Conserved signal peptide recognition systems across the prokaryotic domains.

Sarah J. Coulthurst, Alice Dawson, William N. Hunter and Frank Sargent.

College of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, United Kingdom.

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY TABLE S1. Strains and plasmids used in this study.

Name	Description	Reference / Source
Bacterial strains		
E. coli MC4100	Parental *E. coli* strain (F^-,Δ*lacU169*, *araD139*, *rpsL150*, *relA1*, *ptsF*, *rbs*, *flbB530*1)	(*1*)
E. coli DADE	MC4100 Δ*tatABCD*, Δ*tatE*	(*2*)
E. coli K-38	HfrC, *phoA4*, *pit-10*, *tonA22*, *ompF627*, *relA1*	(*3*)
E. coli MCDSSAC	MC4100 *amiA*$^-$, *amiC*$^-$	(*4*)
E. coli MC4100-A	Arabinose-resistant derivative of MC4100	(*5*)
E. coli DADE-A	Arabinose-resistant derivative of DADE	(*5*)
E. coli BTH101	Host strain for Bacterial Two Hybrid experiments, *cya*$^-$ (F^-, *cya-99*, *araD139*, *galE15*, *galK16*, *rpsL1*, *hsdR2*, *mcrA1*, *mcrB1*)	(*6*)
E. coli BL21(DE3)	Protein overproduction strain (F^-, *omp*T, *hsdS* (rB-, mB-), *dcm*, *gal*, λ(DE3))	(*7*)
E. coli BL21(DE3) Δ*tat*	Δ*tatABC* derivative of BL21(DE3)	B. Ize & T. Palmer; using method of (*8*)
E. coli BL21(DE3) pLysS	As BL21(DE3), also producing T7 lysozyme	Promega Inc.
Plasmids		
pGH-SG1080110	Synthetic *ttrBACD* operon in pUC57 (Amp^R)	This study
pUNIPROM	pT7.5-derived vector allowing constitutive expression under the control of the *E. coli tat* promoter or inducible expression from the upstream T7 promoter (Amp^R)	(*9*)
pUNI-TtrBACD	*ttrBACD* from synthetic operon in pUNIPROM	This study
pUNI-TtrBAC	*ttrBAC* from synthetic operon in pUNIPROM	This study
pUNI-TtrACD	*ttrACD* from synthetic operon in pUNIPROM	This study
pUNI-TtrBC	*ttrBC* from synthetic operon in pUNIPROM	This study
pUNI-TtrAD	*ttrAD* from synthetic operon in pUNIPROM	This study
pUNI-ssTtrA$_{Af}$-AmiA	*ssTtrA*$_{Af}$ (synthetic sequence) fused to *amiA*Δ2-33 in pUNIPROM; to express ssTtrA$_{Af}$ -AmiA (mature AmiA from pTorASP-AmiA, Ize *et al*., 2009)	This study
pSUPROM	pSU40-derived vector allowing constitutive expression under the control of the *E. coli tat* promoter (Kan^R)	(*9*)
pSU-ssTtrA$_{Af}$-Bla	*ssTtrA*$_{Af}$ (synthetic sequence) fused to *bla* (lacking signal peptide) in pSUPROM; to express ssTtrA$_{Af}$ -Bla (derived from pSU-TorAss-Bla, G. Buchanan & T. Palmer,	This study

	unpublished)	
pBAD33	Arabinose-inducible expression vector (Cm^R)	(*10*)
pBAD24	Arabinose-inducible expression vector (Kan^R)	(*10*)
pBAD-TtrD	The *ttrD* gene (synthetic sequence) in pBAD33	This study
pBAD-ssTtrA$_{Af}$-GFP-SsrA	*ssTtrA$_{Af}$* (synthetic sequence) fused to *gfp-ssrA* in pBAD24; to express ssTtrA$_{Af}$ -GFP-SsrA (based on pTGS, DeLisa *et al.*, 2007, and its derivative pBAD24-TorAss-GFP-SsrA, B. Ize & T. Palmer, unpublished)	This study
pETM-11	Protein overproduction vector for the generation of N-terminal His_6-TEV cleavage site-protein fusions (Kan^R)	(*11*)
pETM-TtrD	Synthetic *ttrD* (AF0160) gene in pETM-11	This study
pUT18	Amp^R (ori pMB1) vector for generation of N-terminal T18 fusions for the Bacterial Two Hybrid system (Amp^R)	(*6*)
pT25	Cm^R (p15A) vector for generation of C-terminal T25 fusions for the Bacterial Two Hybrid system (Cm^R)	(*12*)
pUT18-ssTtrA$_{Af}$	ssTtrA$_{Af}$ (natural sequence, codons 1-36) in pUT18	This study
pUT18-ssTtrA$_{Af}$_RA	ssTtrA$_{Af}$ (natural sequence, Region A: codons 1-26) in pUT18	This study
pUT18-ssTtrA$_{Af}$_RB	ssTtrA$_{Af}$ (natural sequence, Region B: codons 10-27) in pUT18	This study
pUT18-ssTtrA$_{Af}$_RC	ssTtrA$_{Af}$ (natural sequence, Region C: codons 11-36) in pUT18	This study
pUT18-ssTtrA$_{Af}$_RD	ssTtrA$_{Af}$ (natural sequence, Region D: codons 7-17) in pUT18	This study
pUT18-ssTtrA$_{Af}$_R6Q	ssTtrA$_{Af}$ (natural sequence) with R6Q substitution in pUT18	This study
pUT18-ssTtrA$_{Af}$_V15Q	ssTtrA$_{Af}$ (natural sequence) with V15Q substitution in pUT18	This study
pUT18-ssTtrA$_{Af}$_S17Q	ssTtrA$_{Af}$ (natural sequence) with S17Q substitution in pUT18	This study
pUT18-ssTtrA$_{Af}$_V20Q	ssTtrA$_{Af}$ (natural sequence) with V20Q substitution in pUT18	This study
pUT18-ssTtrA$_{Af}$_L22Q	ssTtrA$_{Af}$ (natural sequence) with L22Q substitution in pUT18	This study
pUT18-ssTtrA$_{Af}$_G24Q	ssTtrA$_{Af}$ (natural sequence) with G24Q substitution in pUT18	This study
pT25-TtrD	The *ttrD* (AF0160) gene (natural sequence) in pT25	This study
pQE80-GST	Glutathione S transferase coding sequence in pQE80-L (to express His_6-GST)	This study
pQE80-ssTtrA$_{Af}$-GST	*ssTtrA$_{Af}$* (synthetic sequence) in QE80_GST; for production of ssTtrA$_{Af}$ -GST	This study
pQE80-ssTtrA$_{Af}$-GST_TtrD1	The *ttrD* (AF0160) gene (synthetic sequence) in pQE80-ssTtrA$_{Af}$-GST; for co-production of TtrD and ssTtrA$_{Af}$ -GST	This study
pQE80-GST-TtrD1	The *ttrD* (AF0160) gene (synthetic sequence) in pQE80-GST; for co-production of TtrD and His_6-GST	This study

SUPPLEMENTARY TABLE S2. Oligonucleotides used in this study.

Name	Sequence (5'-3')	Details
SC1056	TATAGGATCCATGTCCCCTATACTAGGTTATTGGA	Forward primer to amplify GST from pGEX-5X-1 (BamHI)
SC1057	TATGAAGCTTCAGATCCGATTTTGGAGG	Reverse primer to amplify GST from pGEX-5X-1 (HindIII)
SC1098	TATAGGATCCATGCAACTCTCCCGACGTG	Forward primer to clone ssTtrA_Af (from synthetic operon) into pUNIPROM, pSUPROM and derived vectors (BamHI)
SC1077	TATATCTAGACCGTGGCTTTACCAATCTATCC	Reverse primer to clone ssTtrA_Af (from synthetic operon) into pUNIPROM. pSUPROM and derived vectors (XbaI)
SC1113	TATAGAATTCAGGAGAAATTATCTATGCAACTCTCCCGACG	Forward primer to clone ssTtrA_Af (from synthetic operon) into pQE80-GST (EcoRI)
SC1114	TATAGGATCCCCGTGGCTTTACCAATCTATCC	Reverse primer to clone ssTtrA_Af (from synthetic operon) into pQE80-GST (BamHI)
SC1111	TATAGAATTCAGGAGAAATTACAAATGACAATCGGTCGGG	Forward primer to clone TtrD (from synthetic operon) into pQE80-GST and related vectors (EcoRI)
SC1112	TATAGAATTCTCAGTTGTTATCGGCCCCTTC	Reverse primer to clone TtrD (from synthetic operon) into pQE80-GST and related vectors (EcoRI)
SC1163	TATATCATGACAATCGGTCGGGCTAAAG	Forward primer to clone TtrD_Af (from synthetic operon) into pETM-11 (RcaI)*
SC1112	TATAGAATTCTCAGTTGTTATCGGCCCCTTC	Reverse primer to clone TtrD (from synthetic operon) into pETM-11 (EcoRI)
SC1045	TATAGGATCCAGTGACCATAGGGAGGGCAAAG	Forward primer to clone TtrD_Af into pT25 (BamHI)
SC1046	TATAGGTACCTCAATTGTTGTCAGCTCCCTC	Reverse primer to clone TtrD_Af into pT25 (KpnI)
SC1049	TATAAAGCTTGTAAAGAGGAGGATCCATGCAGCTTAGCAGGAGGG	Forward primer to clone ssTtrA_Af into pUT18 (HindIII)
SC1050	TATAGAATTCGATCTCGGCTTAACCAGCCTG	Reverse primer to clone ssTtrA_Af into pUT18 (EcoRI)
SC1076	TATAGAATTCACCATGCAACTCTCCCGACGTG	Forward primer to clone ssTtrA (Af synthetic) into pBADssTorA-GFP-SsrA (EcoRI)
SC1077	TATATCTAGACCGTGGCTTTACCAATCTATCC	Reverse primer to clone ssTtrA (Af synthetic) into pBADssTorA-GFP-SsrA (XbaI)
SC1078	TATAGAGCTCAGGAGGACAAATGACAATCGG	Forward primer to clone TtrD (Af synthetic) into pBAD33 (SacI)
SC1079	TATAGGTACCTCAGTTGTTATCGGCCCCTTC	Reverse primer to clone TtrD (Af synthetic) into pBAD33 (KpnI)
SC1202	TATAGAATTCGACGAGTAACCTGCGAGGAAAACTG	Reverse primer to clone ssTtrA_Af Region A into pUT18 (EcoRI)
SC1183	TATAAAGCTTGTAAAGAGGAGGATCCATGAAGGGTCTCGTGGCTGTTGG	Forward primer to clone ssTtrA_Af Region B into pUT18 (HindIII)
SC1184	TATAGAATTCGATTCCGAGTAACCTGCGAGG	Reverse primer to clone ssTtrA_Af Region B into pUT18 (EcoRI)
SC1185	TATAAAGCTTGTAAAGAGGAGGATCCATGGGTCTCGTGGCTGTTGGC	Forward primer to clone ssTtrA_Af Region C into pUT18 (HindIII)
SC1203	TATAAAGCTTGTAAAGAGGAGGATCCATGGATTTTATTAAGGGTCTCGTGGCTG	Forward primer to clone ssTtrA_Af Region D into pUT18 (HindIII)
SC1204	TATAGAATTCGACGAGCCAACAGCCACGAG	Reverse primer to clone ssTtrA_Af Region D into pUT18 (EcoRI)
SC1186	CCATGCAGCTTAGCAGGCAGGATTTTATTAAGGGTCTCG	Forward primer for Quikchange ssTtrA_Af R6Q
SC1187	CGAGACCCTTAATAAAATCCTGCCTGCTAAGCTGCATGG	Reverse primer for Quikchange ssTtrA_Af R6Q
SC1188	GGGTCTCGTGGCTCAGGGCTCGGCATCAG	Forward primer for Quikchange ssTtrA_Af V15Q
SC1189	CTGATGCCGAGCCCTGAGCCACGAGACCC	Reverse primer for Quikchange ssTtrA_Af V15Q
SC1190	TCTCGTGGCTGTTGGCCAGGCATCAGTTTTCCTCGC	Forward primer for Quikchange ssTtrA_Af S17Q
SC1191	GCGAGGAAAACTGATGCCTGGCCAACAGCCACGAGA	Reverse primer for Quikchange ssTtrA_Af S17Q
SC1192	CTGTTGGCTCGGCATCACAGTTCCTCGCAGGTTACTCG	Forward primer for Quikchange ssTtrA_Af V20Q
SC1193	CGAGTAACCTGCGAGGAACTGTGATGCCGAGCCAACAG	Reverse primer for Quikchange ssTtrA_Af V20Q
SC1194	GCTCGGCATCAGTTTTCCAGGCAGGTTACTCGG	Forward primer for Quikchange ssTtrA_Af L22Q
SC1195	CCGAGTAACCTGCCTGGAAAACTGATGCCGAGC	Reverse primer for Quikchange ssTtrA_Af L22Q
SC1196	CGGCATCAGTTTTCCTCGCACAGTACTCGGAAACAGTTGAC	Forward primer for Quikchange ssTtrA_Af G24Q
SC1197	GTCAACTGTTTCCGAGTACTGTGCGAGGAAAACTGATGCCG	Reverse primer for Quikchange ssTtrA_Af G24Q

Restriction sites used for cloning are underlined.
* TtrD was amplified as an RcaI-EcoRI fragment and then cloned into pETM-11 cut NcoI-EcoRI
All primers were obtained from Sigma-Genosys, UK.

SUPPLEMENTARY TABLE S3. Structure statistics for TtrD.

PDB Code		2xol	2yjm	2y6y
Crystal Statistics				
Resolution range	(Å)	1.35-64.69	1.84 – 35.01	2.20 – 55.93
Space Group		$P2_1$	$C2$	$P3_221$
Unit cell dimensions:				
a, b, c	(Å)	62.76 52.54 64.95	91.16 37.93 49.12	87.06 | 87.06 | 83.39
β	(°)	91.26	94.16	-
Scaling Statistics				
Data collection location		ESRF ID-29	In – house	ESRF ID-29
No. Measurements		622523 (80665)	27711 (943)	188108 (12266)
Unique reflections		89700 (12764)	13256 (1796)	18960 (2751)
Redundancy		6.9 6.3	2.1 2.1	9.9 | 4.5
Completeness	(%)	96.9 94.9	90.6 84.7	100 | 100
< *I* / σ(*I*) >		17.6 4.2	9.9 4.3	14.1 | 3.5
R_{merge}	(%)	5.4 49.2	5.6 15.9	11.2 | 38.9
Refinement Statistics				
R_{work} (%) / No of reflections		14.85 / 85158	21.41 /12576	21.7 / 17981
R_{free} (%) / No of reflections		17.54 / 4523	24.51 /680	23.8 / 955
R.M.S.D bond lengths	(Å)	0.011	0.009	0.011
R.M.S.D bond angles	(°)	1.27	0.986	1.13
DPI [1]		0.046	0.155	0.165
Ramachandran:				
Favoured region	(%)	98.5	98.8	98.11
Outliers	(%)	0.0	0.0	0.63
Model Statistics				
Protein residues:				
No. in subunits	A/B	168 / 172	165 / -	163 / -
Overall	*B*-factor	20.42 | 20.88	41.72 / -	40.02 / -
Side chain	A / B	22.41 22.02	43.30 / -	42.33 / -
Main chain	A / B	18.20 18.40	40.05 /-	38.37 / -
Additional groups:				
Solvent	No/Av. *B*-factor	393 / 40.7	88 / 33.13	70 / 40.69
Other entities	Identity / *B*-factor	Ethane diol (EDO) / 30.3	CHES (NHE) / 42.7	Cl^- / 31.0

[1] DPI, The diffraction-component precision index (*13*).

SUPPLEMENTARY REFERENCES

1. Casadaban, M. J., and Cohen, S. N. (1979) Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences, *Proc Natl Acad Sci U S A 76*, 4530-4533.
2. Wexler, M., Sargent, F., Jack, R. L., Stanley, N. R., Bogsch, E. G., Robinson, C., Berks, B. C., and Palmer, T. (2000) TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export, *J Biol Chem 275*, 16717-16722.
3. Lyons, L. B., and Zinder, N. D. (1972) Genetic map of filamentous bacteriophage F1, *Virology 49*, 45-&.
4. Ize, B., Stanley, N. R., Buchanan, G., and Palmer, T. (2003) Role of the *Escherichia coli* Tat pathway in outer membrane integrity, *Molecular microbiology 48*, 1183-1193.
5. Ize, B., Gerard, F., Zhang, M., Chanal, A., Voulhoux, R., Palmer, T., Filloux, A., and Wu, L. F. (2002) In vivo dissection of the Tat translocation pathway in *Escherichia coli*, *J Mol Biol 317*, 327-335.
6. Karimova, G., Ullmann, A., and Ladant, D. (2001) Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system, *J Mol Microbiol Biotechnol 3*, 73-82.
7. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J Mol Biol 189*, 113-130.
8. Orriss, G. L., Tarry, M. J., Ize, B., Sargent, F., Lea, S. M., Palmer, T., and Berks, B. C. (2007) TatBC, TatB, and TatC form structurally autonomous units within the twin arginine protein transport system of *Escherichia coli*, *FEBS Lett 581*, 4091-4097.
9. Jack, R. L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins, *EMBO J 23*, 3962-3972.
10. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter, *J Bacteriol 177*, 4121-4130.
11. Dümmler, A., Lawrence, A. M., and de Marco, A. (2005) Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors, *Microb Cell Fact 4*, 34.
12. Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway, *Proc Natl Acad Sci U S A 95*, 5752-5756.
13. Cruickshank, D. W. (1999) Remarks about protein structure precision, *Acta Cryst D, Biol Cryst 55*, 583-601.