

A. Overview

Protein transport across hydrophobic membranes that partition cellular compartments is essential in all cells. The twin arginine translocation (Tat) pathway transports proteins across the prokaryotic cytoplasmic membranes. Distinct from the universally conserved Sec pathway, which secretes unfolded proteins, the Tat machinery is unique in that it secretes proteins in a folded conformation, making it an attractive pathway for the transport and secretion of heterologously expressed proteins that are Sec-incompatible.

During the past 7 years, the DOE-supported project has focused on the characterization of the diversity of bacterial and archaeal Tat substrates as well as on the characterization of the Tat pathway of a model archaeon, *Haloferax volcanii*, a member of the haloarchaea. We have demonstrated that *H. volcanii* uses this pathway to transport most of its secretome.

B. Summary of results from peer-reviewed publications:

1. Prokaryotic Tat substrates: Initial analyses of bacterial (mainly *Escherichia coli*) Tat substrates suggested that proteins secreted via the Tat pathway mainly include cofactor-containing proteins. Extensive *in silico* as well as *in vivo* analysis carried out in our lab over the past 7 years has clearly shown that many prokaryotes secrete a highly diverse set of proteins via this pathway, hence it can be considered a general protein translocation pathway (1, 2, 4, 5, 6). Moreover, we showed that this pathway is used to varying extents among prokaryotes. For example, while some organisms lack the Tat pathway, one group of archaea that thrive in environments with salt concentrations approaching saturation, the haloarchaea, secrete most of their proteins via the Tat pathway (1, 2, 4, 5, 6). The *in silico* analyses initially leading to this conclusion was greatly aided by TatFind, a perl program developed in our laboratory to identify predicted Tat substrates (1, 2).

In vivo confirmation of TatFind positives as bona fide Tat substrates was carried out using the model archaeon *H. volcanii* (1, 4, 6) as well as the bacterium *Streptomyces coelicolor* (5) (the latter was carried out in collaboration with Dr. Tracy Palmer's lab at the Innes Center in England). Moreover, other groups, studying a variety of organisms, have confirmed TatFind positives to be Tat substrates *in vivo*, including *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* proteins that are virulence factors, underlining the usefulness of this publicly available program (signalfind.org).

One particularly interesting set of substrates we confirmed to be secreted via the Tat pathway is archaeal lipoproteins (6). While *in silico* work suggested that *S. coelicolor* and *H. volcanii* Tat substrates include lipoproteins, lipid anchoring of Tat substrates had not been demonstrated in any organism. Moreover, since archaea lack proteins with significant homology to any of the bacterial lipoprotein biogenesis enzymes, the existence of archaeal lipoproteins had not been anticipated.

2. The H. volcanii Tat machinery: We have been studying the Tat pathway in the haloarchaeon *H. volcanii* with the goal to understand how it effectively transports such a large number of proteins. Moreover, since this organism secretes a wide variety of proteins via this pathway, we can use a diverse set of reporter proteins for biochemical as well as genetic analyses in our studies of the Tat pathway.

Consistent with the observation that most secreted proteins are Tat-dependent, we showed that three of the four known Tat components, TatAt, TatCo and TatCt, are essential in *H. volcanii* (4). Moreover, extensive biochemical analyses revealed that distinct from observations in *E. coli*, in addition to the membrane, TatA paralogs (TatAt and TatAo) can be found in the *H.*

volcanii cytoplasmic fraction. Recent analyses of *Bacillus subtilis* TatA paralogs suggest that cytoplasmic TatA may target substrates to the membrane.

While the predicted sizes and topologies of the haloarchaeal TatA homologs are similar to those of other bacterial and archaeal TatA proteins, the predicted size and topologies of their TatC homologs are distinct. TatCo has a greatly extended N-terminus, and TatCt has 14 TMS, rather than the usual 6, likely resulting from a duplication event in the gene that encodes it. These variations may represent adaptations to the extensive use of the Tat pathway in *H. volcanii* (see below, unpublished results).

C. Peer-reviewed publications acknowledging this grant

1. Rose, R.W., T. Brüser, J.C. Kissinger, and **M. Pohlschröder**. 2002. Adaptation of protein secretion to extremely high salt concentrations by extensive use of the twin arginine translocation pathway. *Mol. Microbiol.* **5**: 943-950
2. Dilks, K. W. R. Rose, Enno Hartmann, and **M. Pohlschröder**. 2003. Prokaryotic use of the twin arginine translocation pathway: A Genomic Survey. *J. Bacteriol.* **185**:1478-1483.
3. **Pohlschröder, M.**, K. Dilks, N. Hand, and R. W. Rose. 2004. Protein translocation across the archaeal cytoplasmic membrane. Invited review, *FEMS Microbiology Reviews*. **28**:3-24.
4. Dilks, K. M. I. Gimenez, and **M. Pohlschröder**. 2005. Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea. *J. Bacteriol.* **187**: 8104-8113.
5. Widdick, D., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, **M. Pohlschröder** and Tracy Palmer. 2006. Secretome analysis reveals that the twin-arginine translocation (Tat) pathway is a major route of protein export in *Streptomyces coelicolor*. *PNAS*, **106**:17927-17932
6. Gimenez M. I. Dilks, K. and **M. Pohlschröder**. 2007. *Haloferax volcanii* twin arginine translocation substrates include soluble secreted, C-terminally anchored and lipoproteins. *Mol. Microbiol.* **66**. 1597- 1602.

D. Patent

Dilks, K., Kissinger, J. C., T. Brüser, R.W. Rose, and **M. Pohlschröder**. Heterologous protein production using the twin arginine Translocation Pathway

E. Development of Computer Program

Dilks, K., Kissinger, J. C., T. Brüser, R.W. Rose, and **M. Pohlschröder**. 2002. TATFIND, a perl program that predicts prokaryotic Tat signal sequences.

F. Unpublished data

We have made significant progress in the characterization of one of the *H. volcanii* TatC paralogs. Moreover, we have also initiated attempts to identify components that are involved in *H. volcanii* lipoprotein transport. Finally, we have characterized a haloarchaeal reporter protein PhoD; this protein will be extremely useful for genetic screens and may also be helpful in attempts to determine the functional roles played by each of the two *H. volcanii* TatA paralogs.

1. Analyses of the unique haloarchaeal TatC proteins: TatC has been implicated in the recognition of the Tat signal peptide; once TatC recognizes the signal peptide, the TatC/substrate complex interacts with the proposed pore subunit TatA. TatC forms a multimer, at the very least a homodimer. However, little is known about the molecular basis of these interactions. Interestingly, the structure of one of the haloarchaeal TatC paralog appears to be the result of a

duplication event (TatCt in *H. volcanii*) (Fig. 1A). The duplication within this paralog may allow TatC to function more efficiently, which considering the extensive use of the Tat pathway in haloarchaea may be imperative. To determine the significance of this duplication, we generated 4 TatCt constructs with major N- or C-terminal truncations. While constructs lacking the N-terminal 6 TMS were not able to complement a chromosomal TatCt deletion, a construct containing only the first 6 TMS (N-6) fully complemented the deletion when overexpressed from a plasmid. The fact that N-6 expressed from the chromosome is insufficient to rescue the deletion suggests that the TatC-duplication may be necessary for efficient Tat pathway transport in *H. volcanii*, possibly by increasing the rate of substrate binding. However, the fact that C-6 and C-8 do not complement TatCt deletions indicates that N-6 contains an element lacking in C-6 that is required by TatCt to function properly. This element may facilitate interactions between TatC components or possibly between TatA and TatC. In either case, if the N-terminus (N-6) facilitates interactions with other TatC components or with TatA, within the context of the wild-type TatC the C-terminus would not require a similar motif since it would be brought into close association with these other components via the N-terminus (N-6). Interestingly, sequence alignments with other TatC homologs show that N-6 contains a highly conserved region that is not present in C-6; this conserved region consists of part of the 5th TMS and the 3rd periplasmic loop (Fig. 1B) and may correspond to a unique TatA or TatC binding site. This conserved portion of N-6 is present in the TatC homologues of such distantly related organisms as *Bacillus subtilis* and *Escherichia coli*. Employing Strep-tagged N-6 and C-6 constructs as well as N-6 constructs lacking this conserved domain, we can perform co-purification studies that may generate data that will allow us to determine the functional role played by this conserved TatC domain in protein translocation in both haloarchaea and other prokaryotes.

2. Analysis of Tat component-substrate interaction: We have previously shown that although TatAt can complement TatAo deletions and that overexpressed TatCt can complement TatCo deletions, TatAo cannot complement TatAt and TatCo cannot complement TatCt, suggesting that these paralogs might be involved in the transport of distinct substrates. Indeed, we recently determined that the *Halobacterium* NRC-1 alkaline phosphatase D (PhoD) is transported in a TatAo dependent manner (Fig. 2A). Extracellular PhoD activity can be monitored using enzyme activity assays of supernatant protein fraction or by plating cells on solid medium containing the chromogenic substrate XP. We found that enzyme activity can be observed when PhoD is expressed in a wild-type *H. volcanii* strain. However, when PhoD is expressed in a Δ tatAo strain, only cytoplasmic activity can be detected. Consistent with our interpretation of these results, only cytoplasmic enzyme activity can be detected in cells expressing PhoD constructs lacking a signal peptide or having a signal peptide that contains twin lysines rather than twin arginines exhibit (Fig 2A). Western blot analysis showed that the lack of extracytoplasmic activity was not due to protein instability (2B).

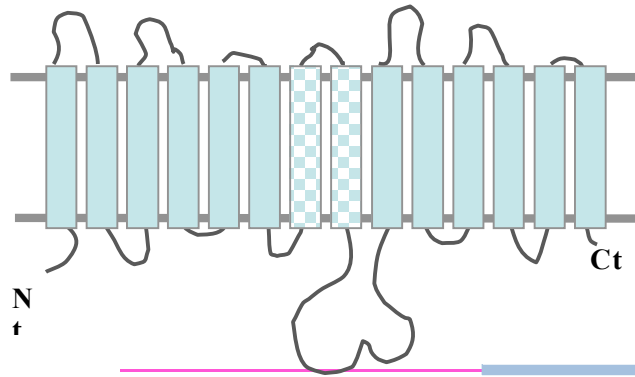
In addition to being the first TatAo-dependent substrate identified, PhoD is a very useful reporter substrate for genetic screens designed to identify Tat mutants because cells that express this protein turn blue on agar plates containing XP.

3. Analyses of *H. volcanii* Tat lipoprotein biogenesis: As noted above, previous work in our lab demonstrated that lipoproteins can be transported via the Tat pathway. Interestingly, globomycin, a non-competitive inhibitor of signal peptidase II (SPase II), the protease responsible for lipoprotein signal peptide processing in bacteria is inhibiting *H. volcanii* lipoprotein processing, although no SPase II homolog has yet been described in archaea. In fact, we have shown that globomycin specifically inhibits the signal peptide processing of *H. volcanii*

lipoproteins. These results represent the first verification of lipoprotein secretion via the Tat pathway and reveal the presence of a lipoprotein biogenesis pathway in archaea (6).

We have now isolated spontaneous and chemically-induced globomycin resistant *H. volcanii* mutants. These mutants are currently being tested for their ability to allow wild-type *H. volcanii* growth on inhibitory concentrations of globomycin in hopes of identifying an archaeal analog or distant homolog of bacterial SPase II. Alternatively, we might identify a novel globomycin target, a result that would be intriguing since globomycin may have potential uses as a therapeutic antibiotic. While it had been thought that globomycin specifically targets SPase II, recent results suggest that it also acts through the inhibition of additional targets.

A

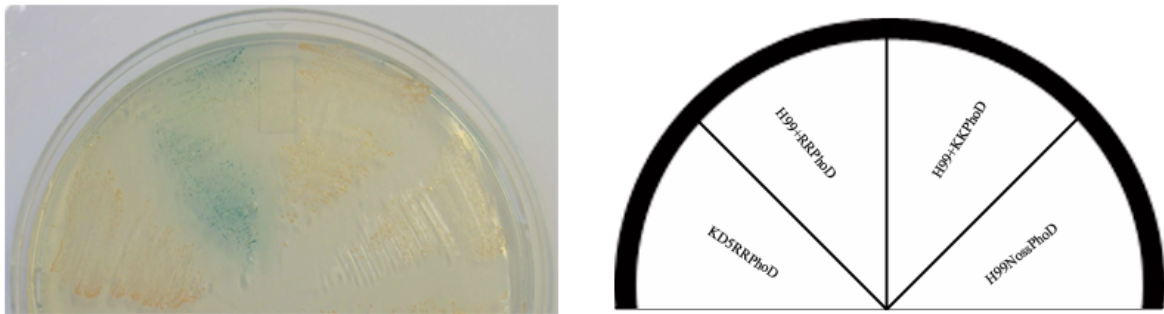


B

6CtHalobact	1	AGM	DAFTEDDTT	EDDIG	YYH	DRF	FDS	LRRAF	TVGGF	HLVM	GVVF	GLWYY	---	GL
6CtNatronom	1	-GV	DAFSEDD	EDDIG	YAY	AF	VDS	TRSF	PCV	PAG	AA	TFE	---	GL
6CtHalofer	1	--F	ELTEGET	EDDIG	YVY	DAF	VDS	TRSF	PCV	PAG	AA	TFE	---	GL
6NtHalofer	1	---	MS---	SA	DEDD	TOQTIA	AGRETAGAM	LRAAQKD	OKVF	TVS	VGFL	DTMA	---	L
6NtNatronom	1	---	MS---	SA	DEDD	VARTVAQGRAS		GVLLRNAQK	LOKVF	TVS	VGFL	DTMA	---	L
6NtHalobact	1	---	MSGSGDL	SPGTAR	IDSGRES		GVLLS	AQDLQ	OKVF	TVS	VGFL	DTMA	---	L
B.subtilis	1	---	---	---	---	---	---	---	---	---	---	---	---	A
E.coliK12	1	---	---	---	---	---	---	---	---	---	---	---	---	A
6CtHalobact	59	ATLR	DEET	---	RIPO	OP	ANSWE	ET	ED	VEAL	FMVK	SVVIA	MAV	PLV
6CtNatronom	58	G	MLHEC	---	---	---	---	---	---	---	---	---	---	---
6CtHalofer	57	RVV	DE	---	---	---	---	---	---	---	---	---	---	---
6NtHalofer	50	RLY	---	---	---	---	---	---	---	---	---	---	---	---
6NtNatronom	50	RLY	---	---	---	---	---	---	---	---	---	---	---	---
6NtHalobact	54	RLF	---	---	---	---	---	---	---	---	---	---	---	---
B.subtilis	34	F	LF	---	---	---	---	---	---	---	---	---	---	---
E.coliK12	43	H	VSAFL	---	---	---	---	---	---	---	---	---	---	---
6CtHalobact	116	S	IS	---	---	---	---	---	---	---	---	---	---	---
6CtNatronom	114	S	IS	---	---	---	---	---	---	---	---	---	---	---
6CtHalofer	114	A	IS	---	---	---	---	---	---	---	---	---	---	---
6NtHalofer	110	A	IS	---	---	---	---	---	---	---	---	---	---	---
6NtNatronom	104	S	IS	---	---	---	---	---	---	---	---	---	---	---
6NtHalobact	107	P	LO	---	---	---	---	---	---	---	---	---	---	---
B.subtilis	87	-F	VAP	---	---	---	---	---	---	---	---	---	---	---
E.coliK12	94	F	VAP	---	---	---	---	---	---	---	---	---	---	---
6CtHalobact	170	V	IS	---	---	---	---	---	---	---	---	---	---	---
6CtNatronom	171	V	IS	---	---	---	---	---	---	---	---	---	---	---
6CtHalofer	168	I	IS	---	---	---	---	---	---	---	---	---	---	---
6NtHalofer	167	TP	IS	---	---	---	---	---	---	---	---	---	---	---
6NtNatronom	160	QP	IS	---	---	---	---	---	---	---	---	---	---	---
6NtHalobact	164	SPK	IS	---	---	---	---	---	---	---	---	---	---	---
B.subtilis	142	ETM	IS	---	---	---	---	---	---	---	---	---	---	---
E.coliK12	146	Q	IS	---	---	---	---	---	---	---	---	---	---	---
6CtHalobact	230	AT	DS	---	---	---	---	---	---	---	---	---	---	---
6CtNatronom	231	TE	DS	---	---	---	---	---	---	---	---	---	---	---
6CtHalofer	227	TE	DS	---	---	---	---	---	---	---	---	---	---	---
6NtHalofer	227	TE	DS	---	---	---	---	---	---	---	---	---	---	---
6NtNatronom	220	TE	DS	---	---	---	---	---	---	---	---	---	---	---
6NtHalobact	224	TE	DS	---	---	---	---	---	---	---	---	---	---	---
B.subtilis	202	TE	DS	---	---	---	---	---	---	---	---	---	---	---
E.coliK12	206	TE	DS	---	---	---	---	---	---	---	---	---	---	---

Figure 1. *H. volcanii* N- and C-terminal TatCt domains are distinct. **A.** Predicted topology of *H. volcanii* TatCt including duplicated TatC domains connected via two non-conserved TMSs (boxed). **B.** Alignment of the N-terminal and C-terminal 6 TMS of haloarchaeal TatCt homologs. The predicted N-terminal (N) and C-terminal (C) TMS were aligned using ClustalW and shaded using Boxshade. The cytoplasmic (pink), transmembrane (blue) and periplasmic (green) domains of *H. volcanii* N-6, as predicted by TMHMM, are indicated. Red rectangle shows conserved domain in N-terminal half of TatCt.

A



B

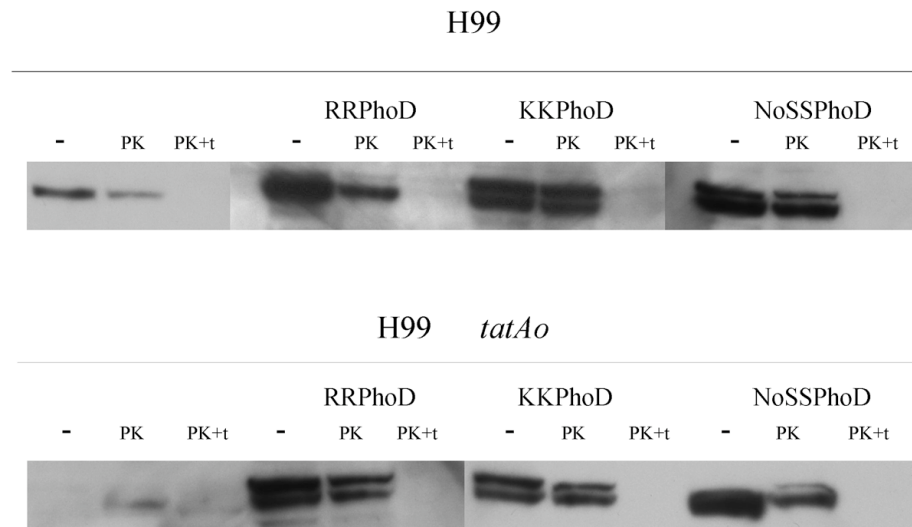


Figure 2. PhoD is exposed to the extracellular environment and its secretion is TatAo-dependent. A. *Halobacterium* NRC-1 PhoD translocation in *H. volcanii*. The *phoD* gene was cloned in an expression vector in order to be used as a secretion reporter. The following PhoD constructs were generated: RRPhoD, wt signal peptide; KKPhoD, signal peptide with the essential twin arginines (RR) mutated to lysines (KK); NossPhoD, no signal peptide. These three constructs were transformed in H99 (wt) or KD5 (TatAo mutant) and streaked in plates containing XP. **B.** Wild type (H99, upper panel) or *tatAo* knock-out (lower panel) *H. volcanii* cells containing pMLH3 with the corresponding PhoD construct under the control of the *fdx* promoter, were grown to early log phase and resuspended in PBS buffer containing 2M NaCl. Cells were treated for 30 min. at 37 C with proteinase K (PK) or proteinase K + triton (PK+t) or incubated with no additives (-) and the reaction was stopped with 1mM PMSF and incubation at 4 C for 30 min.. The samples were subjected to western blot with anti-strep antibody.