


**REVIEW**

 Cite this: *RSC Adv.*, 2021, 11, 32476

# Protein import in mitochondria biogenesis: guided by targeting signals and sustained by dedicated chaperones

 Anna-Roza Dimogkioka, Jamie Lees, Erik Lacko and Kostas Tokatlidis \*

Mitochondria have a central role in cellular metabolism; they are responsible for the biosynthesis of amino acids, lipids, iron–sulphur clusters and regulate apoptosis. About 99% of mitochondrial proteins are encoded by nuclear genes, so the biogenesis of mitochondria heavily depends on protein import pathways into the organelle. An intricate system of well-studied import machinery facilitates the import of mitochondrial proteins. In addition, folding of the newly synthesized proteins takes place in a busy environment. A system of folding helper proteins, molecular chaperones and co-chaperones, are present to maintain proper conformation and thus avoid protein aggregation and premature damage. The components of the import machinery are well characterised, but the targeting signals and how they are recognised and decoded remains in some cases unclear. Here we provide some detail on the types of targeting signals involved in the protein import process. Furthermore, we discuss the very elaborate chaperone systems of the intermembrane space that are needed to overcome the particular challenges for the folding process in this compartment. The mechanisms that sustain productive folding in the face of aggregation and damage in mitochondria are critical components of the stress response and play an important role in cell homeostasis.

 Received 10th June 2021  
 Accepted 25th September 2021

DOI: 10.1039/d1ra04497d

[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

*Institute of Molecular Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, University Avenue, Glasgow G12 8QQ, Scotland, UK. E-mail: Kostas.Tokatlidis@glasgow.ac.uk*

## 1. Introduction

Mitochondria are subcellular organelles that are critically important for cell physiology and development. Mitochondria have a central role in cellular metabolism; they are responsible for the biosynthesis of amino acids, lipids, iron–sulphur



*Anna-Roza Dimogkioka got her BSc in Biochemistry at the University of Glasgow, where she had a 4 month research placement in the group of Professor Kostas Tokatlidis. During her placement she investigated the targeting capacity of a novel mitochondrial targeting peptide. This sparked her interest in mitochondrial biology leading her to pursue a Master's by Research in*

*Biochemistry and Biotechnology in the Tokatlidis group. Her research focused on a potential new protein therapy for mitochondrial disease and the effect of SARS-CoV2 proteins on mitochondrial health.*



*Jamie Lees started his studies at the University of Glasgow in 2015 where he completed a Bachelors in Biochemistry (Hons). From here, Jamie completed a Master's by Research in the Tokatlidis group investigating the oxidative folding pathways in the mitochondrial intermembrane space. While writing his Master's thesis, Jamie worked as a laboratory scientist at Glasgow's*

*COVID-19 Lighthouse testing centre, before moving on to study for a PhD in Cancer Sciences at the Leukaemia Research Centre (University of Glasgow) where he currently works on the roles of FOXO transcription factors in the maintenance and progression of chronic lymphocytic leukaemia.*



clusters, haem, and lipids. In addition, they are crucial in cellular signalling pathways and apoptosis.<sup>1,2</sup> Moreover, mitochondria participate in a range of innate immunity pathways.<sup>3</sup>

Mitochondrial biogenesis requires the integration of two genomes, the nuclear and the mtDNA genome. Despite the plethora of mitochondrial functions, the mitochondrial genome only codes for a small set of proteins, namely 8 proteins in the yeast *S. cerevisiae* and 13 proteins in humans.<sup>4</sup> These proteins are all core subunits of the electron transfer complexes located at the inner mitochondrial membrane. Approximately 99% of mitochondrial proteins are encoded by nuclear genes and are synthesized as protein precursors in the cytosol. Subsequently they are targeted to the surface of mitochondria and then sorted within their specific sub-mitochondrial compartment.<sup>5</sup> The mitochondrial proteome contains about 1000 proteins in yeast and 1500 proteins in humans, fulfilling the mitochondrion's multiple functions.<sup>1,2</sup> Proteomic studies have led to the functional classification of these proteins in yeast and more recently in human mitochondria. An inventory of a little over 1000 genes encoding the mammalian mitochondrial proteome, known as MitoCarta, was created in 2008.<sup>4</sup> This was updated in 2015 to include a total of 1158 human or mouse genes leading to MitoCarta2.0.<sup>6</sup> The latest MitoCarta release in 2020 includes 1136 genes with information on the submitochondrial localisation and the mitochondrial functional pathways they belong to.<sup>7</sup>

According to the inventory, the majority of the proteins are assigned to the matrix (46%) and the inner membrane (32%). 10% are annotated as outer membrane proteins whilst only 5% are designated as intermembrane space (IMS) proteins.<sup>7</sup> Despite making up a small proportion of the mitochondrial

proteome, proteins of the outer membrane and the IMS play a crucial role in the numerous functions of the mitochondrion. These include, apoptosis, phospholipid biosynthesis, haem biosynthesis and critical communications with the cytosol and other organelles. In addition, although the IMS is the smallest mitochondrial subcompartment, it encompasses a very wide range of protein import mechanisms. Protein import is pivotal for the mitochondrial intermembrane space (IMS) since all IMS proteins are encoded by nuclear genes.<sup>2,8</sup>

In this review we discuss the import pathways required to accommodate for the different types of proteins that are imported into mitochondria. Furthermore, we highlight the types and structural features of mitochondrial protein targeting signals that underpin the different import pathways. We then focus on the IMS, which is the only mitochondrial compartment housing a redox modification machinery integral to the import process. We discuss how this machinery works and how it facilitates a chaperone-assisted folding process in this compartment. Finally, we discuss the various chaperone systems in the IMS which are critical for sustaining an efficient folding environment in this compartment.

## 2. General protein import pathways into mitochondria

Several import pathways are needed to accommodate the different types of proteins that are imported into mitochondria (Fig. 1). Each of the import pathways has specialised protein import components, many of which are encoded by genes that are essential for cell viability.<sup>9</sup> The precursors undertaking a specific import pathway are guided by peptide targeting



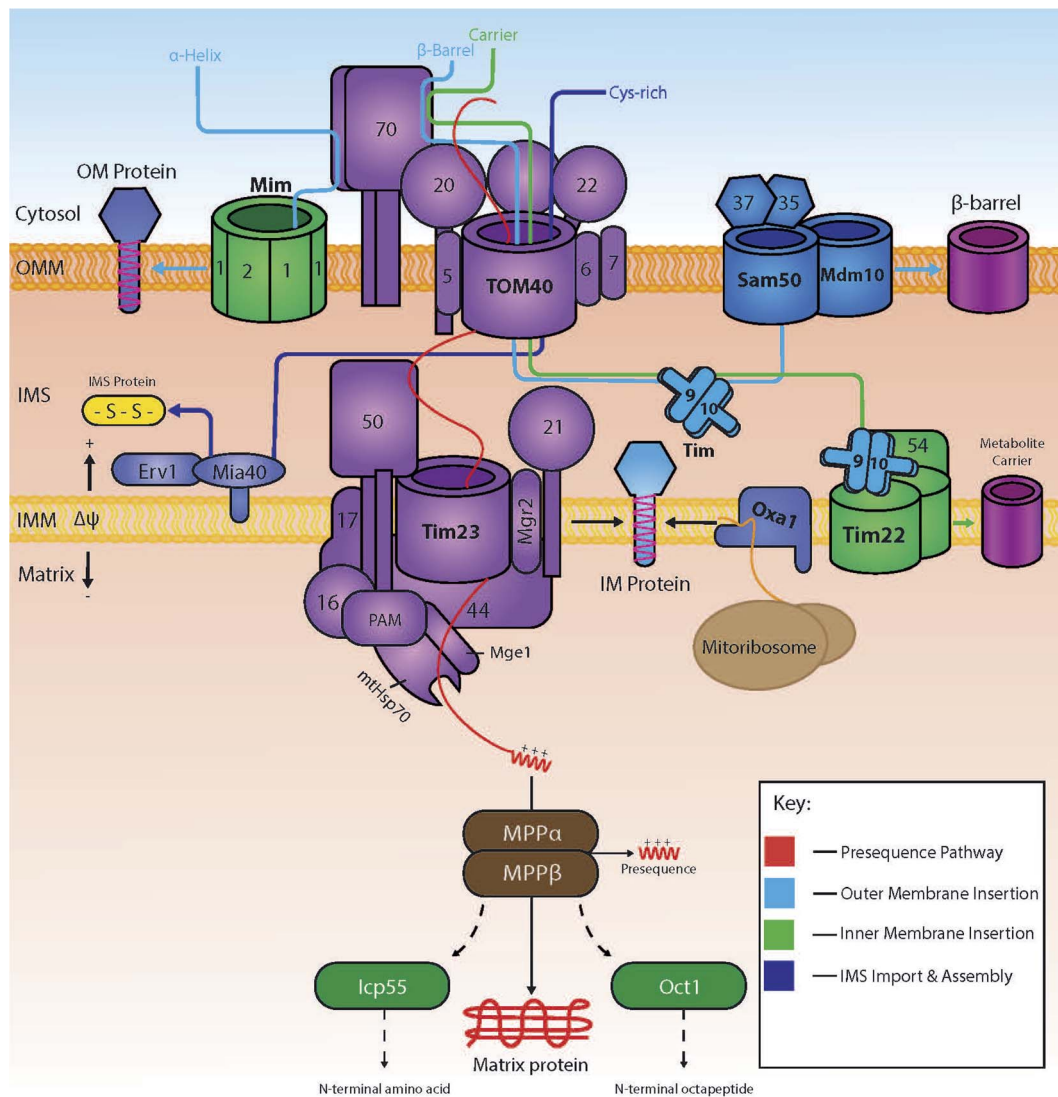
*Erik Lacko did his early academic studies in Denmark, studying Chemical and Biotechnical Science. He is now an Honours graduate of Microbiology and Biotechnology from Edinburgh Napier University. Following a one-year placement at the European Molecular Biology Laboratory (EMBL), he is developing his scientific career and laboratory techniques in protein science and mitochondria biology working as a research technician in the group of Prof Kostas Tokatlidis at the University of Glasgow.*

*ria biology working as a research technician in the group of Prof Kostas Tokatlidis at the University of Glasgow.*



*Professor Kostas Tokatlidis is the Cathcart Chair at the University of Glasgow (2013), a Fellow of the Royal Society of Edinburgh (2016), and of the Royal Societies of Chemistry and Biology an elected EMBO member (2013) and a Royal Society Wolfson research fellow (2013–2018). He graduated Cum Laude from the Aristotle University Thessaloniki, got his PhD from Delaware University*

*and Pasteur Institute (Fulbright/Marie-Curie fellow), then moved to the Biozentrum, Basel (HFSP/EMBO fellow). He became a Professor in 2011 and held previous posts in the Universities of Manchester, and Crete. His research spans over 25 years on mitochondria biogenesis and function and how these underpin cell physiology and human disease. He published extensively and his work has been funded by MRC, Wellcome, Leverhulme Trust, Royal Society, EU, BBSRC and is also very active in public engagement raising awareness for mitochondria research globally working with scientists from different disciplines, artists and designers.*



**Fig. 1** The various import pathways of preproteins into mitochondria. Proteins targeted to the mitochondrial matrix typically follow the pre-sequence pathway (shown in red), where they are finally modified in the matrix by matrix processing proteins (MPP, ICP or Oct). Preproteins destined for the mitochondrial outer membrane are either integrated via the MIM machinery or by the SAM complex (depending on whether the protein is an  $\alpha$ -helical protein or a  $\beta$ -barrel protein respectively, shown in light blue). Hydrophobic proteins destined for the inner membrane, are targeted and carried across the IMS by the Tim9/10 complex, and inserted by the Tim22 complex (shown in green). Proteins destined for the IMS are imported via TOM40 and modified into mature proteins by interactions with the oxidoreductase Mia40 (shown in dark blue).

signals that are variable in length and their chemical properties and, in some cases, not very well understood. In the next paragraphs, we will summarise the main protein import pathways in terms of the components involved; the energetic requirements and the mechanism of interactions that ensure the efficiency of the protein import process which is key to mitochondria biogenesis.

### 2.1 The presequence pathway for import into the matrix

Cytosolically made proteins encoded by nuclear genes but targeted to mitochondria for their function usually require a system of mitochondrial-linked chaperones and a targeting sequences that determine the protein's endpoint within the

mitochondria. Proteins destined for mitochondria typically are translocated as preproteins; preproteins typically contain an N-terminal amino acid presequence that usually forms an  $\alpha$ -helix with positively charged residues on one side.<sup>10</sup> This sequence is commonly referred to as the matrix-targeting signal (MTS), and will ultimately lead to the translocation of the protein into the mitochondrial matrix.<sup>11</sup>

The preprotein is carried within the cytosol in an unfolded state maintained by the cytosolic chaperones Hsp70 and Hsp90.<sup>12</sup> Upon close proximity with the outer membrane of the mitochondria, the preprotein is typically recognised by the membrane-bound receptors Tom20 and Tom22, both of which form part of the main translocase of the outer membrane (TOM) complex, together with Tom40, the main pore-forming subunit

of the TOM complex.<sup>2</sup> In addition to Tom20, Tom70 is one of the first identified import receptors of the mitochondrial OM.<sup>13</sup> Both Tom20 and Tom70 have their own specific preference for substrate, alongside the fact they are non-essential but have the ability to rescue each other's function.<sup>14</sup> Tom40 acts as the main protein import channel and is encoded by an essential gene.<sup>15</sup> Tom40 is endowed with the capacity to allow passage of almost all proteins that are translocated to mitochondria, and it contains a specific region for the binding of different types of preproteins. Also included within the TOM complex are the subunits Tom5, Tom6 and Tom7, all of which are encoded by non-essential genes (in contrast to Tom40 which is encoded by an essential gene) for the optimisation of the function of the complex.<sup>16–18</sup> The matrix targeting signal binds to a groove of Tom20 that recognises its hydrophobic residues;<sup>19</sup> recognition by Tom20 leads to the binding and translocation through the outer membrane *via* Tom40;<sup>20</sup> the preprotein then interacts with the presequence translocase of the inner membrane known as TIM23. This interaction between the preprotein and TIM23 drives preprotein translocation across the inner membrane, in partnership with membrane potential ( $\Delta\psi$ ).<sup>21</sup> Coupled to the TIM23 complex is the presequence translocase-associated motor (PAM), containing the mitochondrial version of heat shock protein 70 (mtHsp70). This motor promotes the unidirectional movement of preproteins into the mitochondrial matrix in an ATP-dependent manner.<sup>22</sup> Following successful import, the positively charged presequence is cleaved by the mitochondrial-processing peptidase (MPP).<sup>23,24</sup> Additional processing enzymes are then available to stabilize and fold the protein into its active form. Examples include: the removal of destabilizing amino acids by 55 kDa intermediate cleaving and octapeptidyl peptidases (Icp55 and Oct1 respectively);<sup>25</sup> and correct folding *via* the Hsp60–Hsp10 chaperonin complex.<sup>26</sup>

## 2.2 Insertion into the inner membrane

If a protein is destined for a location other than the matrix, it will usually include either additional targeting information on top of the matrix targeting signal or an entirely different targeting signal. Some proteins that are destined for the IMS interact first with the inner membrane and are then released into the IMS. In this case, they contain a bipartite targeting signal made of a matrix-targeting sequence followed by a specific hydrophobic sorting signal. These proteins are initially targeted to the TIM23 complex following the matrix-targeting route. However, when in the TIM23 complex they get arrested by the small hydrophobic segment which functions as a stop-transfer signal (stop transfer pathway).<sup>27</sup> Once stalled within the TIM23 complex, they are then released laterally into the inner membrane with the help of a small hydrophobic protein known as Mgr2.<sup>28</sup>

A separate route for insertion of inner membrane proteins exists for very hydrophobic proteins that are subunits of the electron transport complex of the inner membrane and are encoded by the mitochondrial DNA. These proteins are synthesised in the matrix and inserted into the inner membrane from the matrix side by the OXA insertase.<sup>29–31</sup>

A major part of inner membrane proteins are helical multi-spanning proteins that belong to the family of solute transporters of the inner membrane. These have typically six membrane spanning segments and are responsible for the transport of small metabolites across the inner mitochondrial membrane.<sup>32</sup> The insertion of these proteins in the inner membrane is governed by the carrier pathway.<sup>33</sup> In this case, the imported proteins do not contain a presequence, but instead internal peptide sequences found within the more hydrophobic areas of the transmembrane segments of the protein play the role of targeting signals. Proteins with internal signals are carried by Hsp70 and Hsp90 to the outer membrane where they are recognised by the receptor Tom70.<sup>12</sup> Following their release from the cytosolic chaperones, proteins travel into the intermembrane space *via* the Tom40 channel, presumably in a looped formation.<sup>34</sup> Small intermembrane space chaperones known as small TIM chaperones (Tim9 and Tim10 or Tim8 and Tim13) then bind – in a heterohexameric complex – to the protein to prevent any aggregation occurring.<sup>34–37</sup> Here, the proteins are then guided by the chaperones to the translocase of the inner membrane TIM22 complex. The TIM22 complex then inserts the proteins in a process that depends on  $\Delta\psi$ , thus producing a multi-spanning protein within the inner membrane.<sup>38</sup> The insertion of Tim23 is like the carrier pathway as it also uses small TIM chaperones to allow transport across the intermembrane space, albeit with the non-essential chaperones Tim8 and Tim13.<sup>39–41</sup> An example of a preprotein lacking an N-terminal presequence is the ADP/ATP carrier (AAC), as targeting information to the inner membrane is instead found within its three modules.<sup>42</sup>

## 2.3 Insertion into the outer membrane

The insertion of outer membrane proteins also involves the TOM complex, amongst other complexes found within the outer membrane. One such complex is known as the sorting and assembly machinery (SAM) complex and is used for the insertion of outer membrane  $\beta$ -barrel proteins.<sup>43</sup>  $\beta$ -Barrel proteins are characteristic prokaryotic proteins. Interestingly, these proteins are also found within eukaryotic mitochondria and chloroplasts, making their lineage unique to the rest of the cell.<sup>14</sup> There are three main subunits found within the SAM machinery: Sam50 (the main component), Sam35 and Sam37. Mitochondrial preproteins access this machinery by interaction and translocation into the intermembrane space *via* the TOM complex's Tom70 receptor and Tom40 channel. Here, the proteins interact with the small TIM chaperones, allowing for the integration of the precursor proteins into the SAM complex, where they are subsequently inserted into the outer membrane.<sup>44</sup>

Although  $\beta$ -barrel proteins are exclusively localised in the outer and not the inner mitochondrial membrane, the outer membrane also houses  $\alpha$ -helical membrane proteins. These are integrated in the OM by a dedicated machinery, the mitochondrial import complex (MIM complex).<sup>45</sup> Single-spanning proteins with an N-terminal anchor and multi-spanning membrane preproteins are able to bypass the Tom40 channel

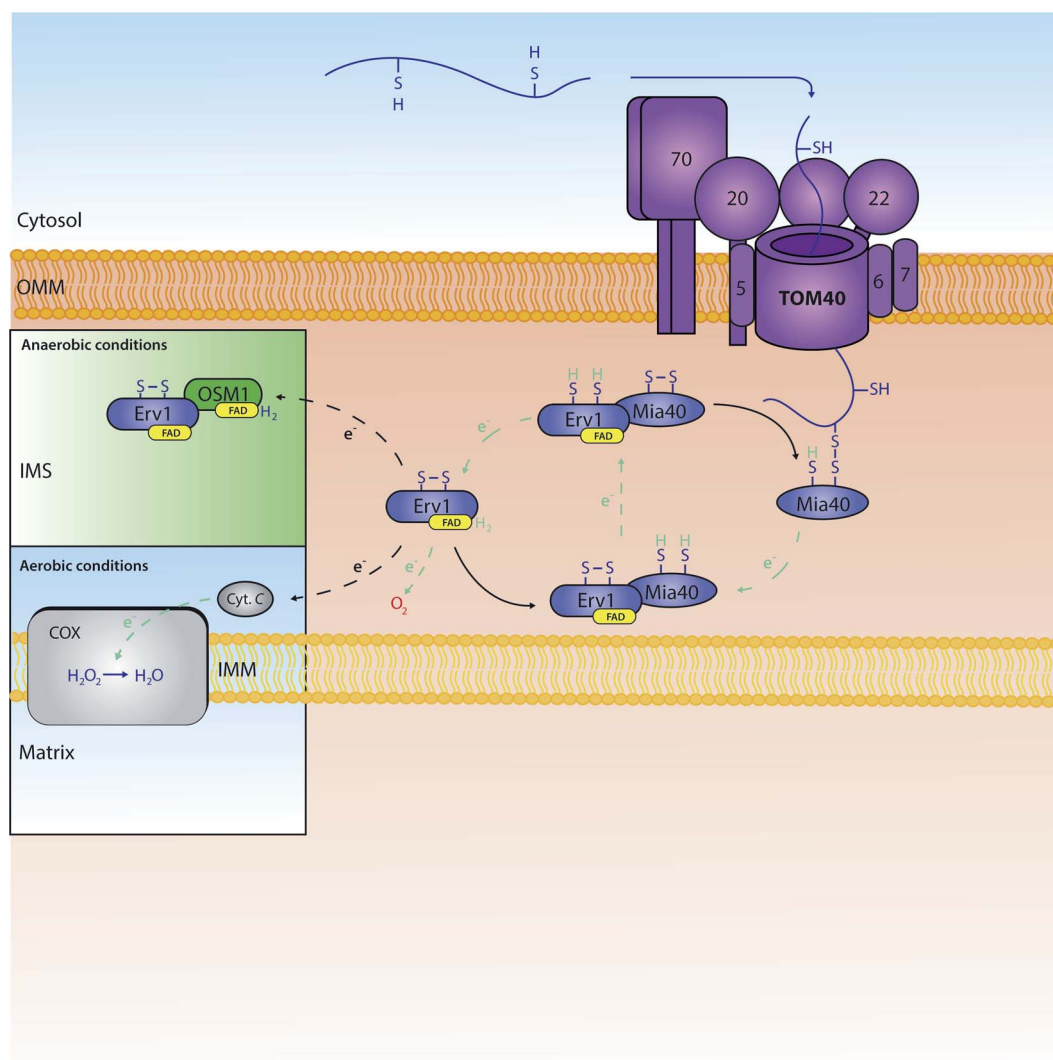


and be imported into the outer membrane directly *via* the MIM complex, usually in partnership with the Tom70 receptor.<sup>46</sup> Proteins Mim1 and Mim2 are crucial for the normal function of the complex; complete absence of Mim2 has been shown to lead to the impairment of mitochondrial protein import, defects in mitochondrial morphology and problems in the assembly of the TOM complex.<sup>47</sup>

#### 2.4 Import into the intermembrane space: chemical modification during import *via* an oxidation-coupled mechanism

All proteins destined for the intermembrane space are encoded by nuclear genes, and thus are synthesised and translocated from the cytosol. For these proteins to reach the intermembrane space, they can include one of two unique properties: (1) they can contain a bipartite sequence as previously discussed, or (2)

they can be retained in the IMS by an interaction with the mitochondrial intermembrane space import and assembly protein 40 (Mia40).<sup>48</sup> Bipartite sequences are targeting signals – much like matrix targeting signals – albeit with an added transmembrane-domain-like hydrophobic region following the MTS as discussed in the previous section. These preproteins following the stop transfer pathway are subject to a double cleavage (i) by the mitochondrial processing peptidase (MPP) to remove the MTS and (ii) by the specific intermembrane space protease (IMP) that removes the stop transfer hydrophobic domain –to discharge the mature protein into the intermembrane space. This pathway is fuelled solely by mitochondrial membrane potential, and does not require ATP-hydrolysis in contrast to the matrix targeting pathway where ATP hydrolysis in the matrix is needed.<sup>11,49–51</sup>



**Fig. 2** The role and recycling of the Mia40 oxidoreductase within the mitochondrial IMS. Proteins are actively oxidised by Mia40 leading to mature disulphide bond formation within the protein. Mia40 is subsequently reduced, and then re-oxidised by the FAD-bound Erv1. Erv1 is then re-oxidised by a number of proteins further down the chain. In aerobic conditions, Erv1 is recycled by O<sub>2</sub> or cytochrome C (Cyt. C). In anaerobic conditions, Erv1 instead interacts with a fumarate reductase known as OSM1 which in turn transports electrons to fumarate.<sup>69</sup> Oxidised Erv1 is then free to re-interact with Mia40 to complete the cycle and free up Mia40 to further interact with imported precursors.

A very substantial part of the intermembrane space proteins contain cysteine motifs – either  $CX_3C$  or  $CX_9C$  – which later become oxidised, forming strong disulphide bonds necessary for intermembrane space protein maturation. This discovery was based upon the presence of internal disulphides found within the TIM intermembrane space chaperones.<sup>36,52</sup> This specific pathway is known as the MIA (mitochondrial protein import and assembly) pathway (Fig. 2), as preproteins involved in this pathway primarily interact with the oxidoreductase protein Mia40.<sup>48</sup> In this interaction, Mia40 is bound to the inner membrane, leaving its C-terminus accessible to the intermembrane space; it is at this C-terminus where substrates interact with the oxidoreductase protein.<sup>10</sup> Mia40 acts by ‘donating’ disulphides to premature intermembrane space proteins, thus leading to correct mature folding and entrapping within the intermembrane space. Previously stated  $CX_3C/CX_9C$  cysteine motifs interact with the specific substrate-binding cleft within the structure of Mia40. Here, the substrates then bind to what is known as a CPC motif, producing a disulphide intermediate between the substrate and Mia40.<sup>53</sup> The specific substrate is then subsequently oxidised with correct disulphide formation leading to protein maturation and thus IMS protein activity. There are many examples of substrates of Mia40 that are vital to the normal function of mitochondria, such as the small TIM chaperones – Tim8, Tim9, Tim10, Tim12 and Tim13.<sup>54</sup> Other examples include COX proteins – containing dual  $CX_9C$  motifs – that are involved in the maintenance of the respiratory chain located in the mitochondrial inner membrane.<sup>55,56</sup> Moreover, it seems that the import of other vital proteins (*e.g.* Mrp10 & Atp23 into the intermembrane space and Tim22 into the inner membrane) is dependent on their interactions with Mia40, albeit without the requirement of their cysteine motifs.<sup>11</sup> Tim22 is a translocase that does not have any  $CX_nC$  motifs but still relies on Mia40 for its correct import and insertion within the inner membrane.<sup>10</sup> These interactions display the importance of this single oxidoreductase protein on the regulation and maintenance of mitochondrial intermembrane space proteins.

It was discovered that a sequence which is not an N-terminal sequence – and is common to many Mia40 substrates – exists within these substrates and is vital for the substrate’s specific targeting capabilities to the mitochondrial intermembrane space. This intermembrane space targeting signal (ITS), consisting of 9 amino acids, is necessary for targeting of intermembrane-space-targeted preproteins to Mia40; the ITS functions by priming a cysteine residue within the substrate for effective docking with Mia40’s active-site-resident CPC motif.<sup>57</sup> Deletion of this ITS has shown to completely abolish the import of Mia40 substrates to the intermembrane space. The ITS functions by forming a helix with its vital hydrophobic and aromatic residues on the same side as this primed cysteine residue. Here, this region of the helix interacts with and induces the interaction of the substrate with the Mia40 oxidoreductase, ultimately leading to further interaction through disulphide bond formation.<sup>57,58</sup>

When Mia40 donates its disulphide to its substrates it stays reduced. To function again it needs to be recycled back to its oxidised state. This re-oxidation is ensured by a second

component of the MIA pathway: a flavin adenine dinucleotide (FAD)-linked sulphhydryl oxidase known as Erv1 (Essential for respiration and viability 1).<sup>59</sup> Erv1 is a Mia40 substrate with a structure unlike any other Mia40 substrate, containing three specific cysteine pairs that are highly conserved between species (residues C30/C33, C130/C133 and C159/C176);<sup>11,56,59,60</sup> Erv1 is unable to oxidise proteins independently, although it can form a higher ternary complex with Mia40 to achieve this effect.<sup>61</sup> Specifically, the first pair of cysteines are involved in the specific interaction of Erv1 with Mia40.<sup>62,63</sup> The third pair of cysteines’ disulphide bond has a more structural role; interestingly, this same disulphide is recognised by the Mia40 oxidoreductase during the import of Erv1 into the intermembrane space,<sup>64</sup> with final folding and maturation of Erv1 occurring with FAD binding.<sup>59</sup> Erv1 is a crucial component of the MIA oxidative folding pathway and functions by removing electrons from Mia40 and relaying them to either molecular oxygen or cytochrome *c*,<sup>65,66</sup> thereby producing either hydrogen peroxide ( $H_2O_2$ ) or water ( $H_2O$ ) respectively; water is produced from cytochrome *c*’s funnelling of electrons to oxygen linking thus this oxidative folding pathway to the ETC. Specifically, Erv1 functions by collecting electrons (through Mia40’s N-terminal  $CX_2C$  motif) in its own  $CX_2C$  motif. Here, the electrons are transferred onto FAD, where they can then be relayed to either molecular oxygen or cytochrome *c*. Although oxygen is generally the final electron acceptor, if electrons are relayed to cytochrome *c*, alternative structures are known to be final electron acceptors – such as the cytochrome *c* heme lyase Ccp1.<sup>67</sup> In essence, through this specific electron relay, Erv1 recycles Mia40 from a substrate-induced reduced state to a functional oxidised state allowing Mia40 to continue its role as a ‘disulphide donor’. We should note that in human mitochondria, the Mia40 homologue (Mia40 or CHCHD4) translocation to the intermembrane space is dependent on the presence of the human homologue of Erv1 (ALR) and an internal targeting signal; this is due to the lack of an N-terminal membrane anchor that is present in the *S. cerevisiae* homologue of Mia40.<sup>68</sup> This further shows the importance of Erv1 to Mia40’s general function within the IMS.

### 3. Types and structural features of protein targeting signals for import into mitochondria

Mitochondrial targeting signals can be considered as ‘post codes’ that are necessary and sufficient when added to proteins for their submitochondrial localisation.<sup>70</sup> About two thirds of all mitochondrial proteins are synthesized in the cytosol with a cleavable presequence. Presequences are of variable length, usually 18–50 amino acid residues, although very short (<10 aa) and very long (up to 100 aa) ones have been observed.<sup>71,72</sup> An important characteristic of mitochondrial presequences is the formation of an amphipathic  $\alpha$ -helix that contains a positively charged face and a hydrophobic face.<sup>73</sup> The elements of the amphipathic helix are specifically recognised by receptors and other import components during preprotein translocation by

**Table 1** Yeast mitochondrial proteins with N-terminal cleavable presequences. The positively-charged amino acid residues are shown in bold, negatively-charged residues are shown in *italic* and hydrophobic residues are shown in regular font. The cleavage sites processed by the mitochondrial processing peptidase (MPP) are mostly found in position R-2 and highlighted in cyan. Sequences were acquired from the *Saccharomyces* Genome Database (SGD)/UniProt and analysed using the MitoFates prediction tool for mitochondrial targeting sequences

Protein	Function	N-terminal presequence	MPP cleavage site
ACO1	Required for the tricarboxylic acid (TCA) cycle	... <b>IKRP</b> IV-	-RGLA...
HSP77	ATPase of the Hsp70 family, involved in protein translocation and folding	...AAKNILNRSS...	...RLQS...
TRX3	Thioredoxin required to maintain redox homeostasis in the cell	...FYK <b>PVM</b> RM <b>AVR</b> PLKSI-	-RFQS...
TUF1	Mitochondrial translation elongation factor	<b>MSALLPRL</b> LTR...	...RTFS...
XDJ1	Chaperone facilitating mitochondrial protein import; associated to Ydj1p	... <i>DRGDR</i> LYDVL...	...RKL <i>A</i> ...
MXR2	Methionine-R-sulfoxide reductase, involved in the response to oxidative stress	...LGK <b>R</b> ICQEAVT...	...RSGK...
SOD2	Manganese superoxide dismutase that protects cells against oxygen toxicity	...FAK <b>T</b> AAANLTK...	...RRTKV...

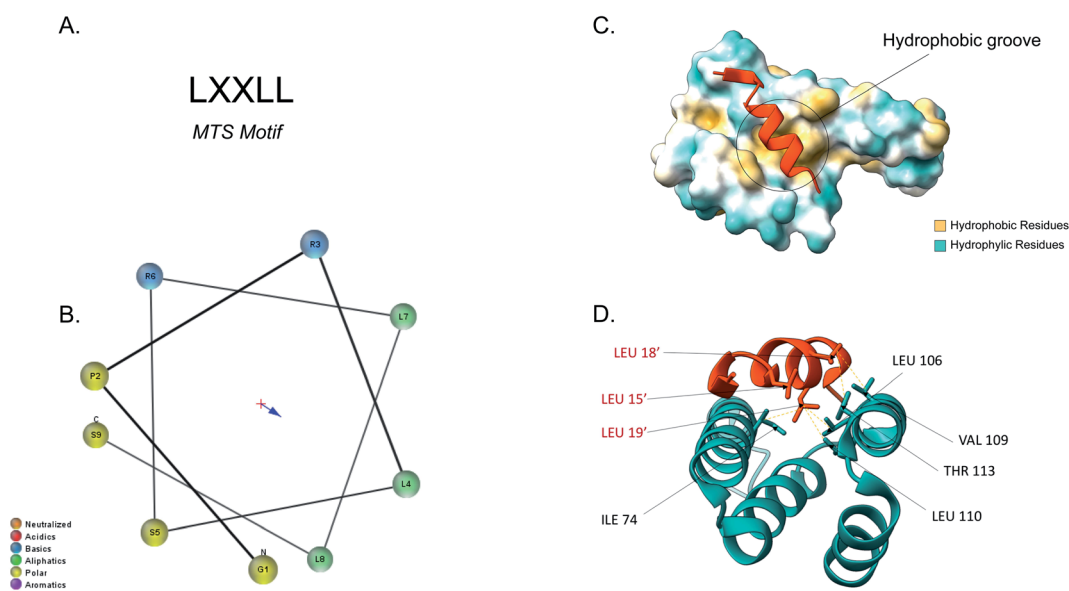
the import machinery.<sup>74</sup> Although the components of the import machinery have been well characterised as mentioned in the first part of this review, the cytosolic targeting steps and the proteins involved in these are less well-understood. Specifically, it is still not clear for many mitochondrial proteins whether their targeting signal can be recognised by one or many dedicated targeting factors, the specificity of such targeting factors and their interplay with the outer membrane receptor subunits that decode the targeting signals.

The main entry gate of precursor proteins to mitochondria is the translocase of the outer membrane (TOM) complex. The TOM complex consists of multiple subunits; Tom20 has been identified as the general receptor for protein sequences while Tom70 may also have a role in signal recognition.<sup>75</sup> Tom20 has the monumental task of recognising roughly up to 1000 mitochondrial proteins, from non-mitochondrial ones, and

recognising at least four different classes of precursor proteins with distinct targeting signals following various import pathways in the organelle.<sup>76,77</sup> The import machinery not only functions as recognitions stations for the mitochondrial targeting signals but also as channels and driving forces for the translocation of the preproteins (Table 1).<sup>74</sup>

### 3.1 Matrix targeting signals

The majority of the matrix proteins contain N-terminal cleavable mitochondrial sequences (MTS, Fig. 3). The MTS is 10–70 amino acids long, it is positively charged with a tendency to form an amphipathic  $\alpha$ -helix (Fig. 3B).<sup>71</sup> A comparative analysis of the N-proteome of mitochondria from mouse, human and yeast revealed that although the N-terminus sequence is poorly conserved, properties such as the length, charge and cleavage



**Fig. 3** Tom20 interaction with the N-terminal presequence of aldehyde dehydrogenase (ALDH). (A) Matrix targeting sequence motif: L stands for leucine while X can be any amino acid. (B) Helical wheel projection of the presequence peptide amphipathic helix. The blue arrow points to the hydrophobic side of the helix. (C) Interaction of the ALDH presequence (orange) with the hydrophobic shallow groove of the helix-turn-helix domains of Tom20. The hydrophobic surface of Tom20 is indicated in yellow while the hydrophilic surface is light blue. (D) Contact residues of ALDH (orange) at the Tom20 interface. ALDH residues are represented in orange. Yellow dashes indicated contact between two atoms (<math><4.0 \text{ \AA}</math> (PDB ID: 2V1S)).

site are well conserved.<sup>78</sup> These presequences are usually removed in the matrix, after import, by the matrix processing peptidase (MPP). MPP cleavage typically occurs at two amino acids C-terminal to an arginine (R-2).<sup>24,71</sup>

Tom20 recognises the hydrophobic side of the  $\alpha$ -helix while Tom22 is presumed to recognise the basic residues of the presequences and thus complement the receptor function of Tom20.<sup>19,79</sup> Previous analyses into the MTS consensus motif (Fig. 3A) revealed the following motif: LSRL where L represents leucine (hydrophobic residue) and R is arginine (positively charged residue). Although this a typical motif, Tom20 has been shown to recognise a variety of this motif in other presequences such as the LRRAY of the mitochondrial heat shock protein 60 (Hsp60). Nevertheless, the three lysine residues appear to be the most important for preprotein recognition.<sup>76</sup> Tom20 consists of two helix-turn-helix domains, that create a shallow groove that acts as the binding site for the presequence; the hydrophobic residues are indicated in yellow while the hydrophilic ones are presented in cyan (Fig. 3C). As seen in Fig. 3D, the three leucine residues (orange) of the presequence interact with residues in the groove (cyan). Leu15' and Leu19' are held in a hydrophobic created by Ile74, Leu106 and Leu110 side chains of Tom20. Leu18' on the other hand, makes contact with the Val109 and Thr113 residues of the translocase.<sup>76</sup>

After recognition the presequence transverses the TOM complex by consecutively binding a number of binding presequence-binding sites spanning the channel formed by the  $\beta$ -barrel protein.<sup>80,81</sup> These binding sites are arranged by in order of increasing affinity for the presequence, with the high-affinity binding site found at the IMS side of the channel.<sup>82-85</sup> Next, the presequence has to interact with the translocation machinery of the inner membrane. The TIM23 channel consists of Tim23, Tim50 and Tim17; Tim23 and Tim17 are structurally similar and form the channel *via* which presequences are translocated.<sup>86-88</sup> Tim50 was later discovered to have an important role in the transfer of proteins from the TOM to the TIM23 channel.<sup>89-91</sup> Additionally, Tim50 regulates the conformation of the Tim23 channel by closing the channel in the absence of the presequence in order to prevent ion leakage. When the presequence binds Tom50, it subsequently leads to a conformational change in Tim23 that causes the channel to open.<sup>77</sup> However, the two membranes are separated by the intermembrane space which is an aqueous sub-compartment. As a result, protein translocation needs to happen at the contact sites where the outer membrane and inner membrane are in close proximity. Once the presequence is translocated *via* the TOM complex, its N-terminus reaches the TIM23 complex of the inner membrane where a translocation intermediate is formed.<sup>92-94</sup> The presequence is now threaded through both the TOM and TIM23 complex. This intermediate happens due to the lack of membrane potential in the outer membrane rendering the TOM complex a passive pore, incapable of completing the translocation of the mitochondrial preprotein. As a result, TIM23 provides a critical driving force for the translocation of the N-terminal presequences. TIM23 requires membrane potential and the activity of the matrix ATP-dependent translocation motor<sup>14,74,95</sup> chaperone heat shock

protein 70 (Hsp70) to drive the precursor protein across the inner membrane and into the matrix.<sup>80,89,94,96</sup>

However, the protein is not functional yet and reaches its mature form after the presequence gets cleaved by the MPP. MPP is a soluble matrix protein that processes the bulk of N-terminal presequences.<sup>71,97</sup> It is a heterodimer consisting of two subunits: Mas1 which is the active site and Mas2 which contains a glycine-rich loop which plays an important role in presequence recognition.<sup>98,99</sup> Although there is not a conserved cleavage site motif, most cleavage sites have an arginine residue situated two amino acids before the C-terminus (R-2, RXX) of the presequence.<sup>71</sup> Another element that helps MPP recognise presequences for cleavage is an aromatic residue, such as phenylalanine or tyrosine, at position +1.<sup>97,100</sup> MPP is not the only peptidase that can cleave N-terminal presequences as recently the human insulin-degrading enzyme (IDE) and its yeast homologue Ste23 have been linked to MTS cleavage and clearance.<sup>101</sup> Although IDE is mainly found in the cytosol, an isoform resulting from the alternative translation of IDE localises to mitochondria and is believed to have an MPP-like function.<sup>102</sup> Ste23 interacts with Cym1, an established peptide-degrading enzyme in the matrix to degrade presequence peptides and promote protein maturation.<sup>101</sup> Presequence cleavage is not only important for protein maturation but also for cell viability. If MPP is impaired the preproteins will not mature and thus the accumulation of unfolded proteins might lead to aggregation in mitochondria.<sup>101,103</sup>

### 3.2 Inner membrane targeting signals

Most inner membrane proteins belong to a family of solute transport proteins (also called metabolite carrier proteins), such as the ADP/ATP carrier, and contain six  $\alpha$ -helical transmembrane domains. Unlike the aforementioned cleavable presequences, carrier precursors contain internal targeting elements that remain part of the mature protein.<sup>12,42</sup> The carrier protein precursors are guided primarily to the Tom70 receptor of the TOM complex by cytosolic chaperones. Nonetheless, there is growing evidence that Tom70 is not a selective receptor protein strictly involved in mitochondrial protein import and it functions as a versatile protein adaptor that facilitates the interaction of mitochondria with different proteins but also other organelles such as the ER.<sup>104</sup> The carrier precursors are inserted into the Tom40 channel in a loop formation with both termini remaining exposed on the cytosolic side.<sup>42,77</sup> Subsequently, chaperones are recruited in the IMS-side of Tom40 that facilitate the translocation of the hydrophobic precursors and protect them from aggregating or misfolding in the aqueous phase. These chaperones are the small Tims, Tim9-10 and Tim8-13, that we will discuss in detail in the next section. The transient binding of the importing carrier proteins to the small Tim chaperones is guided mainly by the transmembrane segments of the carrier proteins that function as the internal targeting signals.<sup>36,40</sup> The presence of the entire sequence of the carrier proteins is presumably necessary to ensure the correct targeting of these proteins, since truncated versions of carrier proteins may be mistargeted to the matrix due to the presence



of cryptic matrix targeting signals that would be rendered non-functional in the context of the entire protein and therefore allow proper targeting to the inner membrane.<sup>105</sup> The small Tim chaperone complexes associate with the TIM22 complex of the inner membrane to aid the transfer of the precursor protein for final insertion into the membrane. The TIM22 complex consists of Tim18, Tim22 and Tim54, with Tim18 and Tim54 stabilising the complex. TIM22 also requires the membrane potential in order to facilitate the insertion of the transmembrane segments of the integral inner membrane proteins.<sup>38,106–108</sup>

### 3.3 Intermembrane space targeting signals

Although the IMS is the smallest mitochondrial subcompartment, it is vital for mitochondrial function. It constitutes a crucial buffer between the cytosol and matrix through the exchange of lipids, proteins, metals and other cofactors necessary for mitochondrial function, redox regulation and apoptosis.<sup>10,71</sup> Additionally, there are numerous protein import pathways into the IMS perhaps indicative of its pleiotropic nature. One well-characterised type of targeting signal for the IMS is the bipartite presequence for a small number of IMS proteins. Additionally, for the majority of IMS proteins contain characteristic cysteine motifs<sup>48,109</sup> and have their own unique targeting signal. The proteins with bipartite presequences contain an N-terminal mitochondrial targeting signal (MTS) as well as a hydrophobic sorting region. The N-terminal part of the presequence is similar to the cleavable mitochondrial targeting presequences, hence it is positively charged and forms an amphipathic helical structure. The positive domain of the bipartite sequence is recognised by the TOM complex receptors Tom20 and Tom22. This signal then allows the presequence to transverse the Tom40 channel. Once in the IMS side of the TOM complex (*trans* site), it is recognised by Tom50 that in turn feeds the presequence to TIM23.<sup>74</sup> The latter region, known as the stop-transfer sequence, arrests translocation to the matrix inserted as the hydrophobic segment gets “stuck” in the Tim23 channel. The transmembrane domain is then laterally diffused and the presequence is anchored to the inner membrane with the N-terminal peptide facing the matrix. The MPP then cleaves the presequence to facilitate protein maturation.<sup>101</sup> There is a secondary cleavage site between the hydrophobic segment and the mature protein facing the IMS. This motif is recognised and cleaved by inner membrane peptidase (IMP).<sup>27,110</sup> The IMP is a heterodimer consisting of two different catalytic subunits; the Imp1 and Imp2 that are anchored to the inner membrane by their N-terminal domain. A third protein Som1 was shown to be involved in the IMP complex and was found to be essential for the proteolytic activity of the Imp1 subunit. The IMP complex cleaves the transmembrane segment leading to the release of the mature protein in the IMS.<sup>49</sup>

The majority of intermembrane space proteins lack these bipartite presequences. Instead, they contain an internal, non-cleavable targeting signal with hydrophobic residues and characteristic cysteine motifs (CX<sub>3</sub>C and CX<sub>9</sub>C) that form intramolecular disulphide bonds as a consequence of their interaction with the MIA machinery. These targeting signals are

termed mitochondrial IMS sorting signal (MISS) or the IMS targeting signal (ITS). The ITS peptide is usually nine amino acids-long and contains conserved residues that are important for both substrate recognition and folding of the substrate protein. The cysteine residues are usually separated by 3 or 9 amino acids, although some MIA substrates contain cysteines not organised in a specific motif. A comprehensive protein engineering and mutagenesis analysis has defined the consensus motif in the ITS signal, CXX[Hy][Hy]XX[Ar]X, where the hydrophobic residues in position –3 and –4 and an aromatic residue in position –7 are strictly conserved (Fig. 4A).<sup>57</sup> The substrates of the Mia pathway are usually small proteins that remain reduced in the cytosol and are translocated to the IMS *via* the TOM channel. Once in the IMS, the ITS targeting signal is recognised by the oxidoreductase Mia40. The ITS forms an amphipathic  $\alpha$ -helix with the hydrophobic face containing the conserved residues that are necessary for recognition and binding by Mia40 (Fig. 4B). Unlike the MTS peptide for targeting to the matrix, the charged residues of the ITS peptide are not functionally important.<sup>57</sup> Recognition of the ITS by Mia40 occurs *via* site-specific hydrophobic interactions which allow a two-step disulphide bond formation event between the CPC site of Mia40 and the substrate's internal targeting signal (Fig. 4C and D).<sup>11,53,57,111</sup> Folding of the substrate occurs *via* two consecutive induced folding events, each one coupled to one disulphide formation:

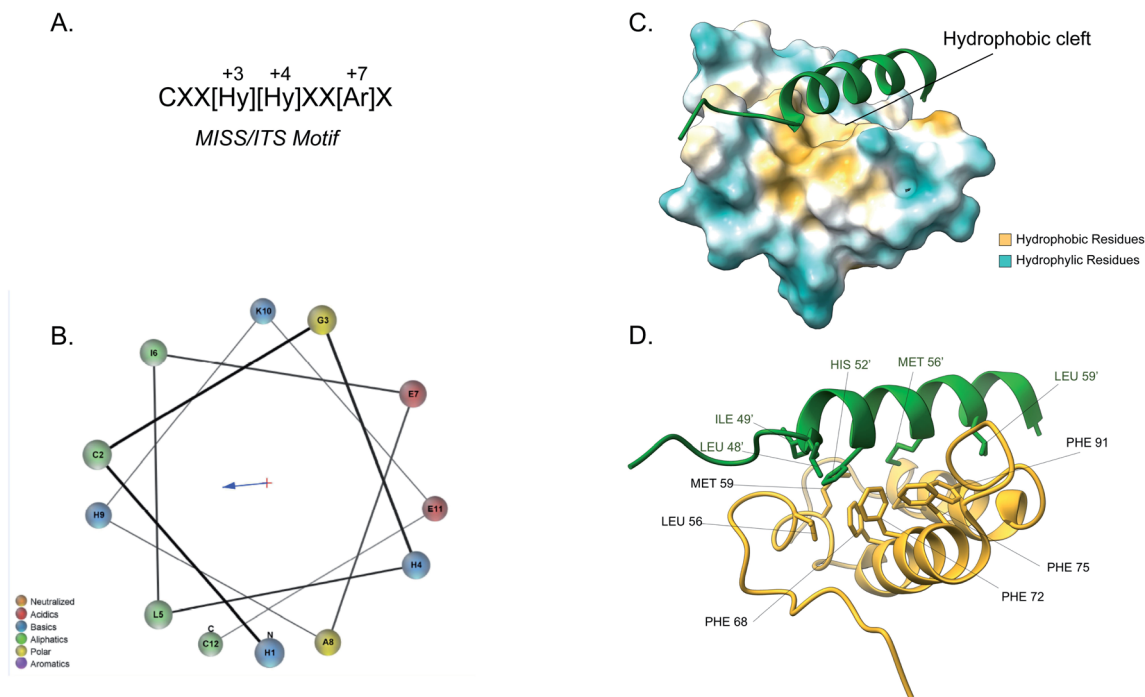
The first ITS helix folding is induced by the substrate cleft of Mia40, coupled to the formation of the intermolecular disulphide. The second helix is induced by the now folded ITS first helix, coupled to the intramolecular disulphide formation between the inner cysteines.

Covalent binding is thermodynamically essential to induce the  $\alpha$ -helical folding of the Mia40 substrates and stabilise the protein–protein interactions. Once the disulphide bond between Mia40 and the substrate is formed, entropic contributions do not prevent the folding anymore and the formation of the  $\alpha$ -helix is thermodynamically favoured. After substrate release, the Mia40 CPC motif remains reduced and gets reoxidised by Erv1 back to its functional state as an IMS receptor.<sup>62,112,113</sup> The Mia40 has an expanded specificity by recognising substrates with unconventional cysteine motifs. For example, Mix23 displays a CX<sub>14</sub>C/CX<sub>13</sub>C and Erv1 is imported *via* a CXXC motif.

### 3.4 Atypical IMS targeting signals

In addition to the majority of IM proteins that contain an ITS targeting signal, there are proteins that have atypical targeting signal which guide them into the IMS.

**The case of cytochrome proteins.** Cytochrome b2 and cytochrome c1 follow the stop-transfer pathway, but their final localisation in the IMS depends additionally on a heme-binding domain (HBD) downstream of the stop-transfer signal. When their presequence is translocated across the outer membrane into the IMS, the HBD domain can fold independently of the presequence. As a result, the mature form of the HBD prevents the backward movement of the presequence and thus the



**Fig. 4** Mia40 interaction with the MISS/ITS peptide of Cox17. (A) Intermembrane space targeting sequence motif: L stands for leucine while X can be any amino acid. (B) Helical wheel projection of the presequence peptide amphipathic helix. The blue arrow points to the hydrophobic side of the helix. (C) Interaction of the Cox17 ITS/MISS (green) with the hydrophobic cleft of Mia40. The hydrophobic surface of Mia40 is indicated in yellow while the hydrophilic surface is light blue. (D) Contact residues of Cox17 (green) at the Tom20 interface. Cox residues are represented in green. Atom distance is  $<4.0 \text{ \AA}$  (PDB ID: 2L0Y).

presequence can only move forward into the IMS.<sup>27,74,114</sup> The HBD domain also drives the import of cytochrome C into the IMS.<sup>115,116</sup> Cytochrome C is one of the Cys containing proteins in the IMS that exceptionally does not depend on the MIA pathway for its import.<sup>48</sup>

**Gpx3.** Gpx3 is the major oxidative stress sensor in the cytosol. A ribosomal profiling study in *S. cerevisiae* provided evidence that under  $\text{H}_2\text{O}_2$  stress conditions, ribosomal binding sites shifted upstream of the endogenous AUG codon leading to alternative translation.<sup>117</sup> Gpx3 localises to the cytosol when translated through its canonical AUG start codon. However, under acute oxidative stress, an isoform of the protein is synthesized, with an 18-amino acid N-terminal peptide, from the translation of a non-AUG codon and is translocated to mitochondria. The N18 peptide is mostly hydrophobic and has positive charges but does not follow any of the conventional protein import pathways into the intermembrane space.<sup>118</sup> Intriguingly, the Gpx3 without the N-terminal extension also gets targeted to mitochondria, albeit in lower yields than the N-extended form, suggesting that an as yet uncharacterised targeting signal must be present even in the mature part of the protein.<sup>8</sup> Gpx3 is not the only protein that is dually localised in the IMS in addition to its cytosolic localisation. Other such proteins are the thioredoxin 1 (Trx1) and thioredoxin 2 (Trx2) and thioredoxin reductase Trr1, which are both targeted to the IMS by yet unknown targeting signals.<sup>119</sup>

### 3.5 Outer membrane targeting signals

The outer membrane possesses two types of integral membrane proteins:  $\beta$ -barrel and  $\alpha$ -helical proteins.  $\beta$ -Barrel proteins are integrated to the outer membrane by multiple  $\beta$ -strands traversing the membrane while  $\alpha$ -helical proteins are anchored to the membrane by one or more transmembrane  $\alpha$ -helices.<sup>2</sup>  $\beta$ -Barrel proteins make up two of the critical translocases of the protein import pathway (the Tom40 and Sam50 channel forming subunits of the TOM and SAM complex) and their function is therefore essential. They are synthesised in the cytosol like 99% of mitochondrial proteins but do not contain a cleavable targeting signal. Instead, the targeting signal of  $\beta$ -barrel proteins is internal to the protein and remains part of the mature protein. At first, the  $\beta$ -strands on the proteins were believed to be the targeting signal for these proteins but no specific motif was identified.<sup>120</sup> However, it was later discovered that the actual targeting signal consists of the most C-terminal  $\beta$ -strand called the  $\beta$ -signal;<sup>121</sup> the properties of the  $\beta$ -signal have been further dissected to include high hydrophobicity and form a beta hairpin element that can extend to include two adjacent  $\beta$  strands that specifically bind the translocation machinery.<sup>122</sup> The  $\beta$ -hairpin is recognised by Tom20 and subsequently translocated through the Tom40 channel. Upon translocation, TIM chaperones in the IMS bind to precursor in order to prevent aggregation in the aqueous sub-compartment. It is then transferred to the SAM complex, which is the central machinery for the  $\beta$ -protein biogenesis pathway. The SAM complex contains Sam50, a  $\beta$ -barrel protein itself, which is integrated in the outer

membrane, as well as Sam35 and Sam37 that are peripheral membrane proteins exposed to the cytosol. Sam35 and Sam50 cooperate in the recognition of the  $\beta$ -hairpin motif and direct membrane insertion. The  $\beta$ -signal is then proposed to induce the opening of the lateral gate of Sam50 which leads to the lateral release of the protein to the outer membrane.<sup>2,121</sup> Sam37 is understood to support the release of the  $\beta$ -barrel protein from the SAM complex to the OM but the exact mechanism of the release is not understood in great detail.<sup>123</sup>

The other group of OM proteins are  $\alpha$ -helical OM proteins that are anchored to the membrane by one or more  $\alpha$ -helices. There are no known examples of OM proteins that will have both  $\beta$ -barrel structure and  $\alpha$ -helical transmembrane domains. They have various functions from protein transport to apoptosis. The biogenesis of these proteins is only understood in part. It appears that the positive charge and moderate hydrophobicity of their membrane anchor acts as a targeting signal.<sup>124</sup> They are split into three main protein classes: signal-anchored, tail-anchored and polytopic outer-membrane proteins. Signal-anchored and tail-anchored proteins contain

an  $\alpha$ -helical transmembrane section at the N- and C-terminus, respectively. The transmembrane domain and its flanking regions function as both a membrane anchor and a targeting signal. The exact targeting signal for polytopic proteins is not known but it may be associated with the multiple transmembrane domains of the proteins. Tom70 is believed to be the receptor for such proteins but it is unclear if they pass through the Tom40 channel for their import.<sup>125</sup> The MIM complex is proposed to have a major role in this pathway; it can mediate the transfer of substrates from the TOM complex as well as promote the import and assembly of all types of  $\alpha$ -helical proteins (Table 2).<sup>2,123</sup>

## 4. Chaperone-assisted protein folding in the IMS

The specific translocon machineries and their interaction with the variety of targeting signals that we have discussed so far ensure the correct targeting and sorting of the imported

**Table 2** Targeting and sorting signals for the import of mitochondrial proteins

Proteins targeted to:	Targeting signal	Description	Import machinery	Example
Matrix	N-terminal cleavable presequence	<ul style="list-style-type: none"> <li>• Amphipathic <math>\alpha</math>-helix</li> <li>• 15–60 residues</li> <li>• Net charge of +3 to +6</li> <li>• High content of hydroxylated residues</li> <li>• MPP cleavage site</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom20, Tom22, Tom70</li> <li>• Translocases: Tom40, Tim23</li> </ul>	TRX3: thioredoxin required to maintain redox homeostasis in the cell
Inner membrane	Internal targeting signal	<ul style="list-style-type: none"> <li>• Six <math>\alpha</math>-helical transmembrane domains</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom70</li> <li>• Translocases: Tom40, Tim22</li> </ul>	OXA1: mitochondrial inner membrane insertase
Outer membrane	Signal-anchored	<ul style="list-style-type: none"> <li>• N-terminal <math>\alpha</math>-helical transmembrane domain</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom70</li> <li>• Translocases: Tom40, MIM complex</li> </ul>	Tom70: component of the TOM complex; acts as a receptor for incoming precursor proteins
	Tail-anchored	<ul style="list-style-type: none"> <li>• C-terminal <math>\alpha</math>-helical transmembrane domain</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom70</li> <li>• Translocases: Tom40, MIM complex</li> </ul>	Tom5: component of the TOM complex; involved in transfer of precursors from the Tom70 and Tom20
	Multiple transmembrane domains	<ul style="list-style-type: none"> <li>• Multiple <math>\alpha</math>-helical transmembrane domains</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom70</li> <li>• Translocases: Tom40, MIM complex</li> </ul>	Tom22: component of the TOM complex; mediates interaction between TOM and TIM complexes
IMS	$\beta$ -Barrels	<ul style="list-style-type: none"> <li>• Multiple transmembrane <math>\beta</math>-strands</li> <li>• Most C-terminal <math>\beta</math>-strand is the targeting signal (<math>\beta</math>-signal)</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom20</li> <li>• Translocases: Tom40, SAM complex</li> </ul>	Tom40: component of the TOM complex; constitutes the core element of the protein conducting pore
	N-terminal cleavable presequence with stop-transfer sequence	<ul style="list-style-type: none"> <li>• Amphipathic <math>\alpha</math>-helix</li> <li>• 15–60 residues</li> <li>• Hydrophobic sorting region</li> <li>• IMP cleavage site</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom20, Tom22</li> <li>• Translocases: Tom40, Tim23</li> </ul>	GUT2: mitochondrial glycerol-3-phosphate dehydrogenase
	Internal targeting signal with characteristic Cys motif	<ul style="list-style-type: none"> <li>• Amphipathic <math>\alpha</math>-helix</li> <li>• 9 amino acids long</li> <li>• Hydrophobic residues in position –3 and –4</li> <li>• Aromatic residue in position –7</li> </ul>	<ul style="list-style-type: none"> <li>• Receptors: Tom20</li> <li>• Translocases: Tom40, Mia40</li> </ul>	Erv1: flavin-linked sulfhydryl oxidase of the mitochondrial IMS, oxidizes Mia40 as part of the disulfide relay system

proteins to the different mitochondrial sub-compartments. However, as the transported proteins are targeted to the organelle in a largely unfolded state (which allows them to thread efficiently through the narrow translocon channels), they need to be folded at the *trans* site inside the organelle after the transport process has been completed. The newly synthesised proteins that are to be transported to mitochondria are kept unfolded by a system of cytosolic chaperones to avoid protein aggregation and premature damage.<sup>126</sup> In mitochondria, dedicated molecular chaperones assist the folding of newly imported proteins and prevent their aggregation while their hydrophobic segments are transported through the aqueous sub-compartments.

In the mitochondrial matrix, various molecular chaperones have been identified that help newly imported proteins to fold. These belong to the well characterised families of Hsp70 Hsp60 and their cochaperones Hsp40 and Hsp10. One key chaperone in the matrix is for example the mitochondrial heat shock protein (mtHsp70) is an ATP-dependent import motor of the TIM23 complex that not only helps drive protein translocation, but also mediates the folding of the imported substrates.<sup>24,127</sup>

By contrast, the IMS presents several unique challenges as a folding environment for the newly import proteins. First, the IMS does not retain its own pool of ATP as the matrix, and in fact there are no ATP-dependent chaperones in the IMS.<sup>128</sup> Second, the IMS is the most constricted mitochondrial sub-compartment, with local concentrations of proteins very high and an increasing potential for aggregation, particularly for proteins with exposed hydrophobic patches which are common for the unfolded imported proteins. Finally, the extended physical interactions between the outer and inner membranes in contact sites and the internal segregation of the cristae lumen separately from the boundary intermembrane space put further constraints on the folding process in this compartment.

In this context, we will discuss in this chapter the three main chaperone systems in the IMS. Namely, we will focus on the structure and chaperone function of Mia40 as a holdase, the

function of the small Tims as dedicated membrane protein chaperones and finally the chaperone role of the protease Yme1.

#### 4.1 The Mia40 holdase

As previously described, several IMS proteins are required to go through oxidative folding during maturation by interaction with an oxidoreductase, namely Mia40 in yeast or its human homolog CHCHD4.<sup>53,56</sup> The most notable difference between the two homologs is that the yeast variant contains a large N-terminal extension of 242 amino acids that tethers the protein into the inner membrane facing the IMS, while this segment is absent in CHCHD4 resulting in the protein being entirely soluble in the IMS. Despite their variance, they share a high sequence identity in the central and catalytically active region of the protein (residues 47–107 in CHCHD4 and residues 290–350 in yeast) which include the conserved cysteine residues in a –CPC–CX<sub>9</sub>C–CX<sub>9</sub>C– motif. The Nuclear Magnetic Resonance (NMR) structure of this segment for the human protein (Fig. 5) revealed the protein core to be composed of 3 helices;  $\alpha$ 1 (residues 56–59) which is a very short helix that contains the redox-active CPC motif and the only defined secondary structure of the N-terminal segment of the protein core;  $\alpha$ 2 (residues 65–77) and  $\alpha$ 3 (residues 88–100) form an antiparallel  $\alpha$ -hairpin that is stabilized by the disulphide pairs Cys64–Cys97 and Cys74–Cys87. The NMR data obtained also suggests that the N-terminal segment varies in flexibility upon disulphide bond formation in the CPC motif, thus otherwise referred to as the “lid”.<sup>53</sup> Mia40 substrates share a coil–helix coiled coil–helix (CHCH) motif that is imported into mitochondria in a reduced and unfolded form.<sup>34,52</sup> The oxidative folding ability of Mia40 depends on the formation of transient intramolecular disulphide bridges between the Cys55 of the CPC motif and specific cysteines of the substrate proteins. Successful binding achieves a rapid disulphide exchange that results in the CHCH domain of substrates being oxidised while Mia40 is reduced. Such exchanges include the Cys45 in the CX<sub>9</sub>C motif of Cox17,<sup>53</sup> and the first cysteine of the CX<sub>3</sub>C motif in Tim9 and Tim10 that

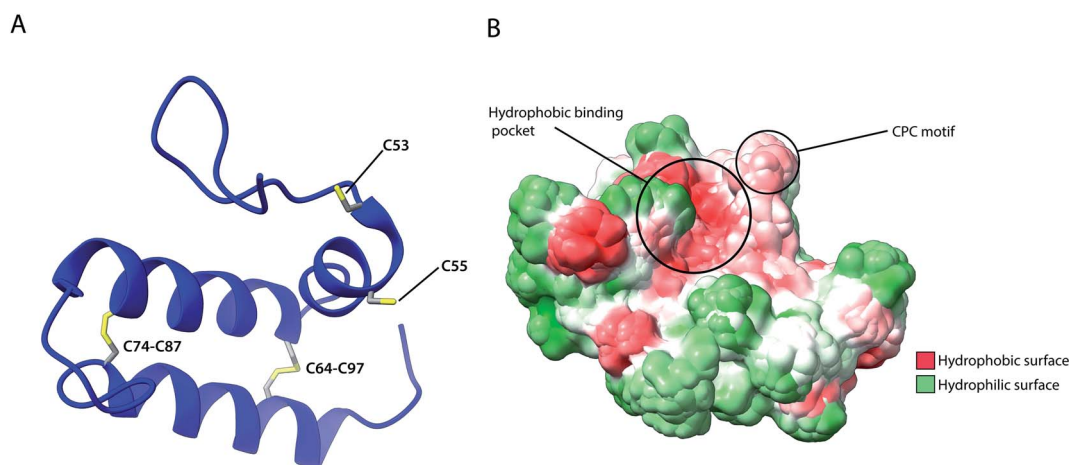


Fig. 5 Structural features of human Mia40. (A) Structure of the protein core (residues 49–109), highlighting intramolecular disulphide bridges as well as Cys53 and Cys55 of the catalytically active CPC motif. (B) Hydrophobicity surface, highlighting the hydrophobic substrate binding cleft and the adjacent CPC motif (PDB ID: 2K3J).



regulates the assembly of the TIM9·10 complex.<sup>37,129</sup> Aside from the oxidative function of Mia40, the orientation of the substrate with the correct Cys to the CPC motif was found to be crucial to achieving efficient disulphide exchange and strictly dependent on hydrophobic packing between the ITS and the Mia40 substrate cleft.<sup>37</sup> The Mia40 structure revealed several conserved hydrophobic residues on the  $\alpha$ -helix opposite the CPC motif made up of residues Leu42, Ile43, Ile49, Trp51, Leu56, Met59, Ala60, Phe68, Phe72, Phe75, Phe91, Met94 and Met98. These residues were collectively shown to be essential for cell survival.<sup>53</sup> The hydrophobic collapse of the ITS hydrophobic residues, that make up the solvent-exposed hydrophobic binding cleft of Mia40, stabilise the unfolded substrate in the IMS milieu, with Mia40 functioning as a holdase.<sup>57,130</sup> This initial step of substrate binding to Mia40 (and hence its chaperone activity) is independent of subsequent oxidation (*i.e.* disulphide bonding) steps.<sup>57,112</sup> In fact, the hydrophobic interactions dictated by the substrate binding cleft of Mia40 are also involved in the recycling of Mia40 by Erv1. In this reaction, Erv1 binds to Mia40 in a manner that mimics the substrate ('substrate mimicry model') and making use of an extended hydrophobic patch in the intrinsically disordered N-terminal of Erv1.<sup>62,63</sup> While Mia40 catalyses disulphide bridges that thermodynamically favour  $\alpha$ -helical folding for its substrates,<sup>131</sup> *in vitro* experiments of Koch *et al.*, with Mia40 and Cox17 proposed

an additional proofreading ability for the chaperone by re-shuffling incorrect disulphides within its substrate protein.<sup>132</sup>

#### 4.2 The small Tim chaperones for polytopic proteins

Among the substrates of Mia40 are the small Tims, a family of  $\sim 10$  kDa chaperones that exist within the IMS of mitochondria. Their most well studied function is the transport metabolite carriers across the IMS, by protecting exposed hydrophobic regions of these integral membrane proteins.<sup>35,36,108</sup> Additionally, the small Tims also assist the integration of  $\beta$ -barrel protein in the OM of mitochondria, making them the only known chaperone system capable of interacting with both types of membrane proteins ( $\beta$ -barrels and  $\alpha$ -helical) the family members include Tim8, Tim9, Tim10, Tim12 and Tim13 in *S. cerevisiae* and are organised in specific oligomeric assemblies. Tim9 pairs with Tim10 and Tim8 with Tim13. They share a conserved  $CX_3CX_nCX_3C$  motif (where  $n \sim 15$ ) or two  $CX_3C$  motifs that are essential for their function.<sup>52,129</sup> Tim9 and Tim10 are essential proteins<sup>35,133</sup> that share 22% sequence identity and 50% sequence similarity. By contrast, Tim8 and Tim13 are not essential for viability. Following their import into the IMS, they assemble into a thermodynamically stable 70 kDa hexameric complex.<sup>37,134</sup> Once assembled, Tim12, an essential subunit of the TIM22 complex binds the Tim9·10 complex to enable (i) substrate insertion into the inner membrane and (ii) dual

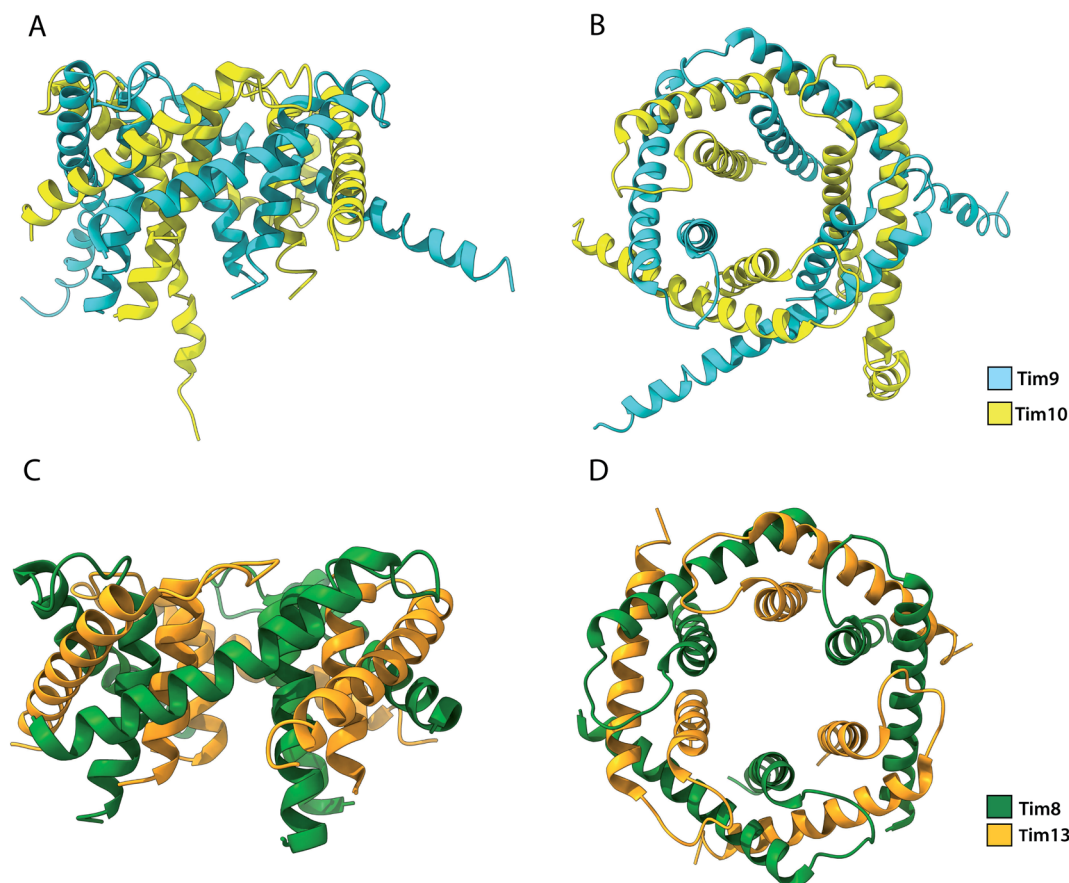


Fig. 6 Structure of the Tim9·10 and Tim8·13 complex. (A) Side view and (B) top view of the Tim9·10 complex (PDB ID: 2BSK). (C) Side view and (D) top view of the Tim8·13 complex (PDB ID: 3CJH).

localisation for the chaperone, in the bulk of the IMS and tethered to the inner membrane in a Tim9·10·12 complex intermediate.<sup>35,135</sup> Folding of the unassembled Tim9 and Tim10 requires their oxidation by Mia40 and it is only the oxidised and fully folded Tim9 and Tim10 that can assemble to form a functional heterohexameric Tim9·10 chaperone complex.<sup>52,129,134</sup> The secondary structure of the complex is exclusively  $\alpha$ -helical, where the two subunits alternate in a 6-fold molecular symmetry to form a closed circular core. Each subunit folds into a helix-loop-helix as disulphide bridges in-between CX<sub>3</sub>C motifs create a central loop. The whole complex comprises 6 inner and 6 outer helices (Fig. 6). In the central part of the subunit, they form a doughnut-shaped substrate-binding core around a hollow pore of 15 Å, while the antiparallel N- and C-termini extend downward of the core.<sup>134</sup> Depending on the length of the transferred protein, multiple chaperones may attach to the same chain for its efficient protection.<sup>136</sup> Sequence analysis of small Tims identified a set of conserved hydrophobic residues that face the binding cleft found in-between the inner and outer helices in both Tim9·10 and Tim8·13 complexes.<sup>137</sup> Mutagenesis of these conserved residues to hydrophilic residues in Tim9·10 resulted in cell death and are thus proposed to be essential for the biogenesis of its substrate proteins.<sup>136</sup> However, it could be argued that replacing multiple residues in the protein core with amino acids of inverse physicochemical properties likely affects the structure of the complex and elicit this cell response. The ADP/ATP carrier protein AAC has been widely reported to bind the Tim9·10 complex during its transport through the IMS to the translocase Tim22.<sup>36,105,134</sup> While the transmembrane domains of AAC bind the Tim9·10 complex core,<sup>36</sup> Tim10 was shown to interact with AAC using its N-terminal domain. The individual subunits' role in substrate binding suggests that Tim9 has more of a structural role within the complex while Tim10 maintains substrate specificity and affinity.<sup>105</sup> Though not essential in yeast, the Tim8·13 complex was shown to be most important in low membrane potential conditions.<sup>39</sup> Import of its primary substrate, Tim23 is heavily dependent on membrane potential. By binding the N-terminal domain of Tim23, the Tim8·13 complex prevents protein aggregation during import as well as prevents its retrograde translocation.<sup>39,41</sup> The Tim8·13 complex shares structural similarities with Tim9·10 but have been shown to have different substrate-binding preferences, as well as additional hydrophilic interactions not reported in the Tim9·10 complex. It appears the two complexes have separate roles within the IMS. However, membrane precursor proteins were shown to be transferrable from one TIM chaperone to the other, indicating a possibility for cooperation in their chaperone activity.<sup>137</sup>

### 4.3 The Yme1: a protease with a chaperone activity

Yme1 is a member of conserved family of AAA+ ATPases that are associated with a plethora of cellular processes.<sup>138</sup> These include: cellular quality control, organelle biogenesis, membrane fission and vesicular transport.<sup>139</sup> Failure to sustain protein quality control may lead to impaired cellular function and as a result the development of numerous human

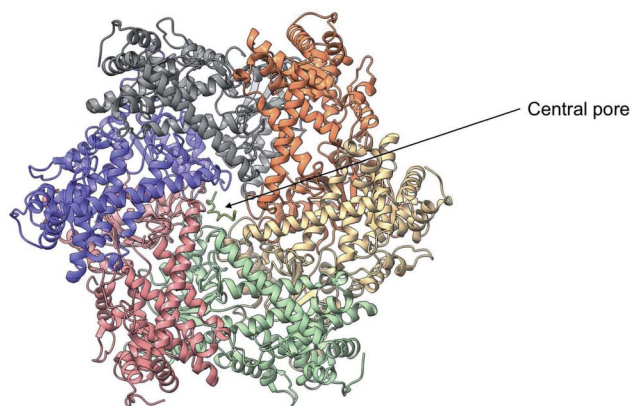


Fig. 7 Cryo-EM structure of Yme1. This is the top view of the protein showing the C6-symmetric protease ring that faces the intermembrane space. A central pore of  $\sim 1.4$  nm diameter is formed where the substrate binds (PDB ID: 6AZ0).

diseases.<sup>140</sup> Yme1 is involved in mitochondrial protein quality control and has been specifically associated with the degradation or unfolded or damaged proteins. However, further research has revealed an additional role of Yme1 as a molecular chaperone assisting the folding of newly imported polypeptides.<sup>141,142</sup>

Despite the crucial function of Yme1, little was known about its structure until Puchades *et al.*, took on the arduous task of resolving its cryo-EM structure.<sup>138</sup> Yme1 is a hexameric protein and every subunit contains an ATPase and peptidase domain (Fig. 7). As seen in Fig. 8, Yme1 is tethered to the inner membrane by a membrane helix and its catalytic domains are facing the IMS.<sup>143</sup> Yme1 assembles into two stacked rings, with the ATPase domains creating an asymmetric spiral staircase above the protease ring. The staircase creates a central pore ( $\sim 1.4$  nm diameter) with a conserved aromatic-hydrophobic motif that is essential for binding substrate proteins. ATPases are ATP-driven machineries that are characterised by a conserved P-looped AAA domain containing Walker A and Walker B motifs. The AAA domain of Yme1 contains a conserved Walker A motif, with the pattern G-X(4)-GK-[TS].<sup>144</sup> Numerous studies show the involvement of the motif in recognising and

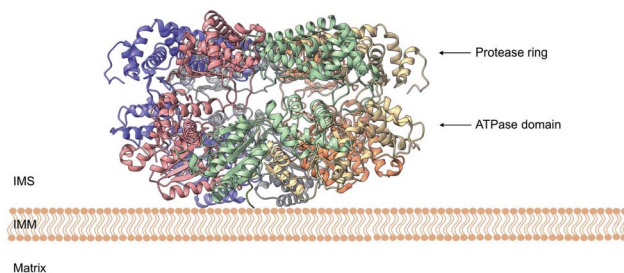


Fig. 8 Side view of the cryo-EM structure of Yme1. The protein is tethered to the inner mitochondrial membrane (IMM) by a membrane helix. The subunits of the ATPase domain create an asymmetric spiral staircase while the protease ring subunits create a planar symmetric C6 ring. The purple and green structures represent the different translocases found on the IMM (PDB ID: 6AZ0).

binding substrates for translocation or degradation. It is also hypothesised that this motif is involved when Yme1 binds substrates to prevent aggregation or mediate protein folding.<sup>141</sup>

The AAA domain of Yme1 was first shown to act as a chaperone by binding to unfolded protein dihydrofolate reductase (DHFR) in *in vitro* experiments using isolated mitochondria.<sup>141,145</sup> Moreover, it was suggested that Yme1 assists with the folding/assembly of Cox2, a subunit part of the catalytic core of cytochrome c oxidase. Thus, supporting that Yme1 acts as a chaperone also *in vivo* conditions.<sup>146</sup> Later research confirmed the role of Yme1 as a folding helper for DHFR in the IMS. Importantly, the chaperone activity of Yme1 was independent from its protein degradation function. The study also identified numerous endogenous substrates that aggregate in the absence of Yme1. These proteins are necessary for the aforementioned functions of the IMS, thus further supporting the importance of Yme1.<sup>142</sup> Interestingly, Erv1 was identified as one of the endogenous substrates of Yme1. As mentioned above, Erv1 is an essential component of the oxidative folding pathway. Therefore, Yme1 presumably has a crucial role in a variety of mitochondrial pathways.

Despite the growing evidence in support of Yme1 as a chaperone in the IMS, the mechanism of its function is not yet clear. It is hypothesized that Yme1 binds newly imported proteins during their translocation and assists with their folding. Another hypothesis is that the AAA domain of Yme1 binds unfolded proteins and helps keep them in a folding-competent state to enable their refolding. If folding is unsuccessful, then the proteins get degraded by the proteolytic domain of Yme1.<sup>142</sup> The cryo-EM structure of Yme1 has given a better picture of the architecture of the protein and the residues involved in substrate binding.<sup>138</sup> Nevertheless, the exact mechanism of the chaperone activity of Yme1 still needs to be determined.

The different chaperone systems in the IMS (Mia40, small Tims and Yme1) seem to display specificity against their substrates with no overlap in their function. However, the small Tims interact both with Mia40 as a substrate and they also seem to be subject to quality control by the Yme1. It is not yet very clear whether other proteins may be processed by more than one of the IMS chaperone systems.

## 5. Conclusions

Mitochondria have developed multiple import pathways to accommodate the import of the different types of proteins needed for mitochondrial function. Most mitochondrial sub-compartments have typically one or two major import pathways, but the intermembrane space seems to divert with a very high variety of pathways, some of which are dictated by targeting signals that are still either completely unknown or poorly understood. This is surprising given that the IMS is the smallest mitochondrial sub-compartment, but it is perhaps indicative of the pleiotropic function of the IMS.

The emergence of new IMS proteins carrying atypical signals is of great interest as not much is known about targeting signal specificity and its recognition by the various import components. Furthermore, there is growing evidence that proteins like

the Mia40, the small Tims and Yme1 have more functions than previously anticipated. Taken together, these might suggest that mitochondria have found ways to evolve in order to recognise novel targeting signals or have developed such alternative import pathways as a backup when mitochondrial function is impaired. Thus, it is important to dissect how these unconventional targeting pathways work in order to better understand mitochondrial function and how it can be altered to prevent disease progression.

## Conflicts of interest

The authors declare no conflict of interests.

## Acknowledgements

Work in our laboratory is supported by UKRI-BBSRC (grants BB/R009031/1 and BB/T003804/1), a BBSRC-EPSC Impact Accelerator grant and a Wellcome Trust early concept development grant (University of Glasgow). Structural biology work in our laboratory benefited from access to the Solution NMR, CERM/CIRMMP, Florence, Italy, an Instruct-ERIC centre. Financial support was provided by Instruct-ERIC (PID: 12148). We thank all members of our lab for useful discussions and comments on the manuscript and Dr Ross Eaglesfield for help with Fig. 3–6.

## References

- 1 N. Wiedemann and N. Pfanner, *Annu. Rev. Biochem.*, 2017, **86**, 685–714.
- 2 N. Pfanner, B. Warscheid and N. Wiedemann, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 267–284.
- 3 A. P. West, G. S. Shadel and S. Ghosh, *Nat. Rev. Immunol.*, 2011, **11**, 389–402.
- 4 D. J. Pagliarini, S. E. Calvo, B. Chang, S. A. Sheth, S. B. Vafai, S. E. Ong, G. A. Walford, C. Sugiana, A. Boneh, W. K. Chen, D. E. Hill, M. Vidal, J. G. Evans, D. R. Thorburn, S. A. Carr and V. K. Mootha, *Cell*, 2008, **134**, 112–123.
- 5 O. Schmidt, N. Pfanner and C. Meisinger, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 655–667.
- 6 S. E. Calvo, K. R. Clauser and V. K. Mootha, *Nucleic Acids Res.*, 2016, **44**, D1251–D1257.
- 7 S. Rath, R. Sharma, R. Gupta, T. Ast, C. Chan, T. J. Durham, R. P. Goodman, Z. Grabarek, M. E. Haas, W. H. W. Hung, P. R. Joshi, A. A. Jourdain, S. H. Kim, A. V. Kotrys, S. S. Lam, J. G. McCoy, J. D. Meisel, M. Miranda, A. Panda, A. Patgiri, R. Rogers, S. Sadre, H. Shah, O. S. Skinner, T. L. To, M. A. Walker, H. Wang, P. S. Ward, J. Wengrod, C. C. Yuan, S. E. Calvo and V. K. Mootha, *Nucleic Acids Res.*, 2021, **49**, D1541–D1547.
- 8 R. Edwards, S. Gerlich, S. Gerlich and K. Tokatlidis, *Biol. Chem.*, 2020, **401**, 737–747.
- 9 K. P. Baker and G. Schatz, *Nature*, 1991, **349**, 205–208.
- 10 A. Chatzi, P. Manganas and K. Tokatlidis, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2016, **1863**, 1298–1306.
- 11 P. Manganas, L. MacPherson and K. Tokatlidis, *Cell Tissue Res.*, 2017, **367**, 43–57.



- 12 J. C. Young, N. J. Hoogenraad and F. U. Hartl, *Cell*, 2003, **112**, 41–50.
- 13 T. Hase, H. Riezman, K. Suda and G. Schatz, *EMBO J.*, 1983, **2**, 2169–2172.
- 14 W. Neupert and J. M. Herrmann, *Annu. Rev. Biochem.*, 2007, **76**, 723–749.
- 15 K. P. Baker, A. Schaniel, D. Vestweber and G. Schatz, *Nature*, 1990, **348**, 605–609.
- 16 K. Dietmeier, A. Hönlinger, U. Bömer, P. J. T. Dekker, C. Eckerskorn, F. Lottspeicht, M. Kübrich and N. Pfanner, *Nature*, 1997, **388**, 195–200.
- 17 P. J. T. Dekker, M. T. Ryan, J. Brix, H. Müller, A. Hönlinger and N. Pfanner, *Mol. Cell. Biol.*, 1998, **18**, 6515–6524.
- 18 E. L. Sherman, N. E. Go and F. E. Nargang, *Mol. Biol. Cell*, 2005, **16**, 4172–4182.
- 19 Y. Abe, T. Shodai, T. Muto, K. Mihara, H. Torii, S. i. Nishikawa, T. Endo and D. Kohda, *Cell*, 2000, **100**, 551–560.
- 20 D. Mokranjac and W. Neupert, *Nature*, 2015, **528**, 201–202.
- 21 M. F. Bauer, C. Sirrenberg, W. Neupert and M. Brunner, *Cell*, 1996, **87**, 33–41.
- 22 P. J. Kang, J. Ostermann, J. Shilling, W. Neupert, E. A. Craig and N. Pfanner, *Nature*, 1990, **348**, 137–143.
- 23 O. Gakh, P. Cavadini and G. Isaya, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2002, **1592**, 63–77.
- 24 D. Mossmann, C. Meisinger and F. N. Vögtle, *Biochim. Biophys. Acta, Gene Regul. Mech.*, 2012, **1819**, 1098–1106.
- 25 M. T. Veling, A. G. Reidenbach, E. C. Freiberger, N. W. Kwiecien, P. D. Hutchins, M. J. Drahnak, A. Jochem, A. Ulbrich, M. J. P. Rush, J. D. Russell, J. J. Coon and D. J. Pagliarini, *Mol. Cell*, 2017, **68**, 970–977.e11.
- 26 M. Y. Cheng, F. U. Hartl, J. Martin, R. A. Pollock, F. Kalousek, W. Neuper, E. M. Hallberg, R. L. Hallberg and A. L. Horwich, *Nature*, 1989, **337**, 620–625.
- 27 B. S. Glick, A. Brandt, K. Cunningham, S. Müller, R. L. Hallberg and G. Schatz, *Cell*, 1992, **69**, 809–822.
- 28 R. Ieva, S. G. Schrempp, Ł. Opaliński, F. Wollweber, P. Höß, A. K. Heißwolf, M. Gebert, Y. Zhang, B. Guiard, S. Rospert, T. Becker, A. Chacinska, N. Pfanner and M. van der Laan, *Mol. Cell*, 2014, **56**, 641–652.
- 29 K. Park, S. C. Botelho, J. Hong, M. Österberg and H. Kim, *J. Biol. Chem.*, 2013, **288**, 1521–1532.
- 30 J. M. Herrmann, W. Neupert and R. A. Stuart, *EMBO J.*, 1997, **16**, 2217–2226.
- 31 S. B. Stiller, J. Höpker, S. Oeljeklaus, C. Schütze, S. G. Schrempp, J. Vent-Schmidt, S. E. Horvath, A. E. Frazier, N. Gebert, M. Van Der Laan, M. Bohnert, B. Warscheid, N. Pfanner and N. Wiedemann, *Cell Metab.*, 2016, **23**, 901–908.
- 32 P. Horten, L. Colina-Tenorio and H. Rampelt, *Biomolecules*, 2020, **10**, 1–13.
- 33 K. N. Truscott and N. Pfanner, *Biol. Chem.*, 1999, **380**, 1151–1156.
- 34 S. P. Curran, D. Leuenberger, E. Schmidt and C. M. Koehler, *J. Cell Biol.*, 2002, **158**, 1017–1027.
- 35 C. M. Koehler, S. Merchant, W. Oppliger, K. Schmid, E. Jarosch, L. Dolfini, T. Junne, G. Schatz and K. Tokatlidis, *EMBO J.*, 1998, **17**, 6477–6486.
- 36 S. P. Curran, D. Leuenberger, W. Oppliger and C. M. Koehler, *EMBO J.*, 2002, **21**, 942–953.
- 37 S. Vial, H. Lu, S. Allen, P. Savory, D. Thornton, J. Sheehan and K. Tokatlidis, *J. Biol. Chem.*, 2002, **277**, 36100–36108.
- 38 C. Sirrenberg, M. F. Bauer, B. Guiard, W. Neupert and M. Brunner, *Nature*, 1996, **384**, 582–585.
- 39 S. A. Paschen, U. Rothbauer, K. Káldi, M. F. Bauer, W. Neupert and M. Brunner, *EMBO J.*, 2000, **19**, 6392–6400.
- 40 K. N. Beverly, M. R. Sawaya, E. Schmid and C. M. Koehler, *J. Mol. Biol.*, 2008, **382**, 1144–1156.
- 41 A. J. Davis, N. N. Alder, R. E. Jensen and A. E. Johnson, *Mol. Biol. Cell*, 2007, **18**, 475–486.
- 42 N. Wiedemann, N. Pfanner and M. T. Ryan, *EMBO J.*, 2001, **20**, 951–960.
- 43 N. Wiedemann, V. Kozjak, A. Chacinska, B. Schönfisch, S. Rospert, M. T. Ryan, N. Pfanner and C. Meisinger, *Nature*, 2003, **424**, 565–571.
- 44 A. I. C. Höhr, S. P. Straub, B. Warscheid, T. Becker and N. Wiedemann, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2015, **1853**, 74–88.
- 45 K. N. Doan, A. Grevel, C. U. Mårtensson, L. Ellenrieder, N. Thornton, L. S. Wenz, Ł. Opaliński, B. Guiard, N. Pfanner and T. Becker, *Cell Rep.*, 2020, **31**, 17593–17605.
- 46 D. Papić, K. Krumpke, J. Dukanovic, K. S. Dimmer and D. Rapaport, *J. Cell Biol.*, 2011, **194**, 397–405.
- 47 K. S. Dimmer, D. Papić, B. Schumann, D. Sperl, K. Krumpke, D. M. Walther and D. Rapaport, *J. Cell Sci.*, 2012, **125**, 3464–3473.
- 48 A. Chacinska, S. Pfannschmidt, N. Wiedemann, V. Kozjak, L. K. Sanjuán Szklarz, A. Schulze-Specking, K. N. Truscott, B. Guiard, C. Meisinger and N. Pfanner, *EMBO J.*, 2004, **23**, 3735–3746.
- 49 J. Nunnari, T. D. Fox and P. Walter, *Science*, 1993, **262**, 1997–2004.
- 50 K. Esser, B. Tursun, M. Ingenhoven, G. Michaelis and E. Pratje, *J. Mol. Biol.*, 2002, **323**, 835–843.
- 51 G. A. McQuibban, S. Saurya and M. Freeman, *Nature*, 2003, **423**, 537–541.
- 52 H. Lu, A. P. Golovanov, F. Alcock, J. G. Grossmann, S. Allen, L. Y. Lian and K. Tokatlidis, *J. Biol. Chem.*, 2004, **279**, 18959–18966.
- 53 L. Banci, I. Bertini, C. Cefaro, S. Ciofi-Baffoni, A. Gallo, M. Martinelli, D. P. Sideris, N. Katrakili and K. Tokatlidis, *Nat. Struct. Mol. Biol.*, 2009, **16**, 198–206.
- 54 D. P. Sideris and K. Tokatlidis, *Antioxid. Redox Signaling*, 2010, **13**, 1189–1204.
- 55 J. M. Herrmann and K. Hell, *Trends Biochem. Sci.*, 2005, **30**, 205–212.
- 56 A. Chatzi, D. P. Sideris, N. Katrakili, C. Pozidis and K. Tokatlidis, *FEBS J.*, 2013, **280**, 4960–4969.
- 57 D. P. Sideris, N. Petrakis, N. Katrakili, D. Mikropoulou, A. Gallo, S. Ciofi-Baffoni, L. Banci, I. Bertini and K. Tokatlidis, *J. Cell Biol.*, 2009, **187**, 1007–1022.



- 58 D. Milenkovic, T. Ramming, J. M. Müller, L. S. Wenz, N. Gebert, A. Schulze-Specking, D. Stojanovski, S. Rospert and A. Chacinska, *Mol. Biol. Cell*, 2009, **20**, 2530–2539.
- 59 E. Kallergi, M. Andreadaki, P. Kritsiligkou, N. Katrakili, C. Pozidis, K. Tokatlidis, L. Banci, I. Bertini, C. Cefaro, S. Ciofi-Baffoni, K. Gajda and R. Peruzzini, *ACS Chem. Biol.*, 2012, **7**, 707–714.
- 60 G. Hofhaus, J. E. Lee, I. Tews, B. Rosenberg and T. Lisowsky, *Eur. J. Biochem.*, 2003, **270**, 1528–1535.
- 61 D. Stojanovski, D. Milenkovic, J. M. Müller, K. Gabriel, A. Schulze-Specking, M. J. Baker, M. T. Ryan, B. Guiard, N. Pfanner and A. Chacinska, *J. Cell Biol.*, 2008, **183**, 195–202.
- 62 L. Banci, I. Bertini, V. Calderone, C. Cefaro, S. Ciofi-Baffoni, A. Gallo, E. Kallergi, E. Lionaki, C. Pozidis and K. Tokatlidis, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4811–4816.
- 63 L. Banci, I. Bertini, C. Cefaro, S. Ciofi-Baffoni, K. Gajda, I. C. Felli, A. Gallo, A. Pavelkova, E. Kallergi, M. Andreadaki, N. Katrakili, C. Pozidis and K. Tokatlidis, *J. Mol. Biol.*, 2013, **425**, 594–608.
- 64 N. Terziyska, B. Grumbt, C. Kozany and K. Hell, *J. Biol. Chem.*, 2009, **284**, 1353–1363.
- 65 V. N. Daithankar, S. R. Farrell and C. Thorpe, *Biochemistry*, 2009, **48**, 4828–4837.
- 66 L. Banci, I. Bertini, V. Calderone, C. Cefaro, S. Ciofi-Baffoni, A. Gallo and K. Tokatlidis, *J. Am. Chem. Soc.*, 2012, **134**, 1442–1445.
- 67 D. V. Dabir, E. P. Leverich, S. K. Kim, F. D. Tsai, M. Hirasawa, D. B. Knaff and C. M. Koehler, *EMBO J.*, 2007, **26**, 4801–4811.
- 68 M. E. Sztolsztener, A. Brewinska, B. Guiard and A. Chacinska, *Traffic*, 2013, **14**, 309–320.
- 69 S. E. Neal, D. V. Dabir, J. Wijaya, C. Boon and C. M. Koehler, *Mol. Biol. Cell*, 2017, **28**, 2773–2785.
- 70 G. Schatz and B. Dobberstein, *Science*, 1996, **271**, 1519–1526.
- 71 F. N. Vögtle, S. Wortelkamp, R. P. Zahedi, D. Becker, C. Leidhold, K. Gevaert, J. Kellermann, W. Voos, A. Sickmann, N. Pfanner and C. Meisinger, *Cell*, 2009, **139**, 428–439.
- 72 S. Backes and J. M. Herrmann, *Front. Mol. Biosci.*, 2017, **4**, 83.
- 73 D. Roise, F. Theiler, S. J. Horvath, J. M. Tomich, J. H. Richards, D. S. Allison and G. Schatz, *EMBO J.*, 1988, **7**, 649–653.
- 74 T. Endo, H. Yamamoto and M. Esaki, *J. Cell Sci.*, 2003, **116**, 3259–3267.
- 75 T. Endo and D. Kohda, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2002, **1592**, 3–14.
- 76 T. Saitoh, M. Igura, T. Obita, T. Ose, R. Kojima, K. Maenaka, T. Endo and D. Kohda, *EMBO J.*, 2007, **26**, 4777–4787.
- 77 N. Bolender, A. Sickmann, R. Wagner, C. Meisinger and N. Pfanner, *EMBO Rep.*, 2008, **9**, 42–49.
- 78 S. E. Calvo, O. Julien, K. R. Clauser, H. Shen, K. J. Kamer, J. A. Wells and V. K. Mootha, *Mol. Cell. Proteomics*, 2017, **16**, 512–523.
- 79 J. Brix, K. Dietmeier and N. Pfanner, *J. Biol. Chem.*, 1997, **272**, 20730–20735.
- 80 T. Bausewein, D. J. Mills, J. D. Langer, B. Nitschke, S. Nussberger and W. Kühlbrandt, *Cell*, 2017, **170**, 693–700.e7.
- 81 W. Wang, X. Chen, L. Zhang, J. Yi, Q. Ma, J. Yin, W. Zhuo, J. Gu and M. Yang, *Cell Discovery*, 2020, **6**, 1–10.
- 82 G. Schatz, *Nature*, 1997, **388**, 121.
- 83 T. Komiyama, S. Rospert, C. Koehler, R. Looser, G. Schatz and K. Mihara, *EMBO J.*, 1998, **17**, 3886–3898.
- 84 D. Rapaport, K.-P. Künkele, M. Dembowski, U. Ahting, F. E. Nargang, W. Neupert and R. Lill, *Mol. Cell. Biol.*, 1998, **18**, 5256–5262.
- 85 T. Kanamori, S. I. Nishikawa, M. Nakai, I. Shin, P. G. Schultz and T. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3634–3639.
- 86 K. N. Truscott, P. Kovermann, A. Geissler, A. Merlin, M. Meijer, A. J. M. Driessen, J. Rassow, N. Pfanner and R. Wagner, *Nat. Struct. Biol.*, 2001, **8**, 1074–1082.
- 87 S. Meier, W. Neupert and J. M. Herrmann, *J. Cell Biol.*, 2005, **170**, 881–888.
- 88 V. Žárský and P. Doležal, *Biol. Direct*, 2016, **11**, 1–13.
- 89 D. Mokranjac, M. Sichtung, D. Popov-Čeleketić, K. Mapa, L. Gevorkyan-Airapetov, K. Zohary, K. Hell, A. Azem and W. Neupert, *Mol. Biol. Cell*, 2009, **20**, 1400–1407.
- 90 C. Schulz, O. Lytovchenko, J. Melin, A. Chacinska, B. Guiard, P. Neumann, R. Ficner, O. Jahn, B. Schmidt and P. Rehling, *J. Cell Biol.*, 2011, **195**, 643–656.
- 91 R. Bajaj, Ł. Jaremko, M. Jaremko, S. Becker and M. Zweckstetter, *Structure*, 2014, **22**, 1501–1511.
- 92 M. Schleyer and W. Neupert, *Cell*, 1985, **43**, 339–350.
- 93 J. Rassow, F. U. Hartl, B. Guiard, N. Pfanner and W. Neupert, *FEBS Lett.*, 1990, **275**, 190–194.
- 94 T. Kanamori, S. I. Nishikawa, I. Shin, P. G. Schultz and T. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 485–490.
- 95 J. Dudek, B. Guiard and P. Rehling, *Enzymes*, 2007, **25**, 387–411.
- 96 T. Shiota, K. Imai, J. Qiu, V. L. Hewitt, K. Tan, H. H. Shen, N. Sakiyama, Y. Fukasawa, S. Hayat, M. Kamiya, A. Elofsson, K. Tomii, P. Horton, N. Wiedemann, N. Pfanner, T. Lithgow and T. Endo, *Science*, 2015, **349**, 1544–1548.
- 97 D. Poveda-Huertes, P. Mulica and F. N. Vögtle, *Cell Tissue Res.*, 2017, **367**, 73–81.
- 98 K. Dvořáková-Holá, A. Matušková, M. Kubala, M. Otyepka, T. Kučera, J. Večeř, P. Heřman, N. Parkhomenko, E. Kutejova and J. Janata, *J. Mol. Biol.*, 2010, **396**, 1197–1210.
- 99 T. Kučera, M. Otyepka, A. Matušková, A. Samad, E. Kutejová and J. Janata, *PLoS One*, 2013, **8**, e74518.
- 100 A. B. Taylor, B. S. Smith, S. Kitada, K. Kojima, H. Miyaura, Z. Otwinowski, A. Ito and J. Deisenhofer, *Structure*, 2001, **9**, 615–625.
- 101 A. A. Taskin, C. Kücükköse, N. Burger, D. Mossmann, C. Meisinger and F. N. Vögtle, *Mol. Biol. Cell*, 2017, **28**, 997–1002.

- 102 M. A. Leissring, W. Farris, X. Wu, D. C. Christodoulou, M. C. Haigis, L. Guarente and D. J. Selkoe, *Biochem. J.*, 2004, **383**, 439–446.
- 103 D. Mossmann, F. N. Vögtle, A. A. Taskin, P. F. Teixeira, J. Ring, J. M. Burkhart, N. Burger, C. M. Pinho, J. Tadic, D. Loreth, C. Graff, F. Metzger, A. Sickmann, O. Kretz, N. Wiedemann, R. P. Zahedi, F. Madeo, E. Glaser and C. Meisinger, *Cell Metab.*, 2014, **20**, 662–669.
- 104 S. Kreimendahl, J. Schwichtenberg, K. Günnewig, L. Brandherm and J. Rassow, *BMC Biol.*, 2020, **18**, 1–23.
- 105 M. A. S. Vergnolle, C. Baud, A. P. Golovanov, F. Alcock, P. Luciano, L. Y. Lian and K. Tokatlidis, *J. Mol. Biol.*, 2005, **351**, 839–849.
- 106 C. M. Koehler, D. Leuenberger, S. Merchant, A. Renold, T. Junne and G. Schatz, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 2141–2146.
- 107 N. Pfanner and A. Geissler, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 339–349.
- 108 R. E. Jensen and C. D. Dunn, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2002, **1592**, 25–34.
- 109 S. M. Gasser, A. Ohashi, G. Daum, P. C. Böhni, J. Gibson, G. A. Reid, T. Yonetani and G. Schatz, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 267–271.
- 110 K. Esser, P. S. Jan, E. Pratje and G. Michaelis, *Mol. Genet. Genomics*, 2004, **271**, 616–626.
- 111 A. Mordas and K. Tokatlidis, *Acc. Chem. Res.*, 2015, **48**, 2191–2199.
- 112 V. Peleh, E. Cordat and J. M. Herrmann, *eLife*, 2016, **5**, e16177.
- 113 M. Habich, S. L. Salscheider, L. M. Murschall, M. N. Hoehne, M. Fischer, F. Schorn, C. Petrungraro, M. Ali, A. J. Erdogan, S. Abou-Eid, H. Kashkar, J. Dengjel and J. Riemer, *Cell Rep.*, 2019, **26**, 759–774.e5.
- 114 M. Esaki, T. Kanamori, S. I. Nishikawa and T. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11770–11775.
- 115 M. A. H. Luttkik, K. M. Overkamp, P. Kötter, S. De Vries, J. P. Van Dijken and J. T. Pronk, *J. Biol. Chem.*, 1998, **273**, 24529–24534.
- 116 S. D. Saladi, F. Boos, M. Poglitsch, H. Meyer, F. Sommer, T. Mühlhaus, M. Schroda, M. Schuldiner, F. Madeo and J. M. Herrmann, *Mol. Cell*, 2020, **77**, 189–202.e6.
- 117 M. V. Gerashchenko, A. V. Lobanov and V. N. Gladyshev, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 17394–17399.
- 118 P. Kritsiligkou, A. Chatzi, G. Charalampous, A. Mironov, C. M. Grant and K. Tokatlidis, *Cell Rep.*, 2017, **18**, 2729–2741.
- 119 M. Cardenas-Rodriguez and K. Tokatlidis, *FEBS Lett.*, 2017, **591**, 2661–2670.
- 120 T. D. Fox, *Genetics*, 2012, **192**, 1203–1234.
- 121 A. I. C. Höhr, C. Lindau, C. Wirth, J. Qiu, D. A. Stroud, S. Kutik, B. Guiard, C. Hunte, T. Becker, N. Pfanner and N. Wiedemann, *Science*, 2018, 359.
- 122 T. Jores, A. Klinger, L. E. Groß, S. Kawano, N. Flinker, E. Duchardt-Ferner, J. Wöhnert, H. Kalbacher, T. Endo, E. Schleiff and D. Rapaport, *Nat. Commun.*, 2016, **7**, 1–16.
- 123 A. Gupta and T. Becker, *Biochim. Biophys. Acta, Bioenerg.*, 2021, **1862**, 148323.
- 124 U. Ahting, T. Waizenegger, W. Neupert and D. Rapaport, *J. Biol. Chem.*, 2005, **280**, 48–53.
- 125 M. Harner, W. Neupert and M. Deponte, *EMBO J.*, 2011, **30**, 3232–3241.
- 126 F. U. Hartl, A. Bracher and M. Hayer-Hartl, *Nature*, 2011, **475**, 324–332.
- 127 E. A. Craig, P. Huang, R. Aron and A. Andrew, *Rev. Physiol., Biochem. Pharmacol.*, 2006, **156**, 1–21.
- 128 N. Petrakis, F. Alcock and K. Tokatlidis, *IUBMB Life*, 2009, **61**, 909–914.
- 129 S. Allen, H. Lu, D. Thornton and K. Tokatlidis, *J. Biol. Chem.*, 2003, **278**, 38505–38513.
- 130 S. Kawano, K. Yamano, M. Naoé, T. Momose, K. Terao, S. I. Nishikawa, N. Watanabe and T. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14403–14407.
- 131 L. Banci, I. Bertini, C. Cefaro, L. Cenacchi, S. Ciofi-Baffoni, I. C. Felli, A. Gallo, L. Gonnelli, E. Luchinat, D. Sideris and K. Tokatlidis, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 20190–20195.
- 132 J. R. Koch and F. X. Schmid, *Nat. Commun.*, 2014, **5**, 1–10.
- 133 A. Adam, M. Endres, C. Sirrenberg, F. Lottspeich, W. Neupert and M. Brunner, *EMBO J.*, 1999, **18**, 313–319.
- 134 C. T. Webb, M. A. Gorman, M. Lazarou, M. T. Ryan and J. M. Gulbis, *Mol. Cell*, 2006, **21**, 123–133.
- 135 N. Gebert, A. Chacinska, K. Wagner, B. Guiard, C. M. Koehler, P. Rehling, N. Pfanner and N. Wiedemann, *EMBO Rep.*, 2008, **9**, 548–554.
- 136 K. Weinhäupl, C. Lindau, A. Hessel, Y. Wang, C. Schütze, T. Jores, L. Melchionda, B. Schönfisch, H. Kalbacher, B. Bersch, D. Rapaport, M. Brennich, K. Lindorff-Larsen, N. Wiedemann and P. Schanda, *Cell*, 2018, **175**, 1365–1379.e25.
- 137 I. Sućec, Y. Wang, O. Dakhlaoui, K. Weinhäupl, T. Jores, D. Costa, A. Hessel, M. Brennich, D. Rapaport, K. Lindorff-Larsen, B. Bersch and P. Schanda, *Sci. Adv.*, 2020, **51**, eabd0263.
- 138 C. Puchades, A. J. Rampello, M. Shin, C. J. Giuliano, R. L. Wiseman, S. E. Glynn and G. C. Lander, *Science*, 2017, **358**(6363), eaao0464.
- 139 M. J. Baker, T. Tatsuta and T. Langer, *Cold Spring Harbor Perspect. Biol.*, 2011, **3**, 1–19.
- 140 M. Opalińska and H. Jańska, *Cells*, 2018, **7**, 163.
- 141 K. Leonhard, A. Stiegler, W. Neupert and T. Langer, *Nature*, 1999, **398**, 348–351.
- 142 B. Schreiner, H. Westerburg, I. Forné, A. Imhof, W. Neupert and D. Mokranjac, *Mol. Biol. Cell*, 2012, **23**, 4335–4346.
- 143 K. Leonhard, J. M. Herrmann, R. A. Stuart, G. Mannhaupt, W. Neupert and T. Langer, *EMBO J.*, 1996, **15**, 4218–4229.
- 144 M. Ammelburg, T. Frickey and A. N. Lupas, *J. Struct. Biol.*, 2006, **156**, 2–11.
- 145 M. Graef, G. Seewald and T. Langer, *Mol. Cell. Biol.*, 2007, **27**, 2476–2485.
- 146 H. L. Fiumera, M. J. Dunham, S. A. Saracco, C. A. Butler, J. A. Kelly and T. D. Fox, *Genetics*, 2009, **182**, 519–528.