya | http://genome.jgi-psf.org/CocheC5_1/CocheC5_1.home.html |
| Cryphonectria parasitica | Fungi; Dikarya | http://genome.jgi-psf.org/Crypa2/Crypa2.home.html |
| Heterobasidion annosum | Fungi; Dikarya | http://genome.jgi-psf.org/Hetan1/Hetan1.home.html |
| Laccaria bicolor | Fungi; Dikarya | http:// genome.jgi-psf.org/Lacbi1 |
| Melampsora laricis-populina | Fungi; Dikarya | http://genome.jgi-psf.org/Mellp1/Mellp1.home.html |
| Mycosphaerella fijiensis | Fungi; Dikarya | http://genome.jgi-psf.org/Mycfi2/Mycfi2.home.html |
| Mycosphaerella graminicola | Fungi; Dikarya | http://genome.jgi-psf.org/Mycgr3/Mycgr3.home.html |
| Nectria haematococca | Fungi; Dikarya | http://genome.jgi-psf.org/Necha2/Necha2.home.html |
| Neurospora discreta | Fungi; Dikarya | http://genome.jgi-psf.org/Neudi1/Neudi1.home.html |
| Neurospora tetrasperma | Fungi; Dikarya | http://genome.jgi-psf.org/Neute_matA2/Neute_matA2.home.html |
| Phanerochaete chrysosporium | Fungi; Dikarya | http://genome.jgi-psf.org/Phchr1/Phchr1.home.html |
| Phanerochaete chrysosporium | Fungi; Dikarya | http://genome.jgi-psf.org/Phchr1/Phchr1.home.html |
| Pleurotus ostreatus | Fungi; Dikarya | http://genome.jgi-psf.org/PleosPC15_1/PleosPC15_1.home.html |
| Postia placenta | Fungi; Dikarya | http://genome.jgi-psf.org/Pospl1/Pospl1.home.html |
| Schizophyllum commune | Fungi; Dikarya | http://genome.jgi-psf.org/Schco1/Schco1.home.html |
| Sporobolomyces roseus | Fungi; Dikarya | http://genome.jgi-psf.org/Sporo1/Sporo1.home.html |
| Sporotrichum thermophile | Fungi; Dikarya | http://genome.jgi-psf.org/Spoth1/Spoth1.home.html |
| Thielavia terrestris | Fungi; Dikarya | http://genome.jgi-psf.org/Thite1/Thite1.home.html |
| Tremella mesenterica | Fungi; Dikarya | http://genome.jgi-psf.org/Treme1/Treme1.home.html |
| Trichoderma atroviride | Fungi; Dikarya | http://genome.jgi-psf.org/Triat2/Triat2.home.html |
| Trichoderma reesei | Fungi; Dikarya | http://genome.jgi-psf.org/Trire2/Trire2.home.html |
| Trichoderma virens | Fungi; Dikarya | http://genome.jgi-psf.org/TriviGv29_8_2/TriviGv29_8_2.home.html |
| Phycomyces blakesleeanus | Fungi; Mucoromycotina | http://genome.jgi-psf.org/Phybl1/Phybl1.home.html |
| Emiliania huxleyi | Haptophyceae; Isochrysidales | http://genome.jgi-psf.org/Emihu1/Emihu1.home.html |
| Capitella teleta | Metazoa; Annelida | http:// genome.jgi-psf.org/capitella |
| Helobdella robusta | Metazoa; Annelida | http:// genome.jgi-psf.org/Helro1/Helro1.home.html |
| Ciona intestinalis | Metazoa; Chordata | http:// genome.jgi-psf.org/ciona |
| Daphnia pulex | Metazoa; Eumetazoa | http://genome.jgi-psf.org/Dappu1/Dappu1.home.html |
| Takifugu rubripes | Metazoa; Eumetazoa | http://genome.jgi-psf.org/Takru4/Takru4.home.html |
| Lottia gigantea | Metazoa; Mollusca | http:// genome.jgi-psf.org/Lotgi1/Lotgi1.home.html |
| Cyanidioschyzon merolae | Rhodophyta; Bangiophyceae | http://merolae.biol.s.u-tokyo.ac.jp/db/cds_view.cgi?locus=CMF062C |
| Galdieria sulphuraria | Rhodophyta; Bangiophyceae | http://genomics.msu.edu/galdieria/ |
| Fragilariopsis cylindrus | Stramenopiles; Bacillariophyta | http:// genome.jgi-psf.org/Fracy1/Fracy1.home.html |
| Phytophthora capsici | Stramenopiles; Oomycetes | http://genome.jgi-psf.org/PhycaF7/PhycaF7.home.html |
| Phytophthora ramorum | Stramenopiles; Oomycetes | http://genome.jgi-psf.org/Phyra1_1/Phyra1_1.home.html |
| Phytophthora sojae | Stramenopiles; Oomycetes | http://genome.jgi-psf.org/Physo1_1/Physo1_1.home.html |
| Aureococcus anophagefferens | Stramenopiles; Pelagophyceae | http:// genome.jgi-psf.org/Auran1/Auran1. home.html |
| Chlorella sp. | Viridiplantae; Chlorophyta | http://genome.jgi-psf.org/ChlNC64A_1/ChlNC64A_1.home.html |
| Chlorella vulgaris | Viridiplantae; Chlorophyta | http://genome.jgi-psf.org/Chlvu1/Chlvu1.home.html |
| Volvox carteri f. nagariensis | Viridiplantae; Chlorophyta | http://genome.jgi-psf.org/Volca1/Volca1.home.html |

**Supplementary Table S3** Amino acid scales used to analyze Tom40 on a residue base

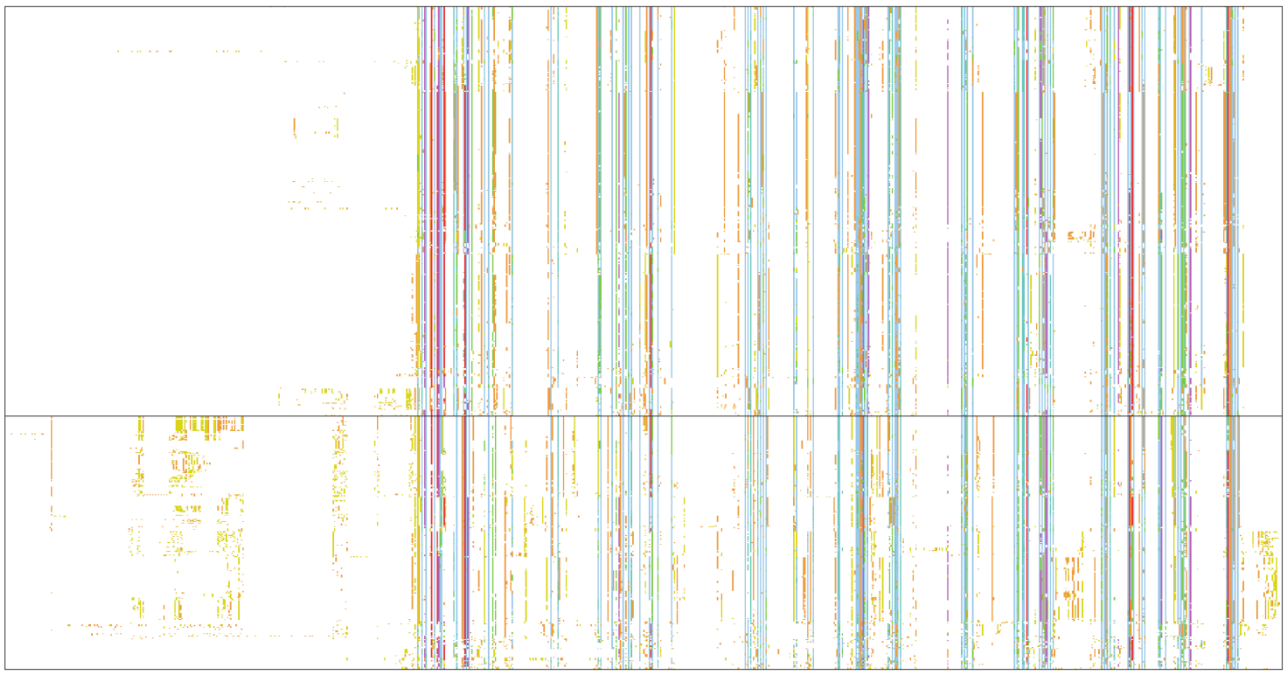
| **scale** | **amino acid** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | F | Y | W | T | G | S | D | N | Q | H |
| K&D | 2,8 | -1,3 | -0,9 | -0,7 | -0,4 | -0,8 | -3,5 | -3,5 | -3,5 | -3,2 |
| size | -5 | -5 | -5 | -2 | 0 | -2 | -3 | -3 | -4 | -4 |
| **scale** | **amino acid** | | | | | | | | | |
| | K | V | I | P | R | A | C | L | M | E |
| K&D | -3,9 | 4,2 | 4,5 | -1,6 | -4,5 | 1,8 | 2,5 | 3,8 | 1,9 | -3,5 |
| size | -5 | -2 | -3 | -2 | -5 | -1 | -2 | -3 | -4 | -4 |

**Supplementary Table S4** Color code for amino acid groups

| **side chain property** | **amino acids in group** | **color** |
|---|---|---|
| large, polar | N, Q | magenta |
| small, polar | C, S, T | violet |
| positively charged | H, K, R | blue |
| negatively charged | D, E | red |
| Aromatic | F, W, Y | yellow |
| small, hydrophobic | A, V | light green |
| large, hydrophobic | I, L, M | green |
| tiny | G | dark violet |
| | P | orange |




**Supplementary Figure S1A**



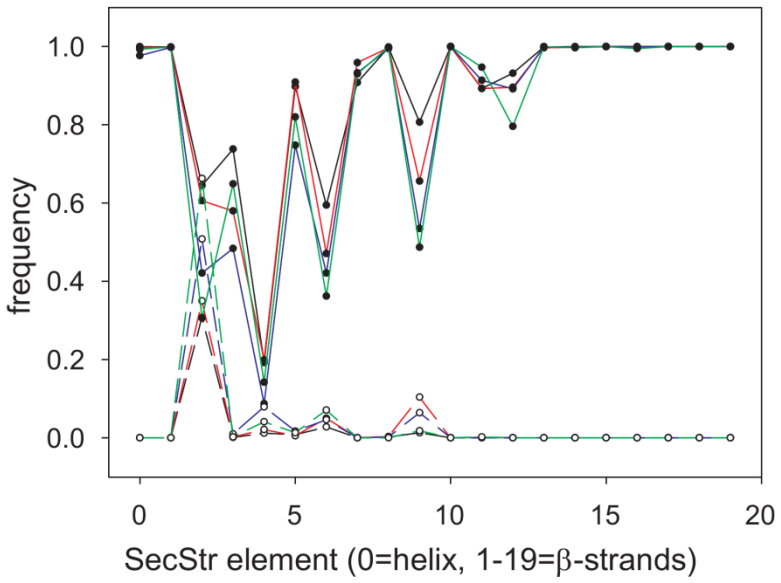
**Supplementary Figure S1B**


**Supplementary Figure S1C**

**Supplementary Figure S1** *Alignment of Tom40 and VDAC*. A compressed version of the multiple sequence alignment (MSA) containing 123 Tom40 and 186 VDAC with a maximum sequence identity of 75% (cd-hit) is shown. Columns occupied by at most 1% of the sequences in the MSA were removed. (A) The MSA is colored with JalView by the hydrophobicity of the residues. Hydrophobic residues are colored in red and hydrophilic residues in blue. (B) The same MSA as in (A) but this time colored by the ClustalX color code (Thompson et al. 1997) as implemented in JalView (Waterhouse et al. 2009) to highlight the conserved positions in both protein families. In more detail, the amino acids are colored according to their degree of conservation in the respective column of the MSA. The more residues of a column are colored with the same color, the stronger the degree of conservation at the respective sequence position. In this representation not all columns visible in (A) are visible here, because their amino acid composition is too diverse to fit into the categories defined at the website http://www.jalview.org/help/html/colourSchemes/clustal.html. But one can clearly see that both Tom40 and VDAC share positions of conserved character. (C) Sequences of VDAC from *Mus musculus*, *Homo sapiens* and of Tom40 from *M. musculus*, *H. sapiens*, *Neurospora crassa*, *Saccharomyces cerevisiae*, and *Arabidopsis thaliana*, were extracted from the alignment shown in Supplementary Fig. S1A, B. Columns with gaps in all five sequences were removed. Residue numbers of the last residue in each line are shown on the right side of the alignment. Please, note that due to the N-terminal His-tag of the sequence of mVDAC-1 one has to subtract 12 from the sequence number to match the residue numbering in Figure 3B. Regions forming α-helices are colored red (VDAC structures) or orange (Tom40), and β-strands are colored green (VDAC structures) or cyan (Tom40). The N-terminal helix and the transmembrane β-strands from the structure of VDAC from *M. musculus* are shown. The β-strands of VDAC are numbered at the first residue of each β-strand. Residues highlighted by blue, green and yellow boxes were detected in here as cutting sites for NcTom40 (see Supplementary Fig. S6). Residues highlighted with a gray box were detected as cutting sites for human VDAC (Engelhardt et al. 2007).

Engelhardt H, Meins T, Poynor M, Adams V, Nussberger S, Welte W, Zeth K. High-level expression, refolding and probing the natural fold of the human voltage-dependent anion channel isoforms I and II. J Membr Biol. 2007 Apr;216(2-3):93-105.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997 Dec 15;25(24):4876-82.

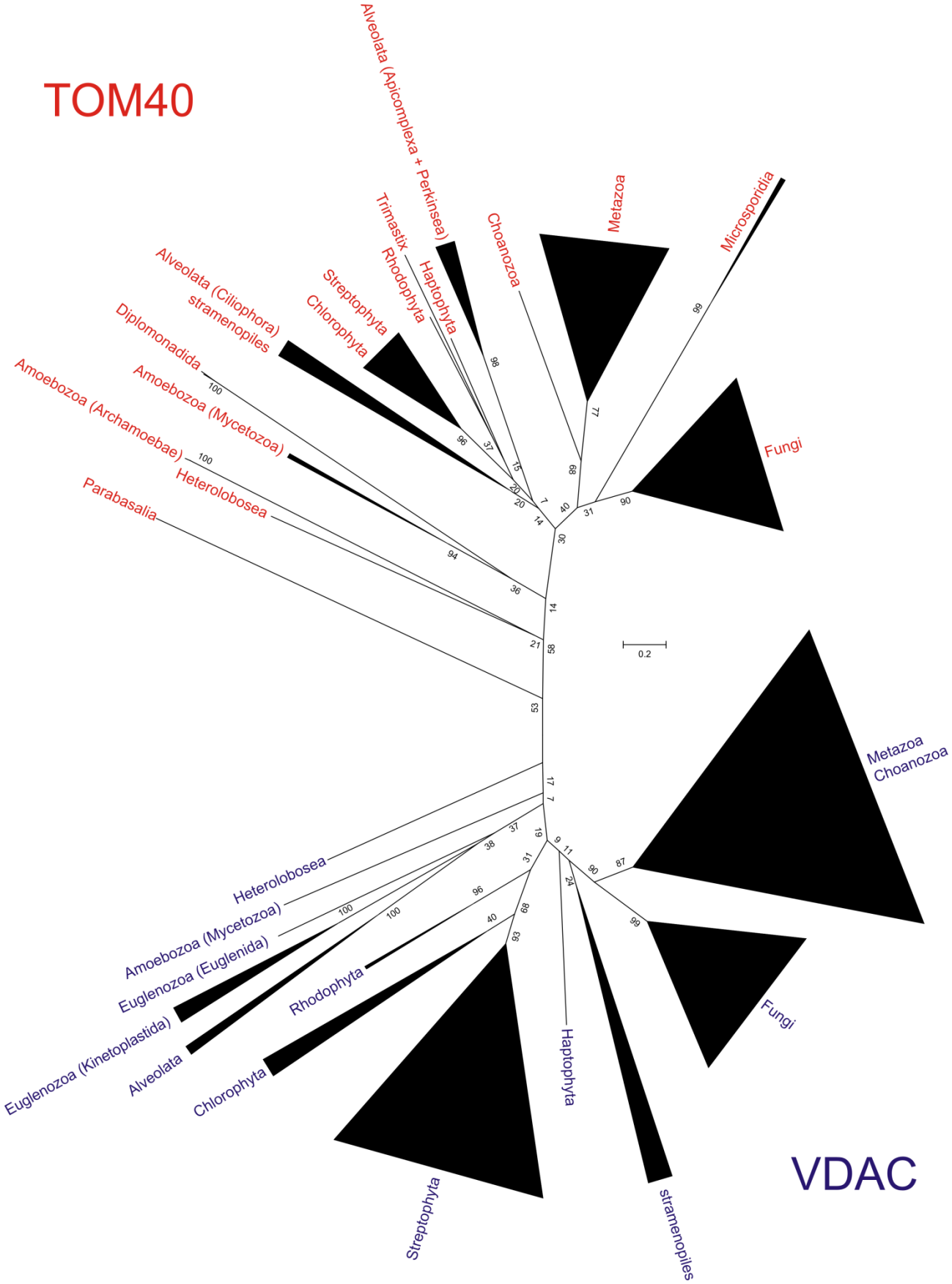



**Supplementary Figure S2** *Assignment of β-strands in Tom40*. The solid black line with filled circles and the dashed black line with open circles refer to the solid black lines with filled black circles and filled gray circles, respectively, in Fig. 1B. In addition we counted how often the most and 2nd most frequently occurring equal alignments of regions of NcTom40 to the helix and the β-strand core regions of mVDAC-1 from the full set (black lines) of Tom40 and VDAC sequences occur in the 90% (red), 80% (blue), and 70% (green) maximal sequence identity sets. In the N-terminal region the frequency of the most frequently occurring β-strand decreases with the maximal allowed sequence identity in the sequence set. Interestingly, at 80% (blue) and 70% (green) the population of the 2nd most frequently occurring equal alignment for β-strand β2 from the full data set (dashed blue and green lines, open circles) is greater than the most frequently occurring equal alignment from the full data set (solid blue and green lines, filled circles).

---

**(*Fig. S1 continued*)**

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009 May 1;25(9):1189-91.

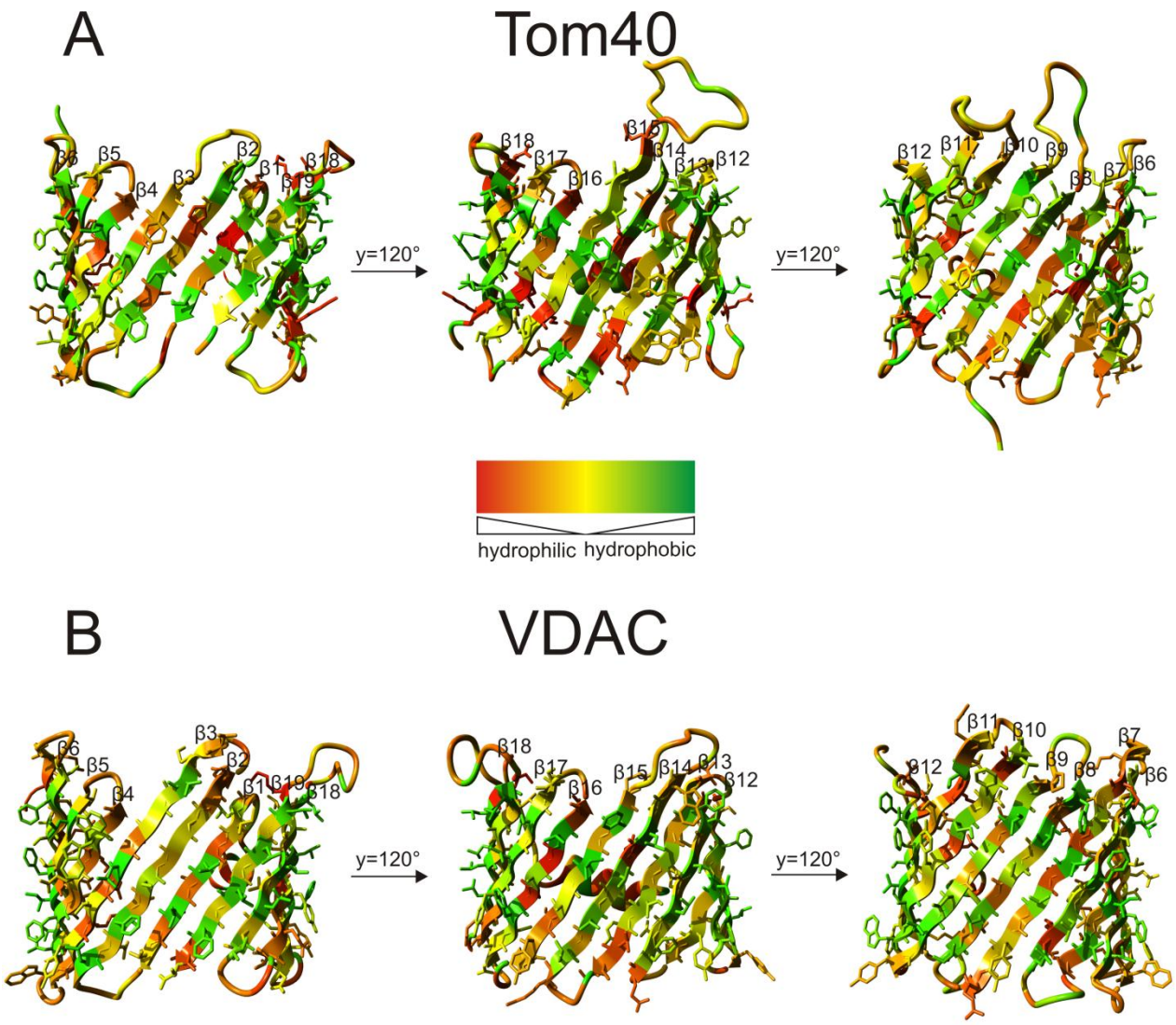

TOM40
Alveolata (Apicomplexa + Perkinsea)
Trimastix
Rhodophyta
Haptophyta
Choanozoa
Metazoa
Microsporidia
Alveolata (Ciliophora)
stramenopiles
Streptophyta
Chlorophyta
Diplomonadida
Amoebozoa (Archamoebae)
Amoebozoa (Mycetozoa)
Heterolobosea
Parabasalia
Fungi
0.2
Metazoa
Choanozoa
Fungi
Heterolobosea
Amoebozoa (Mycetozoa)
Euglenozoa (Euglenida)
Euglenozoa (Kinetoplastida)
Alveolata
Rhodophyta
Chlorophyta
Haptophyta
Streptophyta
stramenopiles
VDAC

**Supplementary Figure S3** *Phylogenetic tree of Tom40 and VDAC*. A maximum likelihood phylogeny of Tom40 and VDAC was reconstructed with RAxML v7.0.4 (Stamatakis 2006) using the WAG model (Whelan and Goldman 2001) and gamma-distributed rate heterogeneity. Branch support values were calculated with RAxML's rapid bootstrap algorithm (Stamatakis et al. 2008). The subtrees were collapsed with MEGA5 (Tamura et al. 2011) to make the tree easier accessible to the reader.

Stamatakis A (2006) "RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic analyses with Thousands of Taxa and Mixed Models", Bioinformatics 22:2688–2690
Stamatakis A, Hoover P, Rougemont J. (2008) A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 57:758-771.

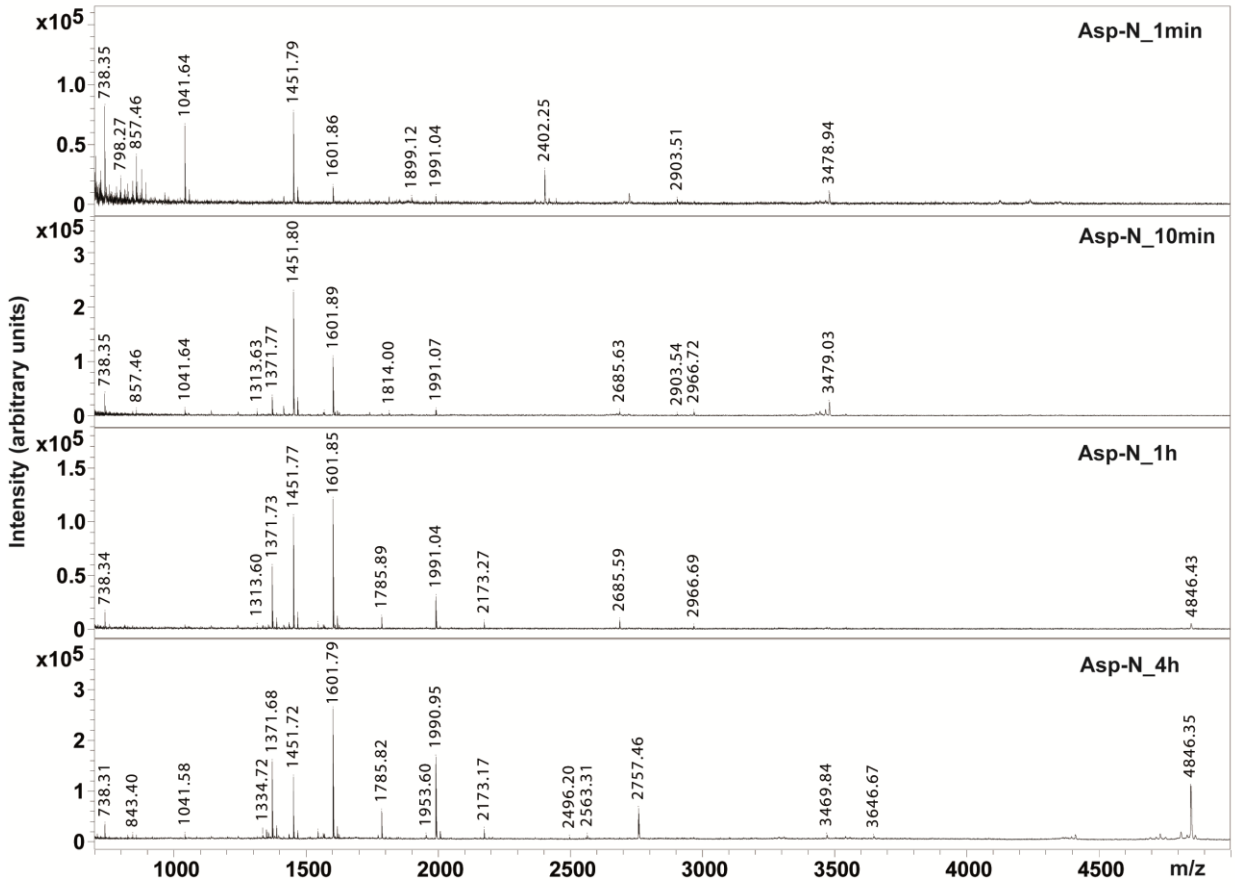
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution (accepted).

Whelan S, Goldman N. 2001 A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. Molecular Biology and Evolution 18:691–699.

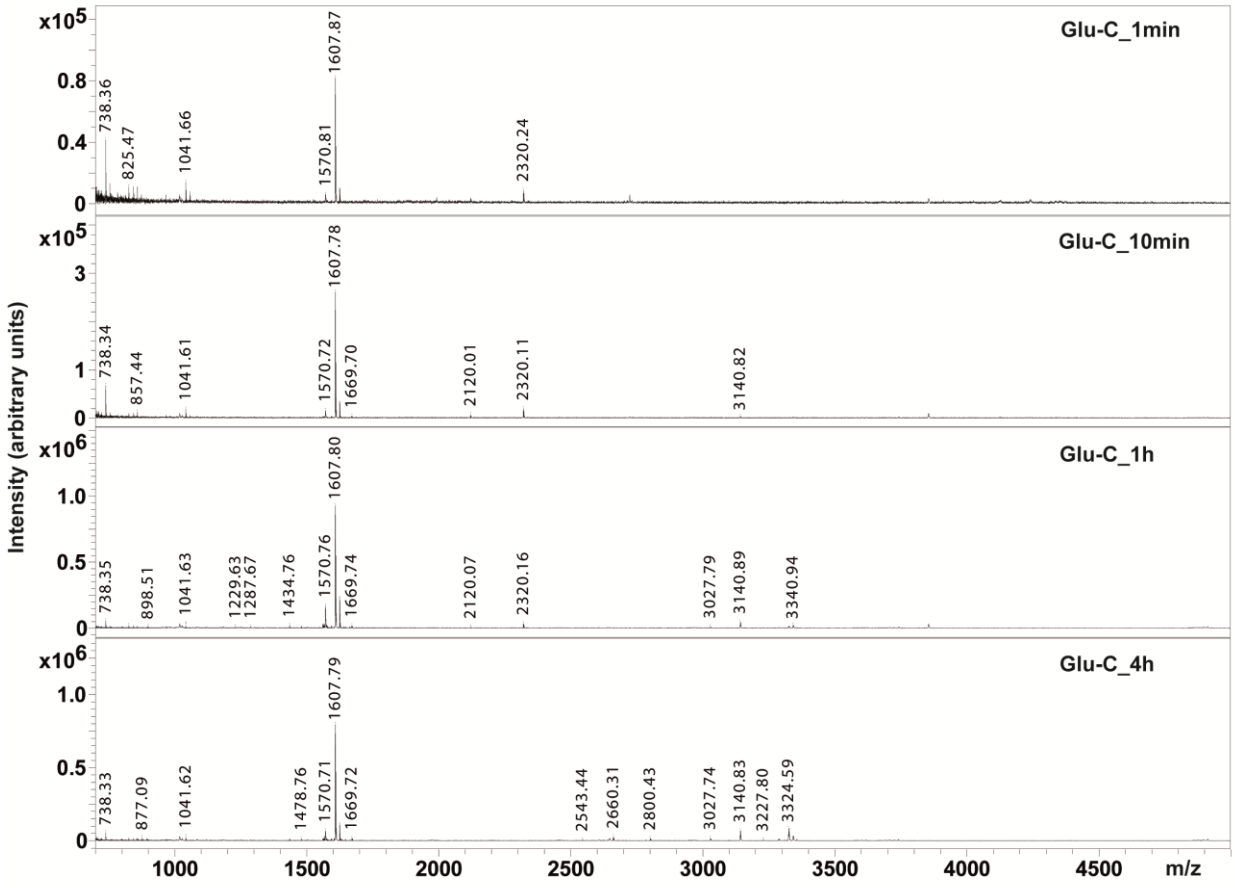



**Supplementary Figure S4** *Hydrophobic properties of the β-barrel of Tom40 and VDAC*
A ribbon representation of (A) the homology model of NcTom40 and (B) VDAC oriented parallel to the membrane is shown. The model is rotated in steps of 120° around the Y axis to show the whole surface of the membrane-embedded β-sheet. The model of NcTom40 (A) and the crystal structure of murine VDAC (B) was colored residue-wise by the average hydrophobicity (Kyte&Doolittle, 1982; Supplementary Table S3) of all residues in the respective column of the multiple sequence alignment of Tom40 and VDAC, respectively, reduced to a maximal sequence identity of 90%.

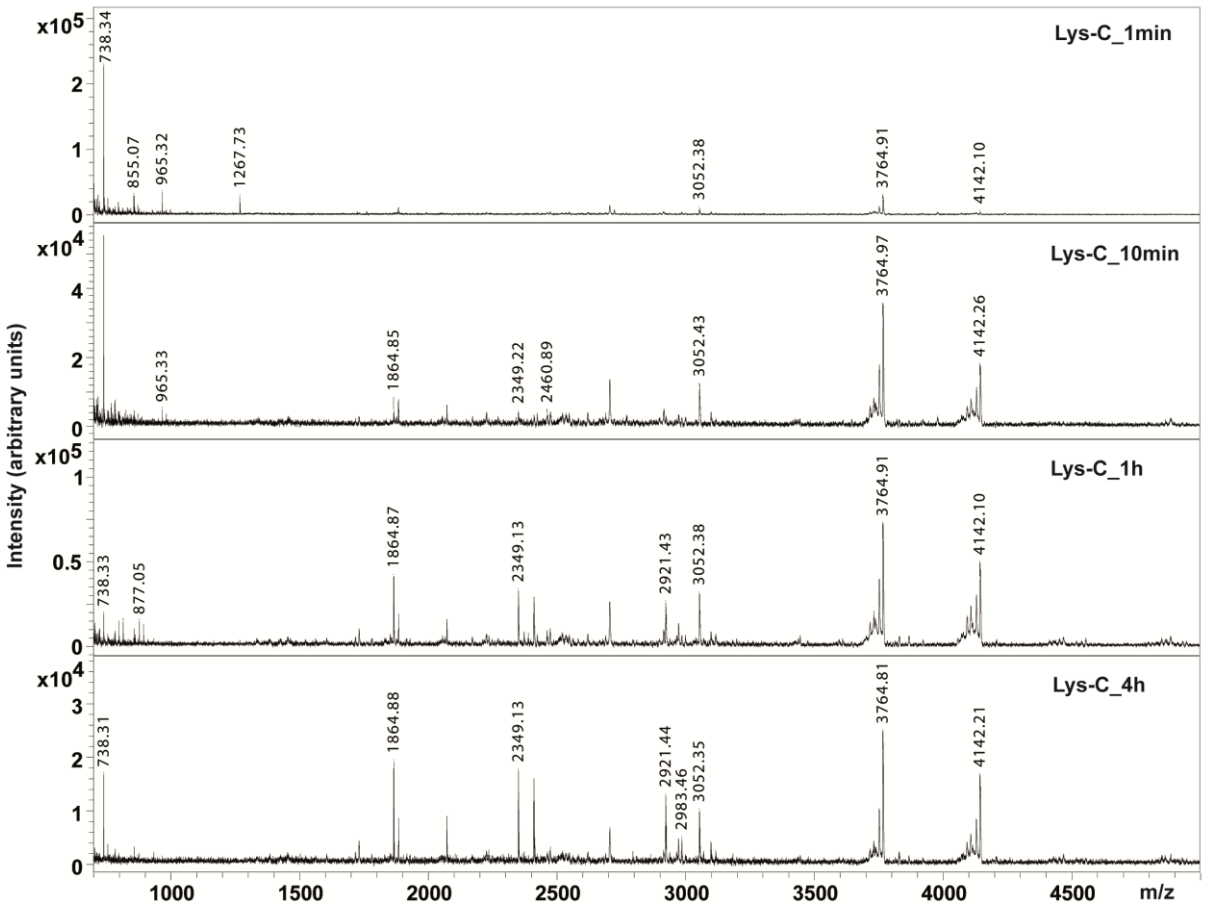
Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982, 157:105-32.



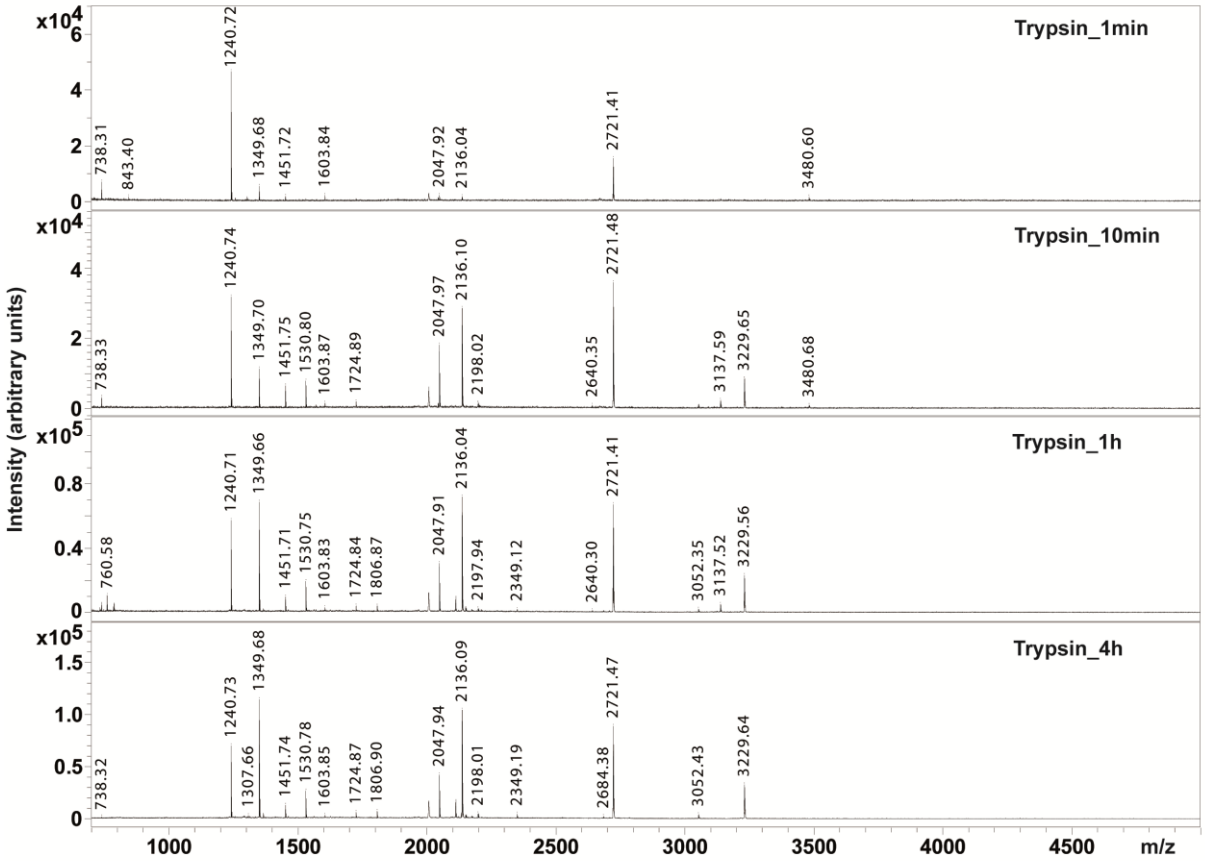

**Supplementary Figure S5A**



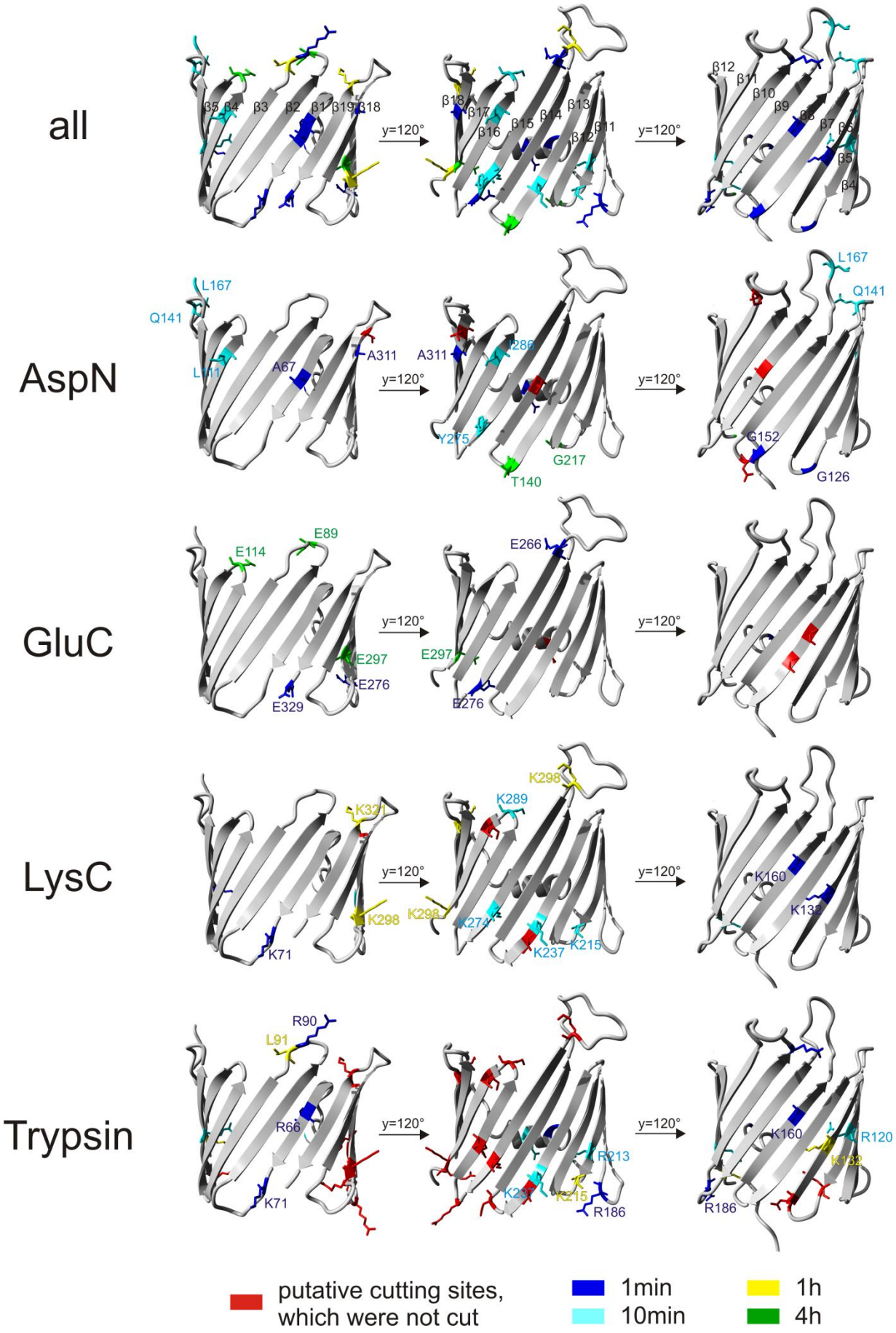

**Supplementary Figure S5B**




**Supplementary Figure S5C**



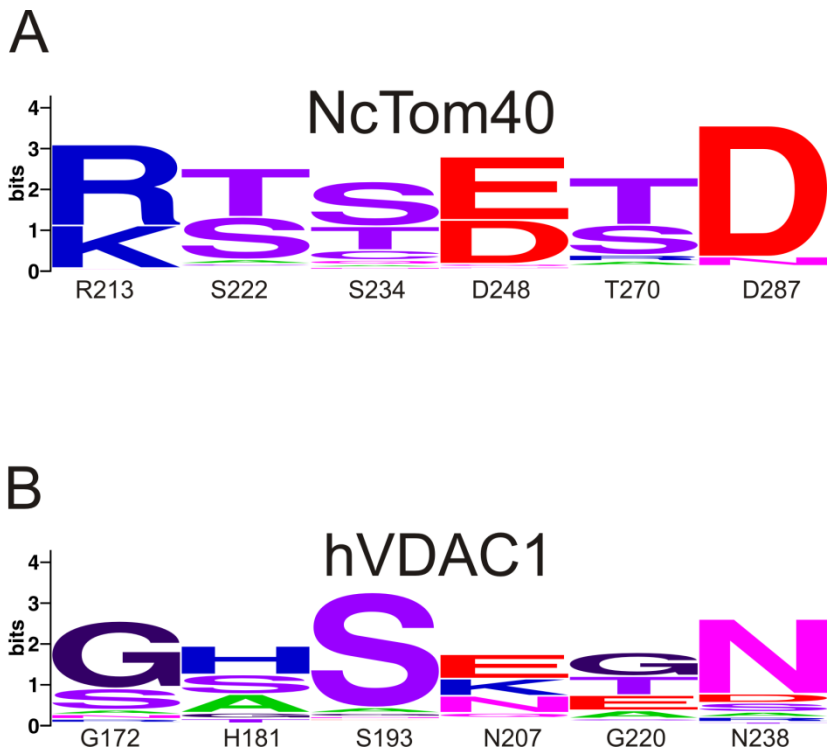

**Supplementary Figure S5D**




**Supplementary Figure S6**

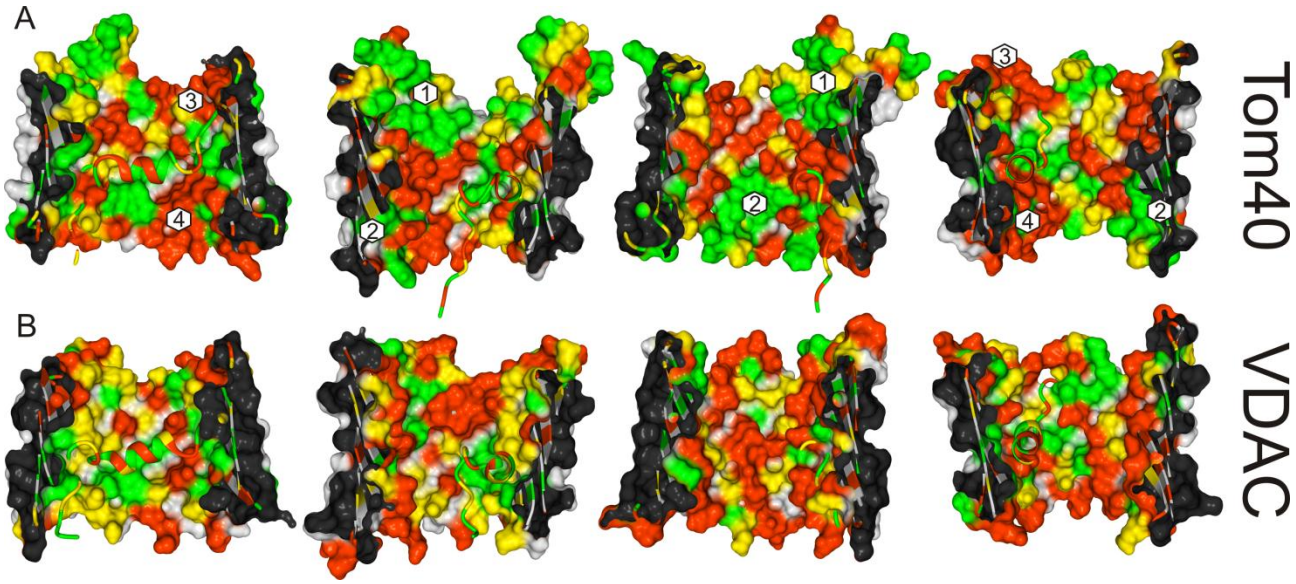
**Supplementary Figure S5** *MALDI-TOF peptide mass fingerprint spectra of NcTom40 time course proteolysis experiments*. Peptide mass fingerprint data were acquired in the reflector MS mode using an accelerating voltage of 21 kV and 2000 laser shots per sample. Four proteases with different cleavage specificity were used: (A) Asp-N (B) Glu-C (C) Lys-C (D) Trypsin. Aliquots of the digests were taken at 1 min, 10 min, 1 h and 4 h as indicated in the figures (A-D). Labeled peptide masses in figures (A-D) are listed in Table 1.
Note that the peptide marked with an asterisk in Supplemental Figure S5b (m/z = 1607.78 Da) corresponds to the acetylated aminoterminus (amino acids 1-14) of NcTom5 which was co-purified with NcTom40. The identity of this peptide was confirmed by MALDI-TOF-MS/MS analysis (spectrum not shown).

**Supplementary Figure S6** *Proteolytic cutting sites in NcTom40*. The cutting sites determined by mass spectrometry from our proteolysis experiments are shown for each proteolytic enzyme separately. Amino acids cut C-terminally are shown and colored according to the time of occurrence during the coupled digestion/mass spectrometry experiment. This means that in case of AspN, which cuts aspartate at its N-terminus, not aspartate itself, but the residue N-terminal to it is marked. For the other enzymes the residue, which is cut C-terminally, is marked.

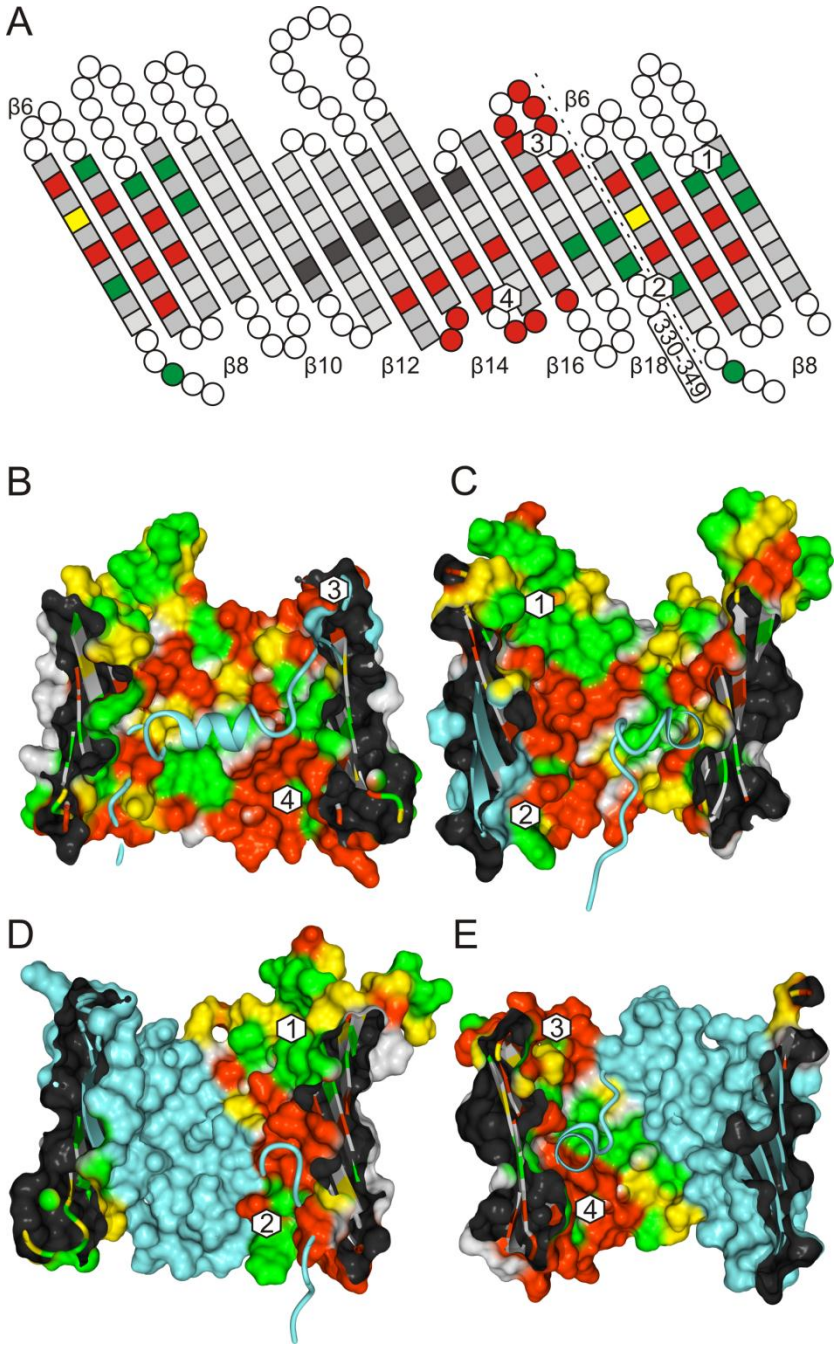



**Supplementary Figure S7** *Conservation in the identified helix anchor region*. Logo plots, created with the Weblogo server (Crooks et al. 2004), of the helix anchor region of (A) Tom40 and (B) VDAC are shown. The residue numbers for (A) NcTom40 and (B) mVDAC1 are given.

Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004 Jun;14(6):1188-90.




**Supplementary Figure S8** *The polar slide of Tom40 is not present in VDAC*. A surface representation of (A) NcTom40 and (B) murine VDAC-1 is shown. Red marks on average strongly hydrophilic positions, yellow a medium average hydrophilicity, and green marks on average hydrophobic positions. The color code is the same as in Fig. 4A-C. The average hydrophobicity was calculated from all amino acids in the respective column of the multiple sequence alignment of all Tom40 sequences reduced to a maximal sequence identity of 90%. From left to right the pores are rotated in steps of 90° clockwise around the longitudinal axis (perpendicular to the membrane) of the β-barrel. In (A) the hydrophobic patches lining the polar slide are labeled 1 and 2, and the hydrophilic patches on the opposite side of the pore interior are labeled 3 and 4.




**Supplementary Figure S9** *The rTom40-Δ(1-165) mutant*. A schematic representation of the homology model of NcTom40-Δ(1-122), which corresponds to rTom40-Δ(1-165), is shown in (A) and a surface representation of the modeled pore cut open in (B-E). In (A) The N-terminal region corresponding to the first 165 residues in Tom40 of *Rattus norvegicus* was removed, but is shown in (B-E) colored in cyan. Red marks on average strongly hydrophilic positions, yellow a medium average hydrophilicity, and green marks on average hydrophobic positions. The color code is the same as in Fig. 4A-C. The average hydrophobicity was calculated from all amino acids in the respective column of the multiple sequence alignment of all Tom40 sequences reduced to a maximal sequence identity of 90%.

# Manuscript III

## Functional refolding and characterization of two Tom40 isoforms from human mitochondria

**Frauke Mager[1,2], Dennis Gessmann[1], Stephan Nussberger[1] and Kornelius Zeth[2]**

[1]Biophysics Department, Institute of Biology, University of Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany

[2]Department of Protein Evolution, Max Planck Institute for Developmental Biology, Spemannstr. 35-39, 72076 Tübingen, Germany

*Running head:* Functional refolding of human Tom40

*Keywords:* Tom40; protein-conducting channel; β-barrel protein; mitochondria; refolding

## Abstract

Tom40 proteins represent an essential class of molecules which facilitate translocation of unfolded proteins from the cytosol into the mitochondrial intermembrane space. They are part of a high-molecular mass complex that forms the protein-conducting channel in outer mitochondrial membranes. This study concerns the recombinant expression, purification and folding of amino-terminally truncated variants of the two human Tom40 isoforms for structural biology experiments. Both CD and FTIR secondary structure analysis revealed a dominant beta-sheet structure and a short alpha-helical part for both proteins together with a high thermal stability. Two secondary structure elements can be denatured independently. Reconstitution of the recombinant protein into planar lipid bilayers demonstrated ion channel activity similar to Tom40 purified from *Neurospora crassa* mitochondrial membranes, but conductivity fingerprints differ from the structurally closely related VDAC proteins.

## Introduction

The first step in mitochondrial protein import is mediated by a multisubunit protein-conducting channel (TOM complex) located in the outer membrane, which comprises up to seven different subunits (Chacinska et al. 2009; Endo and Yamano 2009; Endo et al. 2011; Mokranjac and Neupert 2009; Nussberger and Neupert 2002; Prokisch et al. 2002). Various receptor proteins selectively recognize different substrates. Although some subunits differ or are absent among species (Macasev et al. 2000, 2004; Perry et al. 2006; Poynor et al. 2008), all TOM complexes comprise a 40-kDa protein, termed Tom40, as a major component. Based on single-channel measurements (Ahting et al. 2001; Hill et al. 1998; Poynor et al. 2008; Romero-Ruiz et al. 2010) of isolated Tom40, an ion- and peptide-conducting pore in lipid membranes was demonstrated. It may thus also function as the actual protein-conducting channel that facilitates the transfer of virtually all mitochondrial preproteins synthesized in the cytosol. To date, most Tom40 homologs involved in this translocation process have been described for mitochondria of *Saccharomyces cerevisiae*, *Neurospora crassa*, *Arabidopsis thaliana*, *Homo sapiens* and *Rattus norvegicus* (Kinoshita et al. 2007; Schwartz and Matouschek 1999; Suzuki et al. 2000; Werhahn et al. 2001, 2003; Perry et al. 2006).

Electron microscopy, electrophysiology and biochemical studies measuring the effect of rigid gold labels introduced into precursor proteins during import into mitochondria of *S. cerevisiae* indicated Tom40 pore diameters of ~20 Å (Hill et al. 1998; Kinoshita et al. 2007; Schwartz and Matouschek 1999; Suzuki et al. 2000; Werhahn et al. 2001, 2003; Ahting et al. 2001). This diameter is sufficient to accommodate unfolded protein chains or partially folded domains. Detailed structural and functional studies addressing the interaction of preproteins with purified mammalian Tom40 protein, however, have been hampered by the considerable complexity to isolate the protein from higher eukaryotic organisms.

In this report we explored the possibility of expressing human Tom40 isoforms in *Escherichia coli* and purifying these proteins from inclusion bodies under denaturing conditions. We show that purified, amino-terminally truncated variants of human Tom40A and Tom40B (abbreviated as hTom40AΔN and hTom40BΔN) can functionally be refolded by rapid dilution of denatured protein in different detergent solutions. Secondary structure analyses demonstrated significant β-barrel architecture for both proteins. Reconstitution into planar lipid bilayers and electrophysiology studies showed that both channels are ion-conducting.

## Materials and Methods

### Cloning

The genes of human Tom40 isoforms A and B were synthesized in truncated form. *hTom40A* encoding a gene product missing amino acids 1–82 and *hTom40B* missing amino acids 1–29 were cloned into the NcoI and XhoI cloning sites of expression vector pET24d (Novagen, Darmstadt, Germany) under control of a T7 promotor (Trenyzme Biotechnology, Konstanz, Germany). Both genes encoded proteins with a carboxy-terminal hexahistidinyl tag.

### Bacterial expression and preparation of inclusion bodies

Human Tom40 proteins were expressed as inclusion bodies in *E. coli* BL21-CodonPlus (DE3) cells (Stratagene, Waldbronn, Germany). Cells were grown in LB medium using shaking flasks at 37°C. Tom40 expression was induced with 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) at a cell density corresponding

to an $OD_{600}$ of 0.6. After 6 h of growth, cells were harvested by centrifugation. The expression of human Tom40 was confirmed by SDS-PAGE and Western blotting using mammalian Tom40-specific antibodies.

For purification of Tom40 isoforms, cells were washed twice with PBS buffer and resuspended in 10 ml PBS/g cells supplemented with 2 mg/ml DNaseI from bovine pancreas (Sigma-Aldrich, Hamburg, Germany). After incubation on ice for 20 min, cells were lysed in a French pressure cell (AMINCO; American Instrument Exchange, Haverhill, MA). Inclusion bodies were separated from cell debris by centrifugation at 20,000x *g* for 20 min at 4°C. The inclusion body pellet was washed twice with buffer containing 50 mM Tris and 100 mM NaCl, pH 8. Inclusion bodies were subsequently solubilized in 6 M guanidine hydrochloride, 20 mM Tris, 150 mM NaCl, and 1 mM β-mercaptoethanol (pH 8), using a glass–glass homogenizer. To remove unsolubilized material the homogenate was centrifuged at 30,000x *g* for 30 min at 4°C, and supernatants were recovered and stored at 4°C for further use.

**Purification and refolding of Tom40 isoforms**

Human Tom40A and B were isolated from inclusion bodies under denaturing conditions in 6 M guanidine hydrochloride, 20 mM Tris (pH 8), 150 mM NaCl, 1 mM β-mercaptoethanol using an Äkta Basic system (GE Healthcare, Uppsala, Sweden) and 1 or 20 ml Ni-Sepharose HiTrap columns (GE Healthcare). Then, the purified proteins were diluted to ~5 mg/ml and kept at 4°C.

For refolding, human Tom40A and B proteins (~5 mg/ml) in 6 M guanidine hydrochloride were diluted dropwise 10- to 20-fold in different refolding buffer solutions at 4°C. After 1, 24 and 170 h, the efficiency of protein folding was assessed by evaluation of the amount of precipitated protein after centrifugation at 15,000x *g* for 10 min at 4°C and determination of the protein concentration in the supernatant by UV absorption at a wavelength of 280 nm. The buffer substances used were 50 mM citric acid (pH 5), 50 mM sodium-phosphate buffer (pH 6), 50 mM Tris buffer (pH 7) and 50 mM glycine-sodium-hydroxide (pH 10) and contained 1 mM β-mercaptoethanol. The detergents were either *n*-dodecyl-β-D-maltoside (DDM; Glycon, Luckenwalde, Germany), lauryl-dimethylamine-oxide (LDAO, Sigma-Aldrich), *n*-octyl-polyoxyethylene (oPOE; Bachem, Weil am Rhein, Germany), Brij35 (Sigma-Aldrich), 3-[(3-cholamidopropyl) dimethyl-

ammonio]-1-propanesulfonate (CHAPS; Merck, Darmstadt, Germany) and *n*-octyl-β-D-glucopyranoside (β-OG, Glycon) at concentrations corresponding to ~5x CMC, respectively.

Optimal refolding of human Tom40A and B was achieved by a 10-fold dilution of 5 mg/ml protein in 6 M guanidine hydrochloride into 20 mM Tris (pH 8), 150 mM NaCl, 1 mM β-mercaptoethanol and 1% (w/v) LDAO at 4°C. For further use the LDAO concentration was reduced to 0.5% by dropwise addition of the same buffer without detergent. Final samples were centrifuged at 100,000x *g*, and supernatants were subjected to Ni-NTA chromatography to concentrate the refolded protein. Residual amounts of contaminants were removed by gel filtration using a Superose 12 column (GE Healthcare).

All fractions containing Tom40 were merged and concentrated using spin columns with a 30-kDa cut-off (Vivaspin, GE Healthcare) up to 10 mg/ml. The concentrations of hTom40A and hTom40B were determined by UV absorbance spectroscopy using extinction coefficients $\varepsilon_{W,280nm} \approx 5{,}600\ M^{-1}\ cm^{-1}$ for tryptophan and $\varepsilon_{Y,280nm} \approx 1{,}200\ M^{-1}\ cm^{-1}$ for tyrosine (Pace et al. 1995). All samples were filtered (Rotilabo® filter, pore size 0.22 µm; Carl Roth, Karlsruhe, Germany) and centrifuged at 16,000x *g* for 5 min to remove possible protein aggregates prior to circular dichroism (CD) and Fourier transform infrared spectroscopy (FTIR) measurements.

**Far UV-CD spectroscopy**

CD spectroscopic measurements were performed using a J-815 spectrometer (Jasco, Tokyo, Japan) in quartz cuvettes of 0.1 cm path length. Spectra at 20°C were recorded from 185 to 250 nm with a resolution of 1.0 nm and an acquisition time of 20 nm/min. To avoid high-voltage output from the photomultiplier at high temperature measurements, CD spectra were recorded from 195 to 250 nm. Final CD spectra were obtained by averaging five consecutive scans and corrected for background by subtraction of spectra of protein-free samples recorded under the same conditions. Mean residue ellipticity (Θ) was calculated based on the molar protein concentration and the number of amino residues of the Tom40 isoforms. The protein concentration used for CD spectroscopy was adjusted to 0.2–0.5 mg/ml.

The secondary structure content was estimated using the CDpro package, according to Sreerama and Woody (2000, 2003, 2004), namely, CDSSTR, CONTIN/LL and SELCON 3. Melting curves were recorded at constant wavelength at 216 nm for hTom40AΔ1–82 and BΔ1–29 from 20 to 98°C by applying a temperature ramp of 1°C/min. The percentage of unfolded protein content, $f_U(T)$, was calculated according to $f_U(T)=(\Theta(T)-\Theta_N(T))/(\Theta_U(T)-\Theta_N(T))$ where $\Theta_U(T)=aT+b$ and $\Theta_N(T)=cT+d$ represent the pre- and posttransition baselines, respectively. For evaluating the protein melting temperature, $T_m$, the resulting data $f_U(T)$ were fitted by the Boltzmann equation, $f_U(T)=(f_0-f_{\max})/(1+\exp(T-T_m)/T_s)+f_{\max}$, where $f_0$, $f_{\max}$ and $T_s$ are the minimum and maximum percentages of unfolded protein content and the temperature range over which the transition occurs, respectively (Pace et al. 1998; Fersht 1999).

**FTIR spectroscopy**

Samples for FTIR with a protein concentration of ~5 mg/ml were dialyzed against 20 mM Tris (pH 8), 150 mM NaCl, 1 mM β-mercaptoethanol and 0.5% LDAO at 4°C for 48 h. FTIR spectra were recorded using a Confocheck TENSOR 27 FTIR spectrometer (Bruker Optics, Ettlingen, Germany) equipped with a linear, photovoltaic, mercury–cadmium–telluride detector (Bruker Optics). All spectra were recorded between 4,000 and 800 $cm^{-1}$ for 25 s with a wave number resolution of 4 $cm^{-1}$. To avoid temperature-induced variations of the water signal, the measurement cell was kept at a constant temperature of 20°C. For each spectrum, 32 interferograms were collected and averaged. The aperture setting was 6 mm, and the scanner velocity was at 10 kHz. All procedures were carried out to optimize the quality of the spectrum in the amide I region, between 1,600 and 1,700 $cm^{-1}$. Secondary structure predictions were calculated from baseline corrected spectra with the multivariant pattern recognition method (Fabian and Schultz 2000) of the OPUS Quant 2 software supplied by Bruker Optics. Spectral data were factorized and compared to reference data from a library of more than 40 proteins with known structure (Bruker Optics, source: Protein Data Bank, http://www.pdb.org).

### Electrophysiology

Human Tom40AΔ1–82 and BΔ1–29 in LDAO or o-POE were reconstituted into black lipid membranes, and channel currents were recorded according to standard protocols (Arnold et al. 2007; Poynor et al. 2008; Romero-Ruiz et al. 2010). A 0.5% solution of diphytanoyl phosphatidylcholine (Avanti Polar Lipids, Alabaster, AL) in *n*-decane/butanol (9:1) was spread over a circular hole of 250 μm diameter in the partitioning wall of a cylindrical Delrin cup of a bilayer chamber BCH-13A (Warner Instruments, Hamden, CT). The chamber was filled symmetrically with 1 M KCl and 20 mM Hepes (pH 7.2). Protein solution enriched with cholesterol was added to the *cis* side of the membrane to a final concentration of 10–50 μg/ml. Current fluctuations through single and multiple channels were recorded in voltage-clamp mode using an EPC-8 patch-clamp amplifier (HEKA Electronik, Lambrecht, Germany) connected to the bilayer chambers by a pair of Ag/AgCl pellet electrodes (WPI, Sarasota, FL). Current signals were low pass–filtered at 3 kHz using the built-in Bessel-filter of the amplifier and monitored for channel insertion using an analog oscilloscope. Current and voltage signals from the amplifier were digitized at a sampling rate of 10 kHz per channel using a NI-USB-6251 interface (National Instruments, Austin, TX) controlled by a program of the Strathclyde electrophysiology suite (WinEDR 2.8 or WinWCP 3.6, J. Dempster).

## Results

### Expression and purification

Tom40 isoforms exhibit sequence identities between mammalian species of >91% for Tom40A and 96% for Tom40B. To obtain large amounts of human Tom40 for biochemical and structural studies, human Tom40A and B were cloned as N-terminally truncated proteins (hTom40AΔ1–82 and hTom40BΔ1–29, termed hTom40AΔN and hTom40BΔN; Fig. 1a). Since the N termini of both proteins are predicted to be disordered, we rationalized that they are not essential for formation of the mitochondrial protein translocation pore (see also supplementary Fig. S1 for the secondary structure prediction of hTom40AΔN). Our construction of hTom40AΔ1-82 and hTom40BΔ1-29 was in line with studies on truncated rat Tom40AΔN (Suzuki et al. 2004) and the published mammalian voltage-dependent anion channel 1 (VDAC1) structures (Bayrhuber

et al. 2008; Hiller et al. 2008; Ujwal et al. 2008; Zeth 2010). Hexahistidine-tagged hTom40AΔN and hTom40BΔN were expressed in *E. coli* as inclusion bodies. Bacterial cells grew to a mass of 5 g/l of culture. The wet yield of inclusion bodies ranged between 2 and 3 g/l of culture. Furthermore, testing the functional refolding of recombinant full-length bovine Tom40A (supplementary Fig. S2) revealed a yield of ~5 mg protein/l of *E. coli* culture, which is insufficient for functional and structural studies.

Human Tom40A (hTom40AΔN) was already highly enriched in inclusion bodies (Fig. 1b). The first purification of unfolded protein in chaotropic buffer revealed a rather pure protein. The yield starting from 1 l of culture was ~200 mg of protein. Optimized refolding was achieved using LDAO detergent containing buffer at pH 8 (see Table 1). No or little refolded protein was observed at only slightly decreased pH below 7 or in the presence of detergents such as DDM, oPOE or *n*-octyl-β-D-glucoside (Tables 1, 2). To concentrate refolded human Tom40, the samples were again applied to Ni-NTA affinity and size exclusion chromatography (Fig. 1b, c) for further purification was performed, if needed. When using affinity chromatography before and after refolding, unspecifically bound proteins have been efficiently removed with low imidazole concentrations. After size exclusion chromatography the protein was virtually pure (see Fig. 1b, c).

By contrast, hTom40BΔN was expressed at a lower level (Fig. 1c), and its yield ranged between 20 and 30 mg/l of culture. Nevertheless, the amount of protein obtained was sufficient for electrophysiology, secondary structure determination experiments like CD and FTIR spectroscopy and first crystallization trials. The protocol for pure and refolded protein essentially followed the protocol applied for the Tom40A version (see Fig. 1b).

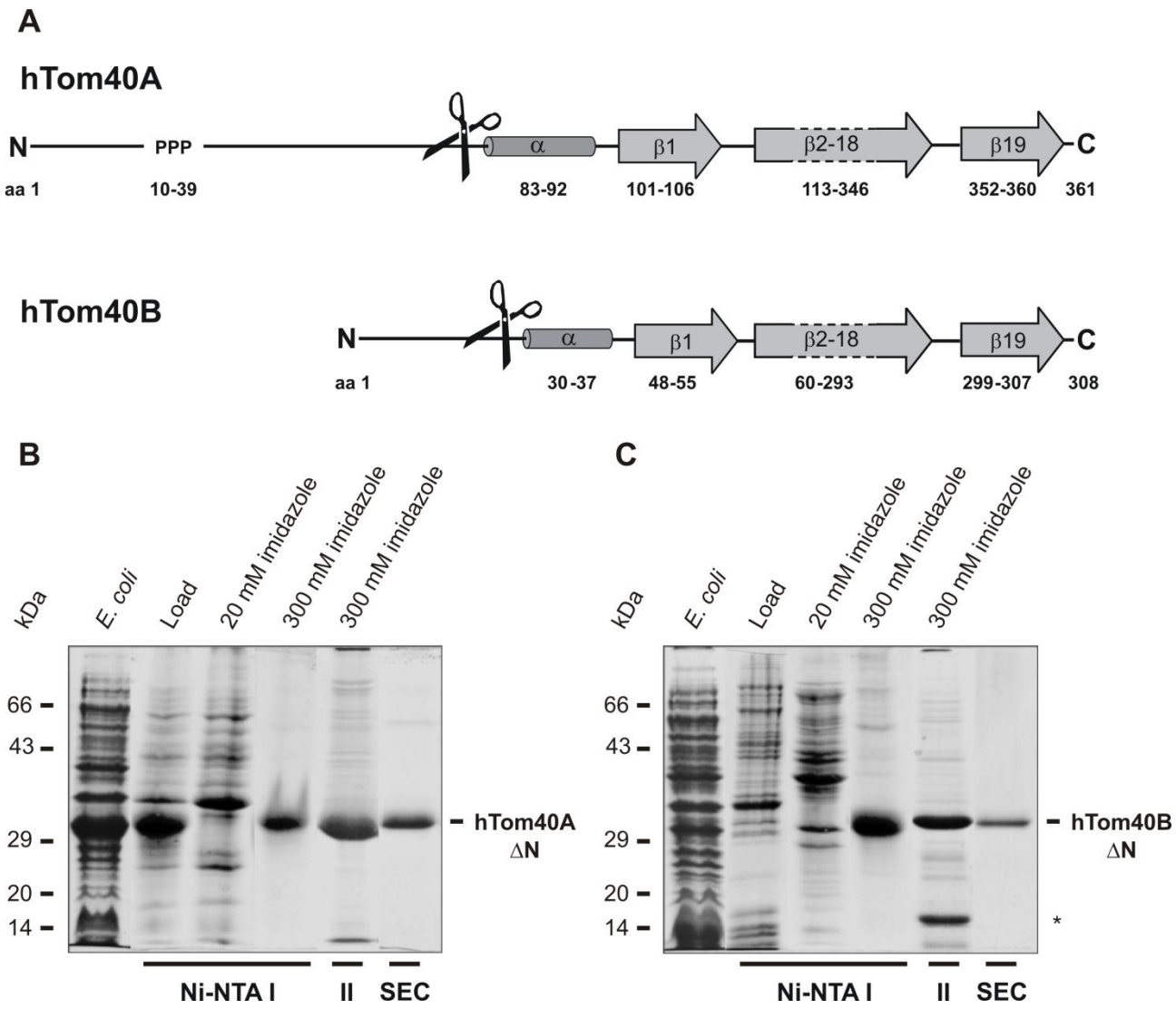



**Fig. 1** Secondary structure prediction and purification of hTom40A and B. **a** PSIPRED-predicted secondary structure of full-length hTom40A and hTom40B. hTom40A and B were truncated by 82 and 29 amino acids, respectively. **b, c** Purification of human Tom40A and B. Inclusion bodies isolated from *E. coli* containing recombinant hTom40AΔ1–82 (hTom40AΔN) and hTom40BΔ1–29 (hTom40BΔN) with C-terminal hexahistidinyl tags were solubilized by 6 M guanidine hydrochloride and loaded onto an Ni-NTA affinity column. hTom40A and B were eluted under denaturing conditions with 300 mM imidazole (Ni-NTA I) and refolded by rapid dilution into 0.5% LDAO. Refolded proteins were subjected to Ni-NTA chromatography (Ni-NTA II) and eventually passed over a Superose12 size-exclusion column (SEC). Aliquots of the resulting column fractions were analyzed by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining. The Ni-NTA II eluate containing hTom40B shows an additional strong band (*), which has been characterized by mass spectrometry as a degradation product. *Lane 1*, *E. coli* cells expressing hTom40A or B; *lane 2*, purified inclusion bodies containing hTom40A or B; *lanes 2–5*, Ni-NTA column fractions and imidazole eluates; *lane 6*, peak fraction of size exclusion column

**Table 1** Yield of soluble hTom40AΔN after refolding

| Detergent | hTom40AΔN 1 h | | | | hTom40AΔN 24 h | | | | hTom40AΔN 170 h | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 |
| 0.05% DDM | 75 | 45 | 95 | 100 | 100 | 40 | 45 | 100 | 60 | 25 | 50 | 60 |
| 0.1% LDAO | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 85 | 70 | 90 |
| 0.5% oPOE | 80 | ≤ 5 | 40 | 95 | 65 | ≤ 5 | ≤ 5 | 35 | 25 | ≤ 5 | ≤ 5 | 15 |
| 0.05% Brij35 | 75 | 100 | 90 | 100 | 100 | 100 | 85 | 100 | 95 | 100 | 70 | 85 |
| 3% CHAPS | 40 | 20 | 10 | 40 | 10 | ≤ 5 | ≤ 5 | 15 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 |
| 1% β-OG | 35 | 15 | 10 | 20 | 25 | ≤ 5 | 10 | 20 | 15 | ≤ 5 | ≤ 5 | ≤ 5 |

hTom40AΔ1-82, termed hTom40AΔN, was refolded in the presence of the detergents DDM, LDAO, oPOE, Brij35, CHAPS and β-OG at different pH values. The amount of folded protein was estimated 1 h, 1 day and 1 week after centrifugation and determination of protein concentration in the soluble phase by UV spectroscopy at 280 nm. The detergent concentrations are given as weight per volume; the protein concentration is given as a percent of total

**Table 2** Yield of soluble hTom40BΔN after refolding

| Detergent | hTom40BΔN 1 h | | | | hTom40BΔN 24 h | | | | hTom40BΔN 170 h | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 | pH 6.0 | pH 7.0 | pH 8.0 | pH 10.0 |
| 0.05% DDM | 25 | 15 | 10 | 25 | 25 | 10 | 20 | ≤ 5 | 15 | ≤ 5 | ≤ 5 | 20 |
| 0.1% LDAO | ≤ 5 | 30 | 10 | 90 | ≤ 5 | 20 | ≤ 5 | 45 | ≤ 5 | 15 | ≤ 5 | 15 |
| 0.5% oPOE | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | 15 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 |
| 0.05% Brij35 | 90 | 60 | 100 | 100 | 85 | 85 | 85 | 75 | 80 | 55 | 55 | 40 |
| 3% CHAPS | ≤ 5 | 10 | ≤ 5 | 15 | ≤ 5 | 10 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 |
| 1% β-OG | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 | 15 | ≤ 5 | ≤ 5 | ≤ 5 | ≤ 5 |

hTom40BΔ1-29, termed hTom40BΔN, was refolded in the presence of different detergents at different pH. The amount of folded protein was estimated as described in Table 1. Detergent concentrations are given as weight per volume; the protein concentration is given as a percent of total

**Secondary structure determination by CD spectroscopy**

The secondary structure content and thermal stability of hTom40AΔN and hTom40BΔN were analyzed in LDAO by CD spectroscopy and compared to spectra of known β-barrel proteins. The spectra showed a clear dominance of beta-sheet content (Fig. 2; Table 3). The minimum of the spectra consistently clustered around 216 nm, with a crossover of the baseline at around 203 nm. At wavelengths >245 nm, the CD spectrum approached ellipticity values close to zero, indicating that the hTom40ΔN preparations were virtually free of any higher-order aggregates which would cause light-scattering effects and interfere with the interpretation of the data. Deconvolution of the data points via CDpro revealed a predominance of β-sheet secondary structure (>30%) for human Tom40s with low α-helical content (<24%) (Table 3). Hence, efficient refolding of truncated hTom40s can be assumed under the conditions chosen. This is further supported by the high thermal stability of hTom40AΔN and BΔN ($T_m$ of 73.1 and 74.1°C; Fig. 2c, d). Thus, the truncated amino acids of the Tom40 isoforms might not be involved in the overall barrel structure. In addition, hTom40AΔN and BΔN exhibit similar CD spectral characteristics with slightly left-shifted spectra for hTom40 isoform B.

An interesting observation was obtained during the analysis of the CD spectra of hTom40s when recorded just below the melting point of the protein at 70°C. We found that both CD spectra of hTom40 isoforms underwent a remarkable right-shift in combination with decreased ellipticity values in comparison to CD spectra recorded at 20°C (Fig. 2a, b) and 50°C (Fig. 2c, d). Determination of secondary structure content displayed a substantial loss in the α-helical content with an increase in β-sheet portion (see supplementary Table 1). For hTom40AΔN and BΔN the α-helical content decreased from 20 to 5% and from 24 to 12%, respectively. The corresponding relative β-sheet content increased from 36 to 43% and from 32 to 37%. The relative amount of random coil and turn structure did not change significantly with temperature. Secondary structure development of hTom40AΔN (Fig. 2c) and BΔN (Fig. 2d) during melting showed that, first, the α-helical portion is reduced during temperature increase, which indicates that the N-terminal helix denatures before the barrel portion follows. Again, hTom40BΔN showed small differences, whereas the

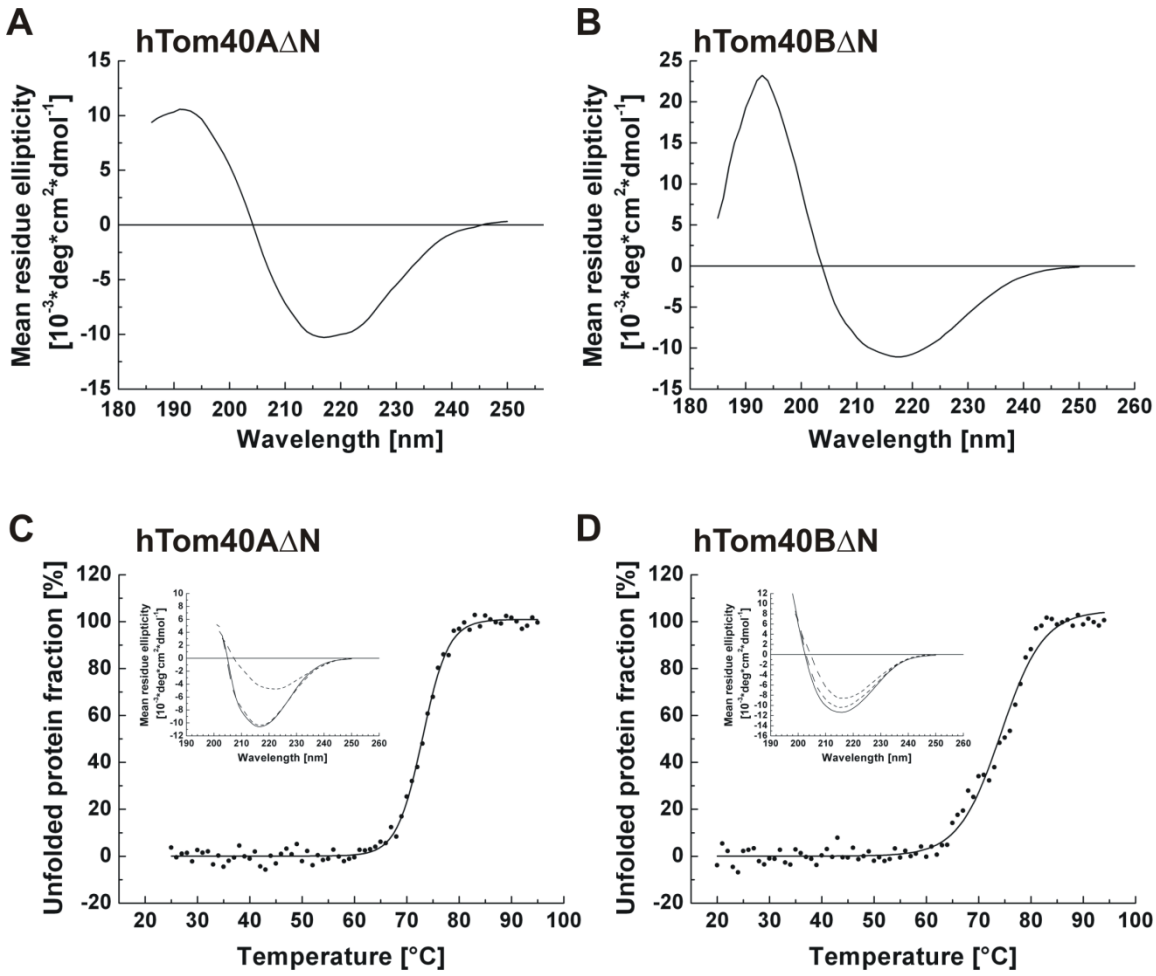



**Fig. 2** Far UV CD spectra and thermal stability of hTom40AΔN and hTom40BΔN. **a, b** hTom40AΔN (~0.2 mg/ml) and hTom40BΔN (~0.2 mg/ml) were solubilized in 0.5% LDAO and 20 mM Tris (pH 8.0). For each CD spectrum, five scans were accumulated at 20°C and background-corrected. Mean residue ellipticity was calculated based on the molar protein concentration and the number of amino acid residues of the according protein. The spectra of hTom40A and B indicate a high ratio of β-sheet with a minimum at 216 nm and a crossover of the baseline at 203 nm. **c, d** Melting curves of hTom40A and hTom40B determined from corrected CD signals measured at a wavelength of 216 nm. The melting temperatures of hTom40AΔN and hTom40BΔN were at 73 and 74°C, respectively. *Inserts* represent spectra recorded at 20°C (*solid line*), 50°C (*dashed line*) and 70°C (*dashed line*)

overall transition was similar to hTom40AΔN (Fig. 2a, b). This might be due to a higher melting point of hTom40BΔN compared to hTom40AΔN. It is noteworthy that similar observations were gained for Tom40 from *Aspergillus fumigatus* in Brij35 and LDAO (data not shown).

## Secondary structure determination by FTIR spectroscopy

To determine a complementary number for the secondary structure and to support the CD spectroscopic data, hTom40AΔN and BΔN were analyzed by FTIR. The amide I band (1,700–1,600 $cm^{-1}$) of the IR spectra of both isoforms is shown in Fig. 3 as it represents the region in which α-helical, β-sheets and random coil structures are activated. The maximum was at 1,632 $cm^{-1}$, which is indicative of β-sheet content. The spectra show a clear dominance of β-sheet contribution, which is comparable to measurements of other β-barrel proteins like human VDAC1 (Engelhardt et al. 2007). As calculated from reference data of more than 40 proteins with known 3D structure, the amount of β-sheet was calculated to be >55% (Table 3). At wave numbers of 1,650 $cm^{-1}$ the spectra showed no significant local maximum. Comparisons with reference spectra indicated an α-helical content of <10%. The additional maximum at 1,656 $cm^{-1}$ indicates antiparallel β-sheet and hints at the organization of the β-sheets in the protein. These measurements support the model of a β-barrel protein with an α-helical elongation at the N terminus. In agreement with CD spectroscopic data, the content of β-sheet is dominant in the structure, while the α-helical amount is relatively low.

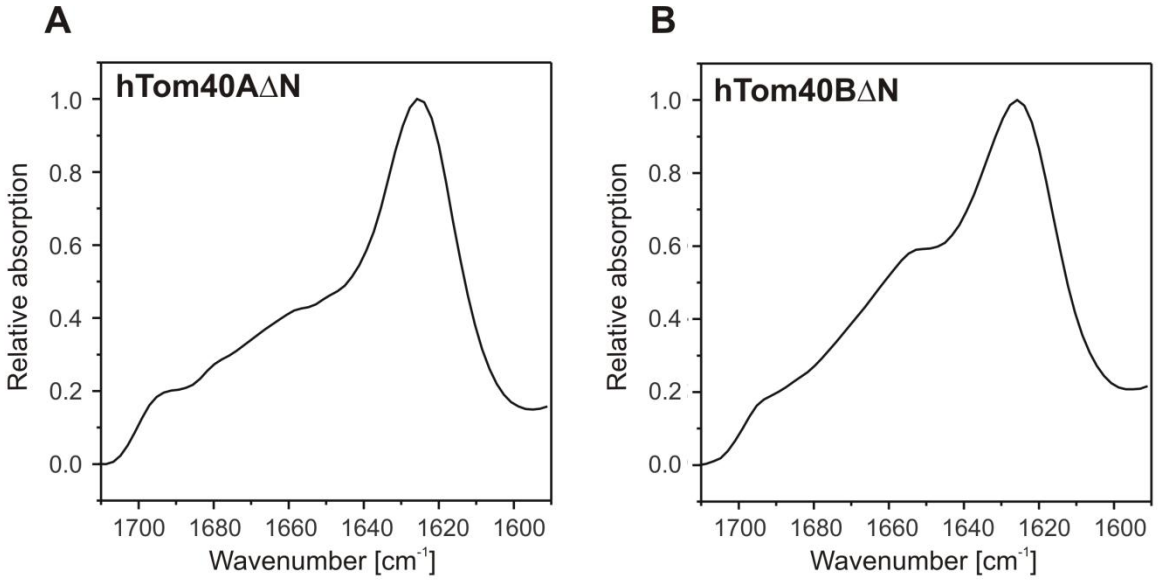



**Fig. 3** FTIR spectra of hTom40AΔN and hTom40BΔN. The IR spectra of hTom40AΔN (~7 mg/ml) and hTom40BΔN (~5 mg/ml) represent an average of three independent spectra with 32 consecutive scans each. Spectral bands assigned to the α-helical structure are centered at a wave number of about 1,650 $cm^{-1}$, random components

**Pore-forming activity of hTom40AΔN and hTom40BΔN**

To test whether purified hTom40ΔN is active and capable of conducting ion flux, the isolated and refolded hTom40AΔN was inserted in an artificial planar lipid bilayer. The channel-forming activity of the protein was tested upon application of different voltages, and the subsequent current traces were recorded. The traces were measured with single channels as well as multiple channels inserted in the artificial lipid layer. For single-channel characterization a planar membrane was reconstituted and the zero current was measured as baseline. Protein was added to the *cis* side of the chamber, while a voltage of 30 mV was applied and current traces were recorded. Single-channel insertions have been observed (Fig. 4a). For human Tom40AΔN the channel showed a conductivity of ~1 nS at 1 M KCl (Fig. 4b). For human Tom40BΔN the single-channel conductivity could not be determined yet.

To test whether the inserted channels show voltage-dependent gating, several Tom40AΔN and Tom40BΔN channels were allowed to reconstitute into planar lipid membranes and steps of different voltages, starting from ±10 mV and ending at ±160 mV, were applied. Baseline currents at 0 mV were always measured in between. For human Tom40AΔN and BΔN the recorded currents showed no voltage-dependent gating at voltages below 120 mV. The conductivity of multiple channels, which was measured at different voltages, remained the same. For hTom40AΔN the measurements showed a slight decrease of conductivity at voltages higher than 120 mV (Fig. 4c). For hTom40BΔN no voltage-dependent gating was observed even at voltages above 120 mV (Fig. 4d). Predominantly, a voltage-independent behavior was observed for both isoforms.

---

**(*Fig. 3 continued*)** at 1,645–1,640 $cm^{-1}$ and β-sheet at 1,630–1,625 $cm^{-1}$. The shoulder at 1,695 $cm^{-1}$ indicates an antiparallel β-sheet with particularly short turns. For all spectra, the baseline was subtracted. The main peak at 1,625 $cm^{-1}$ for hTom40AΔN and hTom40BΔN, respectively, indicates high β-sheet content

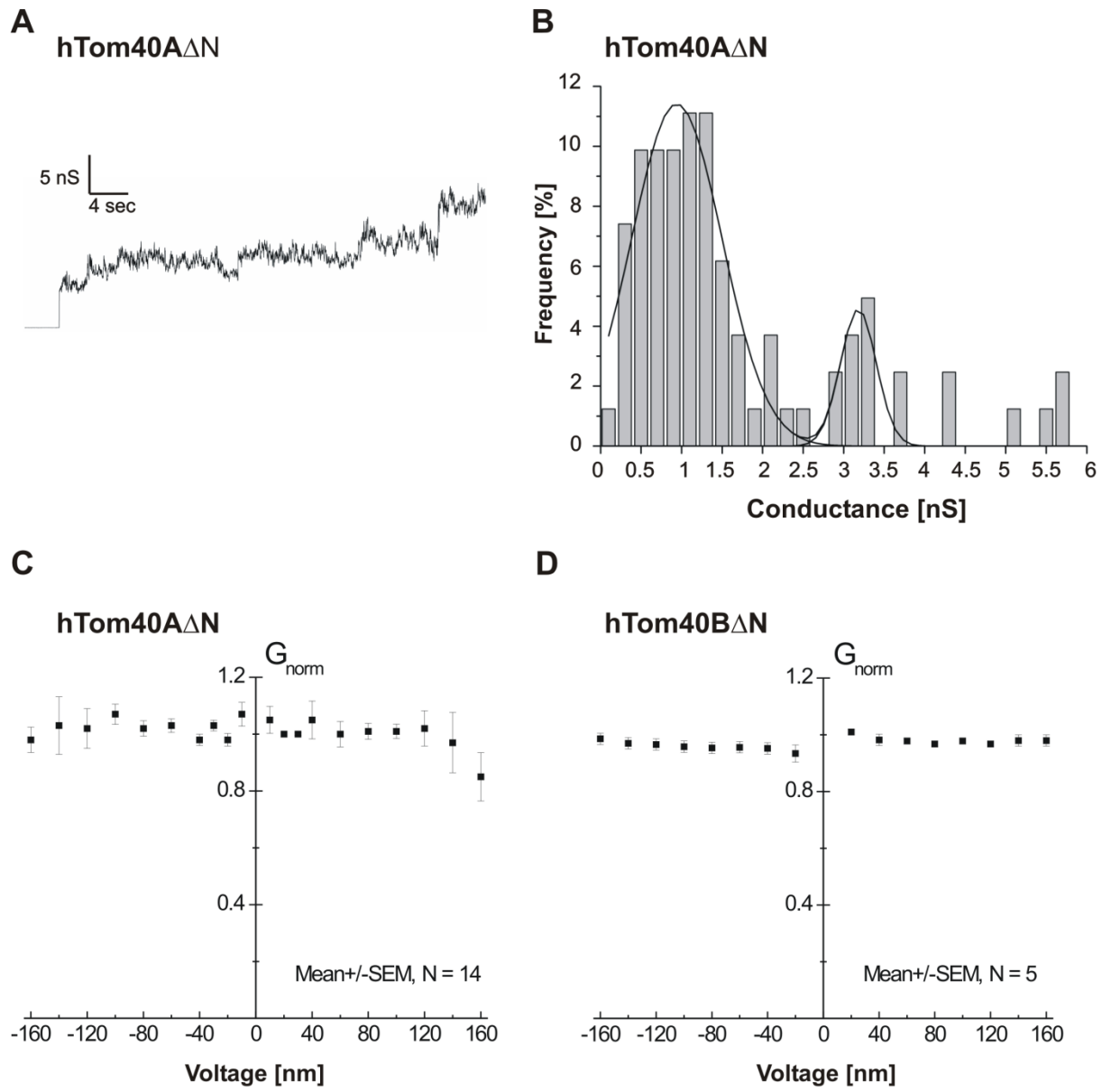



**Fig. 4** Channel conductance of human Tom40AΔN and BΔN. **a** Purified hTom40AΔN was added to both sides of a planar lipid membrane formed by DiphPC/*n*-decane/butanol, and single-channel conductance was measured in 1 M KCl and 50 mM HEPES (pH 7.2) at a membrane potential of 30 mV. **b** Histogram of channel conductance measured at voltages between −50 and +50 mV. A total of 82 single-channel insertions were analyzed. Most channel insertions showed a conductance of 1.2 nS. **c, d** Voltage dependence of hTom40A and hTom40B channel conductance: hTom40AΔN and hTom40BΔN were reconstituted into black lipid membranes, and current traces were recorded at constant voltage between −160 and +160 mV at 10-mV increments. The membrane conductance, $G_{norm}$, was calculated according to Ohm's law and normalized to the mean conductance at +10 mV

## Discussion

The TOM complex of the outer mitochondrial membrane is responsible for the translocation of proteins from the cytosol to the mitochondrial inner membrane space. From electron microscopic analysis it is well established that TOM complexes can show up to three pores (Künkele et al. 1998; Model et al. 2002, 2008), which are based on the presence of single Tom40 molecules. We suppose this holo TOM complex to incorporate three Tom40 monomers to form a trimer. The core complex may incorporate two Tom40 monomers to form a dimer (Mager et al. 2010). How these monomers assemble with the other components of the holo complex; the receptors Tom70, Tom20 and Tom22; as well as the small proteins Tom5, Tom6 and Tom7 to the high-molecular mass complex of 400–500 kDa (Künkele et al. 1998; Model et al. 2002, 2008) is not yet clear.

The presence of several Tom40 isoforms has been shown for other mammalian organisms like *Rattus norvegicus* (Kinoshita et al. 2007) and *Bos taurus* (unpublished data). The C-terminal part of the proteins, which forms the barrel walling, shows an identity of >60% for both human isoforms. Both isoforms can occur in the same tissue in different amounts and form part of the TOM complex (Kinoshita et al. 2007). Whether both isoforms can be present in the same complex is not yet determined.

The N-terminus of Tom40 isoforms differs significantly among several species. In *Arabidopsis thaliana*, the N-terminus is extended to 50 amino acids and in human Tom40A a proline-rich region presents a rather unfolded N-terminal extension. Due to the specific sequence contribution, this extension may contribute to actin binding (Quinlan and Kerkhoff 2008). However, it was reported not to be responsible for targeting and subsequent assembly with preexisting Tom subunits (Humphries et al. 2005). Generally, polyproline-rich sequences result in unstructured parts of proteins and are known to form the basis, e.g., domains known to polymerize actin (FH1, WH2 domains) (Quinlan and Kerkhoff 2008). Nevertheless, the function of the N-terminal helix may take part in the SAM-mediated assembly of the monomer in the complex (Humphries et al. 2005) or may play a role in the interaction with other components of the TOM complex, like the Tom22 receptor (Taylor et al. 2003).

**Table 3** Secondary structure of hTom40AΔN and BΔN in comparison with hVDAC1

| | CD | | FTIR | | Psipred[b] | |
|---|---|---|---|---|---|---|
| | α-Helix | β-Sheet | α-Helix | β-Sheet | α-Helix | β-Sheet |
| hTom40AΔN | 20 | 36 | nd | 62 | 8 | 52 |
| hTom40BΔN | 24 | 32 | 8 | 57 | 8 | 54 |
| hVDAC1[a] | 18 | 40 | 7 | 48 | 0 | 63 |

Values are percentages

[a] Engelhardt et al. (2007); Malia and Wagner (2007)

[b] Jones (1999)

Refolding of the protein was induced by methods which were already successfully applied to, e.g., hVDAC isoforms 1 and 2 (Engelhardt et al. 2007). Several detergents induced refolding but at differing amounts. As for the VDAC proteins, LDAO turned out to be the most successful detergent in terms of protein remaining in solution. Another parameter which was clearly important is the pH of the refolding solution. As for VDAC and, e.g., the BamA receptor protein from *E. coli*, the basic pH increased the refolding excess dramatically, while pH values lower than 7 did in part not reveal any refolded portion of the protein (data on BamA are not published). This behavior may be explained by the close to neutral isoelectric point (pI) of the protein, which is ~6.8 for the truncated version of hTom40A and ~6.4 for the truncated hTom40B protein.

Structural analysis of Tom40 in comparison to VDAC1 with several secondary structure predicting programs revealed a similar pattern of β-sheets. A comparison of the experimentally determined secondary structure of human Tom40 and human VDAC by CD spectroscopy, FTIR and prediction by Psipred (Jones 1999) is given in Table 3 and shows the similarity between the three proteins. This leads to the hypothesis of a similar structure and therefore a similar evolutionary origin, as recently suggested (Bayrhuber et al. 2008; Pusnik et al. 2009; Zeth 2010). According to this topological similarity, Tom40 is predicted to form a β-barrel with 19 strands and an α-helical part at the N-terminus of the protein facing the cytosol (see Fig. S1). Previous studies on Tom40 from fungi, plants and mammals did not address the role of the putative N-terminal α-helical domain of the protein specifically.

This study underlines the predicted β-sheet structure by CD spectroscopy and FTIR and demonstrates the formation of intact channels for both human Tom40 isoforms, which are capable of conducting an ion flux through a lipid membrane. Values for secondary structure are supported by the crystal structure of human VDAC1 and with the secondary structure prediction of other Tom40 proteins. These results are in perfect agreement with former studies on fungal and mammalian Tom40 proteins, suggesting a β-barrel structure for Tom40 in all species (Becker et al. 2005; Hill et al. 1998; Kinoshita et al. 2007; Ahting et al. 2001; Suzuki et al. 2000) and an N-terminal α-helical segment adjacent to the β-barrel (Suzuki et al. 2004). Further, previous studies on mammalian VDAC1 determined similar CD spectrum characteristics: a minimum at ~216 nm, a zero-crossing wavelength at 206 nm and 39.7% β-sheet, together with 17.5% α-helix, 13.3% turn and 28.4% random coil (Malia and Wagner 2007).

Together, these findings suggest the existence of an N-terminal α-helical domain in the 3D structure of human Tom40s that is not necessary for the stabilization of the overall Tom40 β-barrel structure. This fits perfectly with the observed decrease in ellipticity values at increased temperatures as α-helical structures result in much stronger signals than β-sheet structures. Taking into account secondary structure predictions of Tom40 from various organisms, an N-terminal α-helix is predicted in all species in addition to a β-sheet-rich core. In addition, Suzuki et al. (2004) have revealed that the recombinant membrane-embedded C-terminal half of rat Tom40 (Tom40Δ1–165) constitutes the preprotein recognition domain with an enriched β-sheet structure. However, this part of the protein clone consisted of only 13 β-strands and not 19, as for the constructs used in this work. Thus, the location of the α-helix at the N-terminus of the Tom40 protein can be suggested. This is in good agreement with the crystal/NMR structure of mammalian VDAC1, featuring a flexible N-terminal α-helix located inside the VDAC channel (Abu-Hamad et al. 2009; Bayrhuber et al. 2008; Hiller et al. 2008; Ujwal et al. 2008).

Assessment of the secondary structure of Tom40 from *N. crassa* and *S. cerevisiae* suggested about 16 β-strands for the formation of a functional pore (Becker et al. 2005). Studies addressing the location of functional domains in rat Tom40A indicated that the pore is restricted to the C-terminal segment of the

protein, which lacks the first three putative β-strands (Suzuki et al. 2004). Our results predict that Tom40 belongs to the "mitochondrial porin" superfamily with 19 β-strands (Pusnik et al. 2009) that incorporates both the VDACs and Tom40 proteins from mitochondrial outer membranes.

The channel conductance of hTom40AΔN was in the same range as the recombinant yeast Tom40 measured at 250 mM KCl, with states at ~210, ~260, ~440 and ~430 pS (Harsman et al. 2010) as well as native Tom40 from *N. crassa* investigated in 1 M KCl with ~0.4, ~0.8, ~1.4, ~2.0 and ~2.9 nS (Ahting et al. 2001; Poynor et al. 2008; Romero-Ruiz et al. 2010). The channel properties of hTom40AΔN could not yet be compared with those of rat Tom40 (Suzuki et al. 2004) or any other Tom40 protein as they remain to be described.

The electrophysiologic data of both hTom40 isoforms showed no significant voltage dependence. For hTom40AΔN, channel closure has been observed in some experiments and only at voltages above 100 mV. This stands in contrast to electrophysiologic data of hVDAC1, where movement of the N-terminal helix inside the channel has been predicted to result in reduced conductivity of >30% at voltages above 50 mV (Engelhardt et al. 2007; Ujwal et al. 2008). Since our measurements for this publication were performed with protein from different purifications and therefore several refolding experiments, it is possible that the orientation of the flexible N-terminal helix of hTom40 is not definite. If the N-terminal helix is facing into the barrel, it may affect the ion flux and show flickering currents similar to hVDAC1. However, when the N-terminal helix rises out of the barrel, it has no or limited effect on conductance. This would lead to a different behavior of the channel conductance regarding voltage dependence. The channel measurements of isoform B show only little conducting difference. But overall, isoform BΔN shows similar conducting behavior as isoform A and underlines the efficient refolding strategy.

More efforts are necessary to characterize the channel properties of both human Tom40 constructs in more detail. Nevertheless, our method for the expression, purification and refolding of mammalian Tom40 isoforms thus provides the basis for further structural and physiological investigations.

## Acknowledgements

The authors thank Beate Nitschke for expert technical assistance and Tobias Kulschewski and Dr. Simon Stutz for their contributions at the early stage of this work. The work was supported by a Competence Network on Functional Nanostructures grant of the Baden-Württemberg Foundation and the Human Frontier Science Program to S. N. and K. Z., respectively.

## Contribution to this manuscript

FM conducted all experiments, with the contribution of DG to the circular dichroism (CD) measurements and the work on bovine/fungal Tom40. FM analyzed all data sets and made all the graphs, except for CD analysis and CD graphs, which were suggested and generated by DG, respectively. KZ, DG, SN and FM wrote the manuscript. SN and KZ defined and supervised the project.

## References


Abu-Hamad S, Arbel N, Calo D, Arzoine L, Israelson A, Keinan N, Ben-Romano R, Friedman O, Shoshan-Barmatz V (2009) The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins. J Cell Sci 122: 1906–1916

Ahting U, Thieffry M, Engelhardt H, Hegerl R, Neupert W, Nussberger S (2001) Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria. J Cell Biol 153:1151–1160

Arnold T, Poynor M, Nussberger S, Lupas AN, Linke D (2007) Gene duplication of the eight-stranded β-barrel protein OmpX produces a functional pore: a scenario for the evolution of transmembrane β-barrels. J Mol Biol 336:1174–1184

Bayrhuber M, Meins T, Habeck M, Becker S, Giller K, Villinger S, Vonrhein C, Griesinger C, Zweckstetter M, Zeth K (2008) Structure of the human voltage-dependent anion channel. Proc Natl Acad Sci USA 105:15370–15375

Becker L, Bannwarth M, Meisinger C, Hill K, Model K, Krimmer T, Casadio R, Truscott KN, Schulz GE, Pfanner N, Wagner R (2005) Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore. J Mol Biol 353:1011–1020

Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N (2009) Importing mitochondrial proteins: machineries and mechanisms. Cell 138:628–644

Endo T, Yamano K (2009) Multiple pathways for mitochondrial protein traffic. Biol Chem 390:723–730

Endo T, Yamano K, Kawano S (2011) Structural insight into the mitochondrial protein import system. Biochim Biophys Acta 1808:955–970

Engelhardt H, Meins T, Poynor M, Adams V, Nussberger S, Welte W, Zeth K (2007) High level expression, refolding and probing the natural fold of the human voltage-dependent anion channel isoforms I and II. J Membr Biol 216:93–105

Fabian H, Schultz CP (2000) Fourier transform infrared spectroscopy in peptide and protein analysis. In: Meyers RA (ed) Encyclopedia of analytical chemistry. Wiley, Chichester, pp 5779–5803

Fersht A (1999) Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. W.H. Freeman and Company, pp 1–631

Harsman A, Krüger V, Bartsch P, Honigmann A, Schmidt O, Rao S, Meisinger C, Wagner R (2010) Protein conducting nanopores. J Phys Condens Matter 22:454102

Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, Pfanner N (1998) Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature 395:516–521

Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. Science 321:1206–1210

Humphries AD, Streimann IC, Stojanovski D, Johnston AJ, Yano M, Hoogenraad NJ, Ryan MT (2005) Dissection of the mitochondrial import and assembly pathway for human Tom40. J Biol Chem 280:11535–11543

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195–202

Kinoshita J, Mihara K, Oka T (2007) Identification and characterization of a new Tom40 isoform, a central component of mitochondrial outer membrane translocase. J Biochem 141:897–906

Künkele KP, Heins S, Dembowski M, Nargang FE, Benz R, Thieffry M, Walz J, Lill R, Nussberger S, Neupert W (1998) The preprotein translocation channel of the outer membrane of mitochondria. Cell 93:1009–1019

Macasev D, Newbigin E, Whelan J, Lithgow T (2000) How do plant mitochondria avoid importing chloroplast proteins? Components of the import apparatus Tom20 and Tom22 from *Arabidopsis* differ from their fungal counterparts. Plant Physiol 123:811–816

Macasev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, Lithgow T (2004) Tom22´, an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. Mol Biol Evol 21:1557–1564

Mager F, Sokolova L, Lintzel J, Brutschy B, Nussberger S (2010) LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM. J Phys Condens Matter 22:454132

Malia TJ, Wagner G (2007) NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-xL. Biochemistry 46:514–525

Model K, Prinz T, Ruiz T, Radermacher M, Krimmer T, Kühlbrandt W, Pfanner N, Meisinger C (2002) Protein translocase of the outer mitochondrial membrane: role of import receptors in the structural organization of the TOM complex. J Mol Biol 316:657–666

Model K, Meisinger C, Kühlbrandt W (2008) Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase. J Mol Biol 383:1049–1057

Mokranjac D, Neupert W (2009) Thirty years of protein translocation into mitochondria: unexpectedly complex and still puzzling. Biochim Biophys Acta 1793:33–41

Nussberger S, Neupert W (2002) Protein translocation across the outer membrane of mitochondria: structure and function of the TOM complex. In: Kasianowicz JJ, Kellermayer MSZ, Deamer DW (eds) Structure and dynamics of confined polymers. NATO science series. Kluwer Academic, Dordrecht, pp 67–84

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. Protein Sci 4:2411–2423

Pace CN, Hebert EJ, Shaw KL, Schell D, Both V, Krajcikova D, Sevcik J, Wilson KS, Dauter Z, Hartley RW, Grimsley GR (1998) Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. J Mol Biol 279:271–286

Perry AJ, Hulett JM, Likic VA, Lithgow T, Gooley PR (2006) Convergent evolution of receptors for protein import into mitochondria. Curr Biol 16:221–229

Poynor M, Eckert R, Nussberger S (2008) Dynamics of the preprotein translocation channel of the outer membrane of mitochondria. Biophys J 95:1511–1522

Prokisch H, Nussberger S, Westermann B (2002) Protein import into mitochondria of *Neurospora crassa*. Fungal Genet Biol 36:85–90

Pusnik M, Charriere F, Maser P, Waller RF, Dagley MJ, Lithgow T, Schneider A (2009) The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane. Mol Biol Evol 26:671–680

Quinlan ME, Kerkhoff E (2008) Actin nucleation: bacteria get in-Spired. Nat Cell Biol 10:13–15

Romero-Ruiz M, Mahendran KR, Eckert R, Winterhalter M, Nussberger S (2010) Interactions of mitochondrial presequence peptides with the mitochondrial outer membrane preprotein translocase TOM. Biophys J 99:774–781

Schwartz M, Matouschek A (1999) The dimensions of the protein import channels in the outer and inner mitochondrial membranes. Proc Natl Acad Sci USA 9:13086–13090

Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287:252–260

Sreerama N, Woody RW (2003) Structural composition of betaI- and betaII-proteins. Protein Sci 12:384–388

Sreerama N, Woody RW (2004) Computation and analysis of protein circular dichroism spectra. Methods Enzymol 383:318–351

Suzuki H, Okazawa Y, Komiya T, Saeki K, Mekada E, Kitada S, Ito A, Mihara K (2000) Characterization of Rat TOM40, a central component of the preprotein translocase of the mitochondrial outer membrane. J Biol Chem 275:37930–37936

Suzuki H, Kadowaki T, Maeda M, Sasaki H, Nabekura J, Sakaguchi M, Mihara K (2004) Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched beta-structure. J Biol Chem 279:50619–50629

Taylor RD, McHale BJ, Nargang FE (2003) Characterization of *Neurospora crassa* Tom40-deficient mutants and effect of specific mutations on Tom40 assembly. J Biol Chem 278:765–775

Ujwal R, Cascio D, Colletier JP, Faham S, Zhang J, Toro L, Ping P, Abramson J (2008) The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating. Proc Natl Acad Sci USA 105:17742–17747

Werhahn W, Niemeyer A, Jänsch L, Kruft VV, Schmitz UK, Braun HP (2001) Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from *Arabidopsis*. Identification of multiple forms of TOM20. Plant Physiol 125:943–954

Werhahn W, Jänsch L, Braun HP (2003) Identification of novel subunits of the TOM complex of *Arabidopsis thaliana*. Plant Physiol Biochem 41:407–416

Zeth K (2010) Structure and evolution of mitochondrial outer membrane proteins of beta-barrel topology. Biochim Biophys Acta 1797:1292–1299

# Supplementary Material

**Supplementary Table 1:** Secondary structure estimation of human Tom40 AΔN and BΔN at different temperatures

| Protein | T [°C] | $\lambda_{crossover}$ [nm] | $\lambda_{min}$ [nm] | $\Theta_{min}$ [$10^{-3}$ deg cm$^2$ dmol$^{-1}$] | $T_m$ [°C] | α-Helix [%] | β-Sheet [%] | Turn [%] | Random coil [%] |
|---|---|---|---|---|---|---|---|---|---|
| hTom40AΔN | 20 | 206 | 217 | -11.0 | 73.1± | 20 | 36 | 21 | 24 |
| | 30 | 206 | 217 | -11-0 | 0.7 | 23 | 35 | 21 | 23 |
| | 40 | 206 | 217 | -10.9 | | 22 | 34 | 21 | 24 |
| | 50 | 206 | 217 | -10.5 | | 18 | 38 | 21 | 24 |
| | 60 | 206.5 | 217.5 | -10.5 | | 23 | 35 | 21 | 22 |
| | 70 | 208 | 223 | -5.0 | | 5 | 43 | 24 | 27 |
| hTom40BΔN | 20 | 203 | 215 | -10.9 | 74.1± | 24 | 32 | 20 | 26 |
| | 30 | 203.5 | 215 | -10.5 | 0.7 | 24 | 31 | 21 | 25 |
| | 40 | 203.5 | 215 | -10.5 | | 22 | 33 | 22 | 25 |
| | 50 | 204 | 216 | -10.0 | | 22 | 32 | 21 | 26 |
| | 60 | 205 | 216 | -9.5 | | 20 | 33 | 21 | 26 |
| | 70 | 205.5 | 217.5 | -8.5 | | 12 | 37 | 23 | 27 |

Secondary structure percentages were determined from CD spectra via the CDpro package using CDSSTR, CONTIN/LL and SELCON3 (Sreerama & Woody. 2000; Sreerama & Woody. 2003; Sreerama & Woody. 2004).

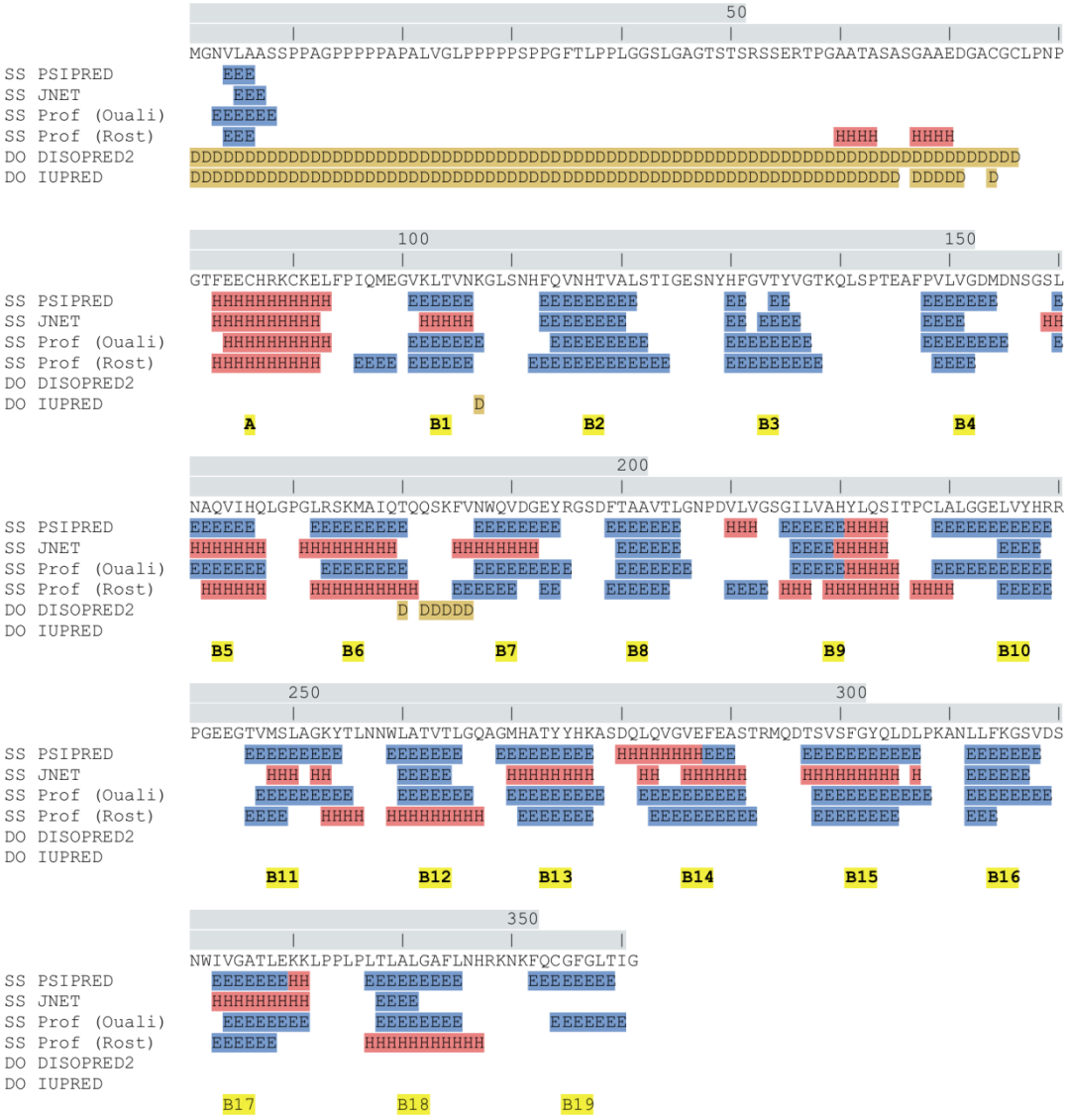



**Fig. S1 Secondary Structure analysis of hTom40 AΔN.** Secondary structure elements were predicted using the Quick2D package of the Toolkit-Server http://toolkit.lmb.uni-muenchen.de/quick2_d.

sp|O96008|TOM40_HUMAN Mitochondrial import receptor subunit (hTom40AΔN)

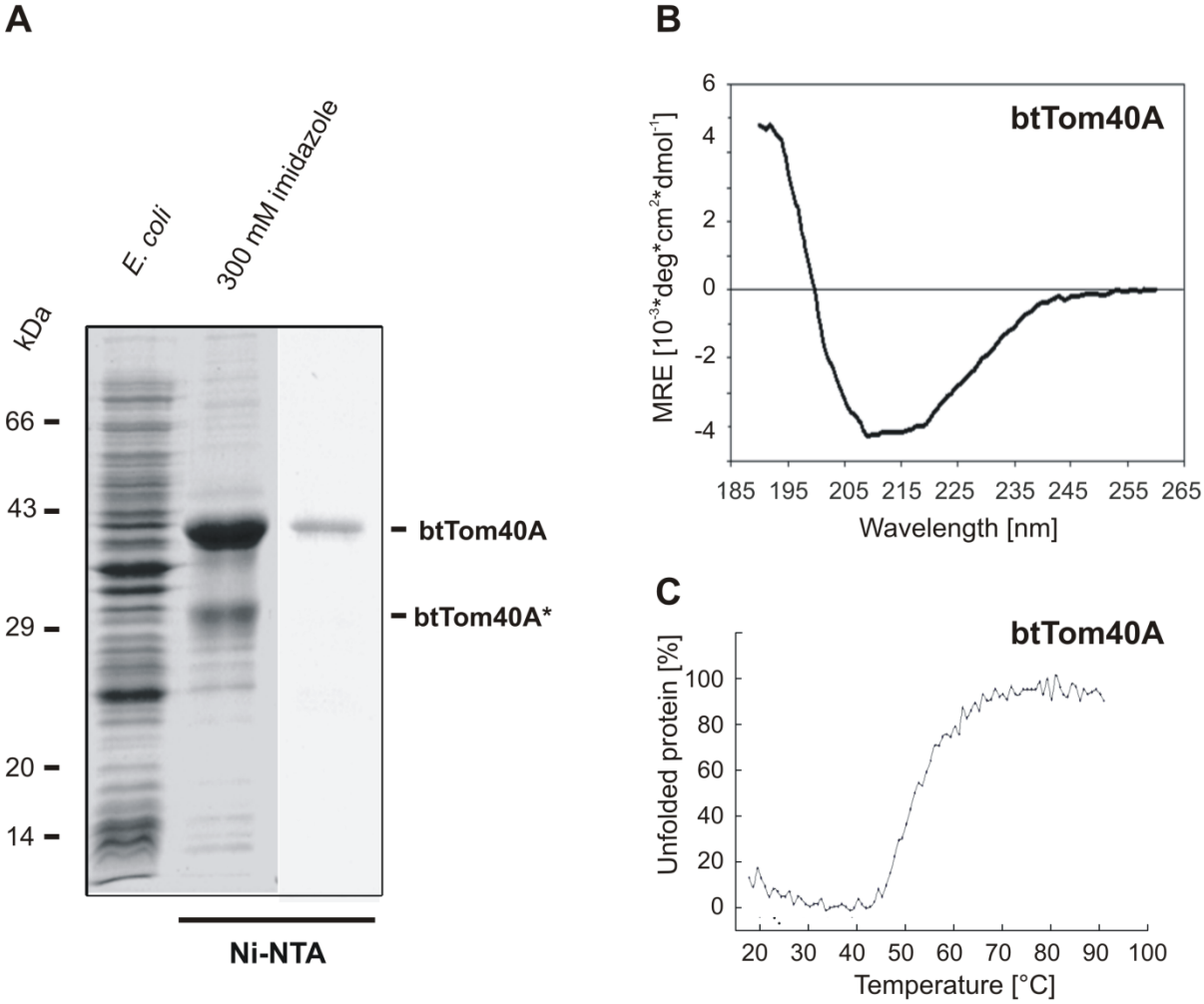



**Fig. S2 Expression, purification and refolding of bovine Tom40A.** (A) Tom40 Isoform A from *Bos taurus* L. (BtTom40A) was isolated from *E. coli* inclusion bodies containing recombinant BtTom40A with an N-terminal hexahistidinyl-tag. Inclusion bodies were solubilized in 8 M urea, 1 mM β-mercaptoethanol, 20 mM Tris (pH 8.0) and loaded onto a Ni-NTA affinity column. Under denaturing conditions BtTom40A was eluted with 300 mM imidazole, dialzyed against 8 M urea, 1 mM β-mercaptoethanol and 20 mM Tris (pH 8.0) and rebound on the same column. Degraded bovine Tom40A (BtTom40A*) was released from the column with 0.5 % polyoxyethylene monolauryl ether (Brij 35), 1 mM β-mercaptoethanol, 20 mM Tris (pH 8.0) at 600 mM imidazole. Full length BtTom40A was subsequently eluted with 8 M urea, 1 mM beta mercaptoethanol, 20 mM Tirs (pH 8.0) and 1 M imidazole and refolded in 0.1 % LDAO by fast dilution. (B and C) UV spectrum and melting curve of BtTom40A recorded as described in Fig. 2.

# Manuscript IV

## *In vivo* and *in vitro* investigations of the N-terminal domain of Tom40 from *Neurospora crassa*

**Dennis Gessmann[1], Franjo Artukovic[1], Thomas Christott[1], Nancy E. Go[2], Sebastian Leptihn[3], Andreas Kuhn[3], Frank E. Nargang[2] and Stephan Nussberger[1]**

[1]Biophysics Department, Institute of Biology, University of Stuttgart, 70550 Stuttgart, Germany

[2]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

[3]Institute of Microbiology and Molecular Biology, University of Hohenheim, 70599 Stuttgart, Germany

*Running head:* Structural role of the Tom40 N-terminal domain

*Keywords:* Tom40, β-barrel, eukaryotic porin, VDAC, TOM complex, protein translocation

*Abbreviations*: LDAO, lauryl-dimethylamine-oxide; Brij35, polyoxyethylene monolauryl ether; MWCO, molecular weight cutoff; IBs, inclusion bodies; Far-UV CD, far ultraviolet circular dichroism; SD, standard deviation; *Nc*Tom40, Tom40 from *Neurospora crassa*; *Sc*Tom40, Tom40 from *Saccharomyces cerevisiae*; hTom40A, Tom40 isoform A from *Homo sapiens*; TOM complex, translocase of the outer mitochondrial membrane; human VDAC1, VDAC isoform 1 from *Homo sapiens*; TM, transmembrane; wt, wild type

## Abstract

The general structure of eukaryotic and prokaryotic β-barrel membrane proteins is a defined network of β-strands. Functional groups exist either as N- or C-terminal extensions or as integrated parts of distinct loops. These groups can be peptide loops, small β-strands or short α-helical segments. Through interaction with specific amino acid residues inside the β-barrel, these functional groups influence for example voltage-dependent gating of the protein pore. Another function is the stabilization of the β-barrel structure or interaction between subunits of oligomeric β-barrel complexes. Here, we used *in vivo* and *in vitro* investigations to identify the structural role of the N-terminal domain of the eukaryotic porin Tom40 from *Neurospora crassa*. Examination of the homology model of *Nc*Tom40 revealed that the N-terminal amino acid residues of Tom40 can be separated into two different domains, the inner-barrel and outer-barrel N-terminus. Both domains prevent heat-induced dysfunction of Tom40 in mitochondria and mutations in both parts resulted in severe growth defects for *N. crassa*, independently. We suggest a specific function for the less conserved outer-barrel N-terminus of Tom40 in between different species. In contrast, a common role for the highly conserved inner-barrel N-terminus is proposed for all species. We further determined that the inner-barrel N-terminus folds into an α-helix but is not necessary for the integrity or the stability of the Tom40 β-barrel. However, a conserved amino acid residue in the inner-barrel N-terminus of Tom40, Ile47, is important for the structural integrity of the N-terminal α-helix. We suggest that proper folding of the N-terminal α-helix is dependent on the crucial Ile47, which then allows the correct function of Tom40 in the mitochondrial outer membrane.

## Introduction

The outer membrane of gram-negative bacteria and mitochondria is coated with integral membrane proteins of the β-barrel type designed to allow passive transport of hydrophilic molecules across the permeation barrier as well as membrane biogenesis (for review see (Zeth and Thein, 2010)). Their inadvertent contribution is important to the development of neurodegenerative diseases in humans due to characteristic dysfunctions in the outer mitochondrial

membrane (OM), as it was shown for e.g. Parkinson's disease recently (Yoshii, Kishi *et al.*, 2011). Therefore, a deeper understanding of their structure and functional properties is essential.

In general, the structure of prokaryotic β-barrels (also termed 'porins') consists of an even number of β-strands (Schulz, 2000; Schulz, 2002; Arnold, Poynor *et al.*, 2007). Eukaryotic porins are grouped into a distinct class of β-barrels, the mitochondrial porins (Pusnik, Charriere *et al.*, 2009), defined by their specific protein fold (for review see (Zeth, 2010)). The eukaryotic porin VDAC (voltage-dependent anion channel) of the outer mitochondrial membrane exhibits a nineteen-stranded β-barrel (Bayrhuber, Meins *et al.*, 2008; Hiller, Garces *et al.*, 2008; Ujwal, Cascio *et al.*, 2008). A similar nineteen-stranded fold is predicted for the general preprotein import pore, Tom40, of the translocase of the outer mitochondrial membrane (TOM) (Bayrhuber, Meins *et al.*, 2008; Zeth, 2010). A second characteristic structural feature of Tom40 and VDAC is an α-helical domain at the N-terminus, which is attached to the inner barrel wall and reaches through the pore from the cytosol to the intermembrane space (Bayrhuber, Meins *et al.*, 2008; Hiller, Garces *et al.*, 2008; Ujwal, Cascio *et al.*, 2008; Zeth, 2010; Gessmann, Flinner *et al.*, 2011).

To this date, there exists a lack of high atomic resolution structures of eukaryotic TM β-barrels (for review see (Fairman, Noinaj *et al.*, 2011)) which provide a detailed explanation of their functional properties, e.g. oligomerization of porin complexes or voltage-dependent gating. VDAC and Tom40 are essential for metabolite flux (Antos, Stobienia *et al.*, 2001), apoptosis (Malia and Wagner, 2007; Ott, Norberg *et al.*, 2007) and preprotein translocation in the OM (Hill, Model *et al.*, 1998), but it is not clear as to how they carry out these functions in particular. Nevertheless, for VDAC, the α-helical domain was shown to be involved in apoptotic events (Abu-Hamad, Arbel *et al.*, 2009), voltage-dependent gating (Popp, Court *et al.*, 1996; Choudhary, Ujwal *et al.*, 2010) and the maintenance of the integrity of the β-barrel (Schneider, Etzkorn *et al.*, 2010). Hence, an equally important function of the N-terminal domain is suggested for Tom40 in preprotein translocation.

In this report, we set about to determine the functional role of the Tom40 N-terminus and its involvement in the overall Tom40 structure. Examinations of the homology model of *Nc*Tom40 revealed that the N-terminus can be separated

into two different domains, the outer-barrel and inner-barrel N-terminus. *In vivo* investigations allowed allocation of an important function to the less conserved outer-barrel N-terminus as well as the highly conserved inner-barrel N-terminus. Moreover, both parts prevent heat-induced dysfunction of Tom40 in *Neurospora crassa* mitochondria independently. We suggest that the outer-barrel N-terminus of Tom40 serves specific functions in between different species, e.g. fungi/mammals. Further, we designed recombinant variants of *Nc*Tom40 to specifically investigate the functional role of the inner-barrel N-terminus *in vitro*. We show that the inner-barrel N-terminus folds into an α-helix in detergent micelles. However, this α-helical domain is not necessary for the stability or the integrity of the Tom40 β-barrel in detergent solutions. Nevertheless, we identified a key amino acid residue in the inner-barrel N-terminus of *Nc*Tom40, Ile47, which is necessary to maintain the structural integrity of the N-terminal α-helix in detergent micelles. We suggest that the presence of the N-terminal α-helix inside the Tom40 pore as well as a specific fold is necessary to enable proper function of Tom40 in the TOM complex of the outer mitochondrial membrane.

## Materials and Methods

### Generation of mutant *Neurospora crassa* strains

Mutations to generate the N-terminal deleted forms of Tom40 were created in a previously described plasmid (Rapaport, Taylor *et al.*, 2001) containing the genomic version of *Neurospora crassa tom40* and a bleomycin resistance cassette. For each deletion, PCR mutagenesis was used to introduce a pair of PmeI sites into the *tom40* gene on each side of the region to be removed. In each case, introduction of the restriction sites resulted in alteration of codons immediately surrounding the area targeted for deletion. The altered plasmids were cut with PmeI and religated. This gave rise to plasmid 40NΔ12-1, encoding a Tom40 protein in which amimo acid residues 5 to 59 were replaced by a Lys and a His residue; plasmid 40NΔ13-1, encoding a Tom40 where residues 5 to 45 were replaced with a Lys and a Gln residue; and plasmid 40NΔ23-1, encoding a Tom40 where residues 43 to 59 were replaced by Phe, Lys, and His residues. The *tom40* gene in each of the plasmids was entirely sequenced to rule out any unintended mutations introduced by PCR.

To develop strains expressing only the mutant versions of Tom40, a sheltered heterokaryon for *tom40* was developed in heterokaryon strain HP1 as described in detail previously (Nargang, Künkele *et al.*, 1995; Nargang and Rapaport, 2007). Briefly, the HP1 heterokaryon contains two different nuclei. One is auxotrophic for pantothenate and carries resistance to benomyl. The other is auxotrophic for histidine and carries resistance to p-fluorophenylalanine. The *tom40* gene was replaced by a hygromycin resistance cassette in the histidine requiring nucleus of HP1 to give the sheltered heterokaryon Tom40KO-5. This strain was transformed with each of the plasmids described above. Transformants were plated on medium containing histidine and p-fluorophenylalanine to select for homokaryons of the *tom40* knockout nucleus that had been rescued by the *tom40* mutant constructs. Transformants were checked for histidine requirement to confirm they were homokaryons expressing the mutant form of the gene. Mutant *N. crassa* strains were termed *N.crassa* Δ13 and Δ23, with KQ replacing amino acids 5-45 and FKH replacing amino acids 43-59, respectively. Both Tom40 wild type and mutant proteins were applied to SDS-PAGE and analyzed through western-blot after isolation of mitochondria according to well established protocols (Nargang and Rapaport, 2007). Tim13 served as a loading control and Tom40 antibodies were specific for a C-terminal peptide of Tom40.

**Growth of *N. crassa* strains on agar plates**

Growth and treatment of *N. crassa* wild type, Δ13 and Δ23 strains was carried out according to standard protocols (Davis and De Serres, 1970), respectively. Rate of growth was investigated by spotting 10 µl of conidia suspension on agar plates using concentrations of 50,000, 5,000, 500 and 50 conidia per ml. Conidia suspensions were made of fresh conidia and agar plates contained Vogel's solution, vitamin solution, sucrose, yeast extract, peptone from casein and glycerol. Agar plates of mutant strains Δ13 and Δ23 were supplemented with histidine. Pictures of plates were taken after one day of growth at 30, 37 and 42 °C, respectively.

## Cloning of Tom40 for expression in *E. coli*

The gene of *tom40* from *N. crassa* was synthesized with truncations at the N- and C-termini in two variations, namely *(ΔN40ΔC15)* and *(ΔN60ΔC15)*, including a C-terminal hexahistidinyl-tag, respectively. The gene encoding for *tom40 (ΔN40ΔC15)* was further mutated to *tom40 (I47PΔN40ΔC15)* (Invitrogen). For better comparison to the homology model of *Nc*Tom40 (Gessmann, Flinner *et al.*, 2011), numbering of the amino acid residues in each *Nc*Tom40 mutant, e.g. I47 and L180 for (ΔN40ΔC15), refers to the amino acid sequence of the full length *Nc*Tom40 protein (Fig. 1). For the production of all *Nc*Tom40 variants, according genes were sub-cloned into pET24d vectors (Invitrogen) and transformed into *Escherichia coli* BL21 (DE3) cells, respectively.

## Expression of recombinant *Nc*Tom40 proteins and isolation of inclusion bodies

By using transformed BL21 (DE3) cells, recombinant *Nc*Tom40 proteins (ΔN40ΔC15), (ΔN60ΔC15) and (I47PΔN40ΔC15) were expressed as described (Gessmann, Mager *et al.*, 2011) with regard to protein expression of *Af*Tom40. Isolation of inclusion bodies (IBs) was conducted according to methods of (Harris, Patel *et al.*, 1986; Engel, 1987) using the protocol for *Af*Tom40 IBs isolation by (Gessmann, Mager *et al.*, 2011).

## Purification and refolding of *Nc*Tom40 proteins

Isolated IBs of *Nc*Tom40 variants (ΔN40ΔC15), (ΔN60ΔC15) and (I47P, ΔN40ΔC15) were purified in urea-containing buffers by using a Ni-Sepharose HiTrap column (1-5 ml; GE Healthcare), respectively, as described (Gessmann, Mager *et al.*, 2011). SDS-PAGE and Coomassie Brilliant Blue staining were performed to assess the purity of the proteins. Protein concentrations were determined by UV-absorbance spectroscopy using extinction coefficients of $\varepsilon_{280}$ = 33920 (ΔN40ΔC15)/ 32430 (ΔN60ΔC15)/ and 33920 (I47PΔN40ΔC15) $M^{-1}$ $cm^{-1}$ (Wilkins, Gasteiger *et al.*, 1999; Gasteiger E., 2005).

Refolding of each *Nc*Tom40 mutant was performed by fast dilution of purified protein into a detergent containing refold solution [20 mM sodium phosphate pH 7, 1 mM β-mercaptoethanol and 0.5 % (w/v) LDAO or 0.15 % (w/v) Brij35] (Gessmann, Mager *et al.*, 2011; Mager, Gessmann *et al.*, 2011). The

final concentration of the protein in the refold solution was adjusted to 0.1-0.5 mg/ml. To remove remaining urea and precipitated proteins, refold solution was dialyzed (6- to 8-kDa MWCO; Spectra/Por®) at 4 °C against 150-fold volumetric excess of refold buffer and further centrifuged at 16,060*x*g for 10 min at 4 °C (Heraeus® Biofuge® fresco).

## Circular dichroism spectroscopy

Far-UV circular dichroism (CD) spectroscopy was performed with a J-715 spectropolarimeter (Jasco Inc) in quartz cuvettes of 0.1 cm path length. *Nc*Tom40 proteins (ΔN40ΔC15), (ΔN60ΔC15) and (I47PΔN40ΔC15) were measured at a protein concentration of 0.1-0.5 mg/ml in 20 mM sodium phosphate pH 7, 1 mM β-mercaptoethanol and 0.5 % (w/v) LDAO/0.15 % (w/v) Brij35, respectively. To remove possible protein aggregates, all samples were centrifuged at 4 °C for 5 min (Heraeus® Biofuge® fresco) and filtered (Rotilabo® filter; pore size, 0.22 µm) prior to measurements.

CD spectra were recorded continuously with a scanning speed of 50 nm/min from 185 to 260 nm at 25 °C. Resolution was set to 1.0 nm with a sensitivity of 100 mdeg. For each spectrum, three consecutive scans were averaged and final spectrum was corrected by subtraction of protein-free buffer measurement. Thermal unfolding was investigated by recording the ellipticity at constant 216 nm. Samples were heated from 25 to 98 °C at 1 °C/min. Data in [mdeg] were converted to mean residue ellipticity (Θ) by using the molar protein concentration and the number of amino acid residues of the regarding *Nc*Tom40 variant. Secondary structure content of individual CD spectrum was determined as described (Mager, Gessmann *et al.*, 2011) using the CDpro package (Sreerama and Woody, 2000; Sreerama and Woody, 2003; Sreerama and Woody, 2004). All melting curves were normalized to fraction of unfolded protein $f_u(T)$ and fitted to a three-state unfolding reaction according to (Zheng and Yang, 2010; Gessmann, Mager *et al.*, 2011).

## Tryptophan fluorescence spectroscopy

Tryptophan fluorescence spectra of purified and refolded *Nc*Tom40 proteins [(ΔN40ΔC15), (ΔN60ΔC15), (I47PΔN40ΔC15)] were measured using a FP-8500 spectrofluorimeter (Jasco). Buffer conditions were as described in the

CD spectroscopy section as well as preparation of samples prior to measurements. Excitation of tryptophans was carried out at 280 nm and emission spectra were recorded from 300 to 420 nm with a scanning speed of 100 nm/min and a data interval of 0.5 nm. Band width of excitation and emission was set to 5 and 2.5 nm, respectively, with a response of 1 sec. Tryptophan fluorescence spectra of *Nc*Tom40 proteins were corrected by subtraction of protein-free background and fitted to the log-normal distribution according to (Ladokhin, Jayasinghe *et al.*, 2000; Winterfeld, Imhof *et al.*, 2009; Gessmann, Mager *et al.*, 2011).

### Computations and protein structure figures

CCP4 program suite (Winn, Ballard *et al.*, 2011) and the molecular replacement tool 'Chainsaw' (Schwarzenbacher, Godzik *et al.*, 2004; Stein, 2008) were used to generate structures of *Nc*Tom40 mutants. *Nc*Tom40 homology model served as template structure (Gessmann, Flinner *et al.*, 2011). Structure figures were created using PyMOL version 0.99 (DeLano 2006).

## Results and Discussion

### The N-terminal domain of Tom40 can be separated into two domains and is crucial for growth of *Neurospora crassa*

Upon investigation of the structural model of Tom40 from *N. crassa* (Gessmann, Flinner *et al.*, 2011), we found that the N-terminus can be separated into two distinct parts. The first part is predicted to be exterior of the pore comprising amino acid residues 1 to 40 (Fig. 1A and B) and secondary structure prediction tools (Jones, 1999; McGuffin, Bryson *et al.*, 2000) predict a random coil structure for fungal Tom40s (data not shown).

The second part is highly conserved among different species (Rapaport, Taylor *et al.*, 2001; Gessmann, Flinner *et al.*, 2011) comprising amino acid residues 41 to 60. It is predicted to form an α-helical domain inside the Tom40 pore with a short random coil towards the N-terminus (Fig. 1A, left panel) (Bayrhuber, Meins *et al.*, 2008; Zeth, 2010). Mutations of amino acid residues 41-60 of Tom40 resulted in growth defects for *N. crassa*. It was further shown that these amino acids are required for proper assembly/stability of Tom40 into the TOM complex (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003).

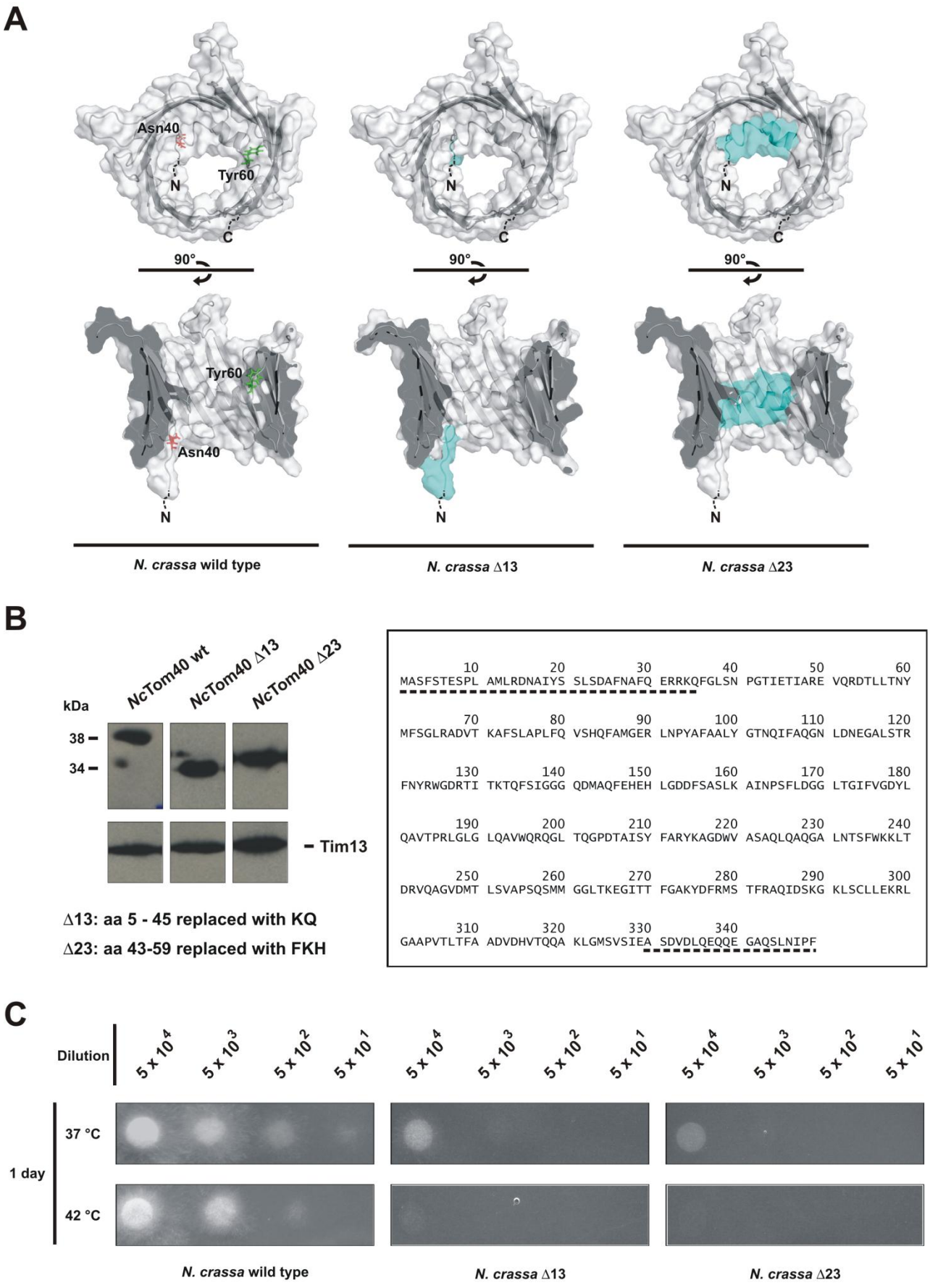



**Fig. 1** Effect of the N-terminal amino acids of Tom40 on the growth of *Neurospora crassa*. (A) Homology model of *Nc*Tom40 wild type (wt) and design of native Tom40

To determine the significance of the two parts of the N-terminal region of Tom40, defined as inner-barrel and outer-barrel N-terminus, on the growth of *Neurospora crassa* independently, we generated two *N. crassa* mutant strains, termed Δ13 and Δ23. The Tom40 of *N. crassa* Δ13 was depleted of most of its outer-barrel domain (Fig. 1A, middle panel) by replacing amino acid residues 5-45 with KQ. Hence, *Nc*Tom40Δ13 exhibits a lower molecular weight compared to wild type *Nc*Tom40 (Fig. 1B). To exclude the formation of a proper inner-barrel domain (Fig. 1A, right panel), we substituted amino acid residues 43-59 with FKH (Fig. 1B).

Conidia suspensions of wild type and mutant (Δ13 and Δ23) *N. crassa* strains were spotted on agar plates at different concentrations and grown at 30, 37 and 42 °C, respectively. Growth rates at 30 °C were similar to incubation at 37 °C for all *N. crassa* strains and are therefore not shown (Fig. 1C). Replacement of the outer-barrel N-terminal amino acids 5-45 with KQ resulted in a severe growth defect compared to the wild type (Fig. 1C). Substitution of the inner-pore N-termini amino acids 43-59 with FKH also revealed significant growth defects in line with (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003). The growth defect of *N. crassa* Δ23 was slightly more pronounced in contrast to the Δ13 mutant (Fig. 1C).

---

***(Fig. 1 continued)*** mutants Δ13 and Δ23. Cyan-colored regions in the *Nc*Tom40Δ13 and Δ23 structures refer to amino acid residues 5-45 and 43-59, respectively. Asn40 is highlighted in red and Tyr60 in green. Structure model of *Nc*Tom40 wt was used from Gessmann *et al.* (2011). (B) Western blot of wild type and mutant Δ13/Δ23 *Nc*Tom40 proteins. Mitochondria were isolated from the strains indicated at the top of the figure. Mitochondrial proteins (30 μg per lane) were subjected to SDS-PAGE and then blotted to nitrocellulose for western blot analysis. The blot was decorated with antibodies raised to a C-terminal peptide of Tom40 and with antibodies to Tim13, as a loading control. Insert in box displays full length *Nc*Tom40 amino acid sequence. Amino acid residues 5-45/43-59 of *Nc*Tom40Δ13/Δ23 were replaced with KQ and FKH, respectively. (C) Growth rates of *tom40* wild type and mutant strains at different temperatures. Fresh conidia of *N. crassa* strains were plated on agar plates, incubated at depicted temperature and photographed after one day of growth. Underlined amino acid residues in (B) represent dashed lines in (A).

At 42 °C, *N. crassa* wild type exhibited a minor growth defect compared to incubation at 37 °C (Fig. 1C). In contrast, *N. crassa* mutant strains Δ13 and Δ23 revealed substantial temperature sensitivity at 42 °C, respectively. Therefore, it can be reasoned that both parts of the N-terminus of *Nc*Tom40 prevent temperature-induced dysfunction of Tom40 in the outer membrane of *N. crassa* mitochondria.

Interestingly, amino acids 1-40 are not well conserved among Tom40s in comparison to the inner-barrel N-terminus (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003; Humphries, Streimann *et al.*, 2005). It is thus tempting to assume a specific function for the outer-barrel N-terminus of Tom40 in certain species and a common function for the inner-barrel N-terminus. For example, in Tom40 isoform A from mammals a highly conserved proline-rich region is located at the far-N-terminus. This region was shown to be not involved in mitochondrial targeting nor TOM biogenesis (Humphries, Streimann *et al.*, 2005). In contrast, depletion of amino acid residues 72-87 from human Tom40 isoform A and corresponding amino acid residues 41-60 from *Nc*Tom40 revealed TOM biogenesis deficiencies, respectively. Hence, we suggest different function for both N-terminal parts, the outer- and inner-barrel N-terminus. Additionally, a role in actin binding was proposed for the proline-rich region of hTom40A (Mager, Gessmann *et al.*, 2011). However, the nature and function of the outer-barrel N-terminus of Tom40 from certain species remains to be shown.

### *Nc*Tom40 comprises an α-helix at the N- and C-terminus

We set about to isolate native *Nc*Tom40Δ23 from outer membrane vesicles (OMVs) of *N. crassa* Δ23 mitochondria. By this means, we wanted to investigate the structure and function of the inner-barrel N-terminus in comparison to wild type Tom40 *in vitro*. However, since this has not been successful, we hypothesize that the composition of the outer mitochondrial membrane of *N. crassa* Δ23 is altered as a characteristic OMV fraction could not be gained via sucrose gradient centrifugation (Poynor, 2008) (data not shown).

With regard to the lack of native *Nc*Tom40 mutants, we used recombinant *Nc*Tom40 to investigate the structure and function of the inner-barrel N-terminus *in vitro*. We truncated amino acids 1 to 40 (ΔN40) and 334 to 349

(ΔC15) of *Nc*Tom40 to be able to specifically investigate the inner-barrel N-terminus, amino acid residues 41 to 60 (Fig. 1A). This mutant was termed *Nc*Tom40ΔN40ΔC15 (Fig. 2A). A second mutant was designed lacking amino acid residues 1 to 60 (ΔN60) in addition to the truncation of amino acid residues 334 to 349 (ΔC15), termed *Nc*Tom40ΔN60ΔC15 (Fig. 2A). Far C-terminal amino acids 334 to 349 are predicted to be not part of the last β-strand of Tom40 (Gessmann, Flinner *et al.*, 2011) and amino acids 1 to 40 comprise the outer-barrel N-terminus (Fig. 1A).

*Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 were over-expressed as inclusion bodies in *E. coli* cells and purified through nickel-affinity chromatography under denaturing conditions. Peak fractions of purified proteins were refolded into LDAO/Brij35-containing buffers and analyzed via SDS-PAGE revealing pure protein samples (Fig. 2B). To address the state of refolding, we applied far-UV circular dichroism and tryptophan fluorescence spectroscopy on both *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 (Fig. 3).

We determined the secondary structure content of *Nc*Tom40ΔN40ΔC15 to ~12 % α-helix and ~33 % β-sheet in LDAO and Brij35, respectively (Table 1). These numbers are in good agreement with previous studies on recombinant *Nc*Tom40 that was reconstituted in detergent micelles (Becker, Bannwarth *et al.*, 2005). Hence, we assume a β-barrel structure for truncated *Nc*Tom40ΔN40ΔC15 due to a predominance of β-sheet content (Table 1). Surprisingly, *Nc*Tom40ΔN40ΔC15 showed reduced α-helical content of about 8 % compared to full length *Nc*Tom40 as reported by (Becker, Bannwarth *et al.*, 2005). Prediction of the secondary structure of the non-barrel C-terminal amino acid residues of Tom40 with PSIPRED (Jones, 1999; McGuffin, Bryson *et al.*, 2000) predicted an α-helical domain for fungal Tom40s exclusively (data not shown). Since amino acid residues 1-40 of fungal Tom40s are predicted to form a random coil, we suggest that fungal Tom40s comprise an α-helix at their C-termini.

Truncation of amino acid residues 41 to 60 resulted in a significant decrease of α-helix secondary structure from ~12 to ~7 % with a slight increase in β-sheet content (Fig. 3A, table 1). In order to account for a detergent-specific effect as shown for eukaryotic porin (Shanmugavadivu, Apell *et al.*, 2007), we

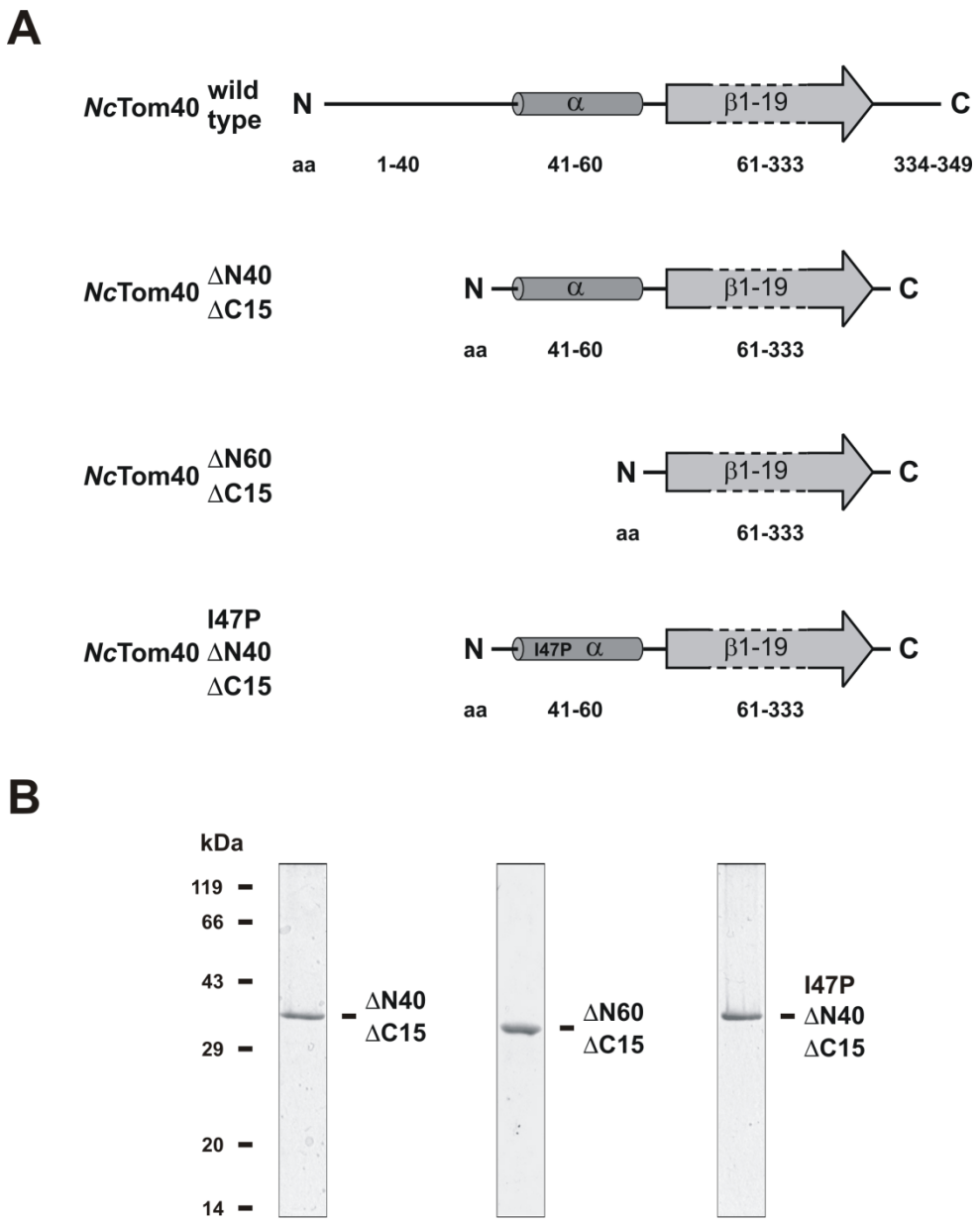



**Fig. 2** Design and purification of recombinant *Nc*Tom40 variants. (A) Schematic representation of secondary structure of *Nc*Tom40 wild type, (ΔN40ΔC15), (ΔN60ΔC15) and (I47PΔN40ΔC15). *Nc*Tom40 (ΔN40ΔC15) was further mutated to (I47PΔN40ΔC15) as depicted. (B) SDS-PAGE of purified and refolded *Nc*Tom40 variants. *Nc*Tom40 proteins were expressed as IBs in BL21 (DE3) cells. IBs were isolated and purified on a Ni-Sepharose HiTrap column under denaturing conditions. Purified proteins were reconstituted in refold buffers containing LDAO/Brij35 by fast dilution. Gels were stained with Coomassie Brilliant Blue and represent refolding of Tom40s in LDAO micelles. Purity of Tom40 proteins in Brij35-containing refold solutions was analog and SDS-PAGE gels are therefore not shown.

refolded both *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 in Brij35-containing buffers (Fig. 3A) revealing same decrease/increase in α-helix/β-sheet content (Table 1). To probe the nature of secondary structure of both these *Nc*Tom40 variants in a native-like environment, we are currently developing a method for the reconstitution of detergent-solubilized *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 into lipid vesicles, respectively.

Nevertheless, our results are in line with similar CD spectroscopy-based investigations on mammalian Tom40 isoform A from *Rattus norvegicus* (Suzuki, Kadowaki *et al.*, 2004). Suzuki *et al.* reported analogue decrease in α-helical secondary structure content upon truncation of the first 165 amino acid residues of the N-terminal domain of recombinant *Rn*Tom40A in Brij35-containing buffers. Further, recent works predicted a 19-stranded β-barrel and an N-terminally located α-helix for human Tom40 based on secondary structure assignment to the 3D structure of human VDAC1 (Bayrhuber, Meins *et al.*, 2008; Zeth, 2010). Furthermore, it was proposed early on that Tom40s adapt a similar architecture as eukaryotic porin in general, due to their common evolutionary origin (Pusnik, Charriere *et al.*, 2009). In line with our results, we conclude that the highly conserved inner-barrel N-terminus (amino acid residues 41 to 60 of *Nc*Tom40) folds into an α-helix.

**Table 1** Secondary structure content of *Nc*Tom40 variants

| Protein | Detergent | α-Helix | β-Sheet | Turn | Random coil |
|---|---|---|---|---|---|
| *Nc*Tom40ΔN40ΔC15 | LDAO | 12 ± 3 | 33 ± 2 | 22 ± 1 | 31 ± 0.2 |
| | Brij35 | 12 ± 2 | 33 ± 3 | 22 ± 2 | 29 ± 1 |
| *Nc*Tom40ΔN60ΔC15 | LDAO | 7 ± 1 | 36 ± 4 | 21 ± 2 | 32 ± 1 |
| | Brij35 | 7 ± 1 | 37 ± 3 | 22 ± 2 | 33 ± 2 |
| *Nc*Tom40I47PΔN40ΔC15 | LDAO | 10 ± 2 | 34 ± 3 | 22 ± 2 | 30 ± 2 |

Secondary structure percentages were calculated from CD data shown in Fig. 3A via CDpro (Sreerama and Woody, 2000; Sreerama and Woody, 2003; Sreerama and Woody, 2004). Values displayed are the average of CDSSTR, CONTIN/LL and SELCON3.

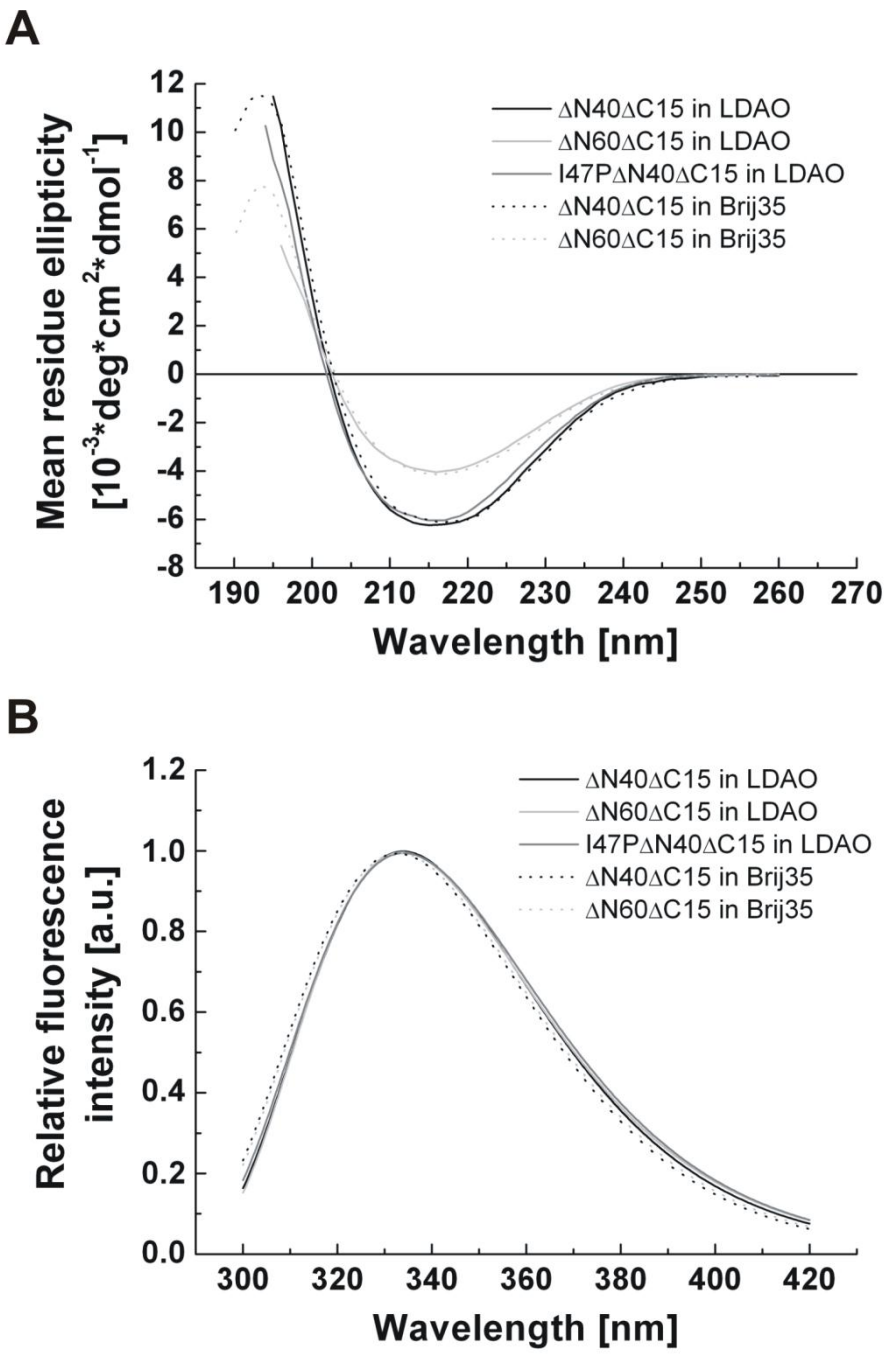



**Fig. 3** Circular dichroism and tryptophan fluorescence spectroscopy of *Nc*Tom40 variants. (A) Far-UV CD spectra of refolded *Nc*Tom40 proteins (ΔN40ΔC15), (ΔN60ΔC15) and (I47PΔN40ΔC15). Purified Tom40 proteins were refolded by fast dilution in either LDAO- or Brij35-containing buffers. Spectra were measured from 190 to 260 nm. For each spectrum, 3 scans were averaged and background-corrected by subtraction of protein-free samples. If necessary, noisy data below ~200 nm was removed. For each *Nc*Tom40 protein, three independent refolding experiments were conducted, whereas only one representative spectrum is shown for each variant. Protein concentrations of independent refold experiments varied in between 0.1-0.5 mg/ml. (B) Tryptophan fluorescence emission spectra of refolded *Nc*Tom40 proteins.

## The N-terminal α-helix of Tom40 is not necessary for β-barrel integrity and stability in detergent micelles

The inner-barrel N-terminus of *Nc*Tom40 is predicted to be present in the pore interior of Tom40 (Fig. 1A) (Zeth 2010; Gessmann, Flinner *et al.* 2011) and seems to affect growth of the complete organism (Fig. 1C). It is therefore likely that the N-terminal α-helix of Tom40 (Fig. 2A) functions as stabilizing element inside the Tom40 pore and maintains β-barrel integrity in terms of an in-plug domain as suggested previously (Naveed, Jackups *et al.*, 2009; Gessmann, Flinner *et al.*, 2011). We set about to address this in-plug hypothesis *in vitro*.

First, we conducted tryptophan fluorescence spectroscopy on recombinant *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 reconstituted in LDAO- or Brij35-micelles. Both Tom40 mutants comprise four Trp-residues which are equally distributed through the amino acid sequence (Fig. 1B). Hence, changes of the tertiary structure in between *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 are likely to result in a different local environment of the Trp-residues, which can be traced by their respective emission spectra. We found that truncation of amino acids 41 to 60 did not change the Trp-fluorescence emission spectrum in either LDAO- or Brij35-containing buffers (Fig. 3B). We therefore conclude that the N-terminal α-helix of *Nc*Tom40 is not necessary for maintaining β-barrel integrity in detergent micelles.

Secondly, we utilized CD spectroscopy to investigate the proposed stabilizing effect of the N-terminal α-helix on *Nc*Tom40. Recombinant *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 reconstituted in LDAO-micelles (Fig. 3A) were measured at constant 216 nm and heated from 25 to 98 °C, respectively. Both Tom40 variants followed three-state unfolding pathways (Fig. 4A) in line with previous works on fungal Tom40 (Gessmann, Mager *et al.*, 2011). *Nc*Tom40ΔN60ΔC15 exhibited a right-shifted melting curve with a higher

---

***(Fig. 3 continued)*** Excitation of tryptophans was performed at 280 nm and emission spectrum was recorded from 300 to 420 nm. *Nc*Tom40 proteins were refolded (*n*=3) and investigated under the same conditions as described in (A). Emission spectra were normalized to relative fluorescence intensity and fitted to the log-normal distribution according to (Ladokhin, Jayasinghe *et al.*, 2000; Winterfeld, Imhof *et al.*, 2009). (a.u.), arbitrary units.

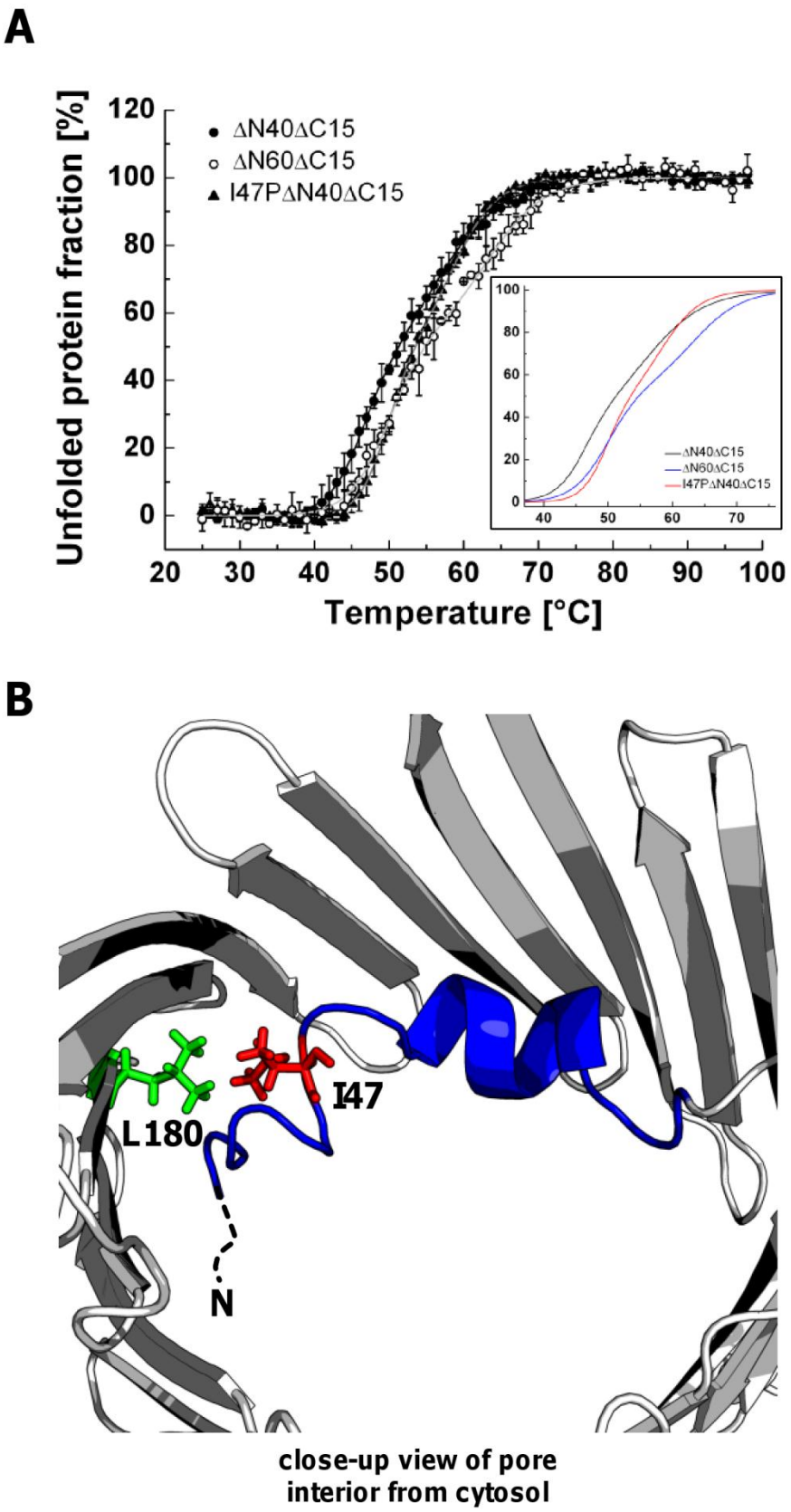



**Fig. 4** Thermal unfolding of *Nc*Tom40 mutants and structural illustration of mutations. (A) Thermal denaturation was investigated through CD by recording the change in ellipticity at 216 nm for wt and mutant *Nc*Tom40 proteins. Conditions were the same as described in Fig. 3A for all proteins. Data points were normalized to minimum and

apparent melting temperature of about 4 °C in comparison to *Nc*Tom40ΔN40ΔC15 (Fig. 4A). Thus, amino acid residues 41 to 60 of *Nc*Tom40 do not stabilize the Tom40 β-barrel in terms of an in-plug domain. Hence, the essential function of the highly conserved inner-barrel N-terminus for the assembly/stability of Tom40 into the TOM complex (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003) remains to be determined.

The truncated from of full length *Nc*Tom40, *Nc*Tom40ΔN60ΔC15, is predicted to comprise the β-barrel solely (Fig. 1A and 4B) and the melting curve is therefore restricted to the unfolding of the β-hairpins (Fig. 4A). Thus, the melting curve of *Nc*Tom40ΔN40ΔC15 displays an overlay of α-helix and β-barrel denaturation (Fig. 4A). Further, the melting curve of *Nc*Tom40ΔN60ΔC15 exhibited a more pronounced three-state unfolding pathway than *Nc*Tom40ΔN40ΔC15, since the intermediate state (IS) at ~50 % of unfolded protein fraction of *Nc*Tom40ΔN60ΔC15 is more defined than the IS of *Nc*Tom40ΔN40ΔC15 (Fig. 4A). This feature clearly discriminates both unfolding curves from each other. Recent works on human Tom40 isoform A and B reported that upon thermal denaturation an α-helical domain unfolds before the β-barrel breaks apart (Mager, Gessmann *et al.*, 2011). We suggest that the left-shift of the melting curve of *Nc*Tom40ΔN40ΔC15, which accounts for loss of secondary structure, is due to unfolding of the N-terminal α-helix before the β-barrel. In other words, we assume that the N-terminal α-helix unfolds independently from the Tom40 β-barrel and that its presence does not destabilize the overall pore structure.

---

***(Fig. 4 continued)*** maximum percentages of unfolded protein fraction (*n*=3, error bars=SD). Unfolding curves were fitted by assuming a three-state unfolding mechanism for *Nc*Tom40 (ΔN40ΔC15; black line), (ΔN60ΔC15; light gray/blue line) and (I47PΔN40ΔC15; gray/red line) and were re-plotted in (insert) for better comparison. (B) Structure model of *Nc*Tom40ΔN40ΔC15. Isoleucine 47 and leucine 180 are shown as stick representation in red and green, respectively. Amino acid residues 41 to 60 are highlighted in blue. I47 is predicted to interact with L180 (Gessmann, Flinner *et al.*, 2011) and amino acid residues 41-60 represent a highly conserved region among Tom40 proteins (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003).

Involvement of the N-terminal α-helix in maintaining the pore stability and integrity was suggested for Tom40 and reported for VDAC (Schneider, Etzkorn *et al.*, 2010; Gessmann, Flinner *et al.*, 2011). However, our *in vitro* investigations of *Nc*Tom40 in detergent micelles did not show such a function for the N-terminal α-helix. Preliminary results gained from electrophysiological investigations of Tom40 mutants in lipid-bilayers, which are currently carried out, point towards a more simple function. We observed high current fluctuations for Tom40 pores that lacked the N-terminal α-helix when constant voltages were applied (data not shown). Similar results were gained for eukaryotic porin VDAC (De Pinto, Reina *et al.*, 2008). We assume that positioning of the N-terminal α-helix inside the Tom40/VDAC pore prevents the β-barrel from fluctuation due to the lateral pressure of the lipid bilayer. In doing so, proper function of Tom40/VDAC in the outer mitochondrial membrane is maintained as the pore becomes more rigid enabling efficient transport of pre-proteins/metabolites, respectively.

Unfolding of solely the Tom40 β-barrel did not follow a two-state unfolding pathway in detergent micelles (Fig. 4A, blue line). This finding contradicts the concerted folding/unfolding pathways reported for prokaryotic β-barrels (Kleinschmidt, den Blaauwen *et al.*, 1999; Kleinschmidt and Tamm, 2002; Conlan and Bayley, 2003). Denaturation of the Tom40 β-barrel (Fig. 4A, blue line) seemed to take place through different intermediate states under the applied conditions. It can be suggested that this happens upon unfolding of individual β-hairpins of the Tom40 β-barrel. This is in line with recent works which revealed via atomic force microscopy that unfolding of the prokaryotic β-barrel OmpG and the refolding of partly unfolded OmpG takes place hairpin after hairpin (Sapra, Damaghi *et al.*, 2009; Damaghi, Koster *et al.*, 2011).

## Ile47 of *Nc*Tom40 maintains the structural integrity of the N-terminal α-helix

Mutation of the single amino acid Leu66 of yeast Tom40 to proline resulted in a temperature-sensitive *Saccharomyces cerevisiae* strain suggesting a crucial role for Leu66 (Rapaport, Taylor *et al.*, 2001). Leu66 of *Sc*Tom40 corresponds to Ile47 of *Nc*Tom40 in the highly conserved inner-barrel N-terminus (Fig. 4B). It has been suggested that the amino acid residue in this position interacts with

another conserved amino acid residue located in the putative β-strand 9 in all Tom40 and VDAC proteins, e.g. Leu180 for *Nc*Tom40 (Fig. 4B) (Schneider, Etzkorn *et al.*, 2010; Gessmann, Flinner *et al.*, 2011). To investigate the role of Ile47 of *Nc*Tom40 *in vitro*, we mutated Ile47 of *Nc*Tom40ΔN40ΔC15 to proline (Fig. 2A). *Nc*Tom40I47PΔN40ΔC15 was expressed as IBs in *E. coli* cells, purified and refolded into LDAO-containing buffers analog to *Nc*Tom40ΔN40ΔC15 (Fig. 2B).

Mutation of Ile47 to Pro resulted in a minor change of the CD spectrum (Fig. 3A) and the secondary structure content displayed a slight decrease of α-helix from ~12 to ~10 % and an increase of ~33 to ~34 % of β-sheet (Table 1). However, the emission spectrum of the tryptophan residues of *Nc*Tom40I47PΔN40ΔC15 was not altered with regard to *Nc*Tom40ΔN40ΔC15 (Fig. 3B). Thus, replacement of Ile47 with Pro altered the α-helix/β-sheet secondary structure content only slightly without influencing the tertiary structure of *Nc*Tom40ΔN40ΔC15 in detergent micelles. However, mutation of Ile47 to Pro resulted in a significant change of the melting curve towards a clearly more sigmoid form (Fig. 4A). The unfolding pathway of *Nc*Tom40I47PΔN40ΔC15 first follows the unfolding of *Nc*Tom40ΔN60ΔC15 until approximately its midpoint and then emerges into the melting curve of *Nc*Tom40ΔN40ΔC15. In other words, the melting curve of *Nc*Tom40I47PΔN40ΔC15 displays a hybrid between *Nc*Tom40ΔN40ΔC15 and *Nc*Tom40ΔN60ΔC15 (Fig. 4A).

Loss of secondary structure of *Nc*Tom40ΔN40ΔC15 in between ~40 and ~55 °C was not seen if Ile47 was deleted or mutated to Pro (Fig. 4A). Moreover, in between ~40 and ~55 °C truncation of the whole inner-barrel N-terminus had the same effect as mutation of Ile47 to Pro. We conclude that the right-shift of the melting curve of *Nc*Tom40I47PΔN40ΔC15 in between ~40 and ~55 °C accounts for decreased α-helical content, in line with truncation of the N-terminal α-helix (Fig. 4A, table 1). Remaining α-helical portion and the β-barrel of *Nc*Tom40I47PΔN40ΔC15 unfold above 55 °C (Fig. 4A), as the melting curve of *Nc*Tom40I47PΔN40ΔC15 crosses over into the melting curve of *Nc*Tom40ΔN40ΔC15 (Fig. 4A).

We therefore propose that Ile47 is necessary to maintain a certain structural integrity of the N-terminal α-helix within the Tom40 β-barrel. We

suggest this is accomplished through hydrophobic contacts of Ile47 to Leu180 on β-strand 9 of *Nc*Tom40 (Fig. 4B) as predicted before (Schneider, Etzkorn *et al.*, 2010; Gessmann, Flinner *et al.*, 2011). This might explain the crucial role reported for Leu66 of yeast Tom40, which corresponds to Ile47 in *Nc*Tom40 (Rapaport, Taylor *et al.*, 2001). Mutation to proline might not only disturb the structure of the N-terminal α-helix (Table 1) but also the interaction to β-strand 9. We thus assume that the structural integrity as well as the presence of the highly conserved amino acid residues of the inner-barrel N-terminus itself are necessary to enable correct assembly of Tom40 into the TOM complex and its stability within the complex (Rapaport, Taylor *et al.*, 2001; Taylor, McHale *et al.*, 2003).

## Conclusion

The N-terminal domain of Tom40 can be separated into two distinct domains, the inner-barrel N-terminus and the outer-barrel N-terminus. *In vivo* investigations revealed that both parts prevent heat-induced dysfunction of Tom40 in the outer membrane of *N. crassa* mitochondria independently. Moreover, our results show an important role for the less conserved outer-barrel N-terminus in addition to the highly conserved inner-barrel N-terminus.

Through *in vitro* investigations of recombinant *Nc*Tom40 variants, we showed an α-helical formation for the inner-barrel N-terminus. This N-terminal α-helix is not necessary for the stability nor the integrity of the Tom40 pore. However, a key residue, Ile47 of *Nc*Tom40, is necessary to maintain the structural integrity of the N-terminal α-helix within the Tom40 pore.

Our results indicate a dynamic conformation for the inner-barrel N-terminus of fungal Tom40. Importantly, this does not exclude the possibility of interactions in between the β-barrel and the N-terminal α-helix in line with the reported helix anchor region inside the Tom40 β-barrel (Gessmann, Flinner *et al.*, 2011). However, the helix anchor region of Tom40 might therefore function more in terms of a ‘guide’ positioning the N-terminal α-helix at its ‘correct’ position inside the pore rather than making it rigid. However, it is important to repeat these experiments in a lipid bilayer to monitor the effect of the N-terminal α-helix as well as Ile47 on the *Nc*Tom40 structure under native-like conditions.

Nevertheless, a dynamic conformation for the N-terminal α-helix might explain the distinct gating behavior reported for *Nc*Tom40 (Poynor, Eckert *et al.*, 2008; Romero-Ruiz, Mahendran *et al.*, 2010) and speak against a stabilizing function. Changes in between different conductance states might be caused by movement of the α-helix inside the Tom40 pore in agreement with eukaryotic porin (Ujwal, Cascio *et al.*, 2008). It should be noted that other works oppose a simple movement of the α-helix of VDAC as the helix might also rotate and/or involve interactions with certain loops (Choudhary, Ujwal *et al.*, 2010). Equal considerations have to be made for Tom40.

## Acknowledgements

The authors thank Beate Nitschke for expert technical assistance. We thank Drs. Jie Liang, Oliver Mirus, Enrico Schleiff, Uwe Gerken and Hammad Naveed for helpful discussions. FEN was supported by a grant from the Canadian Institutes of Health Research and SN by the Competence Network on "Functional Nanostructures" of the Baden-Württemberg Foundation.

## Contribution to this manuscript

FEN and NEG generated *N. crassa* mutant strains. DG and Beate Nitschke carried out growth of wild type and mutant *N. crassa* strains on agar plates. DG conducted expression of *Nc*Tom40 proteins and isolation of inclusion bodies. FA, TC and DG purified and refolded *Nc*Tom40 proteins. DG and FA conducted tryptophan fluorescence measurements. DG, SL and TC performed reconstitution of Tom40 proteins into lipid vesicles as well as CD spectrum measurements. SL and DG carried out thermal unfolding of *Nc*Tom40 variants via CD spectroscopy. DG analyzed and fitted spectroscopic data sets and generated all figures. DG wrote the manuscript with the contribution of FEN and SN. DG and SN defined the experiments. AK, SL, FEN and SN supervised the project.

## References


Abu-Hamad, S., N. Arbel*, et al.* (2009). "The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins." *J Cell Sci* **122** (Pt 11), 1906-16.

Antos, N., O. Stobienia*, et al.* (2001). "Under conditions of insufficient permeability of VDAC1, external NADH may use the TOM complex channel to cross the outer membrane of *Saccharomyces cerevisiae* mitochondria." *J Bioenerg Biomembr* **33** (2), 119-26.

Arnold, T., M. Poynor*, et al.* (2007). "Gene duplication of the eight-stranded β-barrel OmpX produces a functional pore: a scenario for the evolution of transmembrane beta-barrels." *J Mol Biol* **366** (4), 1174-84.

Bayrhuber, M., T. Meins*, et al.* (2008). "Structure of the human voltage-dependent anion channel." *Proc Natl Acad Sci U S A* **105** (40), 15370-5.

Becker, L., M. Bannwarth*, et al.* (2005). "Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore." *J Mol Biol* **353** (5), 1011-20.

Choudhary, O. P., R. Ujwal*, et al.* (2010). "The electrostatics of VDAC: implications for selectivity and gating." *J Mol Biol* **396** (3), 580-92.

Conlan, S. and H. Bayley (2003). "Folding of a monomeric porin, OmpG, in detergent solution." *Biochemistry* **42** (31), 9453-65.

Damaghi, M., S. Koster*, et al.* (2011). "One β-hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane β-barrel protein OmpG." *Angew Chem Int Ed Engl* **50** (32), 7422-4.

Davis, R. H. and F. J. De Serres (1970). "Genetic and microbiological research techniques for *Neurospora crassa*." *Methods Enzymol.* **17**, 79-143.

De Pinto, V., S. Reina*, et al.* (2008). "Structure of the voltage dependent anion channel: state of the art." *J Bioenerg Biomembr* **40** (3), 139-47.

DeLano, W. L. (2006). "The PyMOL Molecular Graphics System, Version 0.99rc6; Schrödinger, LLC."

Engel, J. (1987). "DNA cloning: A practical approach." *Acta Biotechnologica* **9** (5), 254.

Fairman, J. W., N. Noinaj*, et al.* (2011). "The structural biology of β-barrel membrane proteins: a summary of recent reports." *Curr Opin Struct Biol* **21** (4), 523-31.

Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). "Protein Identification and Analysis Tools on the ExPASy Server." *The Proteomics Protocols Handbook*, 571-607

Gessmann, D., N. Flinner*, et al.* (2011). "Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry." *Biochim Biophys Acta*. **1807** (12), 1647-57

Gessmann, D., F. Mager*, et al.* (2011). "Improving the Resistance of a Eukaryotic β-Barrel Protein to Thermal and Chemical Perturbations." *J Mol Biol*. **413** (1), 150-61

Harris, T. J., T. Patel*, et al.* (1986). "Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in *Escherichia coli*." *Mol Biol Med* **3** (3), 279-92.

Hill, K., K. Model*, et al.* (1998). "Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins." *Nature* **395** (6701), 516-21.

Hiller, S., R. G. Garces*, et al.* (2008). "Solution structure of the integral human membrane protein VDAC-1 in detergent micelles." *Science* **321** (5893), 1206-10.

Humphries, A. D., I. C. Streimann*, et al.* (2005). "Dissection of the mitochondrial import and assembly pathway for human Tom40." *J Biol Chem* **280** (12), 11535-43.

Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." *J Mol Biol* **292** (2), 195-202.

Kleinschmidt, J. H., T. den Blaauwen*, et al.* (1999). "Outer membrane protein A of *Escherichia coli* inserts and folds into lipid bilayers by a concerted mechanism." *Biochemistry* **38** (16), 5006-16.

Kleinschmidt, J. H. and L. K. Tamm (2002). "Secondary and tertiary structure formation of the β-barrel membrane protein OmpA is synchronized and depends on membrane thickness." *J Mol Biol* **324** (2), 319-30.

Ladokhin, A. S., S. Jayasinghe*, et al.* (2000). "How to measure and analyze tryptophan fluorescence in membranes properly, and why bother?" *Anal Biochem* **285** (2), 235-45.

Mager, F., D. Gessmann*, et al.* (2011). "Functional refolding and characterization of two Tom40 isoforms from human mitochondria." *J Membr Biol* **242** (1), 11-21.

Malia, T. J. and G. Wagner (2007). "NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-xL." *Biochemistry* **46** (2), 514-25.

McGuffin, L. J., K. Bryson*, et al.* (2000). "The PSIPRED protein structure prediction server." *Bioinformatics* **16** (4), 404-5.

Nargang, F. E., K. P. Künkele*, et al.* (1995). "'Sheltered disruption' of *Neurospora crassa* MOM22, an essential component of the mitochondrial protein import complex." *EMBO J* **14** (6), 1099-1108.

Nargang, F. E. and D. Rapaport (2007). "*Neurospora crassa* as a model organism for mitochondrial biogenesis." *Methods Mol Biol* **372**, 107-23.

Naveed, H., R. Jackups, Jr.*, et al.* (2009). "Predicting weakly stable regions, oligomerization state, and protein-protein interfaces in transmembrane domains of outer membrane proteins." *Proc Natl Acad Sci U S A* **106** (31), 12735-40.

Ott, M., E. Norberg*, et al.* (2007). "The mitochondrial TOM complex is required for tBid/Bax-induced cytochrome c release." *J Biol Chem* **282** (38), 27633-9.

Popp, B., D. A. Court*, et al.* (1996). "The role of the N and C termini of recombinant *Neurospora* mitochondrial porin in channel formation and voltage-dependent gating." *J Biol Chem* **271** (23), 13593-13599.

Poynor, M. (2008). "Dynamics of the preprotein translocase of the mitochondrial outer membrane." (ISBN 978-3-89963-751-9).

Poynor, M., R. Eckert*, et al.* (2008). "Dynamics of the Preprotein Translocation Channel of the Outer Membrane of Mitochondria." *Biophys J.* **95**, 1511-1522.

Pusnik, M., F. Charriere*, et al.* (2009). "The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane." *Mol Biol Evol* **26** (3), 671-80.

Rapaport, D., R. D. Taylor*, et al.* (2001). "Structural requirements of Tom40 for assembly into preexisting TOM complexes of mitochondria." *Mol Biol Cell* **12**, 1189-1198.

Romero-Ruiz, M., K. R. Mahendran*, et al.* (2010). "Interactions of mitochondrial presequence peptides with the mitochondrial outer membrane preprotein translocase TOM." *Biophys J* **99** (3), 774-81.

Sapra, K. T., M. Damaghi*, et al.* (2009). "One β-hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG." *Angew Chem Int Ed Engl* **48** (44), 8306-8.

Schneider, R., M. Etzkorn*, et al.* (2010). "The native conformation of the human VDAC1 N terminus." *Angew Chem Int Ed Engl* **49** (10), 1882-5.

Schulz, G. E. (2000). "β-Barrel membrane proteins." *Curr. Opin. Struct. Biol.* **10**, 443-447.

Schulz, G. E. (2002). "The structure of bacterial outer membrane proteins." *Biochim Biophys Acta* **1565** (2), 308-17.

Schwarzenbacher, R., A. Godzik*, et al.* (2004). "The importance of alignment accuracy for molecular replacement." *Acta Crystallogr D Biol Crystallogr* **60** (Pt 7), 1229-36.

Shanmugavadivu, B., H. J. Apell*, et al.* (2007). "Correct folding of the β-barrel of the human membrane protein VDAC requires a lipid bilayer." *J Mol Biol* **368** (1), 66-78.

Sreerama, N. and R. W. Woody (2000). "Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set." *Anal Biochem* **287** (2), 252-60.

Sreerama, N. and R. W. Woody (2003). "Structural composition of betaI- and betaII-proteins." *Protein Sci* **12** (2), 384-8.

Sreerama, N. and R. W. Woody (2004). "On the analysis of membrane protein circular dichroism spectra." *Protein Sci* **13** (1), 100-12.

Stein, N. (2008). "CHAINSAW: a program for mutating pdb files used as templates in molecular replacement." *Journal of Applied Crystallography* **41** (3), 641-643.

Suzuki, H., T. Kadowaki*, et al.* (2004). "Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched beta-structure." *J Biol Chem* **279** (48), 50619-29.

Taylor, R. D., B. J. McHale*, et al.* (2003). "Characterization of *Neurospora crassa* Tom40-deficient mutants and effect of specific mutations on Tom40 assembly." *J Biol Chem* **278** (2), 765-75.

Ujwal, R., D. Cascio*, et al.* (2008). "The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating." *Proc Natl Acad Sci U S A* **105** (46), 17742-7.

Wilkins, M. R., E. Gasteiger*, et al.* (1999). "Protein identification and analysis tools in the ExPASy server." *Methods Mol Biol* **112**, 531-52.

Winn, M. D., C. C. Ballard*, et al.* (2011). "Overview of the CCP4 suite and current developments." *Acta Crystallogr D Biol Crystallogr* **67** (Pt 4), 235-42.

Winterfeld, S., N. Imhof*, et al.* (2009). "Substrate-induced conformational change of the *Escherichia coli* membrane insertase YidC." *Biochemistry* **48** (28), 6684-91.

Yoshii, S. R., C. Kishi*, et al.* (2011). "Parkin mediates proteasome-dependent protein degradation and rupture of the outer mitochondrial membrane." *J Biol Chem* **286** (22), 19630-40.

Zeth, K. (2010). "Structure and evolution of mitochondrial outer membrane proteins of β-barrel topology." *Biochim Biophys Acta* **1797** (6-7), 1292-9.

Zeth, K. and M. Thein (2010). "Porins in prokaryotes and eukaryotes: common themes and variations." *Biochem J* **431** (1), 13-22.

Zheng, X. Y. and B. S. Yang (2010). "An improved method for measuring the stability of a three-state unfolding protein." *Chinese Science Bulletin* **55** (36), 4120-4124.

# Summary and outlook

In this thesis, new insights into the structural elements of the preprotein translocation channel, Tom40, of the outer mitochondrial membrane were gained. In doing so, new bioinformatics methods were developed and employed that allowed determination of structural elements of Tom40 and the rational design of modified transmembrane β-barrel proteins with desired structural properties.

A recently developed physical interaction model, TmSIP, was verified in this work and allowed the design of Tom40 pores with increased resistance to chemical and thermal perturbations. Further, the oligomeric state was influenced and a modified Tom40 could be retrieved in strictly its monomeric form. This system displays a very useful tool for nanopore technology (Howorka and Siwy, 2009) as it approaches nanopore-engineering from a strictly theoretical basis in contrast to the known strategies (Seitz, Thoma *et al.*, 2010; Grosse, Essen *et al.*, 2011). Further, our strategy might offer a new approach for the crystallization of TM β-barrels since a higher stability and a monomeric form are essential for producing protein crystals diffracting at atomic resolution (Seitz, Bocola *et al.*, 2007; Hocker, Lochner *et al.*, 2009). In addition, TmSIP was recently used to determine the oligomerization contact sites of eukaryotic porin VDAC (Geula, Naveed *et al.*, 2011) and was implemented in a computational method for the prediction of three-dimensional structures of pro-/eukaryotic TM β-barrels (Naveed, Xu *et al.*, 2011) emphasizing its usefulness.

In the absence of a crystal/NMR structure of Tom40, a homology model of *Nc*Tom40 was generated based on the solved crystal structure of mouse VDAC-1 and amino acid sequence information of ~270 different Tom40 and ~480 VDAC amino acid sequences for refinement. Combining mass-spectrometry with time-limited proteolysis of detergent-embedded *Nc*Tom40 further allowed verification of the model structure and assignment of the membrane topology of *Nc*Tom40 to detergent-embedded VDAC-1 from human. In doing so, Tom40 was revealed as 19-stranded β-barrel protein with an N-terminal α-helix inside the pore. Moreover, the *Nc*Tom40 model structure could be used as a mask to

highlight conserved properties between Tom40 and VDAC. In doing so, two conserved domains were found on the inside of the pore, termed 'polar slide' and 'helix anchor' region, clearly discriminating Tom40 from VDAC. The 'helix anchor region' is most likely involved in binding the α-helix, whereas the 'polar slide' might enable preprotein translocation. On the outer surface of the pore, two conserved patches were revealed possibly binding other TOM components or bridging components of the TIM machinery.

On the basis of the developed homology model of *Nc*Tom40 a structure-guided investigation of the N-terminal domain of *Nc*Tom40 was carried out. Examination of the Tom40 model structure revealed two different domains for the N-terminal amino acid residues, termed the inner-barrel and outer-barrel N-terminus. Through *in vivo* techniques it was found that both the inner- and outer-barrel N-terminus prevented heat-induced dysfunction of Tom40 in *N. crassa* mitochondria. Further, using recombinant *Nc*Tom40 variants it was determined that the inner-barrel N-terminus formed the predicted N-terminal α-helix and unfolded independently from the β-barrel. However, the N-terminal α-helix was not necessary to maintain the β-barrel structure or the stability of the Tom40 pore. Based on our structure model of Tom40, a conserved amino acid residue in the inner-barrel domain, Ile47 of *Nc*Tom40, displayed a possible mutation site to investigate the N-terminal α-helix. It was revealed that Ile47 was necessary to maintain the structural integrity of the N-terminal α-helix.

These results further allow a structure-guided approach for the investigation of Tom40. For example, it was discovered recently that two conserved patches in the TM domain of Tom22 interact with Tom40 (Shiota, Mabuchi *et al.*, 2011). These contact sites might be cross referenced to the conserved patches on the outside of our Tom40 model in future works. Further, TmSIP should display a useful tool for the interpretation of these protein-protein interfaces on the basis of weakly stable regions of Tom40 within the TOM complex *in vivo*.

In conclusion, combining theoretical approaches with practical techniques allowed new insights into the structure and function of the eukaryotic porin Tom40 and may serve as basis for future works on other TM β-barrel proteins.

# Abbreviations

| | |
|---|---|
| *Af*Tom40 | Tom40 from *Aspergillus fumigatus* |
| BAM | β-barrel assembly machinery |
| BamA | Omp85 |
| Brij35 | polyoxyethyleneglycol dodecyl ether; or polyoxyethylene monolauryl ether |
| CD | circular dichroism |
| CHAPS | 3-[(3-cholamidopropyl) dimethyl-ammonio]-1-propanesulfonate |
| CMC | critical micelle concentration |
| DDM | *n*-dodecyl-β-D-maltoside |
| DHFR | dihydrofolate reductase |
| DMSO | dimethylsulfoxide |
| ESI | electrospray ionization |
| FTIR | fourier transform infrared spectroscopy |
| GnHCl | guanidine hydrochloride |
| hTom40A | human Tom40 isoform A |
| hVDAC(-)1 | human VDAC isoform 1 |
| IBs | inclusion bodies |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| LB | Luria Bertani |
| LDAO | lauryldimethylamine-oxide; or *N*,*N*-dimethyldodecylamine-*N*-oxide |
| MALDI | matrix-assisted-laser-desorption/ionization |
| ML | maximum likelihood |
| MRE | mean residue ellipticity |
| MS | mass spectrometry |
| MSA | multiple sequence alignment |
| mVDAC(-)1 | VDAC isoform 1 from *Mus musculus* |
| MWCO | molecular weight cutoff |
| *Nc*Tom40 | Tom40 from *Neurospora crassa* |
| Ni-NTA | nickel-nitrilotriacetic acid |
| NMR | nuclear magnetic resonance |

| | |
|---|---|
| OD | optical density |
| OG/β-OG | *n*-octyl-β-D-glucopyranoside |
| OMP | outer membrane protein |
| oPOE | *n*-octyl-polyoxyethylene |
| PAGE | polyacrylamide gel electrophoresis |
| PMSF | phenylmethylsulfonyl fluoride |
| SAM | sorting and assembly machinery |
| *Sc*Tom40 | Tom40 from *Saccharomyces cerevisiae* |
| SD | standard deviation |
| SDS | sodium dodecyl sulfate |
| TFA | trifluoroacetic acid |
| TIM | translocase of the inner (mitochondrial) membrane |
| TM | transmembrane |
| $T_m$ | melting temperature |
| TOB | topogenesis of mitochondrial outer membrane β-barrels |
| TOF | time-of-flight |
| TOM | translocase of the outer (mitochondrial) membrane |
| Tris | tris(hydroxymethyl)aminomethane |
| UV | ultraviolet |
| VDAC | voltage dependent anion channel |
| wt | wild type |

# Acknowledgements

First, I would like to thank my supervisor Professor Stephan Nussberger for giving me the opportunity to work in the exciting field of protein translocation. His continuous support and advice are greatly appreciated. Moreover, he never hesitated to 'shape' and encourage me for the world of science. I am very thankful for that.

A very special thank goes out to Beate for assisting me in every possible way, as well as all my other coworkers at the Institute of Biology in Stuttgart, but especially Frauke, Simon, Mercedes, Lorena, and Melissa. I would like to thank all past and present members of the peer-mentoring group for their support and help.

I am grateful to a number of students with whom I worked with at the biophysical department in Stuttgart, namely Franjo, Thomas, Sara, Andrea, Jan, Tobias, Manami and Mariana.

A very special thank goes out to my project partners Drs. Thomas Arnold and Dirk Linke for all their help and the fruitful discussions.

Equal thanks go out to Drs. Uwe Gerken and Sebastian Leptihn for all their help and suggestions regarding everything there is in spectroscopy. I thank Professor Andreas Kuhn for letting me do work in his laboratory and his kind support during my time as a PhD-student.

I would like to acknowledge Professor Enrico Schleiff for giving me the opportunity to collaborate with his group. A special thank goes out to Dr. Oliver Mirus and Nadine Flinner for working with me on the model structure of Tom40.

I am very thankful to Professor Jie Liang for conducting the project on the stabilization of eukaryotic porins with me. I benefited greatly from our joint research and appreciate all his kindness and the opportunity to travel to Chicago. I had a wonderful time visiting the city and meeting his group and

collaborators. Further, I thank Hammad Naveed for all his nice support and profitable contribution to our research.

I would like to thank the following people and the regarding people of their groups for their productive help on my research: Drs. Jens Pfannstiel, Robin Ghosh, Andrei Lupas, Frank E. Nargang, Ulrich Keyser and Kornelius Zeth.

A special thank goes out to Thomas Christott for designing the cover of the print edition of this work.

I thank my family and my friends for all their support and for always believing in me.

I thank Emily Luiken for love and friendship.

Last but not least, this thesis would not have been possible without the help and contribution of many people. I thank everyone that I forgot to mention here.

# Curriculum Vitae

## Personal details

| | |
|---|---|
| Name: | Dennis Benjamin Gessmann |
| Date and place of birth: | 13.03.1983, Ostfildern |
| Nationality: | German |

## Education

| | |
|---|---|
| Aug. 2009 – Dec. 2011 | PhD student at the Institute of Biology, Department of Biophysics, University of Stuttgart |
| Nov. 2008 – Jul. 2009 | Diploma thesis at the Institute of Biology, Department of Biophysics, University of Stuttgart |
| Jul. 2007 – Apr. 2008 | Pre-diploma thesis at the Bio21 Institute, University of Melbourne and the Institute of Biochemistry, University of Stuttgart |
| Oct. 2003 – Sept. 2008 | Studies in 'Technische Biologie' at the University of Stuttgart |
| Jul. 2003 | Abitur at the Otto-Hahn Gymnasium, Ostfildern |

## Publications

**Improving the resistance of a eukaryotic beta-barrel protein to thermal and chemical perturbations**; Gessmann, D., Mager, F., Naveed, H., Arnold, T., Weirich, S., Linke, D., Liang, J., Nussberger, S.; *Journal of Molecular Biology* (October 2011), Volume 413, Issue 1, Pages 150-161

**Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry**; Gessmann, D., Flinner, N., Pfannstiel, J., Schlosinger, A., Schleiff, E., Nussberger, S. & Mirus, O.; *Biochimica et Biophysica Acta (BBA) – Bioenergetics* (December 2011), Volume 1807, Issue 12, Pages 1647-1657

**Functional refolding and characterization of two Tom40 isoforms from human mitochondria**; Mager, F., Gessmann, D., Nussberger, S., Zeth, K.; *Journal of Membrane Biology* (June 2011), Volume 242, Number/Issue 1, Pages 11-21

***In vivo* and *in vitro* investigations of the N-terminal domain of Tom40 from *Neurospora crassa***; Gessmann, D., Artukovic, F., Christott, T., Go, N.E., Leptihn, S., Kuhn, A., Nussberger, S.; *in preparation*

# Reference List

**Manuscript I**

Reprinted from Journal of Molecular Biology, Volume 413 (Issue 1), Improving the resistance of a eukaryotic β-barrel protein to thermal and chemical perturbations, Gessmann, D., Mager, F., Naveed, H., Arnold, T., Weirich, S., Linke, D., Liang, J., Nussberger, S., Pages 150-161, ©2011, with permission from Elsevier

**Manuscript II**

Reprinted from Biochimica et Biophysica Acta (BBA) - Bioenergetics, Volume 1807 (Issue 12), Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry, Gessmann, D., Flinner, N., Pfannstiel, J., Schlösinger, A., Schleiff, E., Nussberger, S., Mirus, O., Pages 1647-1657, ©2011, with permission from Elsevier

**Manuscript III**

Reprinted from Journal of Membrane Biology, Volume 242 (Number 1), Functional refolding and characterization of two Tom40 isoforms from human mitochondria, Mager, F., Gessmann, D., Nussberger, S., Zeth, K., Pages 11-21, ©2011, with permission from Springer

**Introduction Figure 4**

Reprinted from Journal of Molecular Biology, Volume 162 (Issue 1), Structure of β-sheets: Origin of the right-handed twist and of the increased stability of antiparallel over parallel sheets, Kuo-Chen Chou, Marcia Pottle, George Némethy, Yuzo Ueda, Harold A. Scheraga, Pages 89-112, ©1982, with permission from Elsevier

**Introduction Figure 5**

Reprinted from Journal of Molecular Biology, Volume 128 (Issue 1), Gene duplications in the structural evolution of chymotrypsin, A. D. McLachlan, Pages 49-79, ©1979, with permission from Elsevier

# Bibliography


Abu-Hamad, S., N. Arbel*, et al.* (2009). "The VDAC1 N-terminus is essential both for apoptosis and the protective effect of anti-apoptotic proteins." *J Cell Sci* **122** (Pt 11), 1906-16.

Adams, K. L. and J. D. Palmer (2003). "Evolution of mitochondrial gene content: gene loss and transfer to the nucleus." *Mol Phylogenet Evol* **29** (3), 380-95.

Ahn, V. E., E. I. Lo*, et al.* (2004). "A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin." *Embo J* **23** (15), 2931-41.

Ahting, U., M. Thieffry*, et al.* (2001). "Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria." *J Cell Biol* **153** (6), 1151-60.

Ahting, U., C. Thun*, et al.* (1999). "The TOM core complex: the general protein import pore of the outer membrane of mitochondria." *J Cell Biol* **147** (5), 959-68.

Alberti, S. (1998). "A phosphoinositide-binding sequence is shared by PH domain target molecules--a model for the binding of PH domains to proteins." *Proteins* **31** (1), 1-9.

Aram, L., S. Geula*, et al.* (2010). "VDAC1 cysteine residues: topology and function in channel activity and apoptosis." *Biochem J* **427** (3), 445-54.

Arkin, I. T., A. T. Brunger*, et al.* (1997). "Are there dominant membrane protein families with a given number of helices?" *Proteins* **28** (4), 465-6.

Arnold, T., M. Poynor*, et al.* (2007). "Gene duplication of the eight-stranded β-barrel OmpX produces a functional pore: a scenario for the evolution of transmembrane β-barrels." *J Mol Biol* **366** (4), 1174-84.

Ashida, T., I. Tanaka*, et al.* (1981). " β-pleated sheets in oligopeptide crystals." *Int J Pept Protein Res* **17** (3), 322-9.

Baker, K. P., A. Schaniel*, et al.* (1990). "A yeast mitochondrial outer membrane protein essential for protein import and cell viability." *Nature* **348** (6302), 605-9.

Barlow, D. J. and J. M. Thornton (1988). "Helix geometry in proteins." *J Mol Biol* **201** (3), 601-19.

Basle, A., G. Rummel*, et al.* (2006). "Crystal structure of osmoporin OmpC from *E. coli* at 2.0 Å." *J Mol Biol* **362** (5), 933-42.

Bayley, H. (1997). "Toxin structure: part of a hole?" *Curr Biol* **7** (12), R763-7.

Bayley, H. and P. S. Cremer (2001). "Stochastic sensors inspired by biology." *Nature* **413** (6852), 226-30.

Bayrhuber, M., T. Meins*, et al.* (2008). "Structure of the human voltage-dependent anion channel." *Proc Natl Acad Sci U S A* **105** (40), 15370-5.

Becker, L., M. Bannwarth*, et al.* (2005). "Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore." *J Mol Biol* **353** (5), 1011-20.

Becker, T., L. S. Wenz*, et al.* (2011). "Biogenesis of mitochondria: dual role of Tom7 in modulating assembly of the preprotein translocase of the outer membrane." *J Mol Biol* **405** (1), 113-24.

Bigelow, H. R., D. S. Petrey*, et al.* (2004). "Predicting transmembrane β-barrels in proteomes." *Nucleic Acids Res* **32** (8), 2566-77.

Burger, G., M. W. Gray*, et al.* (2003). "Mitochondrial genomes: anything goes." *Trends Genet* **19** (12), 709-16.

Burgess, N. K., T. P. Dao*, et al.* (2008). "β-barrel proteins that reside in the *Escherichia coli* outer membrane in vivo demonstrate varied folding behavior in vitro." *J Biol Chem* **283** (39), 26748-58.

Cavalier-Smith, T. (2006). "Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium." *Proc Biol Sci* **273** (1596), 1943-52.

Chacinska, A., C. M. Koehler*, et al.* (2009). "Importing mitochondrial proteins: machineries and mechanisms." *Cell* **138** (4), 628-44.

Chan, N. C., V. A. Likic*, et al.* (2006). "The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi." *J Mol Biol* **358** (4), 1010-22.

Chew, O., R. Lister*, et al.* (2004). "A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor." *FEBS Lett* **557** (1-3), 109-14.

Chothia, C. (1973). "Conformation of twisted β-pleated sheets in proteins." *J Mol Biol* **75** (2), 295-302.

Chothia, C., M. Levitt*, et al.* (1977). "Structure of proteins: packing of α-helices and pleated sheets." *Proc Natl Acad Sci U S A* **74** (10), 4130-4.

Chou, K. C., M. Pottle*, et al.* (1982). "Structure of β-sheets. Origin of the right-handed twist and of the increased stability of antiparallel over parallel sheets." *J Mol Biol* **162** (1), 89-112.

Choudhary, O. P., R. Ujwal*, et al.* (2010). "The electrostatics of VDAC: implications for selectivity and gating." *J Mol Biol* **396** (3), 580-92.

Clantin, B., A. S. Delattre*, et al.* (2007). "Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily." *Science* **317** (5840), 957-61.

Conlan, S. and H. Bayley (2003). "Folding of a monomeric porin, OmpG, in detergent solution." *Biochemistry* **42** (31), 9453-65.

Cowan, S. W., T. Schirmer*, et al.* (1992). "Crystal structures explain functional properties of two *E. coli* porins." *Nature* **358** (6389), 727-33.

Crick, F., H., C. (1953). "The packing of α-helices: simple coiled-coils." *Acta Cryst.* (6), 689-697

Dagley, M. J., P. Dolezal*, et al.* (2009). "The protein import channel in the outer mitosomal membrane of *Giardia intestinalis*." *Mol Biol Evol*.

Damaghi, M., S. Koster*, et al.* (2011). "One β-hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane β-barrel protein OmpG." *Angew Chem Int Ed Engl* **50** (32), 7422-4.

De Pinto, V., S. Reina*, et al.* (2008). "Structure of the voltage dependent anion channel: state of the art." *J Bioenerg Biomembr* **40** (3), 139-47.

Deisenhofer, J., O. Epp*, et al.* (1985). "Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution." *Nature* **318** (6047), 618-624.

Deisenhofer, J., O. Epp*, et al.* (1995). "Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodopseudomonas viridis*." *J Mol Biol* **246** (3), 429-57.

Dekker, P. J., M. T. Ryan*, et al.* (1998). "Preprotein translocase of the outer mitochondrial membrane: molecular dissection and assembly of the general import pore complex." *Mol Cell Biol* **18** (11), 6515-24.

DeLano, W. L. (2006). "The PyMOL Molecular Graphics System, Version 0.99rc6; Schrödinger, LLC."

Dembowski, M., K. Künkele*, et al.* (2001). "Assembly of Tom6 and Tom7 into the TOM core complex of *Neurospora crassa*." *J. Biol. Chem.* **276**, 17679-17685.

Dietmeier, K., A. Hönlinger*, et al.* (1997). "Tom5 functionally links mitochondrial preprotein receptors to the general import pore." *Nature* **388** (6638), 195-200.

Dolezal, P., M. J. Dagley*, et al.* (2010). "The essentials of protein import in the degenerate mitochondrion of *Entamoeba histolytica*." *PLoS Pathog* **6** (3), e1000812.

Dolezal, P., V. Likic*, et al.* (2006). "Evolution of the molecular machines for protein import into mitochondria." *Science* **313** (5785), 314-8.

Donohue, J. (1953). "Hydrogen Bonded Helical Configurations of the Polypeptide Chain." *Proc Natl Acad Sci U S A* **39** (6), 470-8.

Eilers, M., A. B. Patel*, et al.* (2002). "Comparison of helix interactions in membrane and soluble alpha-bundle proteins." *Biophys J* **82** (5), 2720-36.

Engelman, D. M. (2005). "Membranes are more mosaic than fluid." *Nature* **438** (7068), 578-80.

Esaki, M., T. Kanamori*, et al.* (2003). "Tom40 protein import channel binds to non-native proteins and prevents their aggregation." *Nat Struct Biol* **10** (12), 988-94.

Fairman, J. W., N. Noinaj*, et al.* (2011). "The structural biology of β-barrel membrane proteins: a summary of recent reports." *Curr Opin Struct Biol* **21** (4), 523-31.

Ferguson, A. D., E. Hofmann*, et al.* (1998). "Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide." *Science* **282** (5397), 2215-20.

Forte, M., D. Adelsberger-Mangan*, et al.* (1987). "Purification and characterization of the voltage-dependent anion channel from the outer mitochondrial membrane of yeast." *J Membr Biol* **99** (1), 65-72.

Freeman, T. C., Jr. and W. C. Wimley (2010). "A highly accurate statistical approach for the prediction of transmembrane β-barrels." *Bioinformatics* **26** (16), 1965-74.

Gabriel, K., B. Egan*, et al.* (2003). "Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins." *Embo J* **22** (10), 2380-6.

Gatzeva-Topalova, P. Z., T. A. Walton*, et al.* (2008). "Crystal structure of YaeT: conformational flexibility and substrate recognition." *Structure* **16** (12), 1873-81.

Geula, S., H. Naveed*, et al.* (2011). "Structure-Based Analysis of VDAC1: Defining Oligomer Contact Sites." *J Biol Chem*.

Girvin, M. E., V. K. Rastogi, *et al.* (1998). "Solution structure of the transmembrane $H^+$-transporting subunit c of the $F_1F_0$ ATP synthase." *Biochemistry* **37** (25), 8817-24.

Gordeliy, V. I., J. Labahn, *et al.* (2002). "Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex." *Nature* **419** (6906), 484-7.

Grosse, W., L. O. Essen, *et al.* (2011). "Strategies and perspectives in ion-channel engineering." *Chembiochem* **12** (6), 830-9.

Hagan, C. L., T. J. Silhavy, *et al.* (2011). "β-Barrel membrane protein assembly by the Bam complex." *Annu Rev Biochem* **80**, 189-210.

Heinz, D. W., W. A. Baase, *et al.* (1993). "How amino-acid insertions are allowed in an α-helix of T4 lysozyme." *Nature* **361** (6412), 561-4.

Henze, K. and W. Martin (2003). "Evolutionary biology: essence of mitochondria." *Nature* **426** (6963), 127-8.

Hill, K., K. Model, *et al.* (1998). "Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins." *Nature* **395** (6701), 516-21.

Hiller, S., R. G. Garces, *et al.* (2008). "Solution structure of the integral human membrane protein VDAC-1 in detergent micelles." *Science* **321** (5893), 1206-10.

Hiller, S. and G. Wagner (2009). "The role of solution NMR in the structure determinations of VDAC-1 and other membrane proteins." *Curr Opin Struct Biol* **19** (4), 396-401.

Hocker, B., A. Lochner, *et al.* (2009). "High-resolution crystal structure of an artificial $(\beta\alpha)_8$-barrel protein designed from identical half-barrels." *Biochemistry* **48** (6), 1145-7.

Howorka, S. and Z. Siwy (2009). "Nanopore analytics: sensing of single molecules." *Chem Soc Rev* **38** (8), 2360-84.

Huysmans, G. H., S. E. Radford, *et al.* (2007). "The N-terminal helix is a post-assembly clamp in the bacterial outer membrane protein PagP." *J Mol Biol* **373** (3), 529-40.

Juin, P., M. Thieffry, *et al.* (1997). "Relationship between the peptide-sensitive channel and the mitochondrial outer membrane protein translocation machinery." *Journal of Biological Chemistry* **272** (9), 6044-50.

Kanamori, T., S. Nishikawa*, et al.* (1999). "Uncoupling of transfer of the presequence and unfolding of the mature domain in precursor translocation across the mitochondrial outer membrane." *Proc Natl Acad Sci U S A* **96**, 3634-3639.

Keefe, L. J., J. Sondek*, et al.* (1993). "The alpha aneurism: a structural motif revealed in an insertion mutant of staphylococcal nuclease." *Proc Natl Acad Sci U S A* **90** (8), 3275-9.

Kiebler, M., P. Keil*, et al.* (1993). "The mitochondrial receptor complex: a central role of MOM22 in mediating preprotein transfer from receptors to the general insertion pore." *Cell* **74** (3), 483-492.

Kinoshita, J., K. Mihara*, et al.* (2007). "Identification and characterization of a new Tom40 isoform, a central component of mitochondrial outer membrane translocase." *Journal of Biochemistry* **141** (6), 897-906.

Kleffel, B., R. M. Garavito*, et al.* (1985). "Secondary structure of a channel-forming protein: porin from *E. coli* outer membranes." *Embo J* **4** (6), 1589-92.

Kleinschmidt, J. H., T. den Blaauwen*, et al.* (1999). "Outer membrane protein A of *Escherichia coli* inserts and folds into lipid bilayers by a concerted mechanism." *Biochemistry* **38** (16), 5006-16.

Kleinschmidt, J. H. and L. K. Tamm (2002). "Secondary and tertiary structure formation of the β-barrel membrane protein OmpA is synchronized and depends on membrane thickness." *J Mol Biol* **324** (2), 319-30.

Korn, E. D. (1966). "Structure of biological membranes." *Science* **153** (743), 1491-8.

Koronakis, V., A. Sharff*, et al.* (2000). "Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export." *Nature* **405** (6789), 914-9.

Kozjak, V., N. Wiedemann*, et al.* (2003). "An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane." *J Biol Chem* **278** (49), 48520-3.

Krissinel, E. and K. Henrick (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." *Acta Crystallogr D Biol Crystallogr* **60** (Pt 12 Pt 1), 2256-68.

Kuhlbrandt, W., D. N. Wang, *et al.* (1994). "Atomic model of plant light-harvesting complex by electron crystallography." *Nature* **367** (6464), 614-21.

Künkele, K. P., P. Juin, *et al.* (1998). "The isolated complex of the translocase of the outer membrane of mitochondria. Characterization of the cation-selective and voltage-gated preprotein-conducting pore." *J Biol Chem* **273** (47), 31032-9.

Lakey, J. H., E. J. Lea, *et al.* (1991). "OmpC mutants which allow growth on maltodextrins show increased channel size and greater voltage sensitivity." *FEBS Lett* **278** (1), 31-4.

Lau, T. L., C. Kim, *et al.* (2009). "The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling." *Embo J* **28** (9), 1351-61.

Lemasters, J. J. and E. Holmuhamedov (2006). "Voltage-dependent anion channel (VDAC) as mitochondrial governator--thinking outside the box." *Biochim Biophys Acta* **1762** (2), 181-90.

Lister, R., C. Carrie, *et al.* (2007). "Functional definition of outer membrane proteins involved in preprotein import into mitochondria." *Plant Cell* **19** (11), 3739-59.

Long, S. B., X. Tao, *et al.* (2007). "Atomic structure of a voltage-dependent $K^+$ channel in a lipid membrane-like environment." *Nature* **450** (7168), 376-82.

Low, B. W. and R. B. Baybutt (1952). "The π helix - a hydrogen bonded configuration of the polypeptide chain." *J. Am. Chem. Soc.* **74** (22), 5806-5807.

Macasev, D., J. Whelan, *et al.* (2004). "Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes." *Mol Biol Evol* **21** (8), 1557-64.

MacKenzie, K. R., J. H. Prestegard, *et al.* (1997). "A transmembrane helix dimer: structure and implications." *Science* **276** (5309), 131-3.

Mager, F., L. Sokolova, *et al.* (2010). "LILBID-mass spectrometry of the mitochondrial preprotein translocase TOM." *J Phys Condens Matter* **22** (45), 454132.

Mahendran, K. R., M. Romero-Ruiz*, et al.* (2012). "Protein translocation through Tom40: kinetics of peptide release." *Biophys J* **102** (1), 39-47.

Malia, T. J. and G. Wagner (2007). "NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-xL." *Biochemistry* **46** (2), 514-25.

Marsh, R. E., R. B. Corey*, et al.* (1955). "An investigation of the structure of silk fibroin." *Biochim Biophys Acta* **16** (1), 1-34.

McLachlan, A. D. (1979). "Gene duplications in the structural evolution of chymotrypsin." *J Mol Biol* **128** (1), 49-79.

Meisinger, C., M. T. Ryan*, et al.* (2001). "Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors." *Mol Cell Biol* **21** (7), 2337-48.

Michel, H. (1982). "Three-dimensional crystals of a membrane protein complex. The photosynthetic reaction centre from *Rhodopseudomonas viridis*." *J Mol Biol* **158** (3), 567-72.

Model, K., C. Meisinger*, et al.* (2008). "Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase." *J Mol Biol* **383** (5), 1049-57.

Model, K., C. Meisinger*, et al.* (2001). "Multistep assembly of the protein import channel of the mitochondrial outer membrane." *Nat Struct Biol* **8**, 361-370.

Mokranjac, D. and W. Neupert (2009). "Thirty years of protein translocation into mitochondria: unexpectedly complex and still puzzling." *Biochim Biophys Acta* **1793** (1), 33-41.

Moon, C. P. and K. G. Fleming (2011). "Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers." *Proc Natl Acad Sci U S A* **108** (25), 10174-7.

Murzin, A. G., A. M. Lesk*, et al.* (1994). "Principles determining the structure of β-sheet barrels in proteins. I. A theoretical analysis." *J Mol Biol* **236** (5), 1369-81.

Murzin, A. G., A. M. Lesk*, et al.* (1994). "Principles determining the structure of β-sheet barrels in proteins. II. The observed structures." *J Mol Biol* **236** (5), 1382-400.

Naveed, H., R. Jackups, Jr.*, et al.* (2009). "Predicting weakly stable regions, oligomerization state, and protein-protein interfaces in transmembrane

domains of outer membrane proteins." *Proc Natl Acad Sci U S A* **106** (31), 12735-40.

Naveed, H., Y. Xu*, et al.* (2011). "Predicting three-dimensional structures of transmembrane domains of β-barrel membrane proteins." *J Am Chem Soc*.

Nishimura, K., N. Tajima*, et al.* (2010). "Autotransporter passenger proteins: virulence factors with common structural themes." *J Mol Med (Berl)* **88** (5), 451-8.

Ostermeier, C., A. Harrenga*, et al.* (1997). "Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody FV fragment." *Proc Natl Acad Sci U S A* **94** (20), 10547-53.

Palade, G. E. (1953). "An electron microscope study of the mitochondrial structure." *J Histochem Cytochem* **1** (4), 188-211.

Paschen, S. A., T. Waizenegger*, et al.* (2003). "Evolutionary conservation of biogenesis of β-barrel membrane proteins." *Nature* **426** (6968), 862-6.

Pauling, L. and R. B. Corey (1951). "Atomic coordinates and structure factors for two helical configurations of polypeptide chains." *Proc Natl Acad Sci U S A* **37** (5), 235-40.

Pauling, L. and R. B. Corey (1951). "The pleated sheet, a new layer configuration of polypeptide chains." *Proc Natl Acad Sci U S A* **37** (5), 251-6.

Pauling, L., R. B. Corey*, et al.* (1951). "The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain." *Proc Natl Acad Sci U S A* **37** (4), 205-11.

Pautsch, A. and G. E. Schulz (1998). "Structure of the outer membrane protein A transmembrane domain." *Nat Struct Biol* **5** (11), 1013-7.

Perry, A. J., K. A. Rimmer*, et al.* (2008). "Structure, topology and function of the translocase of the outer membrane of mitochondria." *Plant Physiol Biochem* **46** (3), 265-74.

Pfanner, N. and A. Geissler (2001). "Versatility of the mitochondrial protein import machinery." *Nat Rev Mol Cell Biol* **2** (5), 339-49.

Phale, P. S., A. Philippsen*, et al.* (1998). "Stability of trimeric OmpF porin: the contributions of the latching loop L2." *Biochemistry* **37** (45), 15663-70.

Phale, P. S., T. Schirmer*, et al.* (1997). "Voltage gating of *Escherichia coli* porin channels: role of the constriction loop." *Proc Natl Acad Sci U S A* **94** (13), 6741-5.

Picot, D., P. J. Loll, *et al.* (1994). "The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1." *Nature* **367** (6460), 243-9.

Popot, J. L. and D. M. Engelman (1990). "Membrane protein folding and oligomerization: the two-stage model." *Biochemistry* **29** (17), 4031-7.

Popot, J. L. and D. M. Engelman (2000). "Helical membrane protein folding, stability, and evolution." *Annu Rev Biochem* **69**, 881-922.

Poynor, M., R. Eckert, *et al.* (2008). "Dynamics of the Preprotein Translocation Channel of the Outer Membrane of Mitochondria." *Biophys J.* **95**, 1511-1522.

Pusnik, M., F. Charriere, *et al.* (2009). "The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane." *Mol Biol Evol* **26** (3), 671-80.

Rada, P., P. Dolezal, *et al.* (2011). "The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*." *PLoS One* **6** (9), e24428.

Ramachandran, R., A. P. Heuck, *et al.* (2002). "Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin." *Nat Struct Biol* **9** (11), 823-7.

Rapaport, D., W. Neupert, *et al.* (1997). "Mitochondrial protein import. Tom40 plays a major role in targeting and translocation of preproteins by forming a specific binding site for the presequence." *J Biol Chem* **272** (30), 18725-31.

Remaut, H., C. Tang, *et al.* (2008). "Fiber formation across the bacterial outer membrane by the chaperone/usher pathway." *Cell* **133** (4), 640-52.

Reumann, S., J. Davila-Aponte, *et al.* (1999). "The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog." *Proc Natl Acad Sci U S A* **96** (2), 784-9.

Rimmer, K. A., J. H. Foo, *et al.* (2010). "Recognition of mitochondrial targeting sequences by the import receptors Tom20 and Tom22." *J Mol Biol* **405** (3), 804-18.

Robert, V., E. B. Volokhina, *et al.* (2006). "Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif." *PLoS Biol* **4** (11), e377.

Romero-Ruiz, M., K. R. Mahendran, *et al.* (2010). "Interactions of mitochondrial presequence peptides with the mitochondrial outer membrane preprotein translocase TOM." *Biophys J* **99** (3), 774-81.

Saitoh, T., M. Igura, *et al.* (2007). "Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states." *Embo J*.

Sanchez-Pulido, L., D. Devos, *et al.* (2003). "POTRA: a conserved domain in the FtsQ family and a class of β-barrel outer membrane proteins." *Trends Biochem Sci* **28** (10), 523-6.

Sapra, K. T., M. Damaghi, *et al.* (2009). "One β-hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG." *Angew Chem Int Ed Engl* **48** (44), 8306-8.

Schirmer, T., T. A. Keller, *et al.* (1995). "Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution." *Science* **267** (5197), 512-4.

Schleiff, E., L. A. Eichacker, *et al.* (2003). "Prediction of the plant β-barrel proteome: a case study of the chloroplast outer envelope." *Protein Sci* **12** (4), 748-59.

Schmitt, S., U. Ahting, *et al.* (2005). "Role of Tom5 in maintaining the structural stability of the TOM complex of mitochondria." *J Biol Chem.* **280**, 14499-14506.

Schneider, R., M. Etzkorn, *et al.* (2010). "The native conformation of the human VDAC1 N terminus." *Angew Chem Int Ed Engl* **49** (10), 1882-5.

Schulz, G. E. (1996). "Porins: general to specific, native to engineered passive pores." *Curr Opin Struct Biol* **6** (4), 485-90.

Schulz, G. E. (2000). "β-Barrel membrane proteins." *Curr. Opin. Struct. Biol.* **10**, 443-447.

Schulz, G. E. (2002). "The structure of bacterial outer membrane proteins." *Biochim Biophys Acta* **1565** (2), 308-17.

Schwartz, M., S. Huang, *et al.* (1999). "The structure of precursor proteins during import into mitochondria." *J. Biol. Chem.* **274**, 12759-12764.

Schwartz, M. and A. Matouschek (1999). "The dimensions of the protein import channels in the outer and inner mitochondrial membranes." *Proc Natl Acad Sci U S A* **9**, 13086-13090.

Seitz, T., M. Bocola, *et al.* (2007). "Stabilisation of a $(\beta\alpha)_8$-barrel protein designed from identical half barrels." *J Mol Biol* **372** (1), 114-29.

Seitz, T., R. Thoma*, et al.* (2010). "Enhancing the stability and solubility of the glucocorticoid receptor ligand-binding domain by high-throughput library screening." *J Mol Biol* **403** (4), 562-77.

Seshadri, K., R. Garemyr*, et al.* (1998). "Architecture of β-barrel membrane proteins: analysis of trimeric porins." *Protein Sci* **7** (9), 2026-32.

Shanmugavadivu, B., H. J. Apell*, et al.* (2007). "Correct folding of the β-barrel of the human membrane protein VDAC requires a lipid bilayer." *J Mol Biol* **368** (1), 66-78.

Shiota, T., H. Mabuchi*, et al.* (2011). "*In vivo* protein-interaction mapping of a mitochondrial translocator protein Tom22 at work." *Proc Natl Acad Sci U S A* **108** (37), 15179-83.

Singer, S. J. (1992). "The structure and function of membranes--personal memoir." *J Membr Biol* **129** (1), 3-12.

Singer, S. J. and G. L. Nicolson (1972). "The fluid mosaic model of the structure of cell membranes." *Science* **175** (23), 720-31.

Sirrenberg, C., M. F. Bauer*, et al.* (1996). "Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22." *Nature* **384** (6609), 582-585.

Snijder, H. J., I. Ubarretxena-Belandia*, et al.* (1999). "Structural evidence for dimerization-regulated activation of an integral membrane phospholipase." *Nature* **401** (6754), 717-21.

Song, J., C. Midson*, et al.* (1998). "The topology of VDAC as probed by biotin modification." *J Biol Chem* **273** (38), 24406-13.

Song, L., M. R. Hobaugh*, et al.* (1996). "Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore." *Science* **274** (5294), 1859-1866.

Stock, D., A. G. Leslie*, et al.* (1999). "Molecular architecture of the rotary motor in ATP synthase." *Science* **286** (5445), 1700-5.

Stutz, S. (2009). "Identifizierung und Charakterisierung der Proteintranslokase der mitochondrialen Außenmembran aus *Bos taurus* L." *http://www.dr.hut-verlag.de/* (ISBN 978-3-86853-031-5).

Subramaniam, S. and R. Henderson (2000). "Molecular mechanism of vectorial proton translocation by bacteriorhodopsin." *Nature* **406** (6796), 653-7.

Suzuki, H., T. Kadowaki*, et al.* (2004). "Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched β-structure." *J Biol Chem* **279** (48), 50619-29.

Thompson, J. R., B. Cronin*, et al.* (2011). "Rapid Assembly of a Multimeric Membrane Protein Pore." *Biophys J* **101** (1), 2679-2683.

Tommassen, J. (2010). "Assembly of outer-membrane proteins in bacteria and mitochondria." *Microbiology* **156** (Pt 9), 2587-96.

Ubarretxena-Belandia, I. and D. M. Engelman (2001). "Helical membrane proteins: diversity of functions in the context of simple architecture." *Curr Opin Struct Biol* **11** (3), 370-6.

Ujwal, R., D. Cascio*, et al.* (2008). "The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating." *Proc Natl Acad Sci U S A* **105** (46), 17742-7.

Ulmschneider, M. B. and M. S. Sansom (2001). "Amino acid distributions in integral membrane protein structures." *Biochim Biophys Acta* **1512** (1), 1-14.

van den Berg, B. (2005). "The FadL family: unusual transporters for unusual substrates." *Curr Opin Struct Biol* **15** (4), 401-7.

van Wilpe, S., M. T. Ryan*, et al.* (1999). "Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase." *Nature* **401** (6752), 485-489.

Vereb, G., J. Szollosi*, et al.* (2003). "Dynamic, yet structured: The cell membrane three decades after the Singer-Nicolson model." *Proc Natl Acad Sci U S A* **100** (14), 8053-8.

Vestweber, D., J. Brunner*, et al.* (1989). "A 42K outer-membrane protein is a component of the yeast mitochondrial protein import site." *Nature* **341** (6239), 205-9 issn: 0028-0836.

Wallin, E. and G. von Heijne (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." *Protein Sci* **7** (4), 1029-38.

White, S. H. (2004). "The progress of membrane protein structure determination." *Protein Sci* **13** (7), 1948-9.

White, S. H., a. S. Ladokhin*, et al.* (2001). "How membranes shape protein structure." *J Biol Chem* **276** (35), 32395-8.

White, S. H. and G. von Heijne (2005). "Transmembrane helices before, during, and after insertion." *Curr Opin Struct Biol* **15** (4), 378-86.

White, S. H. and W. C. Wimley (1999). "Membrane protein folding and stability: physical principles." *Annu Rev Biophys Biomol Struct* **28**, 319-65.

Wiedemann, N., V. Kozjak*, et al.* (2003). "Machinery for protein sorting and assembly in the mitochondrial outer membrane." *Nature* **424** (6948), 565-71.

Wimley, W. C. (2002). "Toward genomic identification of β-barrel membrane proteins: composition and architecture of known structures." *Protein Sci* **11** (2), 301-12.

Wimley, W. C. (2003). "The versatile β-barrel membrane protein." *Curr Opin Struct Biol* **13** (4), 404-11.

Wimley, W. C., K. Hristova*, et al.* (1998). "Folding of β-sheet membrane proteins: a hydrophobic hexapeptide model." *J Mol Biol* **277** (5), 1091-110 issn: 0022-2836.

Winn, M. D., C. C. Ballard*, et al.* (2011). "Overview of the CCP4 suite and current developments." *Acta Crystallogr D Biol Crystallogr* **67** (Pt 4), 235-42.

Xia, D., C. A. Yu*, et al.* (1997). "Crystal structure of the cytochrome $bc_1$ complex from bovine heart mitochondria." *Science* **277** (5322), 60-6.

Yamashita, E., M. V. Zhalnina*, et al.* (2008). "Crystal structures of the OmpF porin: function in a colicin translocon." *Embo J* **27** (15), 2171-80.

Yildiz, O., K. R. Vinothkumar*, et al.* (2006). "Structure of the monomeric outer-membrane porin OmpG in the open and closed conformation." *Embo J* **25** (15), 3702-13.

Zalk, R., A. Israelson*, et al.* (2005). "Oligomeric states of the voltage-dependent anion channel and cytochrome c release from mitochondria." *Biochem J* **386** (Pt 1), 73-83.