

**Table 1. Proteins that are uniquely present in most epsilon proteobacteria
(Campylobacterales)**

Gupta, R. S. (2006) *BMC Genomics* 7, 167.

Wolinella Genome ID No.	Accession Number	Length	Possible/Predicted Function	Comments
WS0030	NP_906303.1	68 aa	Probable periplasmic protein, tat-domain	All significant hits from ϵ -Proteobacteria
WS0086	NP_906354	181 aa	Putative helicase, gnl CDD 14084, COG4951	All significant hits from ϵ -Proteobacteria
WS0133	NP_906397	397 aa	Putative integral membrane protein	All significant hits from ϵ -Proteobacteria
WS0134	NP_906398.1	214 aa	Conserved hypothetical protein	All significant hits from ϵ -Proteobacteria
WS0154	NP_906417.1	336 aa	Probable membrane protein	All significant hits from ϵ -Proteobacteria
WS0159	NP_906422	203 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS0169	NP_906432	92 aa	Possible membrane protein (corresponds to Cj0692c and HP0748)	All significant hits from ϵ -Proteobacteria
WS0172	NP_906435	675 aa	Putative membrane protein, similar to HP0358	All significant hits from ϵ -Proteobacteria
WS0180	NP_906442	74 aa	Related to Cbb3-type cytochrome oxidase, subunit 3, gnl CDD 13876, COG4736	All significant hits from ϵ -Proteobacteria
WS0184	NP_906445.1	205 aa	Probable membrane protein	All significant hits from ϵ -Proteobacteria
WS0185	NP_906446	163 aa	Related to FixH protein, gnl CDD 23975	All significant hits from ϵ -Proteobacteria
WS0216	NP_906474	330 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteo; E value for <i>C. upsaliensis</i> is higher than the threshold.
WS0260	NP_906515.1	142 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria; E value for <i>C. lari</i> is above the threshold.
WS0266	NP_906520	271 aa	Conserved protein, <i>H. pylori</i> homolog may be related to the plasminogen binding protein pgbA.	All significant hits from ϵ -Proteobacteria; Jonsson et al. (2004)
WS0316	NP_906567	163 aa	Conserved protein related to the RDD family; gnl CDD 25144, pfam06271	Besides ϵ -Proteo, three other hits were below the threshold value; Large jump in E value from last ϵ -Proteo (6E-15) to the first of these hits (.003).
WS0447	NP_906689	328 aa	Putative membrane protein, corresponds to antigen P44Hh9 of <i>H. hepaticus</i> .	All significant hits from ϵ -Proteobacteria
WS0448	NP_906690	276 aa	Probable periplasmic protein	All significant hits from ϵ -Proteobacteria
WS0476	NP_906716	77 aa	NuoE, Putative NADH Oxidoreductase I	All significant hits from ϵ -Proteobacteria
WS0480	NP_906720	428 aa	Putative NADH Oxidoreductase I	All significant hits from ϵ -Proteobacteria
WS0490	NP_906728	778 aa	Flagellar functional protein, Pfla	All significant hits from ϵ -Proteobacteria
WS0520	NP_906757.1	247 aa	TonB domain protein	All significant hits from ϵ -Proteobacteria
WS0563	NP_906797	164 aa	Putative integral membrane protein; identified by similarity to PIR:B71953	All significant hits from ϵ -Proteobacteria
WS0575	NP_906809	217 aa	Putative lipoprotein, The <i>C. lari</i> homolog is a secreted protein involved in flagellar motility.	All significant hits from ϵ -Proteobacteria.
WS0604	NP_906835	390 aa	Probable periplasmic protein	All significant hits from ϵ -Proteobacteria
WS0802	NP_907015	333 aa	Probable lipoprotein; identified as plasminogen binding protein pgpB.	All significant hits from ϵ -Proteobacteria; Jonsson et al. (2004)
WS0865	NP_907074.1	126 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria; missing in <i>C. jejuni</i> .
WS1039	NP_907239	156 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1040	NP_907240.1	236 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1235	NP_907415	412 aa	Putative periplasmic protein; COG5659	All significant hits from ϵ -Proteobacteria; not found in <i>H. hepaticus</i> .
WS1244	NP_907424	167 aa	Putative lipoprotein	All significant hits from ϵ -Proteobacteria
WS1329	NP_907504	246 aa	Putative periplasmic protein	All significant hits from ϵ -Proteobacteria; absent in <i>H. pylori</i> .
WS1344	NP_907515.1	123 aa	Putative periplasmic protein	All significant hits from ϵ -Proteobacteria
WS1349	NP_907520.1	110 aa	Probable membrane protein	All significant hits from ϵ -Proteobacteria
WS1485	NP_907639.1	89 aa	Probable integral membrane protein	All significant hits from ϵ -Proteobacteria
WS1495	NP_907647	87 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteo; The E value for <i>C. lari</i> (next best hit) above the threshold.
WS1496	NP_907648	208 aa	Probable periplasmic protein	All significant hits from ϵ -Proteobacteria
WS1640	NP_907771	117 aa	Probable integral membrane protein	All significant hits from ϵ -Proteobacteria; absent in <i>H. pylori</i> .
WS1730	NP_907855	183 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1755	NP_907877	168 aa	Probable lipoprotein	All significant hits from ϵ -Proteobacteria
WS1771	NP_907893	183 aa	Putative membrane protein	All significant hits from ϵ -Proteobacteria; absent

				in <i>H. pylori</i> .
WS1773	NP_907895	351 aa	Putative membrane protein	All significant hits from ϵ -Proteobacteria
WS1777	NP_907899.1	80 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1814	NP_907930.1	85 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1874	NP_907984	352 aa	HolA, DNA polymerase III, delta subunit; gnl CDD 11180, COG1466	All significant hits except one from ϵ -Proteo; E value for <i>Geo. metallireducens</i> hit (.003).
WS1965	NP_908068	121 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS1990	NP_908093	118 aa	Conserved domain DUF 177; COG1399	All significant hits from ϵ -Proteobacteria
WS2120	NP_908218.1	162 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS2123	NP_908221	246 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteobacteria
WS2146	NP_908240	147 aa	Contains Sua5_yciO_yrdC domain involved in binding to dsRNA; gnl CDD 15330	All significant hits except one from ϵ -Proteo; E value changes from 3e-19 to 2e-4 for <i>M. synoviae</i>
WS0230	NP_906487	432 aa	Show significant similarity to deacylase domain; gnl CDD 12932, COG3608	Besides ϵ -Proteo, homologs with very low E values also present in two <i>Desulfovibrio</i> species.
WS2059	NP_908159	259 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteo; not found in <i>T. denitrificans</i> .
WS1752	NP_907874	145 aa	Conserved hypothetical protein, unknown function	All significant hits from ϵ -Proteo; not found in <i>C. fetus</i> and <i>T. denitrificans</i> .
WS1211	NP_907393	621 aa	Homologous to CiaB invasion antigen of <i>C. jejuni</i>	All significant hits from ϵ -Proteo; not found in <i>H. pylori</i> and <i>T. denitrificans</i> . Konkel et al. (1999)

The species distribution of these proteins was determined by BLASTp and PSI-BLAST searches as described in the Methods section. Unless otherwise indicated all of these proteins are uniquely found in the following sequenced genomes: *H. pylori* 26695, *H. pylori* J99, *H. hepaticus* ATCC 51449, *C. jejuni* (various strains: NCTC 11168, RM1221, *HB93-13*, *84-25*, *CF93-6*, *260.94*, *11168* and *81-176*), *C. lari*, *C. coli*, *C. upsaliensis*, *C. fetus*, *W. succinogenes* DSM 1740 and *Thiomicrospira denitrificans* ATCC 33889.