EFFECTS OF MUTATIONS OF THE ARGININE MOTIF OF SIGNAL PEPTIDE IN A TAT SUBSTRATE ON TRANSLOCATION ACTIVITY IN *ESCHERICHIA COLI.*

Prakash Patel[1], Anja Nenninger[2], Colin Robinson[2]

Molecular Organisation and Assembly of Cells[1], and Biological Sciences[2], University of
Warwick, Coventry, CV4 7AL

Running Title: Mutagenesis of TorA-GFP signal peptide RRR motif

The twin-arginine translocase (Tat) system is a recently discovered pathway which transports proteins in bacteria and chloroplasts. This system only translocates proteins with a cleavable N-terminal signal peptide with the distinctive twin-arginine (RR) motif, which is crucial to Tat function. Here, site-directed mutagenesis of the signal peptide of the Tat substrate TorA-GFP - which has a triplet-arginine (RRR) instead of the usual twin arginine - were attempted and the effect on the translocation of this substrate in *Escherichia coli* was measured by SDS-PAGE / Western Blot of the cytoplasmic, periplasmic and membrane fractions of the bacteria. The KRK mutant of the signal peptide of TorA-GFP was found to mediate transport less than that of TorA-GFP.

The synthesis of proteins by the ribosomal translation of messenger RNA takes place in the cytosol of cells, but to function many are required in other compartments of the cell and hence have to be transported in some way, which involves transmembrane translocation. In Gram-negative bacteria, there are only two compartments – the cytoplasm and periplasm, and the transferral of proteins from cytoplasm to periplasm is usually via the Sec (secretory) pathway (1), which transports these proteins in a largely unfolded state.

Another pathway, known as Tat (twin-arginine translocase), is capable of transporting proteins fully-folded proteins within the thylakoid membranes of chloroplasts in plants and within the periplasmic membrane in bacteria such as *Escherichia coli* (Figure 1), and has been actively researched since the discovery of its existence in the early 1990s (2). In *E. coli*, this system for transporting proteins from the cytoplasm to the periplasm is hardly used – it only seems to be an advantage in anaerobic environments - whereas the Sec (secretory) pathway is very much preferred and essential for survival.

The main difference between the Tat and Sec pathways in bacteria is that the Sec pathway transports unfolded proteins in the cytoplasm to the periplasm (where they fold), whilst the Tat pathway transports the proteins fully folded. The Sec pathway in *E. coli* is reasonably well understood (1) compared to the Tat pathway; its ability to let large transition-metal containing proteins such as trimethylamine N-oxidase (TorA) through a tight membrane is still a mystery. Another major difference is that Tat translocation is driven by the change in pH between the cytoplasm and periplasm (ΔpH), whilst the Sec pathway uses the hydrolysis of nucleosides, such as ATP. Also, in the Sec pathway the proteins SecA and SecB act as chaperones which bind to the unfolded protein; there are no known analogues for Tat (although proteins such as TorA have specific enzymes that associate with them to aid cofactor insertions – in this case TorD).

In *E. coli* there are three main components in the Tat system, namely TatA, TatB and TatC. There is another component known as TatE, but this has the same role as TatA and only a very small amount of the protein is expressed in the cell. The removal of any of TatA (or TatE), TatB or TatC inactivates the pathway, which implies each of these is essential for Tat function.

The roles of each of the proteins are unclear at this time, but we are confident that TatA, TatB and TatC are all transmembrane proteins, with A/B containing one transmembrane helix and TatC with four or six. Also it is predicted TatC remains anchored in the periplasmic membrane with TatB (3) which initially attach to the substrate, whilst TatA interacts with TatB (4), and helps to stabilize the complex as a multimer once substrate attachment occurs in TatBC (Figure 2).

For *E. coli* to translocate a particular protein from the cytoplasm to the periplasm, it requires a signal on the protein such that the appropriate machinery can recognise and translocate it. This is in the form of a sequence of amino acids known as a signal

peptide at its N–terminus, which is cleaved after translocation by a enzyme called LepB on the periplasmic side of the membrane(5).

To investigate the transport activities of Tat, a fusion protein containing the signal peptide sequence of trimethylamine N-reductase (TMAO reductase or TorA) with a green fluorescent protein mutant (GFP) was used to produce a *bona fide* Tat substrate whose presence is easily detected either using immunoblotting with widely available antibodies or confocal microscopy to detect its fluorescence in cells, providing a powerful tool for dissection of the Tat pathway. This fusion protein has the advantage that GFP incorrectly folds in the periplasm (6), which helps to reduce the effect of the Sec translocase in studies.

The structure of the Tat signal peptide consists of a hydrophobic domain, proteolytic cleavage site and the twin arginine motif (RR) proceeding the hydrophobic region (2), which occurs in all Tat substrate signal peptides - both in the system in bacteria or in plant thylakoids and gives the translocase its name. The presence of these two amino acids is crucial for the translocation of the substrate by Tat but its importance is still a large mystery.

The effect of the mutagenesis of the twin arginines in the signal peptide of the TorA-GFP fusion protein on Tat transport has already been investigated (7). The first two arginines (counting from the N-terminus of the signal peptide) were changed to lysine (a similarly large positively-charged basic amino acid as arginine) (Figure), which found that the substitution of one lysine does mediate reduced Tat transport, whose rate is dependent on the position of the substituted lysine in the order RR > KR > RK. The KK mutant was unable to mediate transport. Complementary to studies with TorA-GFP, another fusion protein was constructed with the TorA signal peptide with colicin V – a protein which kills *E. coli* only if transported into the periplasm, which confirmed that RR > KR > RK > KK motifs in the signal peptide in the mediation of the Tat transport.

The problem that was not addressed was the TorA signal peptide had three consecutive arginines (RRR) instead of two as in other substrates – this obvious complication seemed to be ignored. In address this issue a comprehensive study was attempted, in which the triple arginines in TorA-GFP were replaced with lysines to form seven different mutants, and the effect on Tat transport was measured using cell fractionation and Western blotting.

MATERIALS + METHODS

The *E. coli* strain used in this study is MC41000, transformed with the ampicillin-resistant arabinose-inducible plasmid vector pJDT1 (pBAD24 with TorA-GFPmut3 insert at 1321bps). All *E. coli* were grown using Luria Broth LB (10:5:10) containing 100μg/ml ampicillin. The pJDT1 plasmid was extracted from the unmutated TorA-GFP MC4100 *E. coli* strain using a Miniprep kit (Qiagen) for the mutagenesis experiments.

Site-directed mutagenesis of the TorA-GFP signal peptide.

The mutations of the TorA-GFP signal peptide region in pJDT1 plasmid was produced using the QuikChange™ mutagenesis system (Stratagene, La Jolla, CA, USA) using the manufacturer's instructions. This is a PCR reaction using the PfuTurbo DNA polymerase enzyme with primers containing the required mutations. This was used to construct the 7 mutants of the RRR motif in the signal peptide KRR, RKR, KKR, RRK, RKK, KRK and KKK.

The parental DNA in the PCR adducts were digested by the enzyme Dpn-1, and were used to transform fully-competent *E. coli* cells using standard procedures, and were cultivated on LB ampicillin (100μg/ml) agar overnight, to be miniprepped to extract the plasmid for sequencing, using the primer seq_pJDT1_TorA.

DNA Primers (Invitrogen):

pJDT1 sequencing primer:

seq_pJDT1_TorA
GGCTAGCAGGAGGAATTCAC

Primers used in QuikChange™ mutagenesis:
(Bases different to TorA-GFP signal peptide in lower case)

QC_TorA_KRR_for
5'-
CGATCTCTTTCAGGCATCAaagCGGCGTTTT
CTGGCACAACTGGCG-3'
QC_TorA_KRR_rev

5'-
CGCCAGTTGTGCCAGAAAACGCCGcttTGAT
GCCTGAAAGAGATCG-3'

QC_TorA_RKR_for
5'-
CGATCTCTTTCAGGCATCACGTaaGCGTTTT
CTGGCACAACTGGCG-3'
QC_TorA_RKR_rev
5'-CGCCAGTTGTGCCAGAAAACGCttACG
TGATGCCTGAAAGAGATCG-3'

QC_TorA_KKR_for
5'-
CGATCTCTTTCAGGCATCAaagaaGCGTTTTC
TGGCACAACTGGCG-3'
QC_TorA_KKR_rev
5'-
CGCCAGTTGTGCCAGAAAACGCttcttTGATG
CCTGAAAGAGATCG-3'

QC_forward_RRK:
5'-
CGCATCTCTTTCAGGCATCACGTCGGaagTT
TCTGGCACAACTCGGCG-3'
QC_Reverse_RRK:
5'-CGCCGAGTTGTGCCAGAAActtCCGACG
GATGCCTGAAAGAGATGCG-3'

QC_forward_RKK
5'-CGCATCTCTTTCAGGCATCACGTAAGaag
TTTCTGGCACAACTCGGCG-3'
QC_Reverse_RKK:
5'-CGCCGAGTTGTGCCAGAAActtCTTACG
TGATGCCTGAAAGAGATGCG-3'

QC_forward_KRK
5'-
CGCATCTCTTTCAGGCATCAAAGCGGaagTT
TCTGGCACAACTCGGCG-3'
QC_Reverse_KRK:
5'-
CGCCGAGTTGTGCCAGAAActtCCGCTTTGA
TGCCTGAAAGAGATGCG-3'

QC_forward_KKK
5'-CGCATCTCTTTCAGGCATCAAAGaaGaag
TTTCTGGCACAACTCGGCG-3'
QC_Reverse_KKK:
5'-
CGCCGAGTTGTGCCAGAAActtCttCTTTGAT
GCCTGAAAGAGATGCG-3'

Cell Fractionation

Cells expressing the correct plasmid and mutation were grown aerobically on LB, and the periplasmic fraction was extracted using the EDTA/lysozyme/cold osmotic shock procedure. The remaining spheroblasts were lysed by sonication, and intact cells and debris removed by centrifugation (5min at 10000g). The membrane and cytoplasmic fractions were separated by centrifugation (30min at 10000g).

SDS-PAGE/Western Blot
To determine the amount of transport of TorA-GFP, Western blots were performed using the ECL Western blotting system, which light is produced on oxidation usually by antibodies conjugated on the protein of interest (in this case GFP) and on an oxidising enzyme (such as a peroxidase). The amount of light produced is roughly proportional to the amount of GFP on the membrane.

An SDS-PAGE was performed on the periplasmic, membrane and cytoplasmic fractions with control ΔABCDE. The resulting gel was electroblotted onto Hybond-P PVDF membrane, which was washed in Milk/PBS-T to block the membrane. Then the membrane was washed in rabbit anti-GFP (Living Colours, Clonetech) in PBS-T (1:1000) for 1h, horseradish peroxidase (HRP) anti-rabbit IgG conjugated antibody in PBS-T (1:6000) for 1h with washes in PBS-T in between (once for 15min, twice for 5min). The membrane is then washed in ECL immunodetection reagent and revealed in the dark to autoradiography film.

RESULTS
Out of the seven attempted QuikChanges™, only two mutants of the RRR signal of TorA-GFP were successfully produced - KRR and RKR. Cultures of *E. coli* with the arabinose-inducible pJDT1 containing native (RRR) TorA-GFP, TorA-GFP KRR mutant, TorA-GFP RKR mutant and ΔABCDE RRR TorA-GFP (*E. coli* without Tat system) were grown overnight in LB.

The cells were induced with 100mM arabinose/ LB and incubated for 1h after the overnight cultures were pelleted, then incubated for 3h in fresh LB. To determine whether TorA-GFP were transported by Tat from the cytoplasm to the periplasm, an SDS-PAGE/Western blot was performed on the cytoplasmic, periplasmic and membrane fractions of the four cultures. The resulting membrane was immunoblotted for GFP-type proteins.

The autoradiograph of the ECL Western blot of the four strains into its fractions (Figure 4) showed that in the RRR strain the wild-type transported the TorA-GFP from the cytoplasm to the periplasm, whereas the ΔTatABCDE strain did not, which was expected. The TorA-GFP KRR mutant produced little TorA-GFP, probably because the plasmid operon did not express, unlike in the three other strains. The most interesting result is the TorA-GFP RKR mutant, which showed that some TorA-GFP was transported by Tat – but a smaller proportion than in the RRR strain.

DISCUSSION OF RESULTS

Comparing the RKR mutant to the RRR mutant, the presence of the lysine residue did not stop translocation of the GFP, but significantly reduced it, which is concurrent with the result in (7), that RKR would mediate substrate transport. This study also suggested the KKR mutant of TorA-GFP, when produced, would not translocate at all. This also showed that the 1h arabinose induction was sufficient for ECL Western blot was sensitive enough to detect these changes and could be used to quantify the differing rates of transport.

The main problem arising was the initial construction of the mutants. In our mutagenesis experiments, the KKR mutant was almost produced, but the sequencing results showed that one of the base pairs out of the five consecutive base pairs was changed to all four bases. This is because the primer has difficulty annealing to the template DNA the greater the number of consecutive mismatched pairs there are. Two attempts at the construction of the other four mutants (RRK, RKK, KRK and KKK) and attempt to complete the mutation the KKR mutant had failed, demonstrating the notoriously temperamental nature of PCR reactions.

From these results, it confirms the knowledge that the twin-arginine motif on Tat substrates is important for recognition by Tat, and is very sensitive to mutations, but the precise importance of that is still a mystery.

In previous studies, TatB was found to interact with the whole signal peptide of the *E. coli* Tat substrate SufI, TatC only interacted with the region around the twin arginine motif (8). There is evidence that an alpha helix forms in the hydrophobic region after the twin arginines in the signal peptide of the protein SufI, in sufficiently hydrophobic conditions (9) such as a lipid membrane.

But to determine whether the crucial factor is the tertiary structure of the signal peptide around the twin arginines or the differing interactions of the arginines compared to lysine is still a question on the structure and mechanism of the Tat system itself.

1. Manting, E. H., and Driessen, A. J. M. (2000) *Molecular Microbiology* **37**(2), 226-236
2. Robinson, C., and Bolhuis, A. (2004) *Biochimica et Biophysica Acta* **1694**, 135-147
3. Mangels, D., Mathers, J., Bolhuis, A., and Robinson, C. (2005) *J. Mol. Biol.* **345**, 415–423
4. Barrett, C., and Robinson, C. (2005) *FEBS Journal* **272**, 2261–2275
5. Berks, B. C., Palmer, T., and Sargeant, F. (2005) *Curr. Op. Microbiol.* **8**, 174 - 181
6. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G. J. (2000) *J. Bacteriol.* **182**(14), 4068–4076
7. Ize, B., Gérard, F., Zhang, M., Chanal, A., Voulhoux, R., Palmer, T., Filloux, A., and Wu, L.-F. (2002) *J. Mol. Biol.* **317**, 327-355
8. Alami, M., Lüke, I., Deitermann, S., Eisner, G., Koch, H.-G., Brunner, J., and Müller, M. (2003) *Molecular Cell* **12**, 937–946
9. San Miguel, M., Marrington, R., Rodger, A., Rodger, P. M., and Robinson, C. (2003) *Eur. J. Biochem.* **270**, 3345–3352

FIGURE LEGENDS

Figure 1: The Tat and Sec translocation pathways in gram-negative bacteria such as *E. coli*. The Sec pathway is by far the most important method of proteins made in the cytoplasm to be transported to the periplasm, whilst the Tat pathway is usually reserved for proteins with unusual cofactors. Sec substrates are usually bound to a soluble component; especially SecB which binds to the protein in an unfolded state. The next step is the unfolded protein is pushed through in stages by the ATP-driven conformational changes of SecA in the membrane-bound SecYEG complex. In contrast, a Tat substrate is transported in a largely folded state through TatABC, with another enzyme which adds a complex cofactor beforehand if necessary. This mechanism is ΔpH driven. In both cases, the signal peptide for Sec and Tat is cleaved once translocation has completed. Figure from (2)

Figure 2: Possible mechanism for the Tat system. The diagram depicts translocation of a substrate comprising a passenger protein with a cleavable Tat signal peptide that forms a partially helical structure in apolar environments. Most likely, TatB and TatC form a functional unit that serves as the primary binding site for substrate (step 1: binding). Some TatA is bound to this TatBC core but most is in the form of a separate homooligomeric complex. Binding of the substrate to the TatABC complex is proposed to trigger the recruitment of this TatA complex (step 2: assembly) and this results in formation of the active translocon and translocation of the substrate (step 3), which the multiple TatA proteins probably form into the pore which the substrate passes. Figure from (2)

Figure 3: Signal peptides of Tat substrates in *E. coli*. The bold characters highlight the twin-arginine motif and other charged residues, and the underlined characters indicate the hydrophobic region after the motif. The TorA signal peptide sequence is highlighted by an arrow. Notice that this signal peptide has three consecutive arginine residues instead of the usual two as in the signal peptides of the other Tat substrates. Figure from (2)

Figure 4: Immunoblot of the periplasmic, membrane and cytoplasmic fractions of four *E. coli* cultures against antibodies to GFP after induction with arabinose of 1h. The RRR represents *E. coli* which expression of the unmutated pJDT1 plasmid; the KRR and RKR have the respective mutants to the TorA-GFP signal peptide, whilst the ΔABCDE is a strain of MC4100 without the tat genes (ΔTatABCDE) with the unmutated pJDT1 plasmid.

1.
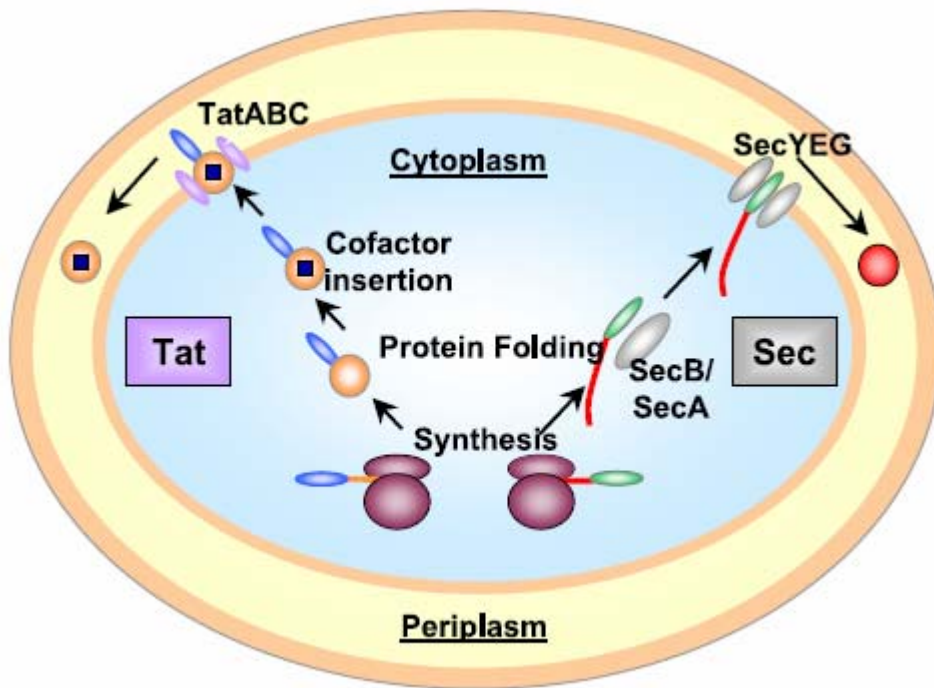
2.
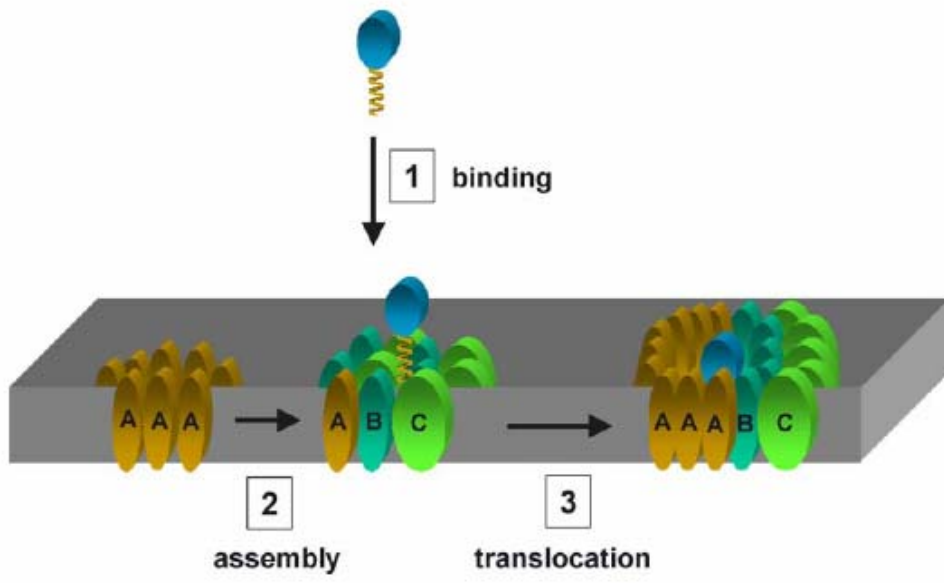
3.

```
E. coli
DmsA    MKTKIPDAVLAAEVSRRGLVKTTAIGGLAMASSALTLPFSRIAHA          DMSO reductase
FdnG            MDVSRRQFFKICAGGMAGTTVAALGFAPKQALA          Formate dehydrogenase
FdoG            MQVSRRQFFKICAGGMAGTTAAALGFAPSVALA          Formate dehydrogenase
HyaA MNNEETFYQAMRRQGVTRRSFLKYCSLAATSLGLGAGMAPKIAWA          Hydrogenase-1, small subunit
HybA                  MNRRNFIKAASCGALLTGALPSVSHA          Hydrogenase-2, small subunit
NapA          MKLSRRSFMKANAVAAAAAAGLSVPGVARA          Nitrate reductase
NrfC          MTWSRRQFLTGVGVLAAVSGTAGRVVA          Nitrite reductase
PcoA        MLLKTSRRTFLKGLTLSGVAGSLGVWSFNARS          Copper resistance protein
SufI          MSLSRRQFIQASGIALCAGAVPLKASA          Suppressor of ftsI mutant
→TorA     MNNNDLFQASRRRFLAQLGGLTVAGMLGPSLLTPRRATAAQA          TMAO reductase
YnfE        MSKNERMVGISRRTLVKSTAIGSLALAAGGFSLPFTLRNAAA          DMSO reductase
YnfF    MKIHTTEALMKAEISRRSLMKTSALGSLALASSAFTLPPSQMVRA          DMSO reductase
```

4.