

Table S1. Protein sequence similarities of deduced gene products encoded by the insert of pLP16

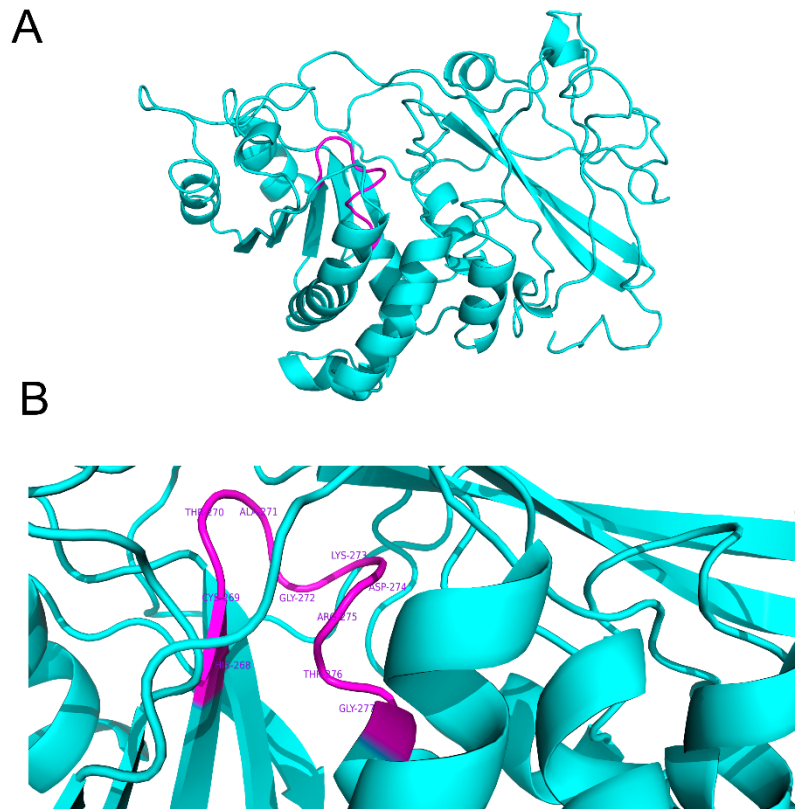
ORF number	No. of encoded amino acids	Accession no, closest similar protein, organism, (no. of encoded amino acids),	Percent identity to the closest similar protein (BlastP)
<i>pho16B</i>	376	WP_042381880.1,protein-tyrosine phosphatase <i>Streptacidiphilus melanogenes</i> , (372)	(79%)
<i>Orf02</i>	325	WP_042417920.1, LacI family transcriptioN/AI regulator <i>Streptacidiphilus anmyonensis</i> , (325)	(82%)
<i>Orf03</i>	117	WP_018557579.1, hypothetical protein <i>Streptomyces sp.</i> BoleA5 (883)	(53%)
<i>Orf04</i>	119	RJW72075, vacuolar protein sorting-associated protein 53 -like protein <i>Clonorchis sinensis</i> ,(1358)	(29%)
<i>Orf05</i>	37	PZN39371.1, sulfoN/Ate ABC transporter ATP-binding protein, Proteobacteria bacterium, (246)	(63%)
<i>Orf06</i>	96	WP_099477000.1, TerC family protein <i>Paenibacillus ihbetae</i> , (221)	(34%)
<i>Orf07</i>	38	AAW27310.1, SJCHGC02811 protein <i>Schistosoma japonicum</i> , (133)	(45%)
<i>Orf08</i>	102	No BLAST hits found	N/A
<i>Orf09</i>	130	XP_016657880.1, uncharacterized protein LOC100159331 isoform X3 <i>Acyrtosiphon pisum</i> , (6853)	(42%)
<i>Orf10</i>	50	No BLAST hits found	N/A
<i>Orf11</i>	818	AWN00235.1, hypothetical protein, uncultured organism, (819)	(99%)
<i>Orf12</i>	35	WP_109505691.1, class I SAM-dependent methyltransferase <i>Nocardioides sp. YIM ART13</i> , (270)	(65%)
<i>Orf13</i>	200	No BLAST hits found	N/A
<i>Orf14</i>	41	No BLAST hits found	N/A
<i>Orf15</i>	109	No BLAST hits found	N/A
<i>Orf16</i>	102	No BLAST hits found	N/A
<i>Orf17</i>	63	OHB94183.1, transcription termination factor NusA <i>Planctomycetes</i> bacterium, (358)	(28%)
<i>Orf18</i>	191	XP_001794145.1, hypothetical protein SNOG_03588 <i>Parastagonospora nodorum SN15</i> (641)	(36%)
<i>Orf19</i>	118	PQE15591.1, Heavy metal tolerance protein <i>Rutstroemia sp. NJR-2017a BBW</i> , (1031)	(47%)
<i>Orf20</i>	25	KMP12518.1, hypothetical protein UZ36_00105 <i>Nitrospina sp.</i> SCGC, (342)	(100%)
<i>Orf21</i>	40	No BLAST hits found	N/A
<i>Orf22</i>	45	No BLAST hits found	N/A
<i>Orf23</i>	410	WP_090595989.1, methyltransferase domain-containing protein <i>Auraticoccus monumenti</i> , (274)	(54%)
<i>Orf24</i>	140	No BLAST hits found	N/A
<i>Orf25</i>	82	No BLAST hits found	N/A
<i>Orf26</i>	133	No BLAST hits found	N/A
<i>Orf27</i>	48	No BLAST hits found	N/A
<i>Orf28</i>	239	ONH24413.1, hypothetical protein BL253_30125 <i>Frankia asymbiotica</i> , (395)	(80%)
<i>Orf29</i>	64	WP_098051725.1, LacI family transcription regulator <i>Pantoea agglomerans</i> , (341)	(39%)
<i>Orf30</i>	105	No BLAST hits found	N/A
<i>Orf31</i>	117	No BLAST hits found	N/A
<i>Orf32</i>	74	No BLAST hits found	N/A
<i>Orf33</i>	49	No BLAST hits found	N/A
<i>Orf34</i>	69	No BLAST hits found	N/A
<i>Orf35</i>	63	No BLAST hits found	N/A
<i>Orf36</i>	110	KQB60501.1, hypothetical protein AE621_04680 <i>Acidovorax sp.</i> SD340 (145)	(42%)
<i>Orf37</i>	186	No BLAST hits found	N/A
<i>Orf38</i>	74	WP_125089351.1, hypothetical protein <i>Saccharopolyspora sp H219</i> , (134)	(52%)
<i>Orf39</i>	26	PYO22201.1, DNA-3-methyladenine glycosylase I , <i>Candidatus Rokubacteria</i> , 194	(100%)
<i>Orf40</i>	186	No BLAST hits found	N/A

N/A: not applicable

Supplementary Table S2. Rank of templates representing the top ten threading PDB templates used by I-TASSER.*

Rank	PDB Hit	Iden1	Iden2	Coverage	Norm. Z-score
1	2oz5	0.33	0.25	0.62	1.43
2	1ywf	0.32	0.23	0.60	3.46
3	5vgr	0.09	0.18	0.93	1.36
4	2oz5	0.33	0.25	0.62	1.87
5	4rkk	0.17	0.16	0.64	1.89
6	1ywf	0.33	0.23	0.60	4.96
7	1ywf	0.33	0.23	0.61	4.15
8	5gnr	0.10	0.17	0.87	1.34
9	2oz5	0.32	0.25	0.62	2.06
10	4rkk	0.14	0.16	0.83	1.12

*Iden1 is the percentage sequence identity of the templates in the threading aligned region with the query sequence. Iden2 is the percentage sequence identity of the whole template chains with query sequence. Coverage of the threading alignment is equal to the number of aligned residues divided by the length of query protein. Norm. Z-score is the normalized Z-score of the threading alignments. The top 10 alignments reported above (in order of their ranking) are from the following threading programs: 1: MUSTER, 2: FFAS-3D, 3: PROSPECT2, 4: SPARKS-X, 5: HHSEARCH, 6: HHSEARCH2, 7: HHSEARCH I, 8: PROSPECT2, 9: Neff-PPAS, 10: SP3.



Supplementary Figure S1. A) Best 3D structure model of the Pho16B protein, predicted by I-TASSER and edited with Pymol 2.2, calculated C score -1.22. B) Amino acids involved in the MptpB-loop in Pho16B.