

Table 2. Exported proteins observed in the cell wall-associated fraction of the M145 strain only identified by either two-dimensional gel electrophoresis and/or multidimensional protein identification technology

Sco number	Method of observation	Tat substrate prediction	Product	Putative signal peptide
Twin-arginine signal peptides				
SCO0432	2D	TatP	Probable secreted peptidase	<u>MRR</u> IRLLAAISAGLALAAGAVAPVPALAS
SCO0474	MudPIT		Putative lipoprotein	MEPTVRTHPT <u>RR</u> LPTALVLTAALLATGCSEQSDDGREPAA
SCO1172	2D	TatFind	Putative amidase (putative secreted protein)	MERERTAPVPT <u>RR</u> LLKGAALATVPYTLLSGTRAAQ
SCO1196	2D	TatP and TatFind	Putative secreted protein	MALT <u>RR</u> RALTTLGAALTGAVALPAGTALAS
SCO1356	MudPIT	TatFind	Putative iron sulfur protein (putative secreted protein)	MTSAPFHPAAGPA <u>RR</u> TVVAAAGAAGLTAVLAACSDSDDGASGD
SCO1432	2D and MudPIT	TatP and TatFind	Putative membrane protein.	MQRRTRPTRTPGGLPLMSRVEQPS <u>RR</u> ALLAAVAAAAVTGGAIPTARAAS
SCO1565	2D and MudPIT	TatP and TatFind	Putative glycerophosphoryl diester phosphodiesterase	MGTQESDERAGGGT <u>RR</u> ALLGAAVLGAGGAVLGLPGTARAA
SCO1590	2D	TatP	Putative secreted protein (transglycosidase)	MGGV <u>SR</u> RAFTVAALSAFTLVPEASAA
SCO1639	2D and MudPIT	TatFind	Putative secreted peptidyl-prolyl cis-trans isomerase protein	<u>MRR</u> RSLLIAVPTGLVTLAACGSDSDSGSSNASESPPDASATSAA

SCO1824	MudPIT		Secreted subtilisin-like protease	MTAPLSRH RR ALAIPAGLAVAASLAF LPGTPAAATPAAEAA
SCO1906	2D		Putative secreted protein (phoX phosphatase)	MSL T RR DFAGRSALTGAGVVLGASV GALATAPNALAS
SCO1948	2D		Putative zinc-binding carboxypeptidase	M RR RARSILAVGALLIGGASFAP IAQAQ
SCO2068	2D and MudPIT	TatP and TatFind	Putative secreted alkaline phosphatase	MTSRHRASENSRTPS RR TVVKAAAAGAVLAA PLAAALPAGAADAAP
SCO2226	2D		Putative bi-functional protein (secreted alpha-amylase/dextrinase)	MPATRRTARV RR VAAVTVTALAAALLPPLAARAD
SCO2286*	2D		Putative alkaline phosphatase	MTPANHQAPTSAPSPAPSQSSHAPELRAAARSLG RRR FLVTGAAAALAFAVNLPAAGTASAA
SCO2383	MudPIT		Putative secreted protein (PF00652 Ricin B lectin)	MRLRRPRRPSDGPPGRSGA RR LRVASAPLIPLTVIGAVVAGLIGPEPPSDG PLGTSAVAD
SCO2637	MudPIT		Putative serine protease (pfam00082, Peptidase_S8, Subtilase family)	MTNSPQREPIPGA RR AARLALATGLAAALAAVGPVPVALAA
SCO2758	2D	TatP and TatFind	Beta-N-acetylglucosaminidase (putative secreted protein)	MHHSSTAGTGSTAQPS RR SVLTATAVVTAALAAGGGTAYAD
SCO2780†	2D and MudPIT	TatFind	Putative secreted protein (cd01146, FhuD, Fe ³⁺ -siderophore binding domain FhuD)	MSHASATHPT RR GILAAGGALGLGAVLAAACGDGDGKSDGA

SCO2786	2D		Beta-N-acetylhexosaminidase (Transglycosidase)	MRPH <u>RR</u> HHRTTPRITRLLGSLLLVAAVGAMTTGAAPVRKAAAE
SCO2821	2D		Putative secreted pectate lyase	<u>MRRP</u> VALRLQAALGTLALAAAAGVVLTMPAAAAA
SCO3456	MudPIT		Putative secreted protein (ABC transporter substrate binding protein)	MRHS <u>GRR</u> VKRATVLAAGAAVAGLLTAACSSPDTTGV
SCO3484	2D	TatP and TatFind	Putative secreted sugar-binding protein	MTKPVVPSGVS <u>RRG</u> FLGSLGVAGVLLAACSGGGNSSQGS
SCO3790	2D	#TatP and TatFind	Conserved hypothetical protein of PhoX type	MRKLLPLIGTPSGSHPGGRSAMTCRFRCGDACFHEVPNTSSNEYVGDVIAGALS <u>RRS</u> MMRAAAVVTVAAAGAGAVGVAGAPSAQAA
SCO4142 [†]	2D		Phosphate-binding protein precursor	MN <u>RR</u> ALALGALAVSGALALTACGSDDTGGNSGSDSSSAAAN
SCO4152	2D and MudPIT		Putative secreted 5'-nucleotidase	MPVTAQPHQPRRR <u>RR</u> TSRLLVVAATVVVTAGVLAALPASASAG
SCO4672	2D and MudPIT	TatP and TatFind	Putative secreted protein	MHTS <u>RRG</u> VLAFAASAATVPFGGAAAASPARATATAASAVDPAAASAA
SCO4884	2D and MudPIT		Putative lipoprotein	<u>MRR</u> TSRLIRVAVGVASLALAATACGGTSGESGD
SCO4885	2D and MudPIT		Putative lipoprotein	<u>MRR</u> ISRITVAGAATASLALALAACGGTSTDSGSE
SCO5461	2D		Putative secreted protein (ADP ribosylation)	MITSLR <u>RR</u> TAAAVLSLSAVLATTAAATAPGAAPAPSAAPAKAA
SCO6009	2D		Solute-binding protein	MNTRM <u>RR</u> AAVAVATTAMAVSLAACGSAK
SCO6052	2D	TatP and TatFind	Putative membrane protein	MSTSRNAAT <u>RR</u> QVLARTGALGAGIAFTGALSELFAGTAAAQ

SCO6198 [†]	MudPIT	TatP and TatFind	Putative secreted protein	MDTTYWS <u>RRR</u> VLTVLGAATAATSIPLAAPSRALAA
SCO6272*	2D	TatP and TatFind	Putative secreted FAD-binding protein	MTEVS <u>RR</u> KLMKGA AVSGGALALPALGAPPATAAPAAGPEDLPGPAAAAAG
SCO6281	2D and MudPIT	TatP and TatFind	Putative FAD-binding protein	MSAIS <u>RRR</u> FIRHGAVAGGAAVALPALGGWAGEAFAA
SCO6418	2D and MudPIT	TatP and TatFind	Hypothetical protein (pectin lyase like)	MD <u>RR</u> ALMLATGGLLGAAGAAQLTAAPSAVAA
SCO6457	2D	TatP and TatFind	Putative beta-galactosidase (Galactose mutarotase-like)	MPHSPVSPAESPAPQPGRPRPVVS <u>RRR</u> LLEGGA AVLGALALSASPLTAQAA VRRAAAD
SCO6580	2D	#TatP and TatFind	Hypothetical protein (Xylose isomerase-like)	MTPFTDSSRTDAGTDPADGPGESLRRALGVN <u>RRR</u> RFLSTCTAVAAGAVAAPVFGA SPALAH
SCO6594	MudPIT	TatP and TatFind	Putative secreted protein	MPSSPTPSPTPSGAPEPSGV <u>RRR</u> SLLA AAAA APVLASLAGAGTAAAD
SCO6644	MudPIT		Putative solute-binding lipoprotein	MRS <u>RR</u> PRL LAPFL VPLMAGCFASGGGDDSASGD
SCO6691	2D	TatP and TatFind	Putative phospholipase C	MAEVN <u>RRR</u> RFLQLAGATTAFSALSASIDRAAAL
SCO7399	2D		Possible binding-protein-dependent transport lipoprotein (FhuD).	<u>MR</u> RLLLTAAATTAALTLAACGTTEPAAD
SCO7631	2D	TatP	Putative secreted protein	MSTEEQSSGRRHPD <u>RR</u> VLLRAAVAVPAAGAAVAGTAAVPAQAADAA
Non twin-arginine signal peptides				
SCO0930	MudPIT		Putative lipoprotein	MKTSWRSASLVAGAAALLALTACGQDGGGPTGSQNVGATAA
SCO1396	MudPIT		Putative D-alanyl-D-alanine dipeptidase	MTRLSAALRGLATTFAALLAVTAVPAASAPATAQAR

			(putative secreted protein)	
SCO1968 [†]	2D and MudPIT		Putative secreted hydrolase	MHVRAVAVTTTALLGVALTAPLSHARAD
SCO3967	MudPIT		Conserved hypothetical membrane protein	MKASRIAAVGAVSGIAVLALSAPAFAHVSVQPEGAAAK
SCO4010	MudPIT		Putative secreted protein (subtilisin inhibitor)	MFQVTRTPAARLLGAAALSAAALASVCAAPSAAVAGAG
SCO5074	2D and MudPIT		Putative dehydratase	MTSSLHHAIRLTTASAIALGGLVTLG TSAHAA
SCO6738	2D		Putative carboxypeptidase (putative secreted protein)	MNRLRTLATTAALAASLLVPSLLTAPAHAD

Proteins are listed by SCO number and the signal peptide sequences are given in the right column. Twin-arginine dipeptides, where present, are shown in bold underline. Where multiple arginine dipeptides are present, only the most plausible one is marked. Method of observation indicates which technique was used to identify a particular protein (2D is two-dimensional gel electrophoresis; MudPIT is multidimensional protein identification technology). The column annotated 'Tat substrate prediction' indicates whether a particular protein has been identified as a putative Tat substrate by either of the two prediction programs TatFind 1.4 or TatP1.0 www.signalfind.org and www.cbs.dtu.dk/services/TatP-1.0/, respectively.

There is some ambiguity about the assigned start codons for SCO3790 and SCO6580. For each of these proteins there is an alternative GTG start codon within the predicted signal peptide coding sequences. These are shown as valine residues that are highlighted in bold underline. Signal peptides initiating at these alternative start codons were tested in the agarase export assay (Fig 3) and in each case shown to mediate higher levels of agarase export than the full length signal peptides.

*These signal peptides have been shown to mediate Tat-dependent export when fused to the mature sequence of *S. lividans* xylanase C (Li H, Jacques PE, Ghinet MG, Brzezinski R, Morosoli R (2005) *Microbiology* 151: 2189-2198).

† These proteins have also been identified in the extracellular fraction of *S. coelicolor* grown in liquid medium (Kim DW, Chater K, Lee KJ, Hesketh, A (2005) *J Bacteriol* 187: 2957-2966).

#These proteins are only detected by TatP and TatFind when truncated *in silico* at the N-terminus to the underlined residue.