

Table S14. Sequence IDs and protein threading results from transcripts with significant internal transcriptional start sites in MED4 under N-limitation. The sequences highlighted in bold contain internal start sites with an NtcA binding site during transcription and a pyrimidine binding site for protein S1 binding during translation. The genes in this table combine for more than 90% of the gene expression during the treatment.

Sequence ID	Structure with best match	Model Confidence	Residues aligned	Amino Acids in Sequence	Difference Between Full length and Internal Protein Length	CDD Domain Affected by Internal TSS	Accession Number of CDD Domain Affected	Result	NtcA Binding	Pyrimidine Rich	Theoretical N-Savings (mol N/mol protein)
PMM0211	<b>c1hvuA</b>	100	301	318							
PMM0211 Int	<b>c1hvuA</b>	100	241	251	67	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	Yes	Yes	79
PMM0214	<b>c5da0A</b>	100	460	554							
PMM0214 Int	<b>c3nu0A</b>	99	110	156	398	Possible specific transcription of HCO <sub>2</sub> cotransp super family domain	c123746	Protein has specific structural element or subunit transcribed	Yes	Yes	474
PMM0370	<b>c246A</b>	100	361	561							
PMM0370 Int	<b>c246A</b>	100	365	406	155	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	Yes	Yes	210
PMM1075	<b>c2b5oA</b>	100	288	373							
PMM1075 Int	<b>c1fbcA</b>	100	234	238	135	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	Yes	Yes	179
PMM1485	<b>c3u0C</b>	100	1032	1100							
PMM1485 Int	<b>c3u0C</b>	100	936	988	112	Possible modification of RNA polymerase beta subunit as almost entire protein is single domain	cd09653	Protein aligns by MSA to major structural elements of PDB structures	Yes	Yes	155
PMM1512	<b>c1lm1A</b>	100	1461	1521							
PMM1512 Int	<b>c2vdcF</b>	100	410	413	1108	Possible specific transcription of GXGXG superfamily domain.	c100239	Protein has specific structural element or subunit transcribed	Yes	Yes	1501
PMM1697	<b>c4igcX</b>	100	338	339							
PMM1697 Int	<b>c4igcX</b>	100	176	187	111	Likely loss of Sigma-70 factor, region 1.2 domain	pfam00140	Protein lacks structural element or structure is incomplete	Yes	Yes	380
PMM0063	<b>c2vc5D</b>	91	79	310							
PMM0063 Int	<b>c4b4aA</b>	89	19	205	105	Possible modification of the Ycf66 protein N-terminus domain	pfam07444	Protein lacks structural element or structure is incomplete	No	No	136
PMM0087	<b>c2k2cA</b>	11	13	112							
PMM0087 Int	<b>c2k2cA</b>	14	13	95	17	No domains detected	N/A	Protein lacks structural element or structure is incomplete	No	No	28
PMM0164	<b>c5knD</b>	100	393	417							
PMM0164 Int	<b>d1v8za1</b>	100	371	377	40	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	No	No	59
PMM0263	<b>c5axwB</b>	100	419	486							
PMM0263 Int	<b>c5axwB</b>	100	403	405	81	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	No	No	113
PMM0368	<b>c3nu0</b>	49	73	105							
PMM0368 Int	<b>d1kp0a1</b>	53	31	84	21	First 16 aas of a Domain of Unknown Function	pfam11746	Protein lacks structural element or structure is incomplete	No	Yes	29
PMM0445	<b>c1m56G</b>	100	512	544							
PMM0445 Int	<b>c1m56G</b>	100	451	463	81	Possible modification of cytochrome c oxidase subunit I as almost entire protein is single domain	TIGR02891	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	105
PMM0545	<b>c3aerC</b>	100	407	421							
PMM0545 Int	<b>c3aerC</b>	100	256	263	158	Possible modification of light-independent protochlorophyllide reductase subunit n as almost entire protein is single domain	PRK02842	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	210
PMM0661	<b>d111a</b>	100	653	805							
PMM0661 Int	<b>d111a</b>	100	599	706	99	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	144
PMM0684	<b>d1smpl</b>	52	23	117							
PMM0684 Int	<b>d2v3za1</b>	40	32	105	12	None	N/A	Protein lacks structural element or structure is incomplete	No	No	15
PMM0687	<b>c3adaA</b>	24	18	111							
PMM0687 Int	<b>c3adaA</b>	40	27	89	22	No domains detected	N/A	Protein lacks structural element or structure is incomplete	Yes	Yes	36
PMM0804	<b>c3a9qB</b>	100	177	184							
PMM0804 Int	<b>c3a9qB</b>	100	160	166	18	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	No	No	24
PMM0872	<b>c4a94D</b>	73	18	82							
PMM0872 Int	<b>c4a94D</b>	73	18	81	1	No domains detected	N/A	Protein lacks structural element or structure is incomplete	No	Yes	1
PMM0970	<b>d1o0oB</b>	100	368	439							
PMM0970 Int	<b>d1o0oB</b>	100	357	381	58	None	N/A	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	72
PMM1077	<b>c4ev8A</b>	100	228	375							
PMM1077 Int	<b>c4ev8A</b>	100	207	226	149	None	N/A	Protein has specific structural element or subunit transcribed	No	Yes	211
PMM1123	<b>d4lgA</b>	33	30	237							
PMM1123 Int	<b>d1o09a</b>	19	13	180	57	None	N/A	Protein lacks structural element or structure is incomplete	No	Yes	88
PMM1158	<b>c2axc</b>	100	446	460							
PMM1158 Int	<b>c2axc</b>	100	416	416	44	Possible modification of the photosystem ii cp43 protein as almost entire protein is single domain	TIGR01153	Protein aligns by MSA to major structural elements of PDB structures	Yes	No	64
PMM1186	<b>3wvfA</b>	100	183	379							
PMM1186 Int	<b>c3wvfA</b>	100	177	288	91	Possible modification of the membrane protein insertase yidC protein as almost entire protein is single domain	PRK02654	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	120
PMM1301	<b>c3n5mb</b>	100	392	417							
PMM1301 Int	<b>c4vndD</b>	100	243	245	225	Possible modification of the acetylornithine aminotransferase protein as almost entire protein is single domain	PRK02627	Protein aligns by MSA to major structural elements of PDB structures	No	No	229
PMM1312	<b>c5f4hf</b>	100	159	369							
PMM1312 Int	<b>c5f4hf</b>	100	159	368	1	Possible specific transcription of (PIT N terminus) domain	cd09877	Protein has specific structural element or subunit transcribed	No	Yes	1
PMM1434	<b>c4my4A</b>	100	500	540							
PMM1434 Int	<b>c4my4A</b>	100	449	474	66	Possible modification of the 2,3-bisphosphoglycerate-independent phosphoglycerate protein as almost entire protein is single domain	cd16010	Protein aligns by MSA to major structural elements of PDB structures	No	No	89
PMM1625	<b>c3dbyC</b>	100	463	475							
PMM1625 Int	<b>c3u58D</b>	100	339	339	136	loss of the beginning amino acids in the homotetramer interface on conserved domain SAHH	PRK05476	Protein aligns by MSA to major structural elements of PDB structures	No	No	178
PMM1636	<b>c5imrX</b>	100	162	190							
PMM1636 Int	<b>c5imrX</b>	100	107	119	71	loss of part of translation initiation factor 3 N terminal domain	pfam05198	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	113
PMM1642	<b>d2gexa1</b>	100	124	146							
PMM1642 Int	<b>d2gexa1</b>	100	94	111	35	loss of the beginning amino acids in the Snoal-like domain	pfam12680	Protein aligns by MSA to major structural elements of PDB structures	No	Yes	111