TOM-complex structure modeling

A.V. DUDKO

Belarusian State University, Minsk, Belarus

e-mail: dav_bel@mail.ru

V.G. VERESOV

Institute of Biophysics and Cell Engineering, Minsk, Belarus.

e-mail: veresov@ibp.org.by

Mitochondria are essential organelles in eukaryotic cells that consist of four subcompartments, the outer and inner membranes, the intermembrane space (IMS), and innermost matrix. Mitochondria contain about 1000–1500 different proteins, most of which are synthesized as precursor proteins in the cytosol and imported into mitochondria. Import and subsequent intramitochondrial sorting of mitochondrial proteins are mediated by membrane-protein complexes called translocators in the outer and inner membranes and soluble factors in the cytosol, IMS, and matrix (1). Most mitochondrial proteins enter mitochondria via the outer-membrane translocator, the TOM complex. The TOM complex consists of the core complex comprised of Tom40, Tom22, Tom5, Tom6, and Tom7, and peripherally associated receptors, Tom20 and Tom70. In the last decade, information on the structural aspects of the TOM complex has been accumulated. In particular, high-resolution structures for many components of the TOM complex as well as single particle electron microscopy (EM) structure of the whole complex have become available (2, 3), which contributed to enhancement of our understanding of the mechanisms of protein import and sorting on the basis of the structures. However, a high-resolution structure of the whole complex is absent. Here, to obtain atomic-resolution structural model of the TOM complex, we use computational structural biology tools and the most recent experimental data on the structure of TOM complex components together with the structure of the whole TOM complex obtained by electron microscopy. The successive docking of TOM components was used. The prediction of the 3D-structures of protein-protein complexes was performed in a stepwise fashion with an initial rigid-body global search using the program Piper (4) and subsequent steps using the combination of the programs ROSETTADOCK$_1$(5) – HADDOCK (6) – ROSETTADOCK$_2$(5) (hereafter for the sake of clarity ROSETTADOCK$_1$ and ROSETTADOCK$_2$ refer to first and final refinement by ROSETTADOCK, respectively) to improve these initial predictions. First, TOM22 and the TOM40 monomer were modeled using the I-TASSER protocol (7). Then modeling (TOM40)$_2$ was performed using the program PIPER (4) with the enhanced electrostatic coefficient option to favor the formation of the dimers with polar groups buried at the monomer-monomer interface in order to avoid energetically

unfavorable exposure of hydrophilic surfaces to a membrane hydrophobic environment. The complex with the greatest electrostatic contribution and parallel arrangement of monomers was selected as the most plausible prediction. One salt bridge and four hydrogen bonds were found between the monomers in this model. These are the Asp296(MI) -Arg293(MII) salt bridge, two hydrogen bonds between Gly285 (MI) and Tyr274(MII), one hydrogen bond between Asp296(MI) and Gln295(MII), and extra hydrogen bond between Asp296(MI) and Arg293(MII). Then, the selected structure was subjected to a refinement procedure by using the combined HADDOCK (5)/ROSETTADOCK (6) protocol and the ROSETTADOCK top-scoring model was chosen as our final prediction. In this final model monomers M1 and M2 were rather tightly bound forming one salt bridge Arg293M1-Asp296M2 and one hydrogen bond Met248M1-His347M2 and possessing good shape complementarity (see Table 1 for the most important docking data). Then we carried out the in silico docking of TOM22 to the $(TOM40)_2$ dimer. To do that, first TOM22 was conditionally divided into three parts: a negatively charged N-terminal cytosol region (residues 1-85, $TOM22_{cyt}$), a putative transmembrane region (residues 86-105, $TOM22_{TM}$), and a C-terminal intermembrane space region (residues 106-142, $TOM22_{is}$), in accordance with experimental data on localization of these three TOM22 domains relative to the membrane (8). Next, we subjected $(TOM40)_2$ to two separate computational dockings with $TOM22_{cyt}$ and $TOM22_{TM}$ using the PIPER protocol. In one case, the $TOM22_{cyt}$ C-terminal end fell near the N-terminal end of $(TOM22)_{TM}$ of the $TOM22_{TM}/(TOM40)_2$ complex. We used this circumstance to combine $TOM22_{cyt}$ with $TOM22_{TM}$ into a single protein ($TOM22_{cyt+TM}$) bound to $(TOM40)_2$. To do this, the two closely-spaced fragments $TOM22_{cyt}$ and $TOM22_{TM}$ were integrated into one structure using the Rosetta loop closure protocol (9). Next, the integral structure was subjected to local docking to $(TOM40)_2$ by using consecutively HADDOCK and ROSETTADOCK. The ROSETTADOCK top-scored complex had two intermolecular hydrogen bonds (Ala73TOM40M2-Asp17TOM22, Cys74TOM40M2-Asp17TOM22) and good shape complementarity (BSA=2480 $Å^2$), resulting in the ROSETTADOCK Isc score of -7.5 (see Table 1). Next, four stage protocol PIPER(4)-$ROSETTADOCK_1$(5) – HADDOCK (6) – $ROSETTADOCK_2$(5) was applied to dock TOM20 to the $(TOM40)_2$-TOM22 complex. The formation of two salt bridges (Arg37TOM20-Glu29TOM22, Arg41TOM20- Glu29TOM22) and six hydrogen bonds (Pro17Tom 40_{MI} – Gln120Tom20, Ser31$Tom40_{MI}$ – Gln120Tom20, Pro30$Tom40_{MI}$ – Ser135Tom20, Asp198$Tom40_{MII}$-Gln75Tom20, Tyr194$Tom40_{MII}$- Lys68Tom 20, Tyr194$Tom40_{MII}$- Glu64Tom20) were detected. Because experimental data show the interaction between TOM20 and TOM70 the study of interaction between TOM70 and the $(TOM40)_2$-TOM22-TOM20 complex was carried out using PIPER(4)-$ROSETTADOCK_1$(5) – HADDOCK (6) – $ROSETTADOCK_2$(5) docking protocol. Five hydrogen bonds (Thr62Tom

40_{MII}-Gly602Tom 70, Glu141TOM20- Tyr601Tom70, Asp142TOM20- Tyr601Tom70, His117Tom40_{MII}-Thr146Tom70, Asn116Tom40_{MII}-Glu181Tom70) were shown to be formed between TOM70 and the $(TOM40)_2$-TOM22-TOM20 complex. Next, the docking of small components TOM5, TOM6, TOM7 to the $(TOM40)_2$-TOM22-TOM20-TOM70 complex was carried out. The 3D-structures of these small components were obtained with the use of I-TASSER protocol. Then these components were docked to the $(TOM40)_2$-TOM22-TOM20-TOM70 complex using four stage PIPER(4)-$ROSETTADOCK_1$(5) – HADDOCK (6) – $ROSETTADOCK_2$(5) docking protocol. In final $ROSETTADOCK_2$ model the following hydrogen bonds were formed: Ala270TOM40_{MII}-Lys47TOM5, Ser317TOM40_{MI}-Arg39TOM5, Val318TOM40_{MI}-Ala36TOM5; Glu289TOM40_{MI}-Arg33TOM6, Ser298TOM40_{MI}-Arg33TOM6, Ser279TOM40_{MII}-Asp22TOM6, Leu282TOM40_{MII}-Asn23TOM6, Asp23TOM5-Leu28TOM6, Tyr274TOM40_{MII}-Arg29TOM6; Thr292TOM40_{MI}-Ser5TOM7, Gln295TOM$40M_I$ -Glu7TOM7, Asp23TOM5-Leu28TOM6, Met16TOM5-Leu28TOM6, Asp23TOM5-Gln20TOM7. Also one salt bridge (His276TOM40_{MII}- Asp26TOM6) was formed. The basic energetic and scoring data are presented in Table 1. The structure of the full TOM complex is shown in Fig.1.

Table 1. The ROSETTADOCK interface energy scores (I_sc), ROSETTADOCK binding score (RDBS), Burried Surface Area (BSA), ROSETTADOCK van der Waals interaction energy score ($\Delta VDWS_{RD}$), ROSETTADOCK desolvation energy score ($\Delta solv_{RD}$), number of intermolecular salt bridges (N_{sb}) and intermolecular hydrogen bonds (N_{HB}) for the highest- rank complexes between different TOM components

Protein pairs	Isc	BSA Å	Evdw kcal/ mol	Edesolv kcal/ mol	N(sb)	N(hb)
Tom40 (M I)-Tom40 (M II)	**-8,8**	**2627,4**	**-78,2**	**-136,9**	**1**	**1**
$(Tom40)_2$-Tom 22	**-7,5**	**1782,9**	**-41,5**	**-94,8**	**0**	**2**
$(Tom40)_2$-Tom22-Tom20	**-7,6**	**3089**	**-81,5**	**-12,9**	**3**	**6**
$(Tom40)_2$-Tom22-Tom20-Tom70	**-8,35**	**1272,2**	**-28,3**	**-49,4**	**1**	**6**
$(Tom40)_2$-Tom22-Tom20-Tom70-Tom5	**-8,94**	**1661,9**	**-43,4**	**-21,2**	**0**	**3**
$(Tom40)_2$-Tom22-Tom20-Tom70-Tom5-Tom6	**-6,1**	**769,6**	**-30,9**	**-3,3**	**1**	**8**
$(Tom40)_2$-Tom22-Tom20-Tom5-Tom6-Tom7	**-6.39**	**962.9**	**-28.7**	**-3,5**	**0**	**3**

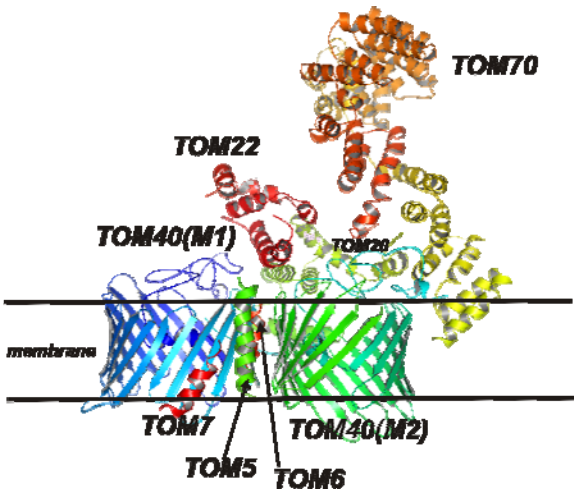

Fig.1. Structural model of the TOM-complex.

1. A. Chacinska, C.M. Koehler, D. Milenkovic, T. Lithgow, N. Pfanner (2009), Importing mitochondrial proteins: machineries and mechanisms, *Cell,* **138,** 628–644.
2. T. Endo, K. Yamano, S. Kawano (2011) Structural insight into the mitochondrial protein import system, *Biochim Biophys Acta.* **1808**, 955-970.
3. K. Model, C. Meisinger, W. Kühlbrandt (2008) Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase, *J. Mol. Biol.* **383,** 1049–1057
4. D. Kozakov, R. Brenke, S. R. Comeau, S. Vajda (2006) PIPER: An FFT-based protein docking program with pairwise potentials, *Proteins*, **65**:392–406
5. S.J. deVries, M. M. vanDijk , A.M.J.J Bonvin (2010) The HADDOCK web server for data-driven biomolecular docking, *Nature Protocols* **5,** 883-897.
6. J. J. Gray, S. Moughon, C. Wang, O. Schueler-Furman, B. Kuhlman, C. A. Rohl, D. Baker (2003). Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations, *J. Mol. Biol.* **331**, 281–299
7. A. Roy, A. Kucukural, Y. Zhang (2010) I-TASSER: a unified platform for automated protein structure and function prediction, *Nat Protoc* **5**, 725-738.
8. F.E. Nargang, D. Rapaport, R.G. Ritzel, W. Neupert, R. Lill (1998) Role of the negative charges in the cytosolic domain of TOM22 in the import of precursor proteins into mitochondria, *Mol. Cell. Biol*. **18,** 3173-3181.
9. D.J. Mandell, E.A. Coutsias, T. Kortemme (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling, *Nat. Methods* **6**, 551-552.