“Multiple partial recognitions in dynamic equilibrium” in the binding sites of proteins form the molecular basis of promiscuous recognition of structurally diverse ligands

Daisuke Kohda¹

Received: 16 September 2017 / Accepted: 19 November 2017 / Published online: 14 December 2017

© International Union of Pure and Applied Biophysics (IUPAB) and Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

Promiscuous recognition of ligands by proteins is as important as strict recognition in numerous biological processes. In living cells, many short, linear amino acid motifs function as targeting signals in proteins to specify the final destination of the protein transport. In general, the target signal is defined by a consensus sequence containing wild-characters, and hence represented by diverse amino acid sequences. The classical lock-and-key or induced-fit/conformational selection mechanism may not cover all aspects of the promiscuous recognition. On the basis of our crystallographic and NMR studies on the mitochondrial Tom20 protein–presequence interaction, we proposed a new hypothetical mechanism based on “a rapid equilibrium of multiple states with partial recognitions”. This dynamic, multiple recognition mode enables the Tom20 receptor to recognize diverse mitochondrial presequences with nearly equal affinities. The plant Tom20 is evolutionally unrelated to the animal Tom20 in our study, but is a functional homolog of the animal/fungal Tom20. NMR studies by another research group revealed that the presequence binding by the plant Tom20 was not fully explained by simple interaction modes, suggesting the presence of a similar dynamic, multiple recognition mode. Circumstantial evidence also suggested that similar dynamic mechanisms may be applicable to other promiscuous recognitions of signal peptides by the SRP54/Ffh and SecA proteins.

Keywords Mitochondrial presequence · Promiscuous recognition · Signal sequence · Targeting signal · Tom20

Introduction

We are interested in molecular recognitions of ligands by proteins with promiscuous specificities. Here, the ligands mean short amino acid sequences. In textbooks, most biological processes are described as depending on the strict specificity of ligand recognitions by proteins, but proteins with promiscuous specificity are equally important. Many examples of promiscuous recognitions exist in nature, but they are still overlooked. An obvious scenario, such as a multi-domain

protein with many binding sites, is possible, but we will focus on a single binding site in a protein molecule in this review. A promiscuous protein can interact with a wide variety of ligands, with similar affinities (Fig. 1). Frequently, the structural or sequence variation of ligands reaches such a high extent that no similarity is detectable visually or even by sophisticated computer programs. It is important to distinguish the promiscuous specificity from a non-specific interaction. The latter is of little biological significance by itself. In many situations, however, non-specific interactions can assist the strict or promiscuous interactions by increasing the binding rate or affinity of ligands.

Promiscuous specificity is usually associated with weak affinity and a fast dissociation rate (Fig. 1). The weak affinity of a typical ligand is reflected by the dissociation constant, K_d , range of 0.1 mM to 10 mM. In the case of diffusion-limited binding, the maximum k_{on} is $10^{10} \text{ M}^{-1} \text{ s}^{-1}$. By using the relationship $K_d = k_{off}/k_{on}$, k_{off} is estimated to be 10^6 to 10^8 s^{-1} . The inverse of k_{off} indicates that a ligand stays in the binding site for a maximum of 1 μs . Such a short dwell time makes the

This article is part of a Special Issue on ‘Biomolecules to Bio-nanomachines — Fumio Arisaka 70th Birthday’ edited by Damien Hall, Junichi Takagi, and Haruki Nakamura.

✉ Daisuke Kohda
kohda@bioreg.kyushu-u.ac.jp

¹ Division of Structural Biology, Medical Institute of Bioregulation, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan

Fig. 1 Trade-off relationships in ligand recognition by a protein molecule. The 3D graphs portray the variety of ligands' shapes and charges as the horizontal plane, and the affinity for a protein as the vertical height

isolation of the complex very difficult, and hence studies of proteins with promiscuous specificity are inherently challenging. Various experimental techniques must be used in an integrated manner, and the development of novel measurement techniques is desired. In contrast, the studies of proteins with strict specificity are relatively easy, because the structure determination of the protein–ligand complex will reveal the details of the interaction between the protein and its unique ligand.

Promiscuous recognition inside cells

Nascent proteins synthesized by ribosomes in the cytosol are transported across the intracellular membranes to the destinations where they function, including the nucleus, mitochondrion, chloroplast, peroxisome, endoplasmic reticulum (ER), Golgi apparatus, lysosome, vacuole, and external environment (Fig. 2). In addition, ligand-receptor complexes on the cell surface are internalized by clathrin-coated vesicles or caveolin-based caveolae, delivered to the endosome, trans-Golgi network (TGN), and lysosome, and then recycled back to the plasma membrane (Fig. 2). The newly synthesized proteins in the cytosol and the cytosolic domains of receptor proteins on the cell surface contain a signal sequence as an address tag. A typical signal sequence is a short linear segment of amino acid residues. The signal sequence either remains as part of the mature protein or is removed from the precursor protein by a specific protease at the final destination. For the latter case, the signal sequence is easily defined by comparing the amino acid sequence deduced from the DNA sequence with that of the mature protein. All of the amino acid residues in the signal sequence are not functionally equal. In general,

two or three amino acid residues are more critical than the other residues. These critical amino acid residues are involved in the recognition by the binding partner/receptor proteins, and collectively define a “targeting signal”. In most cases, the targeting signal is initially identified by the multiple sequence alignment of signal-sequence containing proteins, and then confirmed experimentally by mutagenesis studies. In some cases, however, due to very weak sequence homology, binding experiments with peptide libraries are necessary to discover and test the putative targeting signal. Usually, a targeting signal is represented as a four- to seven-residue amino acid sequence motif. The consensus pattern is conveniently denoted using the PROSITE syntax. An example is AX[BC]X ϕ , where A, B, and C denote one particular amino acid residue, either B or C is selected in the square brackets, and X and ϕ represent any amino acid residue and a hydrophobic amino acid residue respectively. In the case of the targeting signals located at the N- and C-termini, the α -amino and α -carboxylate groups are regarded as an additional charged residue in the motif. When the targeting signals are embedded within the mature protein sequences, the detection and definition of the targeting signals is sometimes difficult. It is possible that the signature residues of the targeting signal are scattered throughout the protein sequence, and the targeting signal is formed only in the (partially) folded states of the protein molecule.

Table 1 summarizes the targeting signals, the targeted processes and pathways, and their cognate binding/receptor proteins. Most of the targeting signals contain wild-characters, X and ϕ , in the consensus motif, which enormously increase the sequence diversity. Thus, the interaction between a targeting signal and its binding/receptor proteins provides a good example of promiscuous recognition by protein molecules. As

Fig. 2 Intracellular transport pathways of nascent proteins synthesized by ribosomes in the cytosol, and endocytosis and recycling of cell-surface receptors. The S-in-a-circle denotes a targeting signal embedded in the nascent proteins and the cytosolic domain of the cell-surface receptors. In general, a targeting signal is a diverse set of short amino acid sequences. The promiscuous recognition of the targeting signals by their cognate receptor proteins in the targeted organelles is the molecular basis of the address during intracellular protein transport

always, there is an exception. The peroxisome targeting signal (PTS) type 1 is the three-residue single sequence, Ser-Lys-Leu, located at the C-terminus of the peroxin proteins (Kunze and Berger 2015).

It should be noted that one signal sequence may contain more than one targeting signal for different targeted processes and pathways. For example, dual localization can occur to mitochondria and chloroplasts (Carrie et al. 2009), or peroxisomes and other organelles (Ast et al. 2013). The targeting signals in the N-terminal presequences of mitochondrial proteins interact with their receptor proteins in a sequential order along the import pathway. The presequence also contains the cleavage-site signals for specific proteases. One targeting signal may exhibit multiple interactions with different binding partner/receptor proteins. The YXX ϕ motif in the cytoplasmic part of the cell surface proteins instructs the internalization of the ligand-receptor complex and the trafficking of the complex to lysosomes (Bonifacino and Traub 2003; Pandey 2010). From a different point of view, a signal sequence is composed of short amino acid motifs that are recognized by several proteins, and importantly, the organization (position, order, and overlapping of the motifs) is unique for each signal sequence. This complicated situation makes the discovery of a consensus sequence by multiple sequence alignment quite difficult.

Molecular basis of promiscuous recognitions

The lock-and-key mechanism is the molecular basis of the strict specificity (Fig. 3). This classical mechanism is based on the high complementarities of shape and charges. The

concept is traceable back to the proposal in 1894 by Emil Fischer (1894). In his analogy, the binding site is a fixed pocket on the protein surface, and only correct ligands can fit into the binding site. In reality, however, protein molecules are intrinsically flexible, and the conformational change of a protein molecule is coupled with the ligand binding. The induced-fit mechanism was proposed by Daniel Koshland (1958). The ligand-binding induces conformational changes of the protein molecule to adjust the pocket to the shape and charges of the ligand in a complementary manner (Fig. 3). The conformational changes range from side-chain rearrangements, main-chain movements, and even domain rearrangements in protein molecules. The conformational selection (or population-shift) mechanism is a similar, but different mechanism. A protein molecule is assumed to be in an equilibrium of two conformational states. The ligand selectively binds to one conformation, and shifts the conformational equilibrium to the bound conformation. The critical difference between the induced-fit and conformational selection mechanisms is whether the conformational change of the protein occurs after or before the ligand binding step. There has been considerable debate about the induced-fit versus conformational selection as a realistic model (Changeux and Edelstein 2011; Gianni et al. 2014; Meyer-Almes 2016; Paul and Weikl 2016; Vogt and Di Cera 2012). However, the two mechanisms are not mutually exclusive, and a mixed mechanism actually operates in many cases (Chakraborty and Di Cera 2017; Csermely et al. 2010; Greives and Zhou 2014; Hammes et al. 2009; Michel 2016). Interestingly, Okazaki and Takada revealed that strong- and long-range interactions between the ligand and the protein favor the induced-fit, and weak- and short-range interactions lead to the conformational selection (Okazaki and Takada

Table 1 Targeting signal motifs embedded in signal sequences involved in various intracellular transport pathways

Signal sequence	Targeting signal motif ^a	Targeted process/pathway	Binding/receptor proteins ^b	Reference
Classical nuclear localization signal (NLS)	cluster of basic residues (two subtypes: monopartite and bipartite NLSs)	nuclear import	importin- α	(Kim et al. 2017)
Nuclear localization signal (PY motif)	hydrophobic region – RX ₂ PY	nuclear import	importin- β	
Non-classical NLS	various sequences containing basic residues	nuclear import	importin- α , importin- β	
Nuclear export signal (NES)	Leu-rich ϕ X ₂ ϕ X ₂ ϕ X ϕ	nuclear export	exportin-1 (CRM1), exportin-2 (CAS)	(Kim et al. 2017)
Nucleocytoplasmic shuttling signal	unknown	nuclear import and export	unknown	(Michael 2000)
Animal and fungal presequences	ϕ XX ϕ ϕ , the 2nd X is preferably a basic residue, amphiphilic α -helix	mitochondrial import targeted to matrix and inner membranes	animal and fungal Tom20	(Muto et al. 2001; Obita et al. 2003)
Plant presequences	short hydrophobic segments containing a basic residue, amphiphilic α -helix	mitochondrial import targeted to matrix and inner membranes	plant and green algal Tom20	(Rimmer et al. 2011)
Internal mitochondrial targeting signals	separate hydrophobic segments	mitochondrial inner membranes	animal Tom70, plant mtOM64	(Brix et al. 1997; Perry et al. 2008)
Transit peptide (TP)	amphiphilic α -helix	plastid (chloroplast)	Toc34/Toc159	(Chotewutmontri et al. 2017)
PTS1	SKL>	peroxisome	Pex5p	(Kunze and Berger 2015)
PTS2	R ϕ XX ϕ XX[HO] ϕ , amphiphilic α -helix	peroxisome	Pex7p	
ER retention signal	KDEL>, but actually [KRHQA][DENQ]EL>	retrotransport from Golgi to ER	KDEL receptor	(Capitani and Sallase 2009)
Golgi retention signal in glycotransferases	[FL][LIV]XX[RK]	Golgi localization of glycotransferases	GOLPH3/Vps74p	(Banfield 2011)
Vacuolar sorting signal (VSS)	[NL][PIL][IP][RNS]	vacuole	BP-80 family	(Jolliffe et al. 2005)
Signal peptide	unidentified motif in the N-terminal 20 to 25 amino acid residues of secretory proteins, hydrophobic α -helix	co-translational and post-translational transport to ER	SRP54/Ffh, SecA	(Shan and Walter 2005; von Heijne 1985)
Tyrosine-based endocytosis signals	NPXY (variants: DPXXY or NPXXY) GDXY	internalization and trafficking of receptor-ligand complexes on the cell surface to endosomes, TGN and lysosomes.	Clathrin, APs, Dab-1, Dab-2	(Braulte and Bonifacio 2009; Pandey 2010)
Dileucine-based endocytosis signal	YXX ϕ DXXLL [DE]XXXL[LJ]		unknown APs GGA APs	

^a The consensus pattern is conveniently denoted using the PROSITE syntax (<http://www.expasy.ch/prosite/>). [ABC], any one of A, B, and C; ϕ , hydrophobic residue; X, any amino acid residue; >, C-terminus

^b APs, adaptor proteins; Dab, AP disabled; Ffh, prokaryotic homolog to the 54 kDa subunit of the eukaryotic SRP54; GGA, Golgi-localizing; gamma-adaptin ear homology domain, ARF-binding; GOLPH, Golgi Phosphoprotein; SRP, signal recognition particle; TGN, trans-Golgi network; Vps, vacuolar protein sorting-associated protein

Fig. 3 Molecular basis of the ligand recognition by a protein molecule, from strict recognition to promiscuous recognition. The physical principles that characterize the recognition modes are shown in the *bottom part* of the figure

2008). They also suggested that large macromolecular ligands tend to connect by induced-fit, and low-molecular weight ligands tend to bind by conformational selection. Even though the conformational flexibility of a protein molecule can accommodate the relaxed specificity of the ligands, many biological processes still require more promiscuous specificity, which cannot be explained by the mixed induced-fit/conformational selection mechanism.

Here, we propose a new mechanism that enables promiscuous recognition. In addition to the complementarities of shape and charges and the protein conformational changes, the large motion of the ligand in the binding pocket of the protein is crucial in the ligand recognition (Fig. 3). First, we assume that the segment containing a targeting signal forms a secondary structure, most likely an α -helix, and the secondary structure behaves as a rigid unit in the bound states. Here, the “pose” of a ligand is defined as the position and orientation of the ligand, relative to the protein molecule. The ligand interacts with the protein in many different bound states, and the pose of the ligand in each bound state is different. In each pose, a subset of the ligand features is recognized by the protein molecule with the mixed induced-fit/conformational selection mechanism. The situation implies that the number of recognizable features of the ligand is larger than the number of recognition sites on the protein molecule. In other words, a mismatch of the numbers exists. To recognize all of the ligand’s features, the rapid exchange between the bound states/poses must occur, presumably without the dissociation of the ligand. The large residual mobility of the ligand in the binding pocket confers an entropic advantage to increase the binding affinity. We refer to this dynamic mechanism as, “multiple partial recognitions in dynamic equilibrium”, or

MPRIDE for short. It should be mentioned here that Mittag and Forman-Kay proposed an apparently similar, but distinct dynamic recognition mode of a phosphorylated intrinsically disordered ligand; i.e., “the polyelectrostatic” model (Mittag et al. 2010).

What is the benefit of the dynamic recognition? The partial recognition mode can increase the chance of one protein adapting to structurally diverse ligands by the mixed induced-fit/conformational selection mechanism. For example, the simultaneous recognition of two hydrophobic side chains using two hydrophobic sites is easier than the simultaneous recognition of three hydrophobic side chains using three hydrophobic sites. There are three ways to choose two side chains from a set of three side chains. Thus, at least three poses in a dynamic exchange are necessary to recognize the three hydrophobic side chains by means of two hydrophobic sites. We think that the MPRIDE mechanism also enables structurally different ligands to have similar affinities. It is frequently asked whether one particular pose can be isolated by removing one feature of the ligand; for example, by replacing one hydrophobic residue with a hydrophilic residue. In reality, however, the loss of the dynamic exchange between the multiple bound states results in a very large affinity reduction due to the entropic loss, and the isolation of any single pose becomes infeasible.

Structural basis of the MPRIDE mechanism

Here, we discuss our current ideas on the structural basis of the MPRIDE mechanism. First, the protein probably adopts a molten globule-like structure in the absence of the ligands.

The considerable flexibility of the protein molecules allows large adaptive conformational changes upon binding to a variety of ligands. Secondly, the binding site is expected to be exclusively composed of aliphatic side chains and contain few aromatic side chains, which makes the hydrophobic binding surface relatively flat and smooth. If bulky and rigid aromatic side chains existed in the binding pocket, then the contacts would be interdigitated and the ligand could not easily move in the binding site. These features will serve as a practical hallmark to identify the protein-ligand interactions operating in the MPRIDE mode.

Mitochondrial protein import system

We will describe our studies of the interactions of mitochondrial signal sequences with the rat and yeast Tom20 proteins, as a typical example of promiscuous recognition. Mitochondrial proteins directed to the mitochondrial matrix or inner membrane are synthesized in the cytosol as precursor proteins with a cleavable N-terminal signal sequence; i.e., a presequence (Fig. 4). The presequences typically consist of 15–40 amino acid residues, and exhibit a biased amino acid composition, with a high frequency of arginine, leucine, serine, and alanine, and few negatively charged residues (von Heijne 1986). Positively charged amino acid residues are sandwiched in between two to four consecutive hydrophobic residues. This pattern tends to form an amphiphilic helical conformation, with opposing hydrophobic and positively charged sides. The targeting signal embedded in the presequences is successively recognized by the receptor proteins along the import pathway across the outer and inner mitochondrial membranes. These receptor proteins are the subunits of the protein assemblies embedded in the two membranes, TOM (translocase of the outer mitochondrial membrane) and TIM (translocase of the inner mitochondrial membrane), respectively (Baker et al. 2007; Endo et al. 2003; Murcha et al. 2014). The 20-kDa TOM subunit, Tom20, is the first presequence receptor along the import pathway (Endo and Kohda 2002). The captured preproteins are passed to the β -barrel pore formed by the Tom40 protein to cross the outer membrane (Fig. 4a). The Tom20 protein is anchored to the mitochondrial outer membrane by the N-terminal hydrophobic segment, and its C-terminal soluble domain, which binds to the presequences, is exposed to the cytosol. In principle, a single Tom20 species is expressed in animal cells, and distinguishes about 1000 mitochondrial proteins from other non-mitochondrial proteins, and sorts them into mitochondria. In the testis, another type of Tom20 is expressed specifically (Likic et al. 2005).

The plant Tom20 lacks amino acid homology with the animal and fungal Tom20s, and is anchored to the outer membrane by the C-terminal hydrophobic segment, suggesting that

the plant Tom20 is a functional homolog of the animal/fungal Tom20s derived from different ancestors by convergent evolution (Perry et al. 2008). The plant Tom20 performs a more difficult job than the animal/fungal Tom20s. It must discriminate between the targeting signals for mitochondria and chloroplasts, which are both encoded in the N-terminal part of the preproteins. The plant mitochondrial presequences are similar enough to the animal/fungal presequences that the yeast TOM complex can import preproteins with plant presequences, and vice versa. In contrast to the single Tom20 in animal/fungal cells, three Tom20 isoforms are expressed in the mitochondria of the plant *Arabidopsis thaliana* (Werhahn et al. 2001).

Identification of the Tom20-binding consensus motif

To identify the targeting signal for Tom20 in the presequences, we used the chemical shift perturbation of the NMR signals of ^{15}N -labeled presequence peptides, induced by the addition of the cytosolic domains of the rat and yeast Tom20s (Muto et al. 2001). The perturbed regions occupy different positions, either near the N-terminus or at the C-terminus, in the five presequence peptides derived from the rat and yeast mitochondrial proteins. We found a common 5-residue pattern for the Tom20 binding. To refine the pattern, we devised a new peptide library approach (Obita et al. 2003). A mixture of presequence variants was tethered to Tom20 in a competitive manner with an intermolecular disulfide bond, and the relative amounts in the covalent complexes were measured by MALDI-TOF mass spectrometry (Fig. 4b). We concluded that Tom20 recognizes a 5-residue common motif, $\phi\text{XX}\phi\phi$, where ϕ means a hydrophobic residue and X is any amino acid residue. It is frequently stated in the literature that no obvious common sequence motifs exist among the presequences, but in fact, the presequences contain the Tom20-binding consensus at various positions (Fig. 4c). This complicated situation can explain the reason why a simple alignment of the signal sequences often does not work, without understanding the encryption position of the target signal.

Structural analyses of the Tom20–presequence interaction

We determined the NMR structure of the cytosolic domain of rat Tom20, in a complex with an 11-residue presequence peptide derived from rat aldehyde dehydrogenase (ALDH) (Fig. 4d), which contains the LSRL sequence as the Tom20-binding motif (Abe et al. 2000). The cytosolic domain of rat Tom20 is an all α -helical structure, with a groove to accommodate the presequence peptide. The bound peptide

Fig. 4 Tom20 is the receptor that recognizes the targeting signal embedded in the N-terminal presequences of diverse mitochondrial preproteins. **a** The membrane topology and functions of the Tom20 and Tom40 subunits of the TOM complex in the mitochondrial outer membrane. Tom20 is the presequence receptor and Tom40 is the channel-forming protein. This figure is purposely oversimplified to explain the function of Tom20. The actual TOM complex is a macromolecular assemblage consisting of many membrane-spanning protein subunits. **b** Sequence logo representation of the Tom20-binding

motif revealed by the peptide library experiment, using the competitive formation of an intermolecular disulfide bond and the subsequent MALDI-MS quantification [adapted from Fig. 6 of Obita et al. 2003, Copyright (2003), with permission from Elsevier]. **c** The mitochondrial targeting signal (i.e., Tom20-binding motif, denoted by the *S-in-a-circle*) is embedded in various positions of the N-terminal presequences (colored zigzag lines). **d, e** NMR structures of the cytosolic domains of animal (Rat) Tom20 and plant (*A. thaliana*) Tom20. The presequence peptide bound to the Rat Tom20 is colored blue

formed an amphiphilic helical structure with the three leucine side-chains aligned on one side, to interact with the hydrophobic surface in the Tom20 groove. As expected, the substitution of one of the leucine residues to a glutamine residue resulted in the loss of binding, indicating that the presequence binding to Tom20 was mediated mainly by hydrophobic rather than ionic interactions (Abe et al. 2000). We also monitored the effects of various mitochondrial presequences on the ^1H - ^{15}N cross peaks of ^{15}N -labeled Tom20. The different presequence peptides induced very similar patterns of chemical shift changes in the ^1H - ^{15}N cross peaks of Tom20, indicating that the binding site on Tom20 was unique. Before our NMR experiment, the acidic-residue rich cytosolic domain of

another TOM subunit protein, Tom22, was a strong candidate for the presequence receptor, and the ionic interactions were assumed to be important (“acid bristle” or “acid chain” hypothesis) (Lithgow et al. 1994). Contrary to the common belief at that time, our NMR experiments clearly indicated that Tom20 was the general receptor for the presequences of mitochondrial proteins, and the hydrophobic residues in the presequences were the determinants for the interactions. It is interesting to mention the possible role of Tom22. The cytosolic acidic region of Tom22 may recognize the basic side of the amphiphilic helix formed by presequences through ionic interactions, so that Tom20 and Tom22 may recognize opposite sides of the same presequence, with Tom20 recognizing

the hydrophobic surface and Tom22 the hydrophilic surface (Yamano et al. 2008). Finally, in consideration of the first structure of a signal sequence recognized by its receptor protein in an amphiphilic helical conformation, many biochemistry textbooks have adopted the figure depicting the Tom20–presequence structure to explain the structural basis of the signal sequence recognition.¹

The much smaller number of intermolecular nuclear Overhauser effects (NOEs) than that of the NOEs within the Tom20 protein implied that residual large motions of the presequence peptide occur in the binding site (Abe et al. 2000). Such large amplitude motions should be considered seriously, since the highly non-linear nature of the NOE information (i.e., proportional to the one-sixth power of the distance) could distort the NMR-derived structures. If large motions exist, then the NOE-based structure is not even an average structure. Thus, the structure of Tom20 was reliable, but the conformation and pose of the presequence were not. To gain realistic insights into the Tom20–presequence interactions, we crystallized the Tom20–presequence complex (Saitoh et al. 2007, 2011). We expected that the crystal contacts would effectively cease the presequence motions in the binding site and provide a snapshot structure of the presequence in the binding site at an atomic resolution. Due to the weak affinity, the simple mixing of Tom20 and the presequence peptide did not produce any cocrystals. Therefore, we used two biochemical techniques to shift the free-bound equilibrium toward the bound state (Fig. 5). The first method was the covalent-bond tethering of the peptide to Tom20, to increase the local concentration of the presequence. The position of the disulfide bond and the number of spacer residues between the presequence and the anchoring cysteine residue were optimized beforehand, by the aforementioned peptide library experiment (Obita et al. 2003). The second method was molecular stiffening. By introducing a disulfide bond between D-Cys and L-Cys at the *i* and *i* + 3 positions in the presequence respectively, the α -helical conformation in the unbound state was stabilized entropically (Saitoh et al. 2011). With different spacer sequences and lengths and under various crystallization conditions, three different poses of the peptide in the binding site were obtained from six crystals (Fig. 6a) (Saitoh et al. 2007, 2011). To our surprise, Tom20 was equipped with only two sites for the recognition of the three hydrophobic side chains in the Tom20 binding motif. The two binding sites on Tom20 are colored blue and red on

the surface representations, and are depicted by flat circles in the schematic drawing (Fig. 6b). In each pose, the two binding sites recognize one particular combination of two leucine side chains of the three leucine residues. Our working hypothesis is that the MPRIDE mechanism enables Tom20 to recognize the three hydrophobic side chains in the $\phi XX\phi\phi$ motif, using only two binding sites.

NMR relaxation study and MD calculation in solution

To analyze the dynamic aspects of the Tom20–presequence interactions, we measured the ^{15}N relaxation rates of the ^{15}N -labeled Tom20 or ^{15}N -labeled presequence peptide within the disulfide-tethered Tom20–presequence complex (Saitoh et al. 2007). First, we estimated the spatial distribution of the large exchange-induced spin-relaxation rates, R_{ex} , using the ^{15}N -labeled Tom20 in the *absence* of the presequence peptide. Relatively large R_{ex} values were obtained over the entire Tom20 structure, suggesting that, in the free form, the Tom20 structure is a loosely packed four-helix bundle (Fig. 7a). In contrast, in the presence of the presequence, the limited spatial distribution of the ^{15}N spins with the large R_{ex} values suggested motions on the sub-millisecond time scale at the interface between Tom20 and the presequence peptide (Fig. 7b). Upon binding to a presequence, the motions converge to a small number of states that are necessary for the recognition of presequences.

We also performed a replica-exchange molecular dynamics (MD) simulation. The simulation of a single Tom20 molecule in the complex with the presequence *without tethering* was performed in water for about 6 microseconds (Komuro et al. 2013). The MD calculation revealed that the presequence peptide remained in the binding site in an α -helical conformation, and successfully reproduced the dynamic equilibrium between the multiple poses of the presequence peptide. Recently, we performed a new MD calculation of a single Tom20 molecule in the complex with the presequence *with tethering* in water, and obtained essentially the same results (unpublished results, Miyashita et al.). We hope that MD calculations in different situations will yield insights into the dynamic, multiple recognition modes of the Tom20–presequence interactions, and complement the crystallographic and NMR studies.

Comparisons with plant Tom20 and nuclear receptors

It is interesting to compare the structure of the plant Tom20 with that of the animal/fungal Tom20 (Fig. 4e).

¹ Fig. 19-1 in Pollard TD, Earnshaw WC (2002) Cell biology, 1st edn. Elsevier Science, Philadelphia, page 298; Fig. 12-69 in Voet D, Voet JG (2003) Biochemistry, Vol. 1, 3rd edn. John Wiley & Sons, New Jersey, page 437; Fig. 3.50 in Lewin B, Cassimeris L, Lingappa VR, Plopper G (eds.) (2007) Cells, 1st edn. Jones and Barlett Publishers, Sudbury MA, page 139; Fig. 12-22 in Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (eds.) (2008) Molecular biology of the cell, 5th edn. Garland Science, New York, page 714.

Fig. 5 Two methods to shift the free-bound equilibrium to the bound state, without disturbing the motions in the binding site. **a** Molecular tethering and **b** molecular stiffening. Intermolecular and intramolecular

disulfide bonds were used in our studies, but other types of covalent bonds are also usable. For example, a stapled helix peptide has been used for molecular stiffening (Walensky et al. 2004)

The NMR structure of the cytosolic domain of the plant Tom20 (an isoform from *A. thaliana* chromosome 3) consisted of seven α -helices (Perry et al. 2006). Chemical shift perturbation experiments using ^{15}N -labeled AtTom20 indicated that the concave surface of the

helical bundle structure contains two distinct binding sites. In addition, chemical shift perturbation experiments using ^{15}N -labeled plant presequence peptides suggested that the plant presequences contained two or more separate binding regions, with each presequence binding

Fig. 6 Three poses of the presequence helix in the binding site on Tom20. **a** Three crystallographic snapshots of the presequence conformation and relative position when bound to Tom20. The three essential hydrophobic side chains in the Tom20-binding motif are depicted as *cyan, green, and orange sticks*. The contacting Tom20 surface is colored *red and blue*. **b** Schematic diagram of the three poses. The presequence helix is drawn as a *cylinder*, and the three hydrophobic side chains are depicted as *ovals*. The three poses are assumed to exist in a dynamic equilibrium. **c** Single pose of the LXXLL sequence when bound to a nuclear receptor protein. The coloring schemes in **b** and **c** are the same as in **a** [adapted from Fig. 6 of Saitoh et al. 2007, Copyright (2007), with permission from John Wiley and Sons, and Fig. 5 of Saitoh et al. 2011, Copyright (2011), with permission from American Chemical Society]

Fig. 7 NMR relaxation study of the presequence motions in the binding site of Tom20. ^{15}N nuclear spins with large R_{ex} values ($> 1 \text{ s}^{-1}$), due to the exchange of at least two states on the sub-millisecond time scale, are depicted as *spheres* in the free state (**a**), and in the bound state (**b**) in an open-book representation [adapted from Fig. 5 of Saitoh et al. 2007, Copyright (2007), with permission from John Wiley and Sons]. In the free state, the amide nitrogen atoms with large R_{ex} values (*green spheres*) are widely distributed in the Tom20 structure. In the complex state, the

amide nitrogen atoms with large R_{ex} values (*blue spheres*) are located at the hydrophobic contact surface (*red patches on the ribbons*), whereas the amide nitrogen atoms (*cyan spheres*) correspond to the conformational changes accompanying the motion of the presequence peptide. The *yellow sphere* represents the amide nitrogen atom of the residue next to the anchoring cysteine residue. This could arise from the tethering to the presequence peptide

region able to interact with either binding site on AtTom20 (Rimmer et al. 2011). Long yeast mitochondrial presequences also contain multiple Tom20-binding regions, although these binding regions are recognized sequentially by the single presequence binding site on yeast Tom20 (Yamamoto et al. 2011). Paramagnetic spin labels introduced at different positions in the presequence peptide affected the intensity of the ^1H - ^{15}N cross peaks in the plant Tom20, with strikingly similar patterns (Rimmer et al. 2011). These results suggest multiple and dynamic modes of presequence recognition, similar to the rat Tom20–presequence interaction.

The second comparison is with the strict recognition of the LXXLL motif in the coregulator (i.e., coactivator and corepressor) proteins by nuclear receptor proteins (Fig. 6c). The intolerance to the replacement with a valine residue indicates the absolute requirement of the leucine side chains (Heery et al. 1997). The nuclear receptors are transcription factors that are activated by the binding of small hydrophobic compounds, such as steroid hormones, and their transcriptional responses are regulated by the interactions with the coregulators (McKenna et al. 1999; Plevin et al. 2005; Savkur and Burris 2004). The structural basis of the strict specificity toward the three leucine residues in the LXXLL motif was established by several crystal structures (Darimont et al. 1998; Nolte et al. 1998; Shiau et al. 1998). The three binding sites on the nuclear receptor protein recognize the three leucine side chains simultaneously, in addition to the two helix-capping interactions (Savkur and Burris 2004). The typical lock-and-key mechanism of the nuclear receptors–LXXLL interaction provides a sharp contrast with the MPRIDE mechanism of the Tom20– $\phi\text{XX}\phi\phi$ interactions.

Signal-peptide recognitions by SRP54 and SecA

The N-terminal signal peptide specifies the targeting of a protein to the ER, and eventually leads to secretion from a cell (Table 1 and Fig. 2). The signal peptides are typically 15–30 amino acid residues. Although no apparent consensus motifs have been found, the signal peptide consists of a positively charged N-terminal region, a hydrophobic core region, and a less hydrophobic region, which ends with the cleavage site (von Heijne 1985). The typical length of the central hydrophobic region is seven to 15 residues. The signal peptides are recognized by the eukaryotic and archaeal SRP54 proteins and their eubacterial counterpart, Ffh (Rapoport et al. 2017; Shan and Walter 2005). The structural information of the signal peptide bound to SRP54 was obtained by crystallography. Two archaeal SRP54 proteins or their C-terminal Methionine-rich domain (M domain) were C-terminally fused with a native or artificial signal peptide after 9–11 residue glycine-serine linker sequences, and crystallized alone or with SRP19 and SRP-RNA (Hainzl et al. 2011; Hainzl and Sauer-Eriksson 2015; Janda et al. 2010). The signal peptides were bound to the same position, but the poses of the signal peptides obtained by the two research groups were somewhat different. Due to the intertwined dimer formation and the direct crystal contacts between two signal peptides in neighboring molecules in the crystals, however, it is difficult to define a clear pose of the binding mode. A recent cryo-electron microscopy structure of SRP in the complex with the ribosome particle was determined at 3.8 Å resolution (PDB 4JAJ) (Voorhees and Hegde 2015). The cryo-EM single particle analysis is superior to X-ray crystallography in terms of the freedom from crystal contact artifacts and the tolerance of

multiple states. The density supposedly corresponding to the bound TMD region (a transmembrane-domain sequence, alternative to a signal peptide) was roughly consistent with the position of the signal peptides in the crystal structures, but the modeling of the peptide portion in the density was incomplete: all of the side-chain atoms were modeled as leucine, and their temperature factors were inadequately small ($20 \text{ \AA}^2$). Thus, the precise pose and dynamics of the signal peptide bound to SRP54 are still elusive. Interestingly, the contacting surface of the SRP54 protein is abundant in aliphatic residues, similar to the animal/fungal and plant Tom20s.

In eubacteria and the organelles in eukaryotic cells, the SecA protein also recognizes signal peptides, and translocates proproteins across the inner membranes with the aid of ATP hydrolysis. Although the exact signal peptide binding site on the SecA protein remained a controversial subject for a long time (Clerico et al. 2008), a solution NMR experiment revealed that the hydrophobic core region of the signal peptide adopted an α -helical conformation and bound to a groove in the PBD (preprotein-binding domain), and simultaneously, the positively charged N-terminal region electrostatically interacted with acidic residues in the groove (Gelís et al. 2007). The pose of the signal peptide was calculated by docking the signal peptide onto the fixed crystal structure of the free SecA structure, to satisfy the NMR distance information. The intramolecular distance constraints within the signal peptide were used to form the α -helical structure of the hydrophobic core region of the signal peptide, and the intermolecular distance constraints that determine the pose were obtained from the PRE (paramagnetic relaxation enhancement) effects on the ^{13}CH methyl groups of SecA, caused by spin labels site-directly introduced to the signal peptide. Note that the PRE effect is also proportional to the one-sixth power of the distance. The complex structure (PDB 2VDA) is surely an achievement facilitated by state-of-the-art NMR techniques, but the dynamic nature of the signal peptide in the bound state might be overlooked. In summary, for the cases of both SRP and SecA, we suggest that the concept of the MPRIDE mechanism will be useful for the design of future experiments.

Conclusions

In this review, the promiscuous recognition of mitochondrial presequences by the mitochondrial presequence receptor protein, Tom20, was chosen as a typical example (Fig. 4). Promiscuous recognition is naturally weak in affinity and has a fast on–off rate, which makes the isolation of complexes difficult (Fig. 1). To apply the currently available methods, such as crystallography and NMR, to the promiscuous interactions, we need some technical development. A basic concept is the shift of the free-bound equilibrium toward the bound state, without disturbing the motions in the binding site

(Fig. 5). The tethering of ligands with an intermolecular disulfide bond to increase the local concentration, and the stiffening with an intramolecular disulfide bond to decrease the conformational freedom of the free state, are our solutions. We applied the two methods to a new type of peptide library experiment for the analysis of consensus sequences, the crystallization of the complex states, and the NMR relaxation study. Based on our long-term studies of the rat Tom20–presequence interactions, we proposed the MPRIDE mechanism as the molecular basis for the promiscuous recognition (Figs. 3 and 6b). The MPRIDE mechanism should be applicable to other promiscuous proteins, including plant Tom20, SRP54, and SecA.

Acknowledgements This review is the achievement of our long-term research project for more than 20 years, conducted at the Biomolecular Engineering Research Institute with Drs. Yoshito Abe and Takanori Muto, and at the Medical Institute of Bioregulation, Kyushu University, with Drs. Takayuki Obita, Takashi Saitoh, Toyoyuki Ose, Nobuo Maita, Reiko Kojima, Mayumi Igura, Rei Matsuoka, and Atsushi Shimada, Mr. Keisei Izumi, and Ms. Han Xiling. We thank Professor Toshiya Endo (Kyoto Sangyo University) for fruitful discussions on the biochemical functions of the TOM and TIM proteins, and Drs. Yasuaki Komuro and Yuji Sugita (RIKEN Advanced Science Institute), and Dr. Naoyuki Miyashita (RIKEN Quantitative Biology Center) for their MD calculations and stimulating discussions.

Compliance with ethical standards

Conflict of interest Daisuke Kohda declares that the author has no conflicts of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by the author.

Abbreviations ER, Endoplasmic reticulum; MD, Molecular dynamics; MPRIDE, Multiple partial recognition in dynamic equilibrium; NOE, Nuclear Overhauser effect; TIM, Translocase of the inner mitochondrial membrane; TOM, Translocase of the outer mitochondrial membrane; Tom20, 20-kDa subunit of the TOM complex; PRE, Paramagnetic relaxation enhancement

References

- Abe Y et al (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* 100:551–560
- Ast J, Stiebler AC, Freitag J, Bolker M (2013) Dual targeting of peroxisomal proteins. *Front Physiol* 4:297. <https://doi.org/10.3389/fphys.2013.00297>
- Baker MJ, Frazier AE, Gulbis JM, Ryan MT (2007) Mitochondrial protein-import machinery: correlating structure with function. *Trends Cell Biol* 17:456–464. <https://doi.org/10.1016/j.tcb.2007.07.010>
- Banfield DK (2011) Mechanisms of protein retention in the Golgi. *Cold Spring Harb Perspect Biol* 3:a005264. <https://doi.org/10.1101/cshperspect.a005264>
- Bonifacino JS, Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem* 72:395–447. <https://doi.org/10.1146/annurev.biochem.72.121801.161800>

- Braulke T, Bonifacino JS (2009) Sorting of lysosomal proteins. *Biochim Biophys Acta* 1793:605–614. <https://doi.org/10.1016/j.bbamer.2008.10.016>
- Brix J, Dietmeier K, Pfanner N (1997) Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *J Biol Chem* 272:20730–20735
- Capitani M, Sallèse M (2009) The KDEL receptor: new functions for an old protein. *FEBS Lett* 583:3863–3871. <https://doi.org/10.1016/j.febslet.2009.10.053>
- Carrie C, Giraud E, Whelan J (2009) Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. *FEBS J* 276:1187–1195. <https://doi.org/10.1111/j.1742-4658.2009.06876.x>
- Chakraborty P, Di Cera E (2017) Induced fit is a special case of conformational selection. *Biochemistry* 56:2853–2859. <https://doi.org/10.1021/acs.biochem.7b00340>
- Changeux JP, Edelstein S (2011) Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Reprod* 3:19. <https://doi.org/10.3410/B3-19>
- Chotewutmontri P, Holbrook K, Bruce BD (2017) Plastid protein targeting: preprotein recognition and translocation. *Int Rev Cell Mol Biol* 330:227–294. <https://doi.org/10.1016/bs.ircmb.2016.09.006>
- Clerico EM, Maki JL, Gierasch LM (2008) Use of synthetic signal sequences to explore the protein export machinery. *Biopolymers* 90:307–319. <https://doi.org/10.1002/bip.20856>
- Csermely P, Palotai R, Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* 35:539–546. <https://doi.org/10.1016/j.tibs.2010.04.009>
- Darimont BD et al (1998) Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev* 12:3343–3356
- Endo T, Kohda D (2002) Functions of outer membrane receptors in mitochondrial protein import. *Biochim Biophys Acta* 1592:3–14
- Endo T, Yamamoto H, Esaki M (2003) Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles. *J Cell Sci* 116:3259–3267. <https://doi.org/10.1242/jcs.00667>
- Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges* 27:2984–2993
- Gelis I et al (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131:756–769. <https://doi.org/10.1016/j.cell.2007.09.039>
- Gianni S, Dogan J, Jemth P (2014) Distinguishing induced fit from conformational selection. *Biophys Chem* 189:33–39. <https://doi.org/10.1016/j.bpc.2014.03.003>
- Greives N, Zhou HX (2014) Both protein dynamics and ligand concentration can shift the binding mechanism between conformational selection and induced fit. *Proc Natl Acad Sci U S A* 111:10197–10202. <https://doi.org/10.1073/pnas.1407545111>
- Hainzl T, Huang S, Meriläinen G, Brannstrom K, Sauer-Eriksson AE (2011) Structural basis of signal-sequence recognition by the signal recognition particle. *Nat Struct Mol Biol* 18:389–391. <https://doi.org/10.1038/nsmb.1994>
- Hainzl T, Sauer-Eriksson AE (2015) Signal-sequence induced conformational changes in the signal recognition particle. *Nat Commun* 6:7163. <https://doi.org/10.1038/ncomms8163>
- Hammes GG, Chang YC, Oas TG (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc Natl Acad Sci U S A* 106:13737–13741. <https://doi.org/10.1073/pnas.0907195106>
- Heery DM, Kalkhoven E, Hoare S, Parker MG (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* 387:733–736. <https://doi.org/10.1038/42750>
- Janda CY, Li J, Oubridge C, Hernandez H, Robinson CV, Nagai K (2010) Recognition of a signal peptide by the signal recognition particle. *Nature* 465:507–510. <https://doi.org/10.1038/nature08870>
- Jolliffe NA, Craddock CP, Frigerio L (2005) Pathways for protein transport to seed storage vacuoles. *Biochem Soc Trans* 33:1016–1018. <https://doi.org/10.1042/BST20051016>
- Kim YH, Han ME, Oh SO (2017) The molecular mechanism for nuclear transport and its application. *Anat Cell Biol* 50:77–85. <https://doi.org/10.5115/acb.2017.50.2.77>
- Komuro Y, Miyashita N, Mori T, Muneyuki E, Saitoh T, Kohda D, Sugita Y (2013) Energetics of the presequence-binding poses in mitochondrial protein import through Tom20. *J Phys Chem B* 117:2864–2871. <https://doi.org/10.1021/jp400113e>
- Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci U S A* 44:98–104
- Kunze M, Berger J (2015) The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Front Physiol* 6:259. <https://doi.org/10.3389/fphys.2015.00259>
- Likic VA et al (2005) Patterns that define the four domains conserved in known and novel isoforms of the protein import receptor Tom20. *J Mol Biol* 347:81–93. <https://doi.org/10.1016/j.jmb.2004.12.057>
- Lithgow T, Junne T, Suda K, Gratzer S, Schatz G (1994) The mitochondrial outer membrane protein Mas22p is essential for protein import and viability of yeast. *Proc Natl Acad Sci U S A* 91:11973–11977
- McKenna NJ, Lanz RB, O'Malley BW (1999) Nuclear receptor coregulators: cellular and molecular biology. *Endocr Rev* 20:321–344. <https://doi.org/10.1210/edrv.20.3.0366>
- Meyer-Almes FJ (2016) Discrimination between conformational selection and induced fit protein-ligand binding using integrated global fit analysis. *Eur Biophys J* 45:245–257. <https://doi.org/10.1007/s00249-015-1090-1>
- Michael WM (2000) Nucleocytoplasmic shuttling signals: two for the price of one. *Trends Cell Biol* 10:46–50
- Michel D (2016) Conformational selection or induced fit? New insights from old principles. *Biochimie* 128–129:48–54. <https://doi.org/10.1016/j.biochi.2016.06.012>
- Mittag T, Kay LE, Forman-Kay JD (2010) Protein dynamics and conformational disorder in molecular recognition. *J Mol Recognit* 23:105–116. <https://doi.org/10.1002/jmr.961>
- Murcha MW, Wang Y, Narsai R, Whelan J (2014) The plant mitochondrial protein import apparatus - the differences make it interesting. *Biochim Biophys Acta* 1840:1233–1245. <https://doi.org/10.1016/j.bbagen.2013.09.026>
- Muto T, Obita T, Abe Y, Shodai T, Endo T, Kohda D (2001) NMR identification of the Tom20 binding segment in mitochondrial presequences. *J Mol Biol* 306:137–143. <https://doi.org/10.1006/jmbi.2000.4397>
- Nolte RT et al (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* 395:137–143. <https://doi.org/10.1038/25931>
- Obita T, Muto T, Endo T, Kohda D (2003) Peptide library approach with a disulfide tether to refine the Tom20 recognition motif in mitochondrial presequences. *J Mol Biol* 328:495–504
- Okazaki K, Takada S (2008) Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms. *Proc Natl Acad Sci U S A* 105:11182–11187. <https://doi.org/10.1073/pnas.0802524105>
- Pandey KN (2010) Small peptide recognition sequence for intracellular sorting. *Curr Opin Biotechnol* 21:611–620. <https://doi.org/10.1016/j.copbio.2010.08.007>
- Paul F, Weikl TR (2016) How to distinguish conformational selection and induced fit based on chemical relaxation rates. *PLoS Comput Biol* 12:e1005067. <https://doi.org/10.1371/journal.pcbi.1005067>

- Perry AJ, Hulett JM, Likic VA, Lithgow T, Gooley PR (2006) Convergent evolution of receptors for protein import into mitochondria. *Curr Biol* 16:221–229. <https://doi.org/10.1016/j.cub.2005.12.034>
- Perry AJ, Rimmer KA, Mertens HD, Waller RF, Mulhern TD, Lithgow T, Gooley PR (2008) Structure, topology and function of the translocase of the outer membrane of mitochondria. *Plant Physiol Biochem* 46:265–274. <https://doi.org/10.1016/j.plaphy.2007.12.012>
- Plevin MJ, Mills MM, Ikura M (2005) The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem Sci* 30:66–69. <https://doi.org/10.1016/j.tibs.2004.12.001>
- Rapoport TA, Li L, Park E (2017) Structural and mechanistic insights into protein translocation. *Annu Rev Cell Dev Biol* 33:369–390. <https://doi.org/10.1146/annurev-cellbio-100616-060439>
- Rimmer KA et al (2011) Recognition of mitochondrial targeting sequences by the import receptors Tom20 and Tom22. *J Mol Biol* 405:804–818. <https://doi.org/10.1016/j.jmb.2010.11.017>
- Saitoh T, Igura M, Miyazaki Y, Ose T, Maita N, Kohda D (2011) Crystallographic snapshots of Tom20-mitochondrial presequence interactions with disulfide-stabilized peptides. *Biochemistry* 50:5487–5496. <https://doi.org/10.1021/bi200470x>
- Saitoh T et al (2007) Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. *EMBO J* 26:4777–4787. <https://doi.org/10.1038/sj.emboj.7601888>
- Savkur RS, Burris TP (2004) The coactivator LXXLL nuclear receptor recognition motif. *J Pept Res* 63:207–212. <https://doi.org/10.1111/j.1399-3011.2004.00126.x>
- Shan SO, Walter P (2005) Co-translational protein targeting by the signal recognition particle. *FEBS Lett* 579:921–926. <https://doi.org/10.1016/j.febslet.2004.11.049>
- Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95:927–937
- Vogt AD, Di Cera E (2012) Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochemistry* 51:5894–5902. <https://doi.org/10.1021/bi3006913>
- von Heijne G (1985) Signal sequences. The limits of variation. *J Mol Biol* 184:99–105
- von Heijne G (1986) Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J* 5:1335–1342
- Voorhees RM, Hegde RS (2015) Structures of the scanning and engaged states of the mammalian SRP–ribosome complex. *Elife* 4. <https://doi.org/10.7554/eLife.07975>
- Walensky LD et al (2004) Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* 305:1466–1470. <https://doi.org/10.1126/science.1099191>
- Werhahn W, Niemeyer A, Jansch L, Kruft V, Schmitz UK, Braun H (2001) Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from Arabidopsis. Identification of multiple forms of TOM20. *Plant Physiol* 125:943–954
- Yamamoto H, Itoh N, Kawano S, Yatsukawa Y, Momose T, Makio T, Matsunaga M, Yokota M, Esaki M, Shodai T, Kohda D, Hobbs AE, Jensen RE, Endo T (2011) Dual role of the receptor Tom20 in specificity and efficiency of protein import into mitochondria. *Proc Natl Acad Sci U S A* 108:91–96. <https://doi.org/10.1073/pnas.1014918108>
- Yamano K, Yatsukawa Y, Esaki M, Hobbs AE, Jensen RE, Endo T (2008) Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import. *J Biol Chem* 283:3799–3807. <https://doi.org/10.1074/jbc.M708339200>