Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Structural biology of Tat protein transport

Ben C Berks[1], Susan M Lea[2] and Phillip J Stansfeld[1]

The Tat protein transport system is found in the cytoplasmic membrane of prokaryotes and the thylakoid membrane of plant chloroplasts. Unusually, the Tat system translocates proteins only after they have folded. Proteins are targeted to the Tat system by specific N-terminal signal peptides. High resolution structures have recently been determined for the TatA and TatC proteins that form the Tat translocation site. These structures provide a molecular framework for understanding the mechanism of Tat transport. The interactions between TatC and the signal peptide of the substrate protein can be provisionally modelled. However, the way that TatA and TatC combine in the active translocation site remains to be definitively established.

Addresses
[1] Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom
[2] Sir William Dunn School of Pathology, South Parks Road, Oxford OX1 3RE, United Kingdom

Corresponding authors: Berks, Ben C (ben.berks@bioch.ox.ac.uk) and Lea, Susan M (susan.lea@path.ox.ac.uk) and

Current Opinion in Structural Biology 2014, **27**:32–37

This review comes from a themed issue on **Membranes**

Edited by **Tamir Gonen** and **Gabriel Waksman**

0959-440/$ – see front matter, Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.sbi.2014.03.003

Introduction

The Tat (twin-arginine translocation) system is a general protein transport pathway found in the cytoplasmic membrane of prokaryotes and conserved in plant chloroplasts as a thylakoid import pathway essential for the biogenesis of the photosynthetic apparatus [1–3]. The most distinctive feature of the Tat pathway is that it transports proteins that have already attained a folded state. This contrasts with other protein transport systems in ion-impermeable membranes which only transfer unstructured protein substrates. The mechanistic challenge faced by the Tat system is that it must provide a transmembrane passageway for large proteins with a wide variety of shapes and sizes (in the range 25–70 Å in diameter) that does not allow ion leakage either during transport or in the resting state when substrates are not present. Surprisingly, this difficult feat is achieved using small membrane proteins from just two structural families named TatA (containing a single transmembrane helix) and TatC (containing six transmembrane helices). Minimal Tat systems contain only one type of TatA (termed Tha4 in plants) and one type of TatC polypeptide. However, the majority of Tat systems, including the best-studied pathways found in *Escherichia coli* and plants, contain a second, functionally distinct, member of the TatA structural family called TatB (Hcf106 in plants).

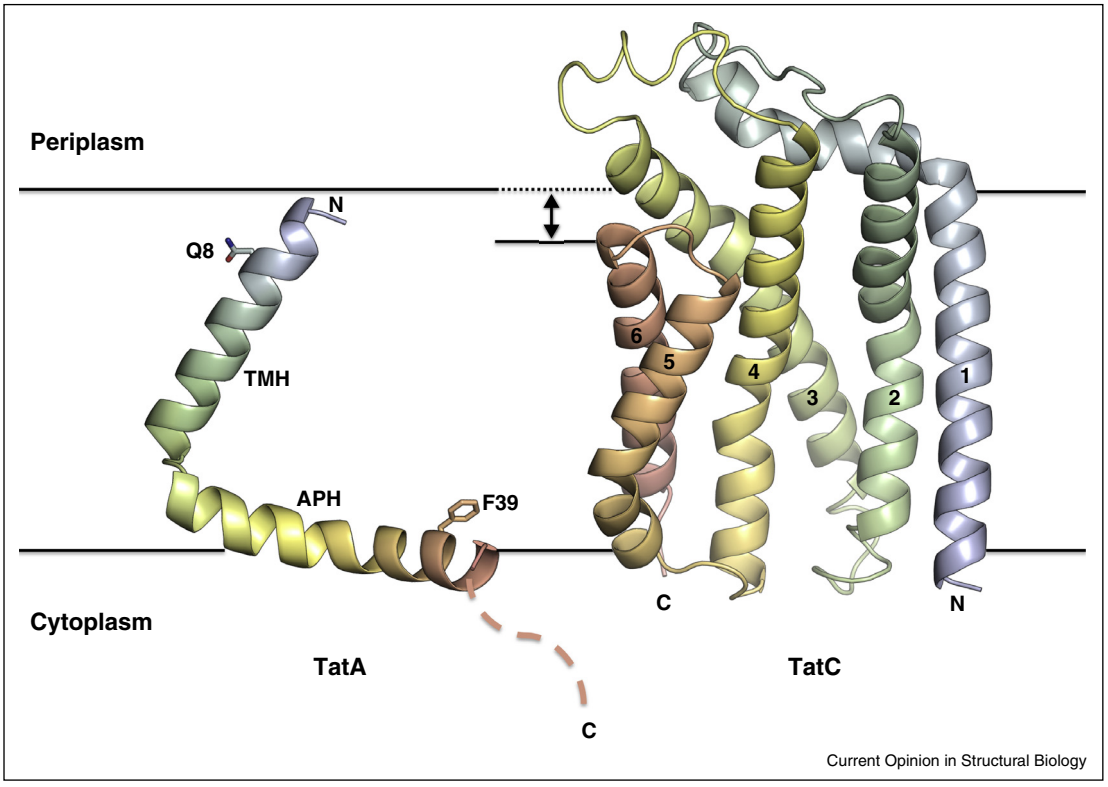
Proteins are targeted to the Tat pathway by means of an N-terminal signal peptide containing a pair of adjacent arginine residues (the 'twin-arginines') [4]. The signal peptide is recognized by a TatBC complex in the membrane [5] and this binding event triggers the protonmotive force-dependent recruitment and oligomerization of TatA protomers to form the active TatABC-containing translocation site [6,7,8•,9•]. Thus, the Tat transporter is assembled on demand, potentially solving the problem of sealing the transporter between translocation events. Within the TatABC complex the substrate is in close contact with TatA [10•] suggesting that TatA forms the protein-translocating element of the Tat system. However, in the absence of structural information on the Tat proteins and their organization in the assembled translocation site the molecular basis of the transport mechanism has been obscure. In this review we describe how recently determined high resolution structures of both TatC and TatA proteins are now providing a basis for a mechanistic understanding of Tat transport.

For the purposes of this review we will refer to the two sides of the Tat-containing membrane as the cytoplasm and periplasm. The corresponding compartments for the plant chloroplast Tat system are, respectively, the stroma and thylakoid lumen.

Structure of TatC

Within the last year the structure of TatC from the hyperthermophilic bacterium *Aquifex aeolicus* has been solved in three crystal environments. The structure of the native protein was determined at 3.5 Å resolution using experimental phasing [11••]. This model was then used to solve two lower resolution crystal forms derived from proteins engineered either to reduce surface entropy or fused to lysozyme (resolutions of 4.0 Å and 6.8 Å, respectively) [12••]. These different forms of the protein were purified in three different detergents (lauryl maltose neopentyl glycol [11••], diheptanol phosphatidylcholine, and dodecyl maltoside [12••]) and crystallized from very different solutions at pHs between 4.5 and 7.5. Despite these experimental differences there are no significant differences between the three structures at the resolutions obtained suggesting that the structure adopted by TatC on detergent extraction is stable and likely a

Figure 1



Structures of the Tat system components. Cartoon representations of the structures of *E. coli* TatA determined by solution NMR in dodecylphosphocholine [25••] and of *A. aeolicus* TatC crystallized in lauryl maltose neopentyl glycol [11••]. The proteins are positioned in the membrane bilayer as suggested by MD simulations [11••,25••] and solid state NMR measurements [26]. The transmembrane (TMH) and amphipathic (APH) helices of TatA are indicated, while the TatC transmembrane helices (TM) are numbered from the N-terminus of the protein. The functionally important TatA residues *Ec*E8 and *Ec*F39 are shown in sticks representation. Membrane thinning induced by TM5 and TM6 in MD simulations is indicated by a dashed line and arrow.

good representation of the structure of the protein in the lipid bilayer. In support of this suggestion it has been observed that the structure is well-maintained in atomistic molecular dynamics (MD) simulations in a bilayer environment [11••]. The *A. aeolicus* TatC protein (*Aa*TatC) shares 40% amino acid identity with *E. coli* TatC (*Ec*TatC) enabling the structural data to be interpreted in the light of previous biochemical and genetic experiments carried out in the *E. coli* Tat system.

The TatC fold is unique (Figure 1). While TatC contains the expected six transmembrane (TM) helices [13] these helices are kinked and tilted within the membrane leading to little exposure of the protein outside the predicted location of the membrane bilayer. In particular, TM5 and TM6 are seen to be too short to fully span the membrane bilayer. Indeed, MD simulations show that these helices are likely to induce membrane distortion leading to bilayer thinning around the helix ends [11••]. The overall shape of the molecule resembles a cupped hand with the curved transmembrane helices forming a cavity overhung by a periplasmic cap that locks the helices in place (Figure 2).

Now that the structure of the TatC protein has been resolved much current effort is being devoted to understanding where and how it interacts with substrates and with other Tat components. In this context it is notable that, for a protein of its size, TatC has an unusually large surface area available for interactions with other proteins [~16,500 Å^2, [11••]].

Signal peptide binding to TatC

Tat signal peptides (Figure 3a) are normally around 30 amino acids in length [4] and lack secondary structure in aqueous solution [14]. The N-terminal part of the signal peptide (the n-region) bears a S_1-R_2-R_3-x$_4$-F_5-L_6-K_7 consensus motif [4] containing the invariant and functionally essential arginine pair. The central part of the signal peptide is composed of hydrophobic amino acids (the h-region) while a short C-terminal portion (the c-region) usually contains the target sequence for a signal peptidase