

Supplementary Material for “Expansion of the mycobacterial PUPylome”

Jeramie Watrous^a, Kristin Burns^e, Wei-Ting Liu^a, Anand Patel^d, Vivian Hook^c, Vineet Bafna^d, Clifton E. Barry 3rd
^e, Steve Bark^c, and Pieter C. Dorrestein^{abc*}.

^a: Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093

^b: Department of Pharmacology, University of California, San Diego, La Jolla, California 92093

^c: Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093

^d: Department of Computer Science, University of California, San Diego, La Jolla, California 92093

^e: Tuberculosis Research Section, NIAID, National Institutes of Health, Bethesda, Maryland 20892

*: To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Ave.,
BSB 4090 MC 0636, La Jolla, Ca 92093. Tel.: 858-534-6607; Fax: 858-822-0041; E-mail: pdorrestein@ucsd.edu.

Table of Contents

I.	<u>Summary Results</u>	<u>3</u>
	a. Table ST1: Summary of PUPylated proteins	
	b. Table ST2: Summary of candidate proteins	
	c. Table ST3: Summary of non-covalent interacting proteins	
	d. Table ST4: Summary of non-candidate proteins	
II.	<u>Possible Correlation Between Nitrosylation and PUPylation</u>	<u>8</u>
	a. Table ST5: Comparison of Nitrosylated and PUPylated proteins	
III.	<u>Manual Confirmation of PUPylated Proteins</u>	<u>10</u>
	a. Figures SF1-SF48: Spectral annotations for PUPylated proteins	
IV.	<u>Sequence Alignments</u>	<u>59</u>
	a. Figure SF49: Sequence alignment of all PUPylated proteins centered at site of modification	
	b. Figure SF50: Full length sequence alignment of all PUPylated proteins using ClustalW algorithm	
	c. Figures SF51-SF102: Individual sequence alignments of PUPylated proteins showing extent of evolutionary sequence conservation at site of modification	
V.	<u>Structural Annotations</u>	<u>120</u>
	a. Figures SF103-SF123: Indication of site of modification on crystal structures of homologous proteins	
VI.	<u>Pathway Annotations</u>	<u>142</u>
	a. Figures SF124-SF136: Annotated pathway maps showing location of PUPylated and candidate proteins in pathway	

I. Summary Results

a. Table ST1: Summary table of all PUPylated proteins.....	4
b. Table ST2: Summary table of all candidate proteins.....	5
c. Table ST3: Summary table of all non-covalent interacting proteins.....	6
d. Table ST4: Summary table of all non-candidate protein.....	7

ST1: Summary table of all 52 PUPylated proteins

Protein Name	GI Number	Enzyme Code	Locus Tag (MSMEG_#)	% Coverage	# Peptides	Spectral Counting			
						Total Spectra	% Spectra in Non-tagged spl.	% Spectra in Native spl.	% Spectra in His-Tag spl.
[Mn] superoxide dismutase	118469767	1.15.1.1	6427	93.2	79	937	6.0	1.8	92.2
2,5-diketo-D-gluconic acid reductase A*	118173010		2408	28.7	3	5	NA	NA	NA
30S ribosomal protein S17	118171465		1445	56.1	5	8	0.0	0.0	100.0
30S ribosomal protein S6	118171449		6897	66.7	6	34	0.0	2.9	97.1
50S ribosomal protein L11	118473904	2.2.1.6	1346	26.8	4	24	0.0	4.2	95.8
Acetolactate synthase 3 regulatory subunit	118470606		2373	54.7	6	47	0.0	15.7	84.3
Acetyl-CoA acetyltransferase*	118170768		4920	72.3	19	58	NA	NA	NA
Acetylornithine aminotransferase	118472927		3773	35.6	11	103	0.0	42.2	57.8
Aconitate hydratase 1	118471676	4.2.1.3	3143	55.6	50	474	5.3	61.9	32.8
Acyl carrier protein	118471527		4326	34.3	5	61	21.9	20.5	57.5
Acyl-CoA synthetase	118470672		5435	42.6	25	168	43.1	51.7	5.2
Adenylosuccinate synthetase	166232470		0759	47.6	24	163	0.0	26.2	73.8
Alcohol dehydrogenase, zinc-containing	118468910	1.1.1.2	1037	46.7	15	116	4.1	17.1	78.9
Alkylhydroperoxide reductase*	118172773		4891	64.6	18	59	NA	NA	NA
Anthranilate synthase component I*	118472702		3217	40.3	14	27	NA	NA	NA
Aspartate transaminase	118468930		6286	48.8	27	178	0.0	16.3	83.7
ATPase, AAA family protein	118470801	2.6.1.42	3902	43.7	18	110	0.0	44.1	55.9
Branched-chain amino acid aminotransferase	118472496		4276	31.5	13	46	0.0	28.9	71.1
Carbonic anhydrase	118169435		6082	58.7	7	39	0.0	10.0	90.0
Carveol dehydrogenase*	118169959		6031	25.1	5	6	NA	NA	NA
Chaperonin GroEL	118469510	1.1.1.49	0880	57.1	28	361	27.5	21.8	50.6
Conserved hypothetical protein*	118172453		4692	52.9	5	7	NA	NA	NA
Conserved hypothetical protein	118171476		6518	78.3	15	88	0.0	5.7	94.3
Cyclopropane-fatty-acyl-phospholipid synthase	118173720		6284	67.5	33	447	41.6	21.9	36.4
DNA topoisomerase IV subunit B	118468752	6.3.1.2	0457	33.8	15	41	0.0	80.5	19.5
Elongation factor Ts	166221229		2520	67.6	20	178	1.6	37.4	61.0
Ferritin family protein	118468959		6422	65.2	14	168	1.7	30.7	67.6
Glucose-6-phosphate 1-dehydrogenase	118171568		0314	23.4	8	15	0.0	93.3	6.7
Glutamine synthetase, type I	118474021	6.3.1.2	4294	31.8	12	61	42.2	54.7	3.1
Glutamine synthetase, type I	118169243		4290	51.7	24	189	7.4	48.9	43.7
Glyceraldehyde-3-phosphate dehydrogenase, type I	118467799		3084	54.4	15	81	0.0	55.3	44.7
Glyoxalase family protein*	118173889		6630	65.5	4	4	NA	NA	NA
Hypothetical protein	118468878	1.2.1.12	2782	37.7	9	17	0.0	94.1	5.9
Hypothetical protein	118473705		6159	65.7	2	6	0.0	0.0	100.0
Inositol-5-monophosphate dehydrogenase	118473142		1603	24.8	5	12	0.0	63.6	36.4
Malonyl CoA-acyl carrier protein transacylase	118471928		4325	23.8	4	19	0.0	15.8	84.2
Myo-inositol-1-phosphate synthase	118470644	2.3.1.39	6904	68.3	33	543	3.1	44.0	52.9
N-carbamoyl-L-amino acid amidohydrolase	118473294		3995	34.5	11	39	0.0	48.1	51.9
Phosphoenolpyruvate carboxykinase*	118173553		0255	64.3	30	NA	NA	NA	NA
Prolyl-tRNA synthetase	118472270		2621	56.6	29	147	0.0	100.0	0.0
Putative adenosylhomocysteinase	116266924	6.1.1.15	1842	60.2	55	538	1.9	61.8	36.3
Putative thiosulfate sulfurtransferase	118472788		5789	60.3	18	152	11.9	45.6	42.5
Pyruvate dehydrogenase subunit E1	118470577		4323	66.0	75	454	2.5	80.2	17.3
Ribonucleoside diphosphate reductase, alpha subunit*	118170953		2299	41.2	23	58	NA	NA	NA
Rieske 2Fe-2S family protein	118474008	1.2.4.1	6410	33.4	12	35	12.8	59.6	27.7
S-adenosylmethionine synthetase	118169294		3055	77.2	61	656	3.3	6.1	90.7
Serine hydroxymethyltransferase	118468818		5249	69.8	40	386	4.1	38.5	57.4
Short chain dehydrogenase	118473731		3619	86.5	39	818	44.6	1.1	54.4
SseC protein	118468699	4.2.3.1	5790	98.9	10	229	4.3	17.4	78.3
Threonine synthase	118468890		4956	66.7	18	148	0.7	5.9	93.4
Transaldolase	118171487		3102	57.4	18	80	0.0	40.2	59.8
Universal stress protein family protein	118471745		3950	39.5	6	20	0.0	14.3	85.7

* Proteins found using clustered mass spectra data set and therefore spectral counting cannot be done

ST2: Summary table of all 103 candidate proteins

Protein Name	GI Number	Enzyme Code	Locus Tag (MSMEG_#)	% Coverage	# Peptides	Spectral Counting			
						Total Spectra	% Spectra in Non-tagged spl.	% Spectra in Native spl.	% Spectra in His-Tag spl.
[Mn] superoxide dismutase	118173832	1.15.1.1	6636	46.6	7	38	5.49	3.05	91.46
[NADP+] succinate-semialdehyde dehydrogenase	118470915	1.2.1.16	5538	21.4	7	32	0.00	51.43	48.57
1-aminocyclopropane-1-carboxylate deaminase	118467529	3.5.99.7	6740	24.3	5	17	0.00	17.65	82.35
2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	118471051	5.4.2.1	0935	34.4	7	14	0.00	30.77	69.23
2-isopropylmalate synthase	118470831	2.3.3.13	6271	37.0	16	60	0.00	68.33	31.67
2-oxoglutarate dehydrogenase, E2 component	118172342	2.3.1.61	4283	19.3	7	30	23.40	0.00	76.60
30S ribosomal protein S10	118473457		1012	36.6	3	10	20.00	30.00	50.00
3-deoxy-7-phosphoheptulonate synthase	118468924	2.5.1.54	4244	30.6	11	54	1.72	87.93	10.34
3-isopropylmalate dehydratase small subunit	118171809	4.2.1.33	2388	47.7	7	27	0.00	40.70	59.30
3-ketoacyl-CoA thiolase	118169455	2.4.1.18	0373	15.5	4	18	0.00	0.00	100.00
3-oxoacyl-(acyl carrier protein) synthase II	118467655	2.3.1.179	4328	38.8	9	61	0.00	50.79	49.21
3-oxoacyl-(acyl carrier protein) synthase II	118471172	2.3.1.41	4327	52.9	18	136	4.26	31.21	57.45
4-hydroxy-2-ketovalerate aldolase	118469867	4.1.3.-	5937	64.5	20	202	15.61	21.95	62.44
50S ribosomal protein L1	118473196		1347	31.5	6	27	0.00	25.00	75.00
50S ribosomal protein L2	118173550		1439	52.5	13	77	41.50	5.20	53.30
50s ribosomal protein L4/L1	118171375		1437	38.1	4	29	0.00	10.30	89.70
Acetyl-CoA acetyltransferase	118467515	2.3.1.9	5273	26.9	7	47	4.08	20.41	75.51
Acyl-ACP thioesterase superfamily protein	118173664		0909	20.9	4	12	0.00	23.08	76.92
Acyl-CoA synthetase	118473450	2.3.1.86	4301	13.0	6	13	0.00	91.67	8.33
Acyl-CoA synthetase	118471852	2.3.1.86	6393	24.3	10	12	0.00	83.33	16.67
Adenosine kinase	118470653	2.7.1.20	4270	28.6	6	12	0.00	0.00	100.00
Adenylate kinase	118473144	2.7.4.3	1484	44.2	6	15	0.00	28.57	71.43
Alcohol dehydrogenase, iron-containing	118471656	1.1.1.1	6242	43.7	20	151	4.43	60.13	35.44
Aldehyde dehydrogenase (NAD) family protein	118171071	1.2.1.-	1665	25.4	8	11	0.00	36.36	63.64
Allantoicase	118470580	5.4.3.8	5727	17.8	5	11	0.00	45.45	54.50
Arabitol-phosphate dehydrogenase	118473835		3265	25.1	6	11	9.09	18.18	72.73
Argininosuccinate synthase	118468824	6.3.4.5	3770	65.9	34	252	8.66	43.70	47.64
ATP synthase subunit alpha	158512473	3.6.3.14	4938	12.6	5	14	0.00	7.10	92.90
Bifunctional 3,4-dihydroxy-2-butanone 4-phosphate	118470936	3.5.4.25	3072	32.6	9	39	0.00	26.32	73.68
Bifunctional phosphopantothienoylcysteine decarboxylase	118473275		3054	15.0	4	10	30.00	0.00	70.00
Bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase	118471289	2.1.2.3	5515	36.1	11	25	16.00	4.00	80.00
Catalase/oxidase HPI	118471704	1.11.1.6	3461	21.8	10	17	0.00	50.00	50.00
Chaperonin GroEL	118469216		1583	87.4	114	4147	23.80	44.50	31.70
Chaperonin GroS	118171718		1582	38.0	3	19	0.00	15.79	84.21
Chorismate synthase	171472980	4.2.3.5	3030	37.8	10	22	0.00	66.67	33.33
Citrate Synthase I	118171778	2.3.3.1	5672	40.3	15	96	8.00	78.00	14.00
Conserved hypothetical protein	118173839		5246	42.6	8	14	0.00	0.00	100.00
CTP synthetase	118473994	6.3.4.2	3746	14.4	5	11	9.09	18.18	72.73
Cysteine desulfurase IscS	118471674	4.4.1.-	1242	12.4	4	12	0.00	40.00	60.00
Dihydrolipoamide dehydrogenase	118173883	1.8.1.4	0903	42.2	16	70	11.27	1.41	87.32
Dihydropyrimidinase	118471068	3.5.2.2	3996	23.4	12	25	0.00	75.00	25.00
Dihydroxy-acid dehydratase	118471241	4.2.1.9	0229	11.6	5	15	0.00	31.25	68.75
DNA-binding protein HU	118471339		2389	19.2	3	28	25.93	33.33	40.74
DNA-binding response regulator MtrA	118467592		1874	13.6	2	16	0.00	18.75	81.25
Electron transfer flavoprotein, alpha subunit	118471409		2352	42.5	9	67	0.00	31.88	68.12
Electron transfer flavoprotein, beta subunit	118469254		2351	26.6	5	16	0.00	76.47	23.53
Elongation factor Tu	166222872	3.6.5.3	1401	67.9	26	285	32.17	12.59	55.24
Enoyl-CoA hydratase	118471625	4.2.1.17	5639	60.9	10	24	4.35	82.61	13.04
F0F1 ATP synthase subunit beta	118471849	3.6.3.14	4936	64.2	23	118	16.95	15.25	67.80
F420-dependent glucose-6-phosphate dehydrogenase	3056729		0777	15.8	3	14	0.00	50.00	50.00
Ferredoxin sulfite reductase	118471944	1.8.7.1	4527	24.8	9	38	0.00	15.79	84.21
Fumarate hydratase	118472968	4.2.1.2	5240	23.3	6	13	0.00	75.00	25.00
Geranylgeranyl reductase	118467508		2308	22.5	6	12	41.67	25.00	33.33
Glutamate synthase family protein	118471237		6263	20.4	6	18	0.00	64.71	35.29
Glutamate synthase subunit beta	118473425	1.4.1.13	6458	23.0	7	15	0.00	33.33	66.67
Glycerol kinase	118469921	2.7.1.30	6759	54.7	30	311	18.46	28.31	53.23
Glycine dehydrogenase	118171574	1.4.4.2	3642	13.7	11	23	0.00	73.68	26.32
Histidinol dehydrogenase	118471811	1.1.1.23	3205	22.9	7	41	0.00	33.33	66.67
Histidinol-phosphate aminotransferase	118169422		6351	42.7	13	43	0.00	86.10	13.90
Holo-(acyl carrier protein) synthase	118174107	2.7.8.7	4756	44.6	8	93	32.65	3.06	64.29
Homoserine dehydrogenase	118473179	1.1.1.3	4957	13.2	4	10	0.00	63.64	36.36
Hypothetical protein	118469419		3952	44.0	9	40	16.67	2.38	80.95
Hypothetical protein	118471602		4200	14.9	4	18	0.00	0.00	100.00
IMP dehydrogenase family protein	118173825	1.3.1.10	3634	20.9	6	13	0.00	66.67	33.33
Integration host factor	118473958		3050	45.7	4	11	0.00	0.00	100.00
Iron-dependent repressor IdeR	118169560		2750	76.5	30	347	44.13	46.09	9.78
Isocitrate dehydrogenase, NADP-dependent	118467432	1.1.1.42	1654	35.8	19	95	32.32	27.27	40.40
Isocitrate lyase	118471509	4.1.3.1	0911	24.1	5	13	0.00	75.00	25.00
Ketol-acid reductoisomerase	118173469	1.1.1.86	2374	40.1	14	85	9.20	28.74	62.07
Magnesium chelatase	118172834		5512	23.0	8	10	0.00	40.00	60.00
Malate synthase G	118469725	2.3.3.9	3640	21.4	12	48	16.00	4.00	80.00
Methylmalonate-semialdehyde dehydrogenase	118469836	1.2.1.27	2449	21.7	8	18	0.00	53.85	38.46
N-acetyl-gamma-glutamyl-phosphate reductase	118471377	1.2.1.38	3776	48.1	10	20	16.67	10.00	73.33
Nucleoside-diphosphate-sugar epimerase	118471043		6142	67.0	20	164	3.59	60.48	35.93
O-acetylhomoserine aminocarboxypropyltransferase	118469926	2.5.1.49	1652	23.0	5	19	22.22	27.78	50.00
Oxidoreductase alpha (molybdopterin) subunit	118473746		4668	22.3	11	18	0.00	70.59	29.41
PhoH family protein	118473775		5247	43.0	17	73	0.00	48.68	51.32
Phosphoglycerate kinase	166219328	2.7.2.3	3085	29.2	8	14	7.14	0.00	92.86
Phosphopyruvate hydratase	118471749	4.2.1.11	5415	36.8	9	29	0.00	68.97	31.03
Phosphoribosylformylglycinamidine synthase II	118473588	6.3.5.3	5824	32.7	20	107	1.75	43.86	54.39
Polynucleotide phosphorylase/polyadenylase	118471281	2.7.7.8	2656	14.9	8	14	35.71	14.29	50.00
Proline dipeptidase	118173682		3881	34.1	8	43	0.00	25.58	74.42
Propionyl-CoA carboxylase beta chain	118471917	6.4.1.3	1813	22.9	8	27	23.08	26.92	57.69
Proteasome component	118469231		3897	51.0	30	164	0.00	90.20	9.80
Putative acyl-CoA dehydrogenase	118472712		5199	31.2	7	19	0.00	42.86	52.38
Putative elongation factor G	116266976		6535	31.5	15	81	24.14	26.44	49.43
Pyridoxal Biosynthesis lyase protein	166980429		2937	30.4	7	24	0.00	65.22	34.78
Pyruvate kinase	118470856	2.7.1.40	3227	32.8	11	44	0.00	40.91	59.09
Ribose-phosphate pyrophosphokinase	118470710	2.7.6.1	5427	25.8	7	29	0.00	85.71	14.29
S30AE family protein, ribosome asosociated	118473146		1878	30.9	7	43	0.00	23.40	76.60
Steroid delta-isomerase	118471742	5.3.3.1	5258	57.3	4	16	0.00	6.67	93.33
Succinate dehydrogenase flavoprotein subunit	118169365	1.3.5.1	1670	27.4	12	49	9.09	0.00	90.91
Succinyl-CoA synthetase subunit alpha	118469852	6.2.1.5	5524	18.3	3	10	0.00	0.00	100.00
Taurine-pyruvate aminotransferase	118173357		1662	56.4	19	126	0.00	48.75	51.25
Thiazole synthase	118469300	730	0793	25.4	5	39	0.00	47.62	52.38
Thioredoxin-disulfide reductase	118473970	1.8.1.9	6933	46.6	10	94	0.00	22.45	77.55
Transketolase	118473102	2.2.1.1	3103	41.1	20	104	29.52	26.67	43.81
Tryptophan synthase subunit beta	118175161	4.2.1.20	3220	16.7	7	34	0.00	67.65	32.35
UDP-N-acetylglucosamine 1-carboxyvinyltransferase	118171565	2.5.1.7	4932	14.1	4	10	20.00	20.00	60.00
Universal stress protein family protein	118472903		3811	89.8	11	74	0.00	52.00	48.00
Universal stress protein family protein	118473140		5733	36.3	7	41	0.00	11.90	88.10
Universal stress protein family protein	118168988		3945	34.8	6	15	0.00	37.50	62.50
Urease subunit alpha	118473986	3.5.1.5	3625	16.8	6	10	10.00	0.00	90.00

ST3: Summary table of all 12 non-covalent interacting proteins

Protein Name	GI Number	Enzyme Code	Locus Tag (MSMEG_#)	% Coverage	# Peptides	Spectral Counting			
						Total Spectra	% Spectra in Non-tagged spl.	% Spectra in Native spl.	% Spectra in His-Tag spl.
2-oxoglutarate dehydrogenase, E1 component	118173342	1.2.4.2	5049	36.5	31	98	0.00	98.00	2.00
Acetaldehyde dehydrogenase	118471479	1.2.1.10	5939	45.8	10	39	0.00	100.00	0.00
Acetolactate synthase, large subunit	118174208	2.2.1.6	2372	22.7	9	34	0.00	100.00	0.00
Acyl-CoA-dehydrogenase	118473445		0406	24.4	11	23	0.00	100.00	0.00
DNA gyrase subunit A	118474051	5.99.1.3	0456	20.9	10	18	0.00	100.00	0.00
Eptc-inducible aldehyde dehydrogenase	118467594		1543	22.2	8	11	0.00	100.00	0.00
Excinuclease ABC subunit A	118471228		3808	12.1	8	14	0.00	100.00	0.00
Ferredoxin-dependent glutamate synthase 1	118169347	1.4.7.1	6459	10.2	11	24	16.00	80.00	4.00
Propionyl-CoA carboxylase beta chain	118471103	6.4.1.3	6391	16.4	6	18	0.00	100.00	0.00
Putative 3-hydroxyacyl-CoA dehydrogenase	118473807	1.2.1.3	5720	14.7	7	10	0.00	100.00	0.00
Pyruvate carboxylase	118473087	6.4.1.1	2412	17.1	14	23	0.00	100.00	0.00
Pyruvate synthase	118473956		4646	12.4	5	17	0.00	93.75	6.25

ST4: Summary table of all 78 non-candidate proteins

Protein Name	GI Number	Locus Tag (MSMEG_#)	% Coverage	# Peptides	Spectral Counting			
					Total Spectra	% Spectra in Non-tagged spl.	% Spectra in Native spl.	% Spectra in His-Tag spl.
1,4-alpha-glucan branching enzyme	118169575	4918	43.3	22	111	45.90	0.00	54.10
16S rRNA-processing protein	118469073	2437	34.7	4	21	71.43	28.57	0.00
2,3-dihydroxybiphenyl 1,2-dioxygenase	118170987	2006	8.8	2	20	94.74	0.00	5.26
29 kDa antigen Cfp29	118470265	5830	52.8	10	52	94.00	6.00	0.00
30S ribosomal protein S1	118472813	3833	51.6	17	78	45.57	5.06	49.37
30S ribosomal protein S11	118467427	1522	32.6	6	65	69.12	4.41	26.47
30S ribosomal protein S15	118471582	2654	59.6	17	150	67.33	2.00	30.67
30s ribosomal protein S19	118471461	1440	54.8	8	40	87.50	5.00	7.50
30s ribosomal protein S3	118174447	1442	34.9	8	52	75.00	9.62	15.38
30s ribosomal protein S4	118169588	1523	18.9	3	21	85.71	4.76	9.52
30S ribosomal protein S5	189045482	1472	40.2	7	32	93.94	3.03	3.03
3-ketoacyl-(acyl-carrier-protein) reductase	118471722	0372	33.6	10	43	76.74	11.63	11.63
50S ribosomal protein L13	118471450	1556	47.6	4	11	50.00	0.00	50.00
50S ribosomal protein L15	118470728	1474	39.5	4	16	50.00	0.00	50.00
50s ribosomal protein L16	118173349	1443	47.1	17	110	58.80	0.00	41.20
50S ribosomal protein L18	118471867	1471	49.6	5	34	47.83	0.00	52.17
50S ribosomal protein L21	118473704	4625	38.8	6	59	69.49	0.00	30.51
50S ribosomal protein L27	166225014	4624	50.0	4	18	63.79	6.90	29.31
50S ribosomal protein L3	118473190	1436	32.3	7	71	63.01	12.33	24.66
50S ribosomal protein L7/L12	118468616	1365	36.2	5	37	61.54	7.69	30.77
ABC-type molybdenum transport system, ATPase component	118468212	2326	44.0	10	58	79.31	20.69	0.00
Acetyl-/propionyl-coenzyme A carboxylase alpha chain	118469108	1807	51.5	18	200	65.17	31.34	3.48
Acyl-CoA oxidase	118473351	4474	19.2	9	45	79.55	6.82	13.64
AmiB	118469679	1679	29.6	8	45	46.67	6.67	46.67
Amidohydrolase 3	118470881	3861	40.6	22	164	81.33	0.60	18.07
Antigen 85-A	118170031	6398	11.0	2	23	60.87	0.00	39.13
Asparagine synthase (glutamine-hydrolyzing)	118173395	2594	13.1	5	13	46.15	53.85	0.00
Cell division ATP-binding protein FtsE	118473811	2089	24.9	5	17	90.91	9.09	0.00
Chaperone protein DnaK	118173710	0709	46.1	21	91	53.68	29.47	16.84
Conserved hypothetical protein	118171429	1013	24.9	6	14	71.40	28.60	0.00
Conserved hypothetical protein	118169679	5933	16.9	4	10	70.00	30.00	0.00
Cupin domain-containing protein	118472920	5707	78.6	20	148	60.78	0.00	39.22
Diaminopimelate decarboxylase	118169223	4958	66.5	38	287	53.92	46.08	0.00
DNA-binding protein	118468644	0906	16.7	6	12	91.67	8.33	0.00
DNA-directed RNA polymerase subunit alpha	118469556	1524	35.7	8	30	63.33	16.67	20.00
DNA-directed RNA polymerase subunit beta	118169416	1367	16.4	11	25	56.00	28.00	16.00
DNA-directed RNA polymerase subunit beta'	118469768	1368	9.2	8	12	68.75	18.75	12.50
Extracellular solute-binding protein, family protein	118171622	0643	26.1	8	16	66.67	13.33	20.00
FAD-dependent thymidylate synthase	118473148	2683	35.2	6	25	88.24	11.76	0.00
Ferric uptake regulation protein	118470711	4487	44.1	7	80	94.19	1.16	4.65
Glycerol kinase	118469746	6756	28.1	14	91	64.89	35.11	0.00
Glycerol-3-phosphate dehydrogenase 2	118173731	6761	21.8	7	29	66.67	25.93	7.41
GTP-binding protein TypA/BipA	118473304	5132	11.5	6	20	63.16	0.00	36.84
Heparin-binding hemagglutinin	118468465	0919	27.2	4	13	92.31	7.69	0.00
Homoserine O-acetyltransferase	118171770	6772	47.1	16	74	81.82	18.18	0.00
Hydrogenase accessory protein HypB	118467565	2271	63.3	17	289	58.28	18.87	22.85
Hypothetical protein	118469659	5010	33.5	5	20	71.43	28.57	0.00
Hypothetical protein	118473453	6590	42.7	3	12	90.91	0.00	9.09
Hypothetical protein	118472727	4029	15.4	2	17	50.00	50.00	0.00
L-aspartate oxidase	118174288	3200	14.3	5	25	93.10	6.90	0.00
Metallo-beta-lactamase family protein	118470931	1334	48.6	7	19	63.16	0.00	36.84
Methyltransferase	118171848	4303	58.1	19	138	52.52	48.20	0.00
Methyltransferase type 11	118174667	6483	26.7	5	52	46.15	0.00	53.85
PhoH family protein	118468558	4497	31.9	7	23	95.65	4.35	0.00
Phosphoglycerate mutase family protein, putative	118169676	6338	43.6	5	21	80.95	19.05	0.00
Phosphomethylpyrimidine kinase	118171480	0825	56.5	11	203	53.40	10.19	36.41
Primosome assembly protein PriA	118473637	3061	18.3	8	20	75.00	25.00	0.00
Protein RecA	118472151	2723	28.9	6	16	96.15	3.85	0.00
Puromycin N-acetyltransferase	118471592	4563	29.0	5	16	93.75	0.00	6.25
Putative acyl-CoA dehydrogenase	118470260	1387	31.2	7	18	77.78	11.11	11.11
Putative acyl-CoA dehydrogenase	118170119	4084	34.0	9	42	81.40	18.60	0.00
Pyridoxamine 5'-phosphate oxidase family protein	118472674	6519	66.2	21	124	62.20	37.80	0.00
Regulatory protein	118472940	4304	27.1	5	18	88.24	11.76	0.00
Regulatory protein	118472886	1068	17.4	5	12	100.00	0.00	0.00
Ribosomal large subunit pseudouridine synthase D	118471435	3175	20.4	5	22	85.71	14.29	0.00
Serine/threonine dehydratase family protein	118170386	3532	21.5	5	12	83.33	16.67	0.00
SUF system FeS assembly protein, NifU family protein	118470835	3126	29.3	3	23	72.73	0.00	27.27
TetR-family protein transcriptional regulator	118473939	5838	39.8	4	23	63.64	36.36	0.00
Transcription termination factor Rho	118471197	4954	15.1	5	12	85.71	7.14	7.14
Transcriptional regulator, ArsR family protein	118169693	4486	20.2	3	17	94.12	5.88	0.00
Transcriptional regulator, GntR family protein	118471885	4042	66.7	15	151	65.43	6.79	27.78
Transcriptional regulator, TetR family protein	118469406	1397	29.8	4	47	56.25	43.75	0.00
Transcriptional regulator, TetR family protein	118173887	2898	13.6	2	13	85.71	14.29	0.00
tRNA (adenine-N(1)-)-methyltransferase	118473840	3906	30.5	5	38	78.38	8.11	13.51
Tryptophan synthase subunit beta	118469854	3309	11.3	4	17	62.50	0.00	37.50
Universal stress protein family protein	118471715	4362	67.3	16	80	59.49	40.51	0.00
Urease subunit gamma	166227122	3627	54.0	3	11	71.43	28.57	0.00
Zinc finger, UBP-type	118473531	0765	20.0	2	15	100.00	0.00	0.00

II. Possible Correlation Between Nitrosylation and PUPylation

a. Table ST5: Comparison of Nitrosylated and PUPylated proteins.....9

The following table shows a comparison between nitrosylated proteins found in Mtb by Rhee *et al* and PUPylated proteins found in Msm in the present study. Out of the 29 nitrosylated proteins found in Mtb, 15 common proteins with high percent identity were also found to be PUPylated. Two additional common proteins were found with lower percent identity (acyl-CoA synthetase and oxidoreductase).

ST5: Correlation of Nitrosylated and PUPylated Proteins

Nitrosylated Proteins Found in Study by Rhee <i>et al</i>		PUPylated Proteins Found in Current Study		%Identity
Mtb Locus Tag	Mtb Protein Name	Msm Locus Tag	Msm Protein Name	
RV0211	Phosphoenolpyruvate Carboxykinase	MSMEG_0255	Phosphoenolpyruvate Carboxykinase	90
RV0873	Acyl CoA Dehydrogenase			
RV1023	Enolase	MSMEG_5415	Enolase (Phosphopyruvate hydratase)	87
RV1308	ATP Synthase	MSMEG_4938	F0F1 ATP Synthase (alpha subunit)	84
RV1449c	Transketolase	MSMEG_3103	Transketolase	81
RV1475c	Aconitase	MSMEG_3143	Aconitate Hydratase 1	82
RV1837c	Malate Synthase	MSMEG_3640	Malate Synthase G	80
RV1876	Bacterioferritin			
RV2201	Asparagine Synthase			
RV2220	Glutamine Synthetase	MSMEG_4290	Glutamine Synthetase (type I)	84
RV2280	Oxidoreductase	MSMEG_5037	Oxidoreductase	35
RV2467	Lysyl aminopeptidase			
RV2996c	3-phosphoglycerate dehydrogenase			
RV3248c	S-adenosyl-L-homocysteine Hydrolase	MSMEG_1842	Putative Adenosylhomocysteinase	87
RV3509c	Acetohydroxyacid Synthase			
RV3859	Glutamine Synthase	MSMEG_6459	Glutamate Synthase 1 (Ferredoxin-dependent)	85
RV1925	Acyl CoA Synthetase	MSMEG_4301	Acyl CoA Synthetase	59
RV2940c	Mycocerosic Acid Synthase			
RV3800c	Polyketide Synthase			
RV3801c	Fatty Acyl-AMP Ligase	MSMEG_6393	Acyl CoA synthetase	76
RV0462	Dihydrolipoamide Dehydrogenase	MSMEG_0903	Dihydrolipoamide Dehydrogenase	85
RV1908c	Catalase	MSMEG_3461	Catalase/Peroxidase	66
RV2115c	Mycobacterial Proteasome-Associated ATPase	MSMEG_3902	ATPase, AAA family protein	91
RV0350	HSP70			
RV0440	Chaperonin	MSMEG_0880	Chaperonin GroEL	93
RV2299c	HSP 90			
RV0685	EF-Tu	MSMEG_1401	EF-Tu	96
RV0667	RNA Polymerase Beta Subunit			
RV0668	RNA Polymerase Beta Prime Subunit			

III. Manual Confirmation of PUPylated Proteins

a. Figures SF1-SF48: Spectral annotations for PUPylated proteins.....11

The following figures show the results of manual annotation of all 43 PUPylated proteins found using the unclustered data set. PUPylated proteins found using the cluster data sets (from MSCluster) could not be annotated manually due to inability to convert the data file back to a spectral format.

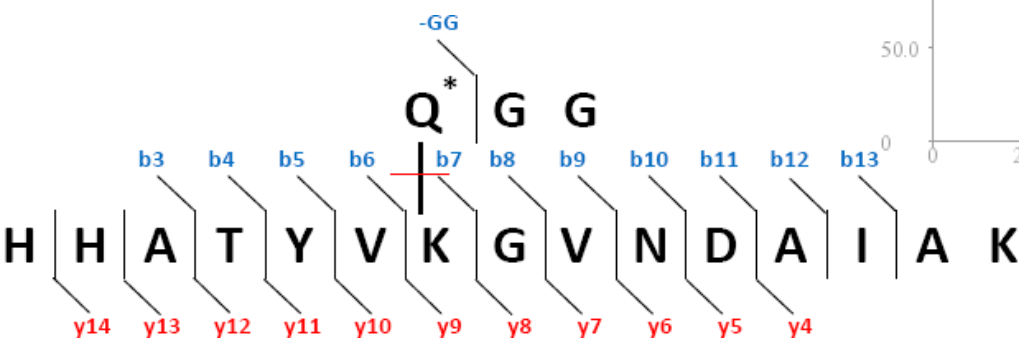
Each figure in this section is set up with the protein name, modified peptide found using InSpecT and the calculated mass and charge in the upper left. The plot in the upper right is the spectra annotated by InSpecT. Toward the middle of each figure is the peptide sequence again showing the ions found by manual annotation with the corresponding mass spectra on the bottom of the page.

SF1: [Mn] Superoxide dismutase (modification site 1)

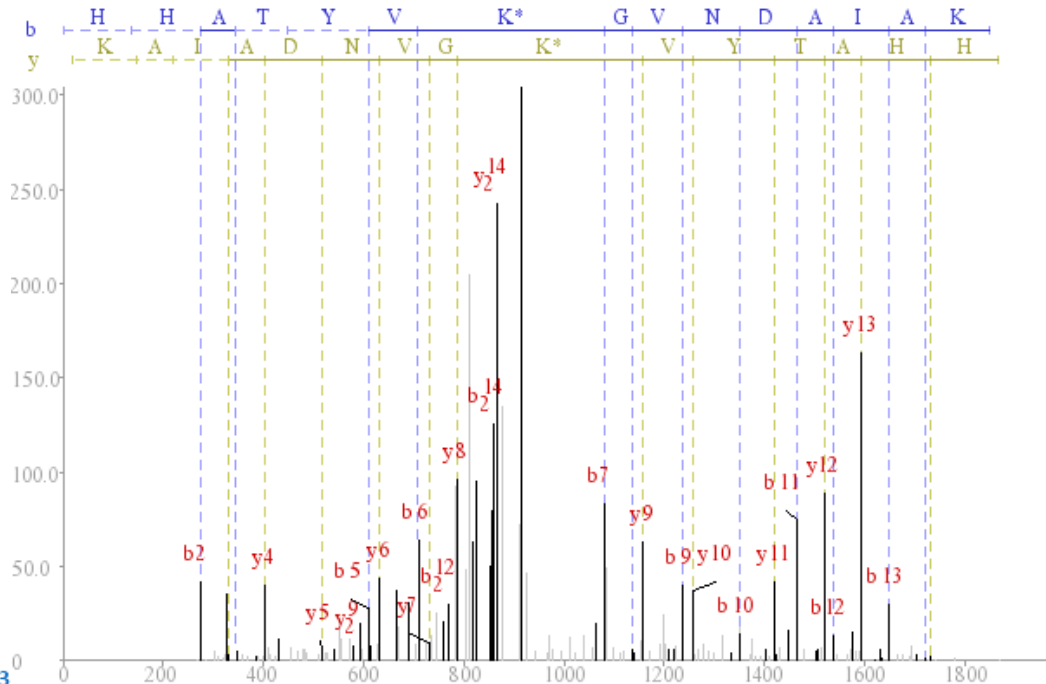
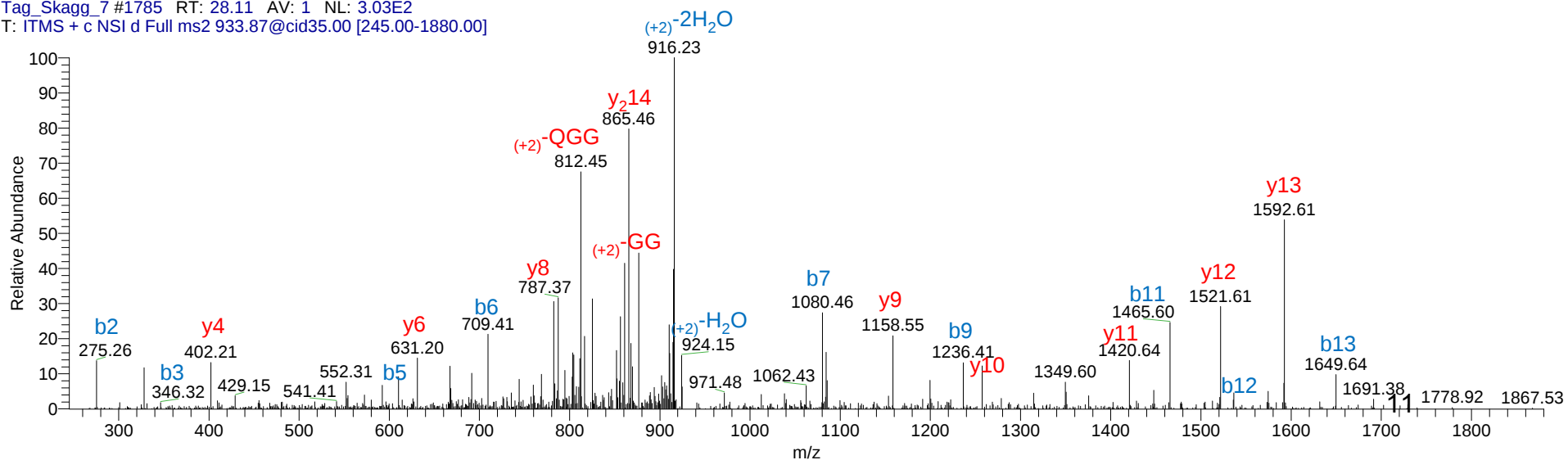
[K.HHATYVK+243GVNDIAIK.L](#)

Calc MW: 1866.94 Da

Charge = +2



Tag_Skagg_7 #1785 RT: 28.11 AV: 1 NL: 3.03E2
T: ITMS + c NSI d Full ms2 933.87@cid35.00 [245.00-1880.00]

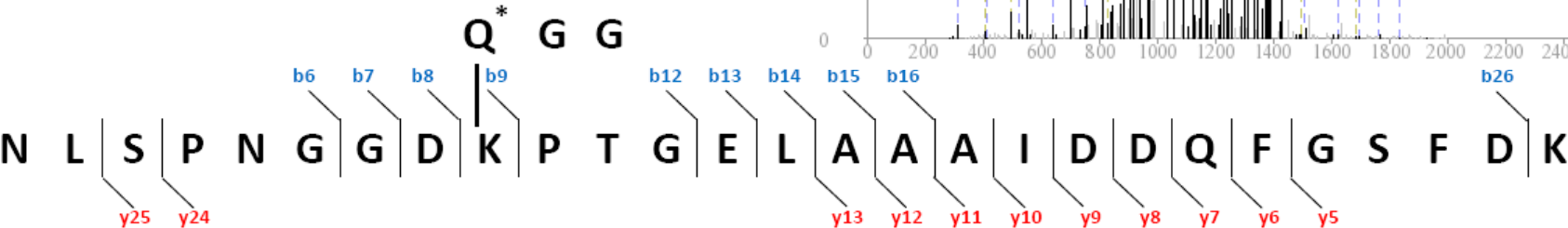


SF2: [Mn] Superoxide dismutase
(modification site 2)

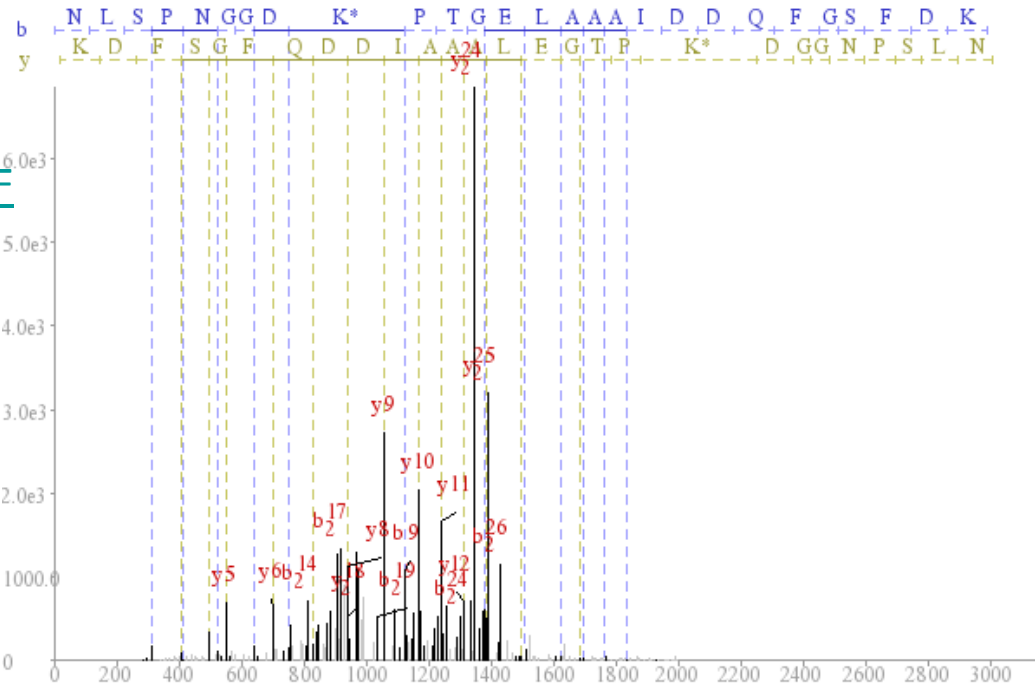
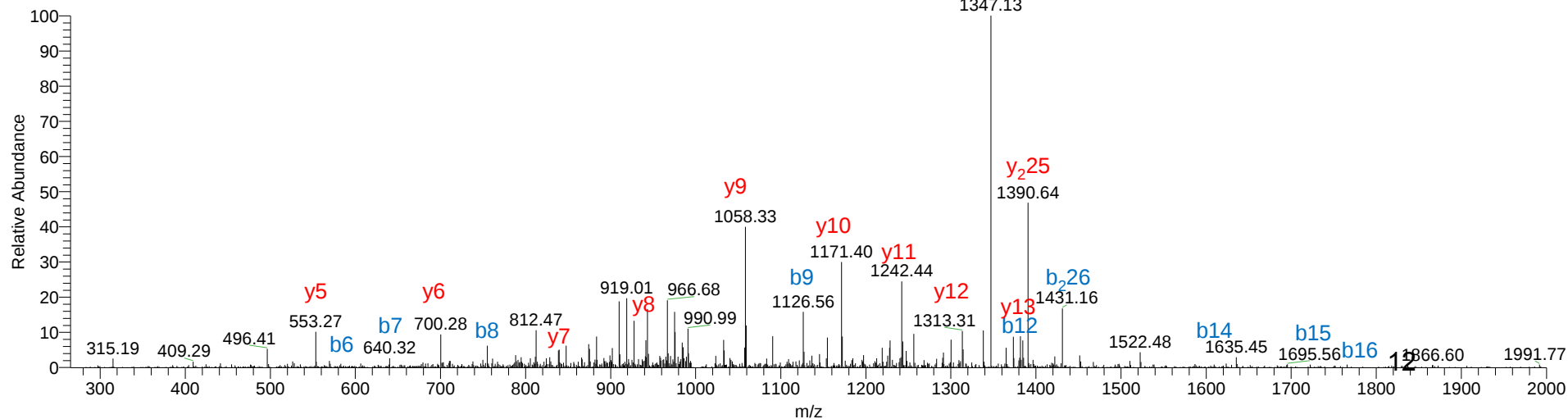
K.NLSPNGGDK+243PTGELAAAIDDQFGSFDK.F

Calc MW: 3007.38 Da

Charge = +3



LTQ_Tag2 #17937 RT: 85.88 AV: 1 NL: 6.83E3
T: ITMS + c NSI d Full ms2 1003.14@cid35.00 [265.00-2000.00]

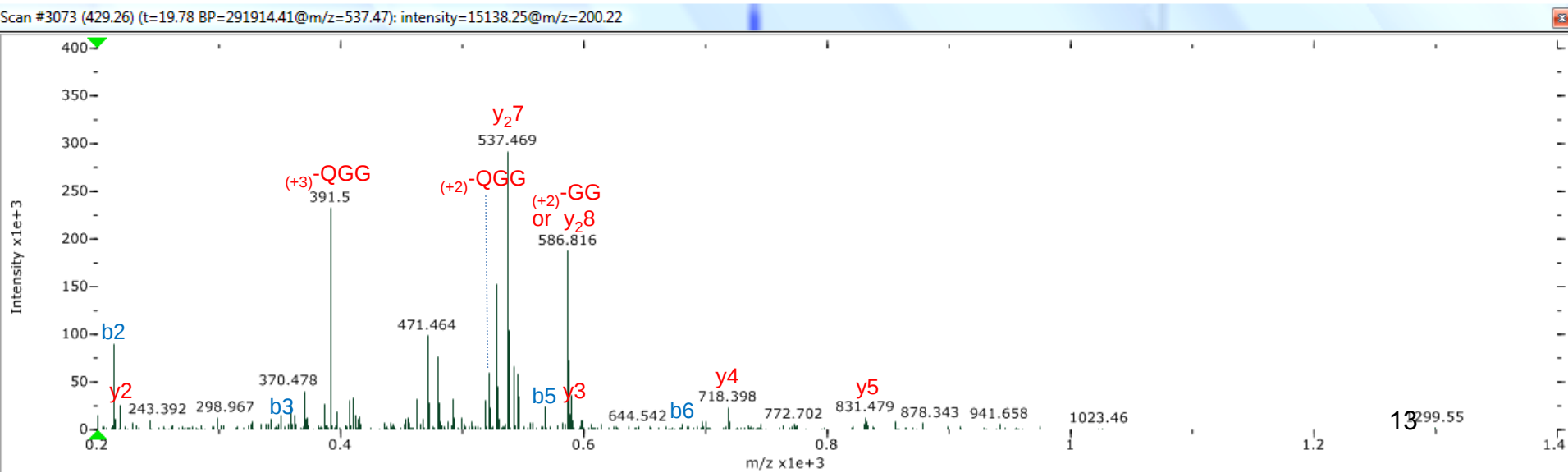
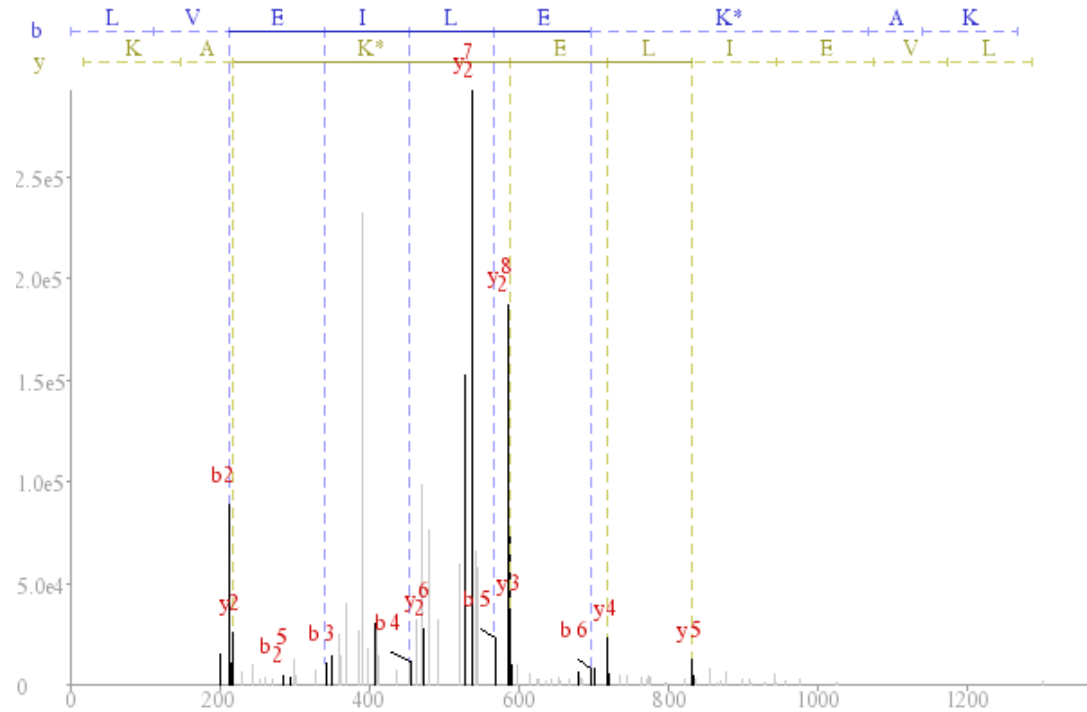
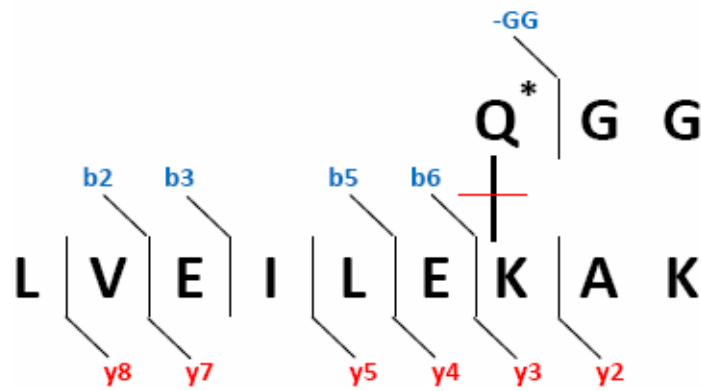


SF3: 30S ribosomal protein S17

R.LVEILEK+243AK.*

Calc MW: 1284.73 Da

Charge = +3

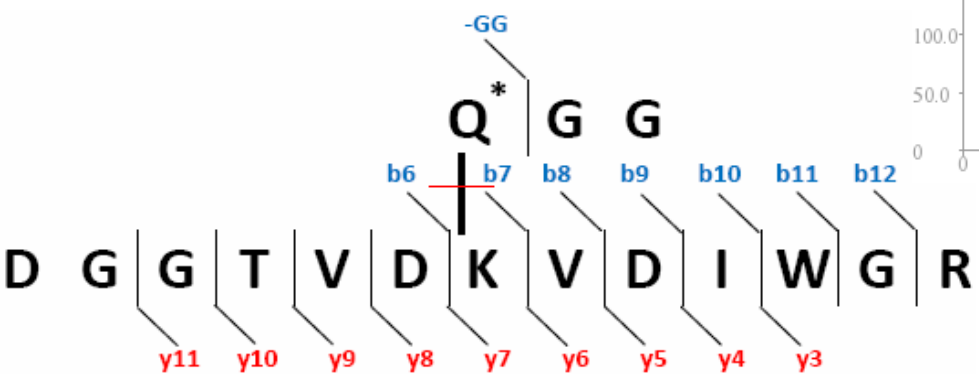


SF4: 30S ribosomal protein S6

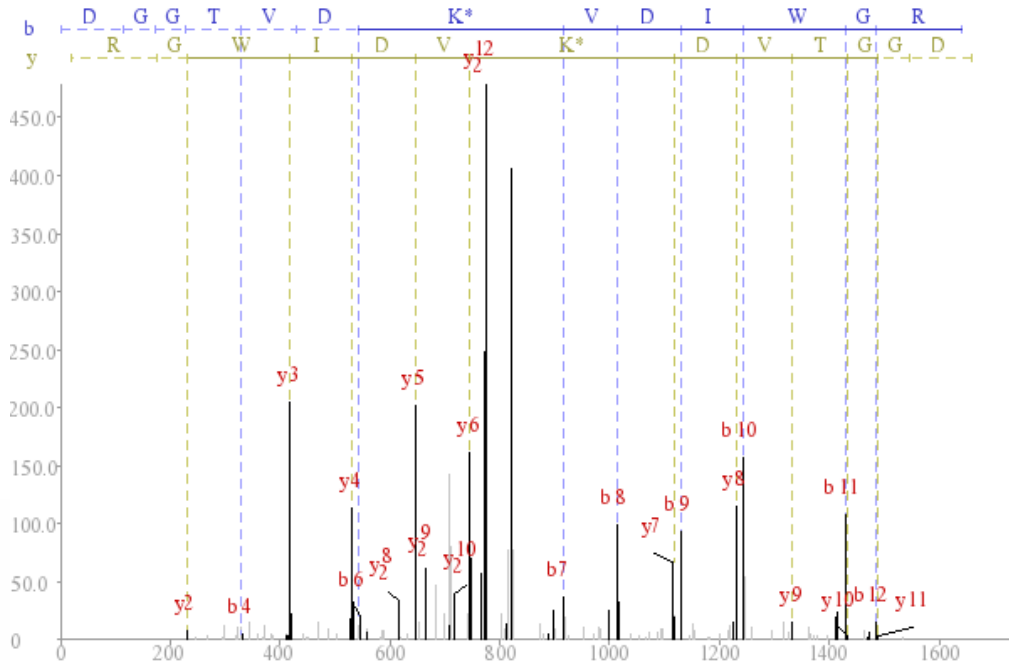
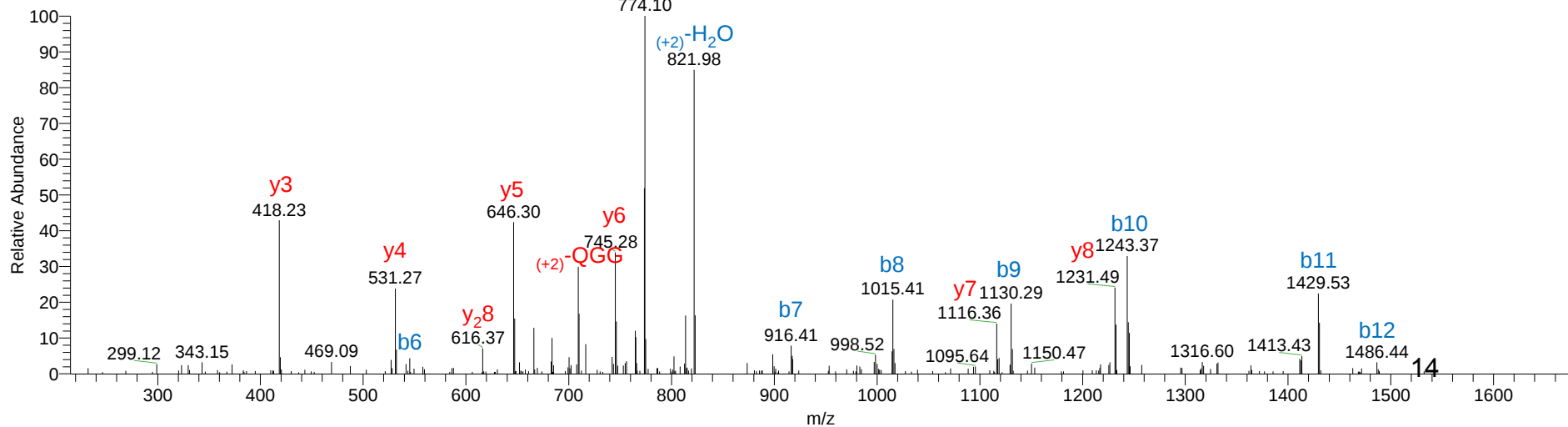
K.DGGTVDK+243VDIWGR.R

Calc MW: 1659.79 Da

Charge = +2



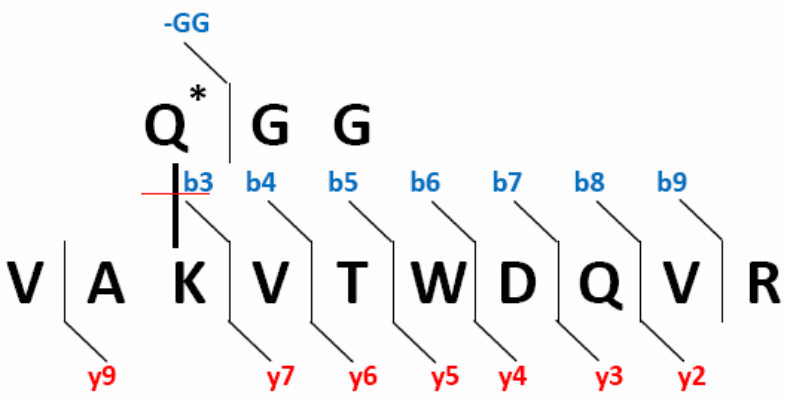
LTQ_Tag2 #12273 RT: 60.01 AV: 1 NL: 4.77E2
T: ITMS + c NSI d Full ms2 831.52@cid35.00 [215.00-1675.00]



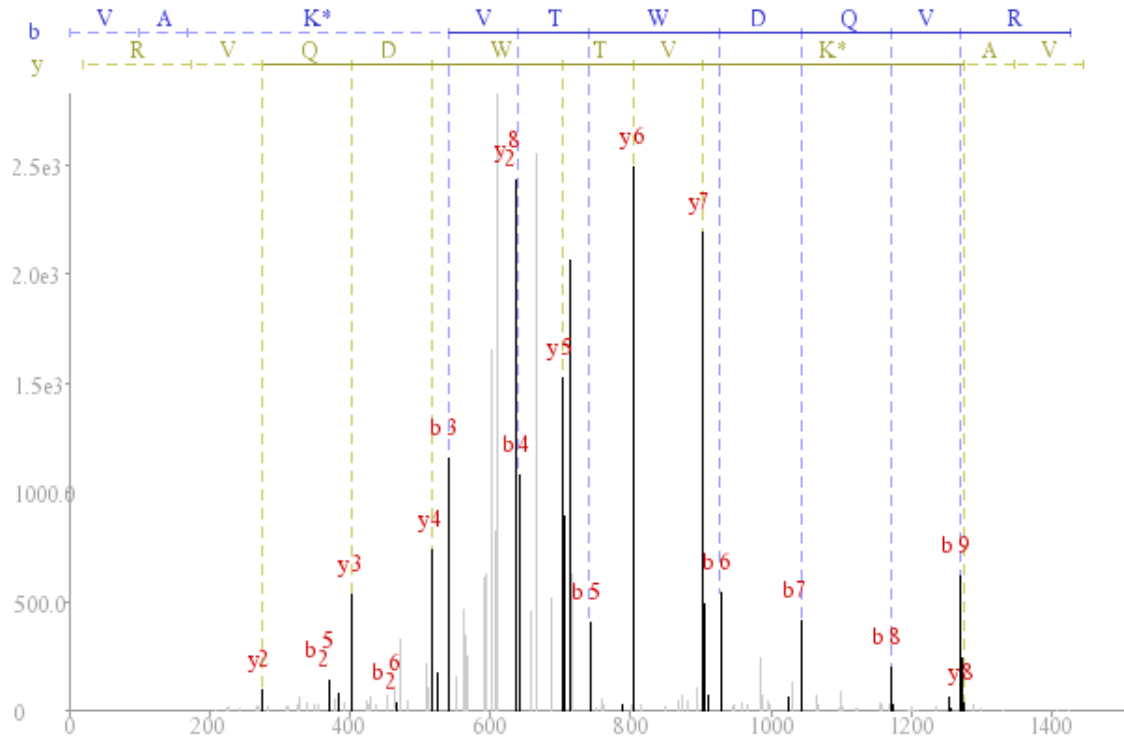
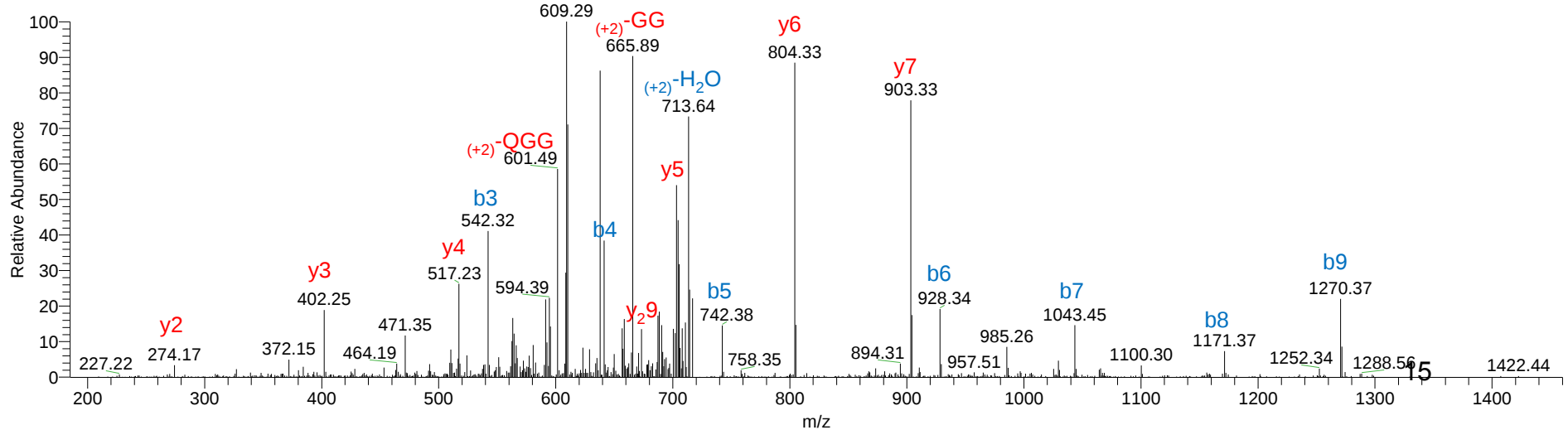
SF5: 50S ribosomal protein L11

K.VAK+243VTWDQVR.E

Calc MW: 1443.66 Da
Charge = +2



LTQ_Tag3 #6871 RT: 35.57 AV: 1 NL: 2.82E3
T: ITMS + c NSI d Full ms2 723.14@cid35.00 [185.00-1460.00]

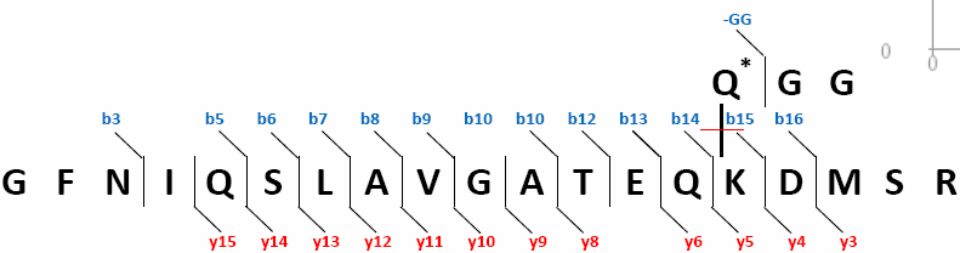
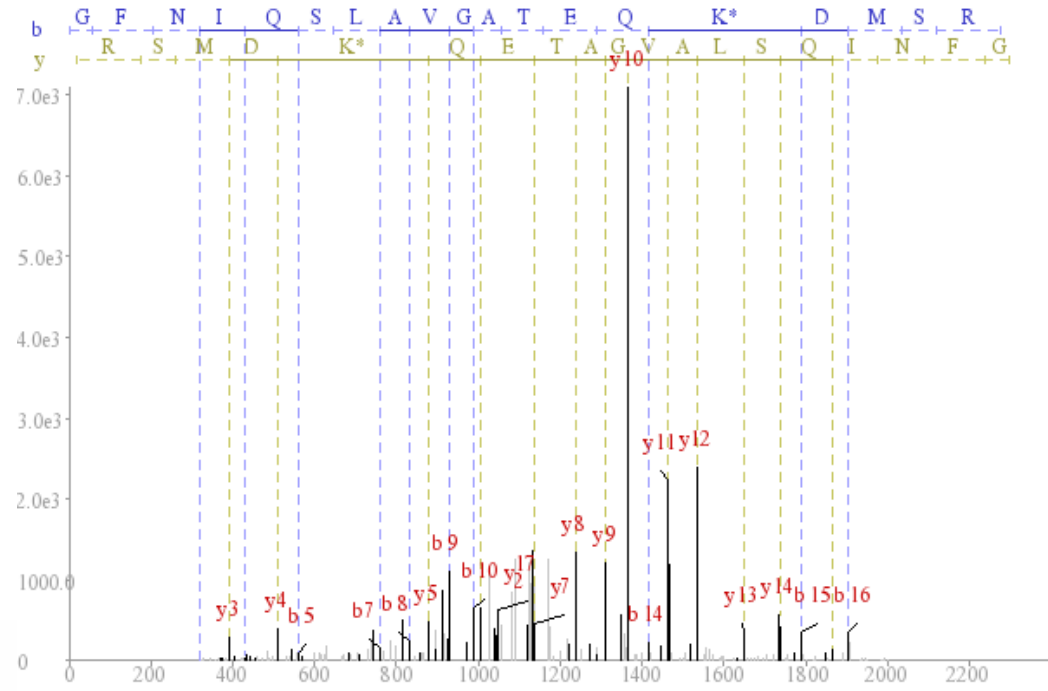


SF6: Acetolactate synthase 3 regulatory subunit

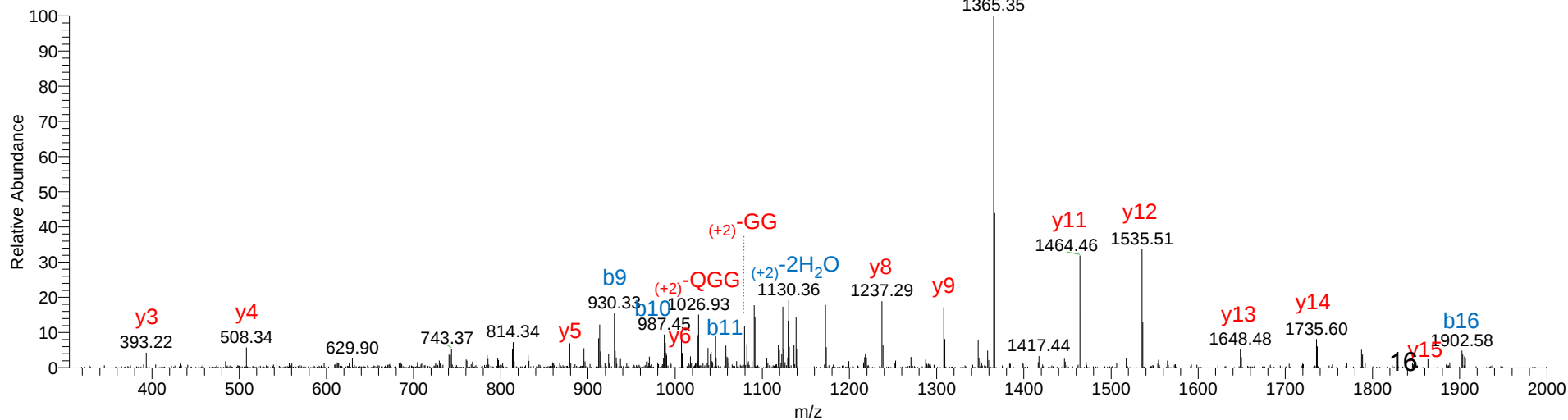
R.GFNIQSLAVGATEQK+243DMSR.M

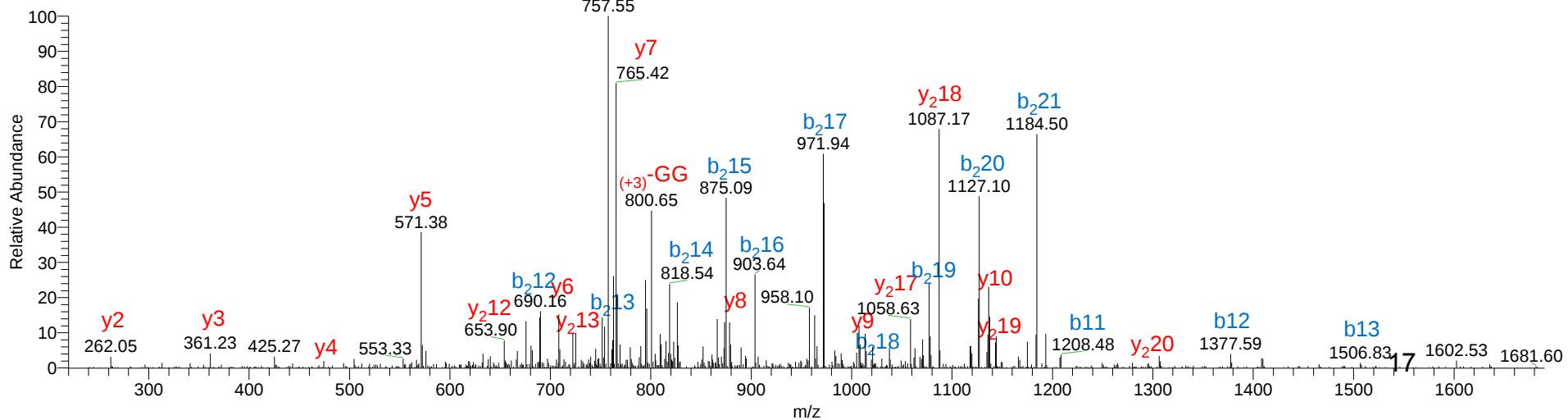
Calc MW: 2294.10 Da

Charge = +2



LTQ_Tag2 #16269 RT: 78.59 AV: 1 NL: 7.08E3
T: ITMS + c NSI d Full ms2 1148.26@cid35.00 [305.00-2000.00]



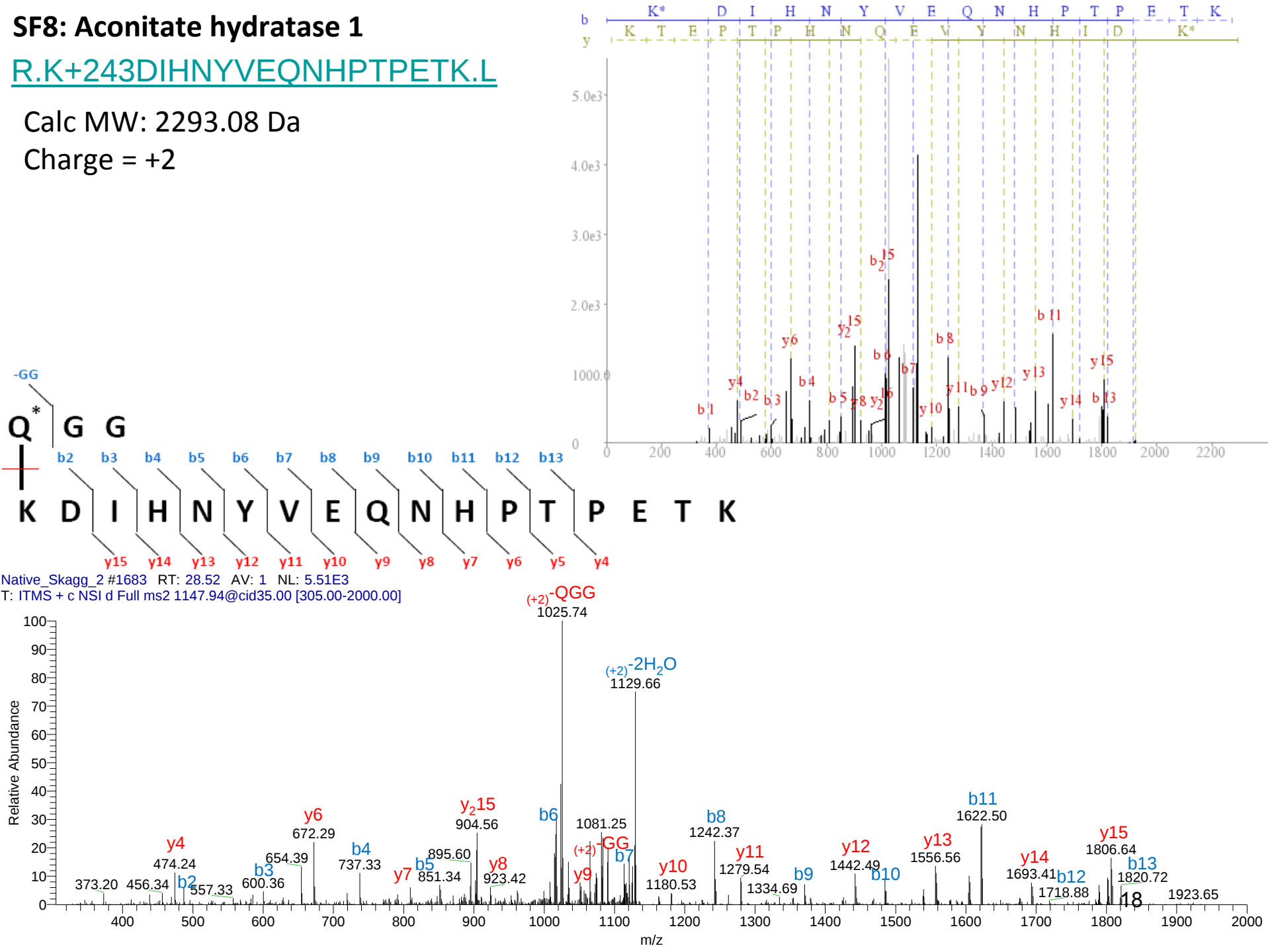


SF8: Aconitate hydratase 1

R.K+243DIHNYVEQNHPETPK.L

Calc MW: 2293.08 Da

Charge = +2

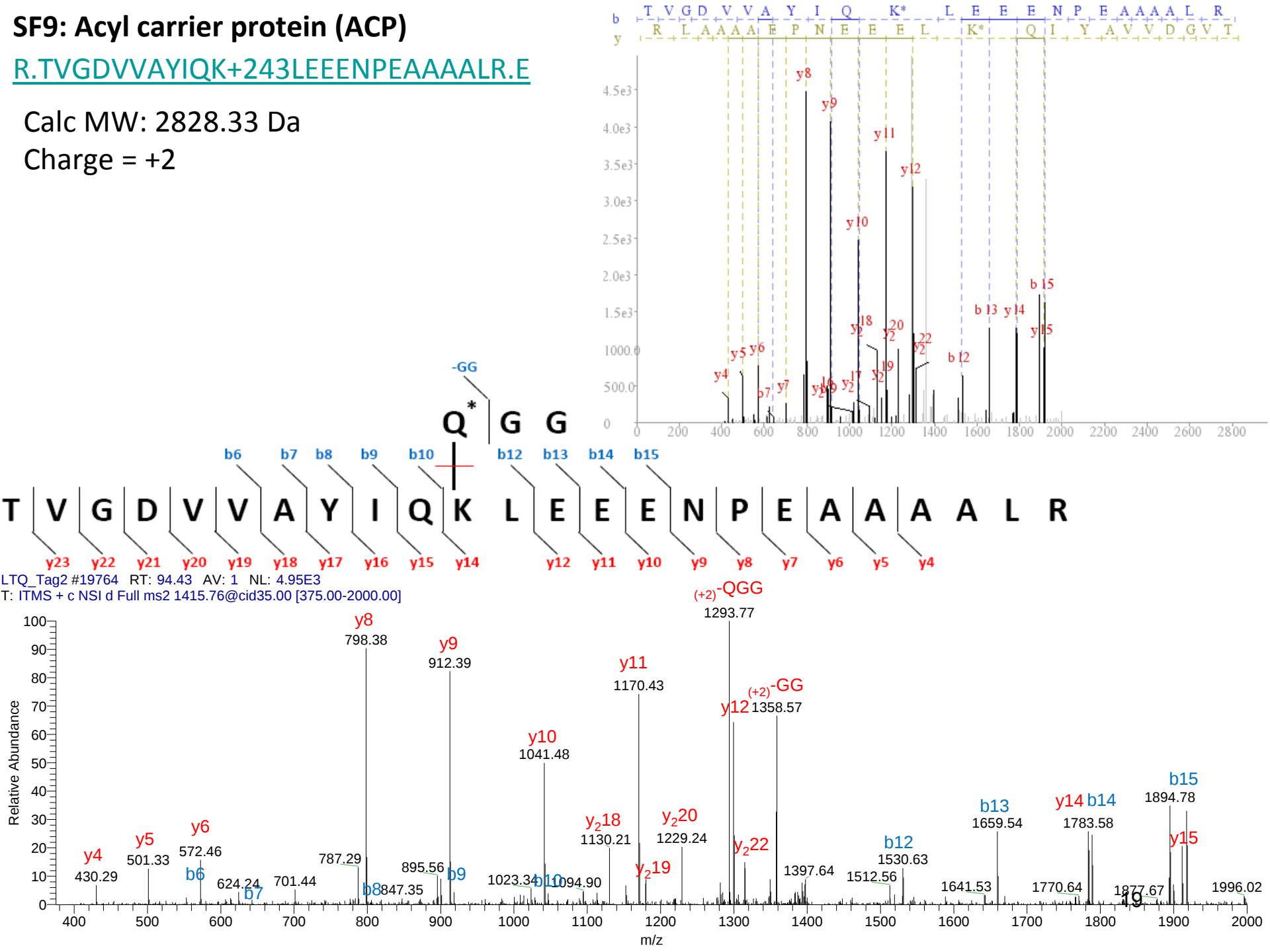


SF9: Acyl carrier protein (ACP)

R.TVGDVVAYIQK+243LEENPEAAAALR.E

Calc MW: 2828.33 Da

Charge = +2

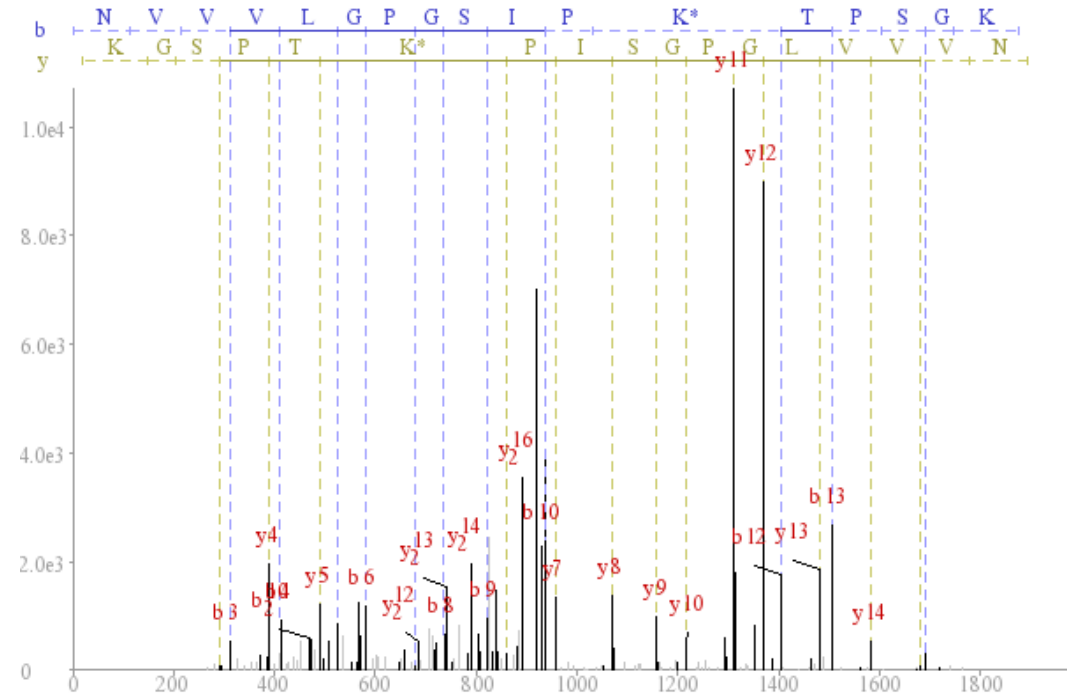
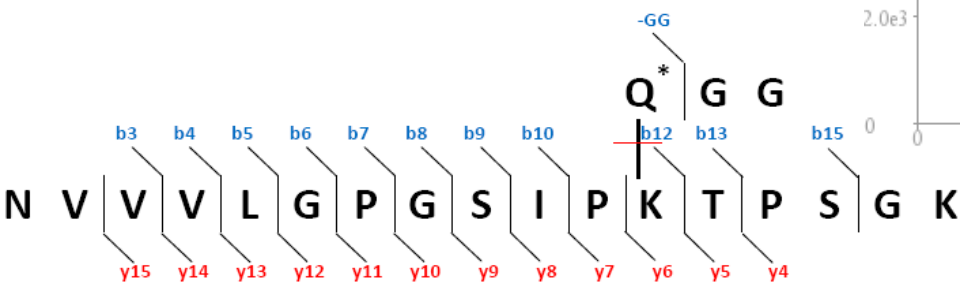


SF10: Acyl-CoA synthetase

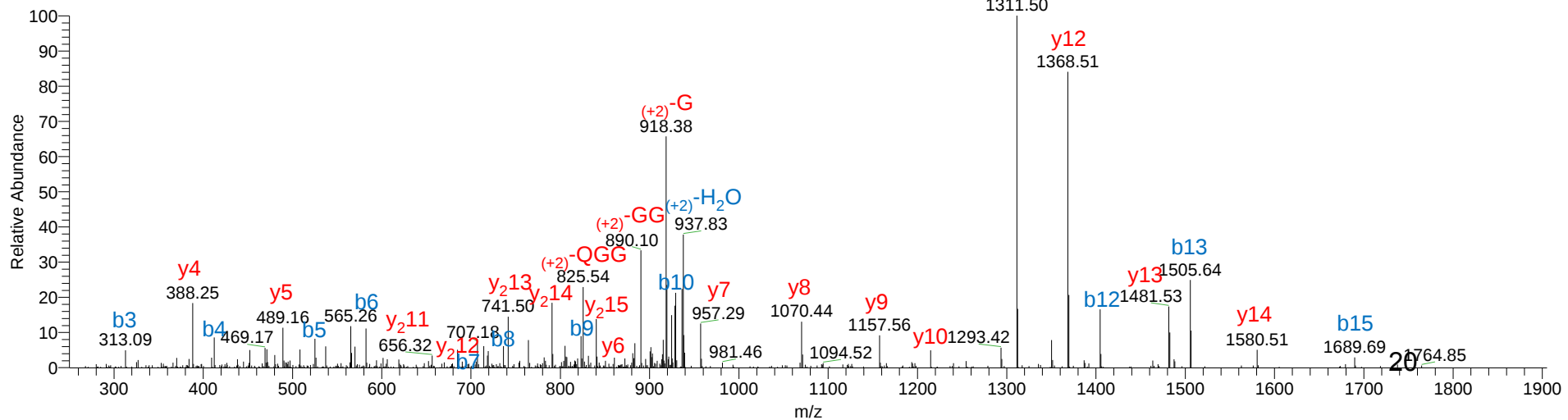
R.NVVVLGPGSIPK+243TPSGK.L

Calc MW: 1892.04 Da
Charge = +2

-Manual annotation puts
modification on indicated
lysine, as opposed to the
terminal residue.



LTQ_Native3 #7660 RT: 39.51 AV: 1 NL: 1.07E4
T: ITMS + c NSI d Full ms2 946.94@cid35.00 [250.00-1905.00]

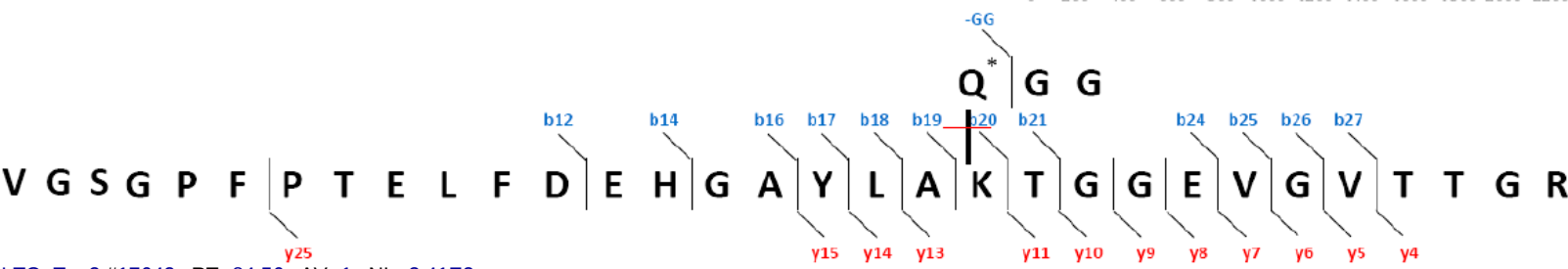
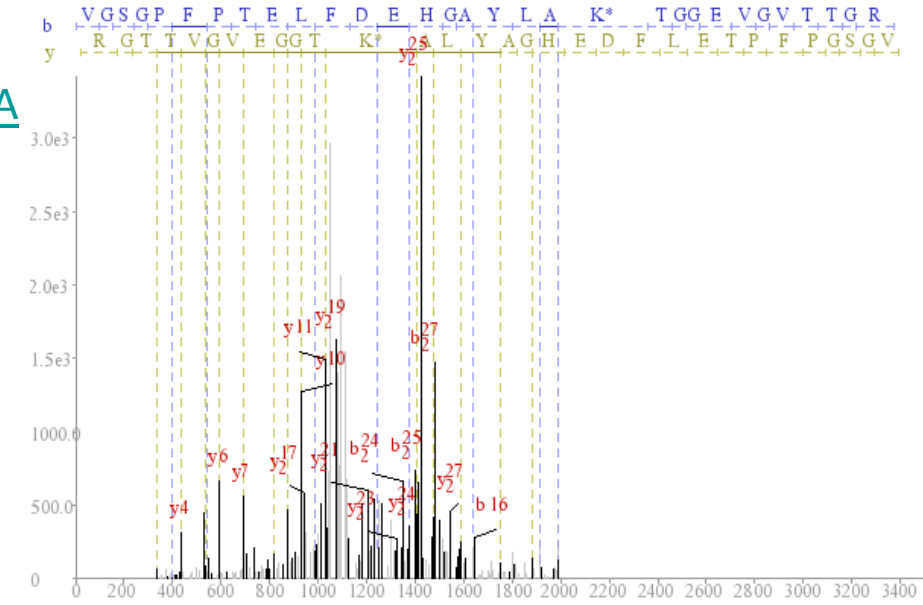


SF11: Adenylosuccinate synthetase

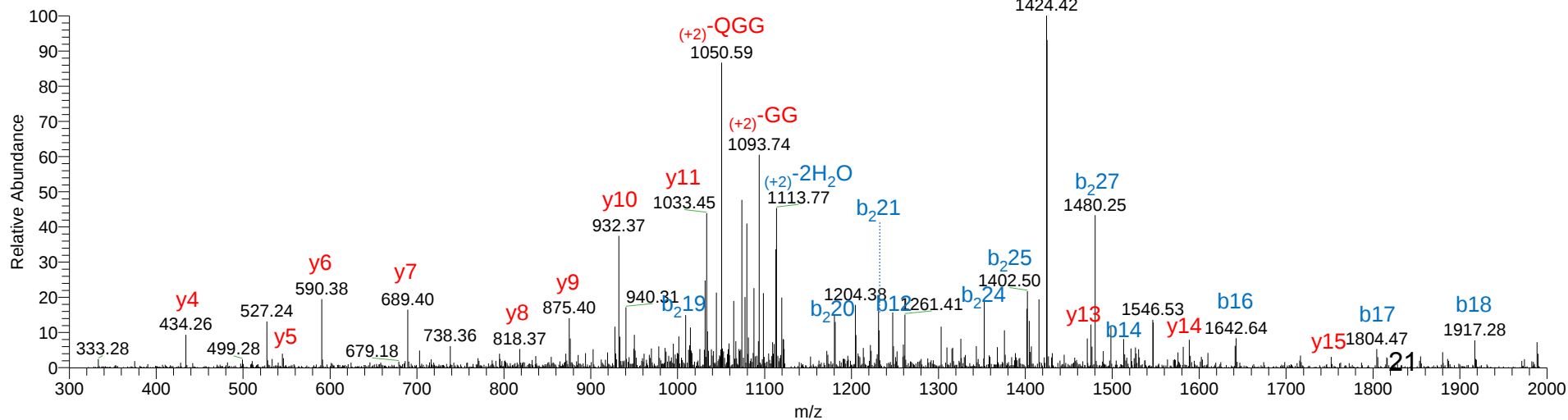
R.VGSGPFPTELFDEHGAYLAK+243TGGEVGVTTGR.A

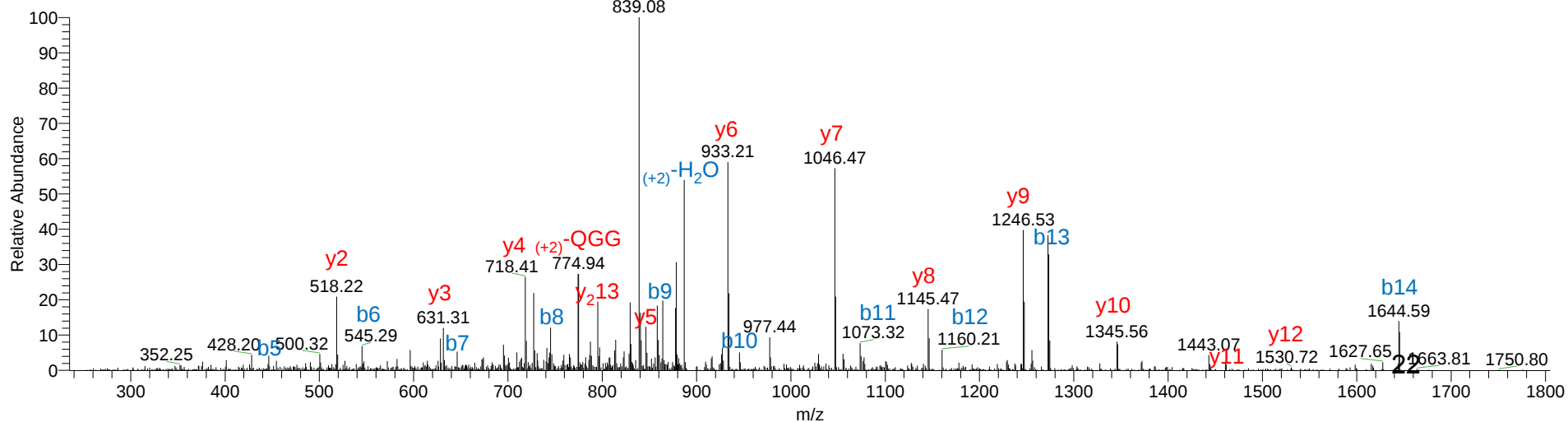
Calc MW: 3391.63 Da

Charge = +3



LTQ_Tag2 #17642 RT: 84.56 AV: 1 NL: 3.41E3
T: ITMS + c NSI d Full ms2 1131.58@cid35.00 [300.00-2000.00]



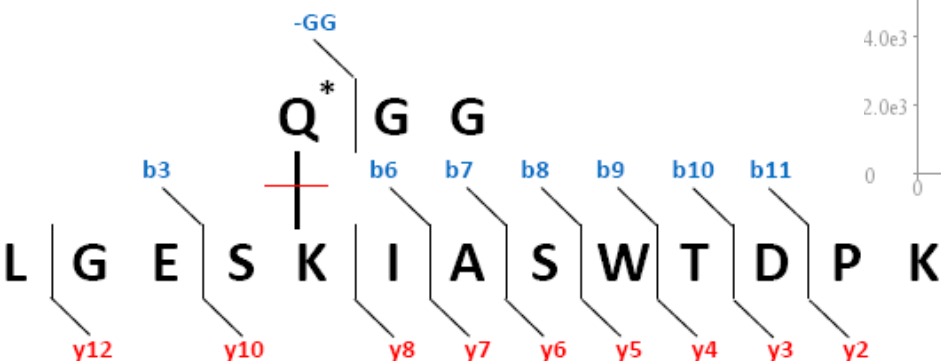


SF13: Aspartate transaminase

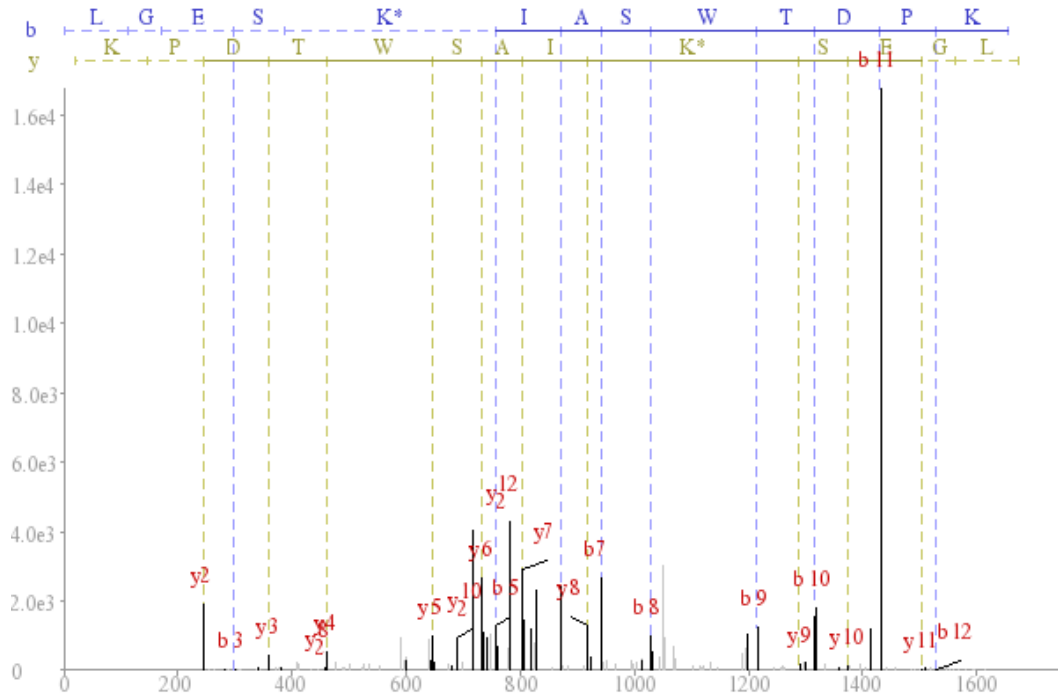
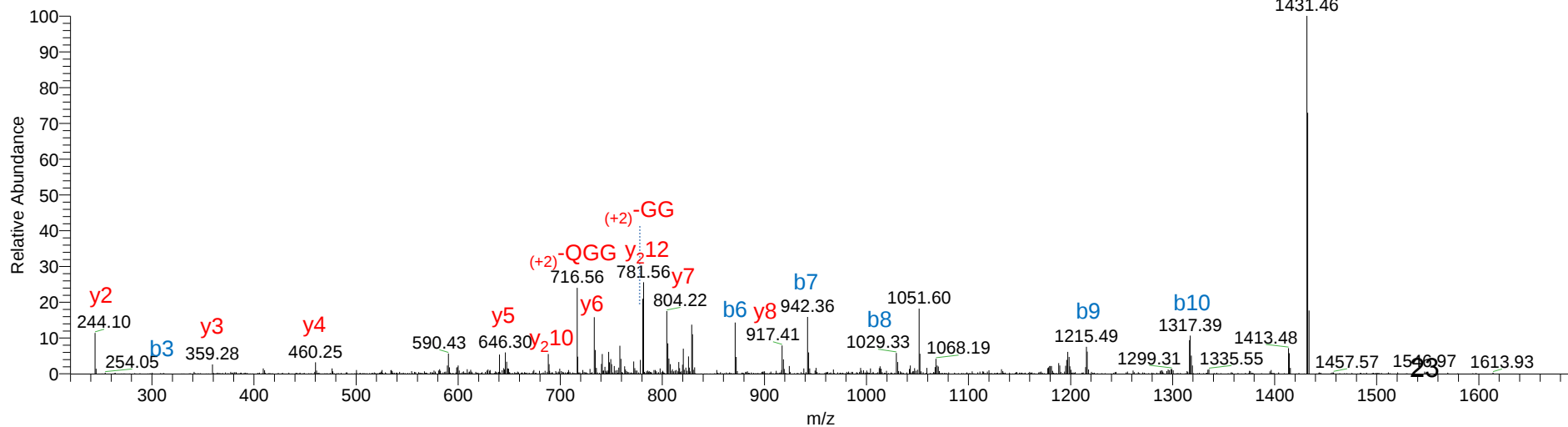
R.LGESK+243IASWTDPK.G

Calc MW: 1673.82 Da

Charge = +2



LTQ_Tag2 #9716 RT: 47.64 AV: 1 NL: 1.67E4
T: ITMS + c NSI d Full ms2 838.46@cid35.00 [220.00-1690.00]

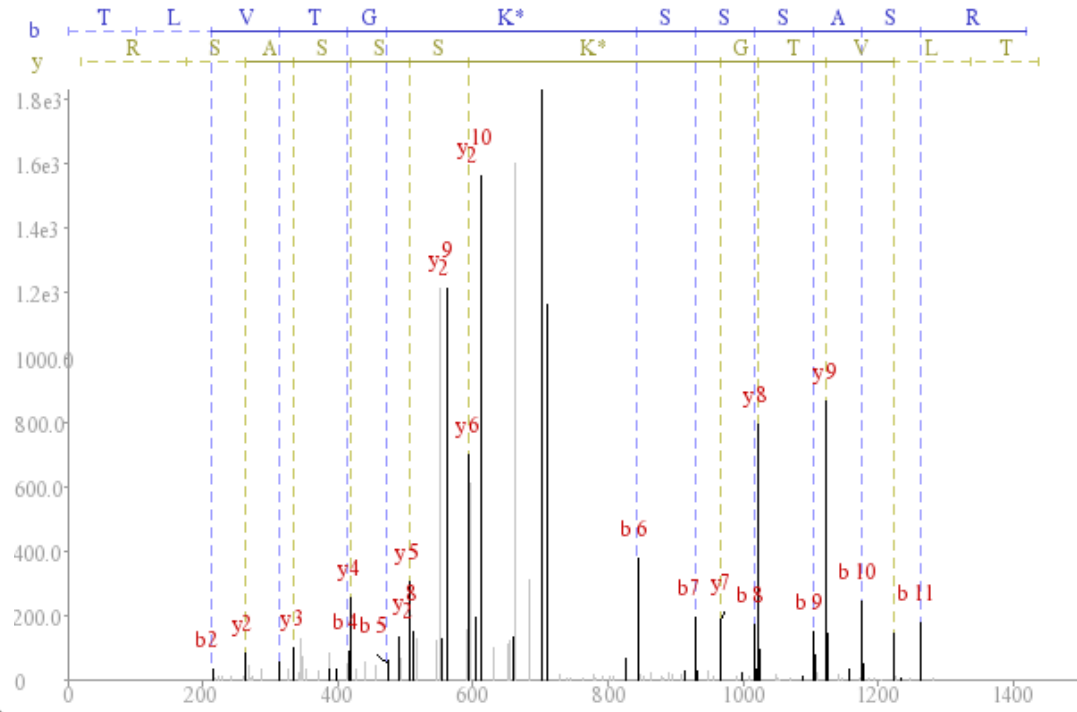
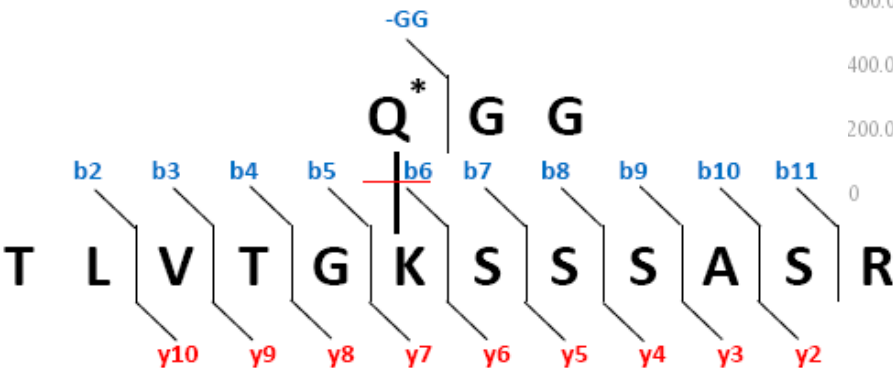


SF14: ATPase, AAA family protein

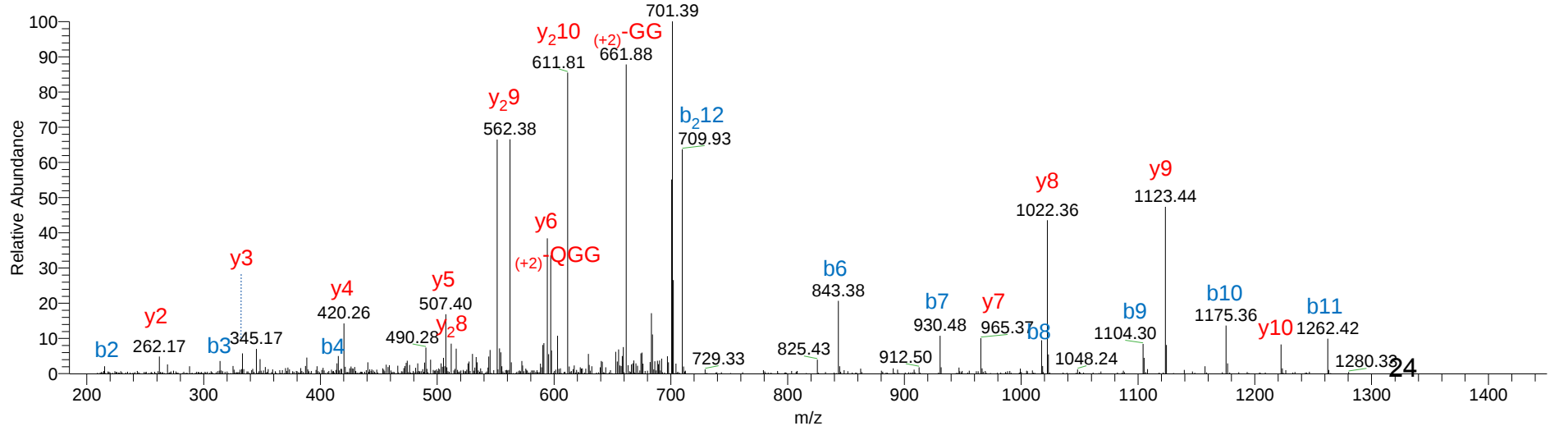
R.TLVTGK+243SSSASR.A

Calc MW: 1435.64 Da

Charge = +2



LTQ_Native2 #6828 RT: 34.94 AV: 1 NL: 1.82E3
T: ITMS + c NSI d Full ms2 719.21@cid35.00 [185.00-1450.00]



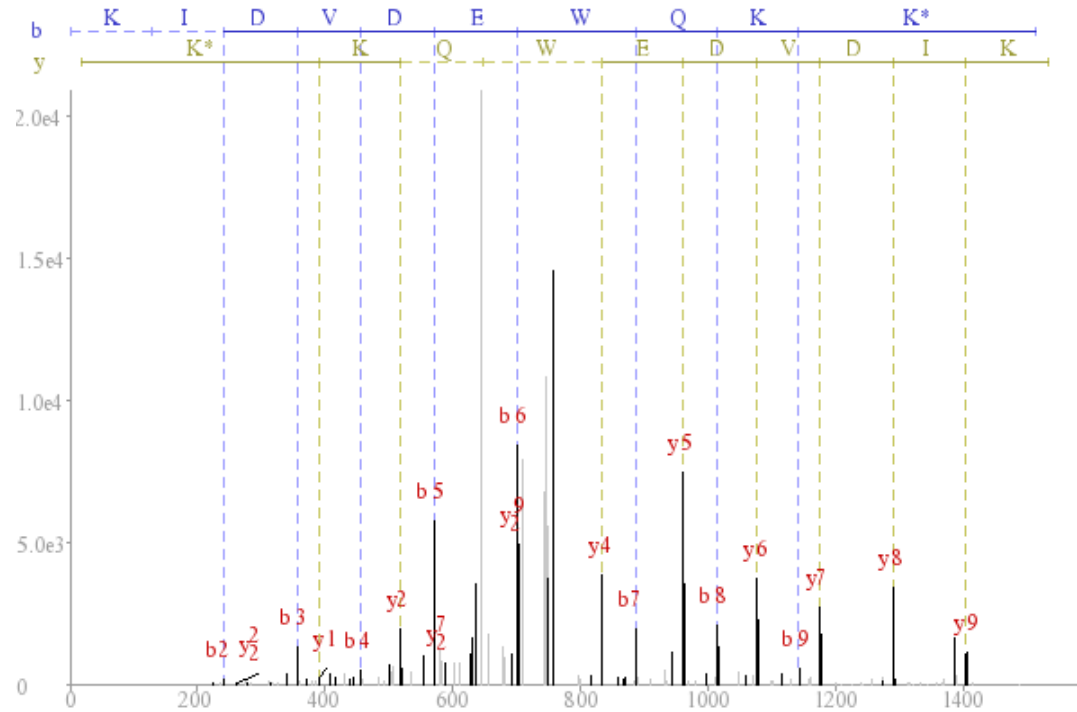
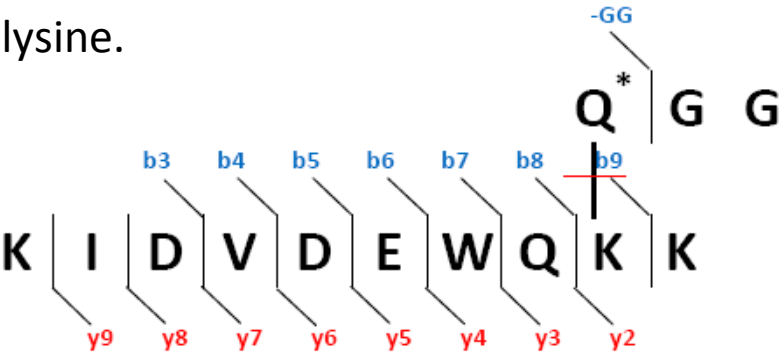
SF15: Branched-chain amino acid
aminotransferase

R.KIDVDEWQK+243K.A

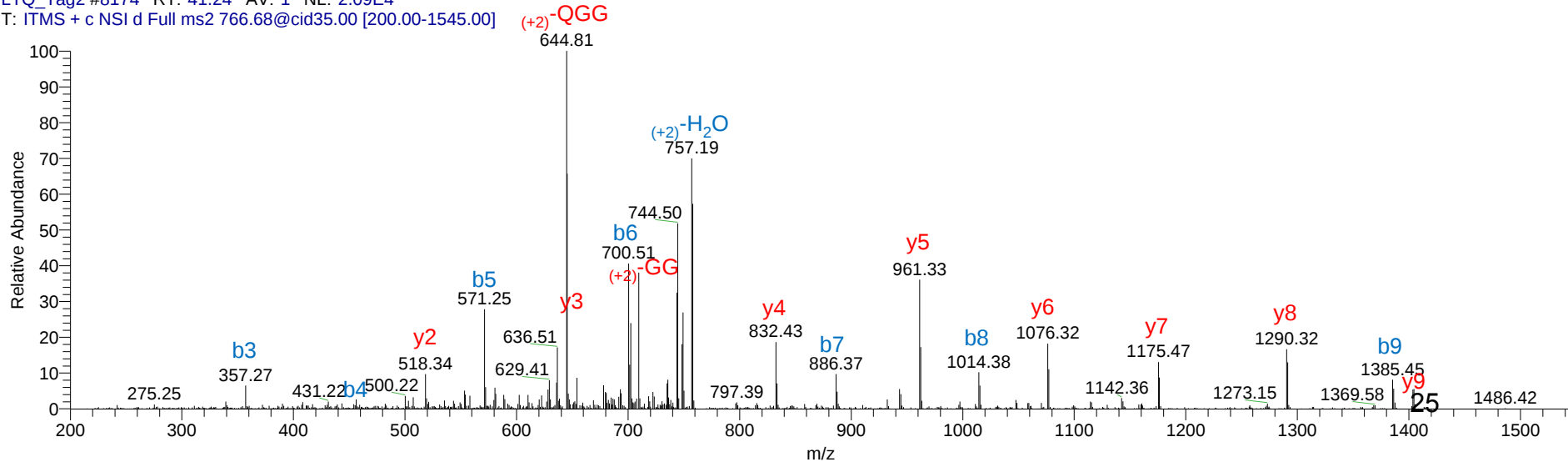
Calc MW: 1531.977Da

Charge = +2

-Modified b9 ion and no
modified y1 ion suggests
modification in on first
lysine.



LTQ_Tag2 #8174 RT: 41.24 AV: 1 NL: 2.09E4
T: ITMS + c NSI d Full ms2 766.68@cid35.00 [200.00-1545.00]

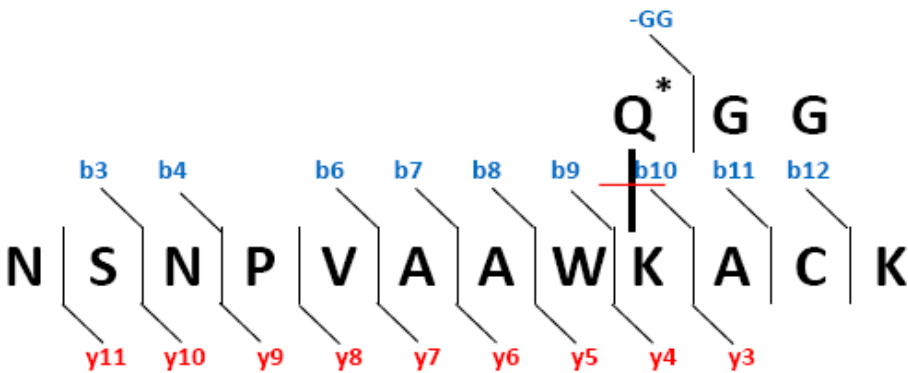


SF16: Carbonic anhydrase

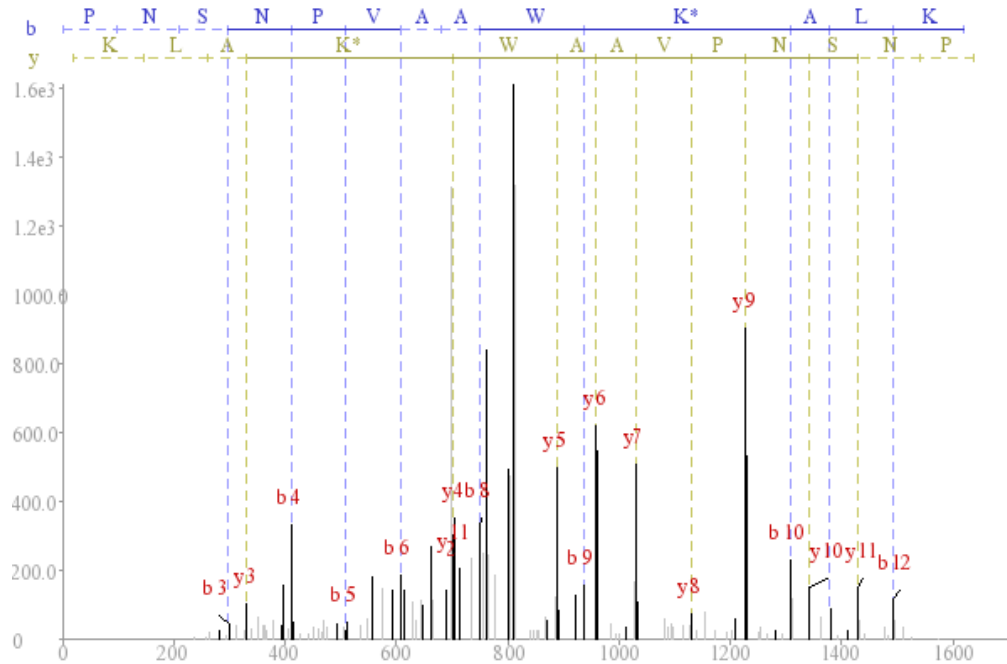
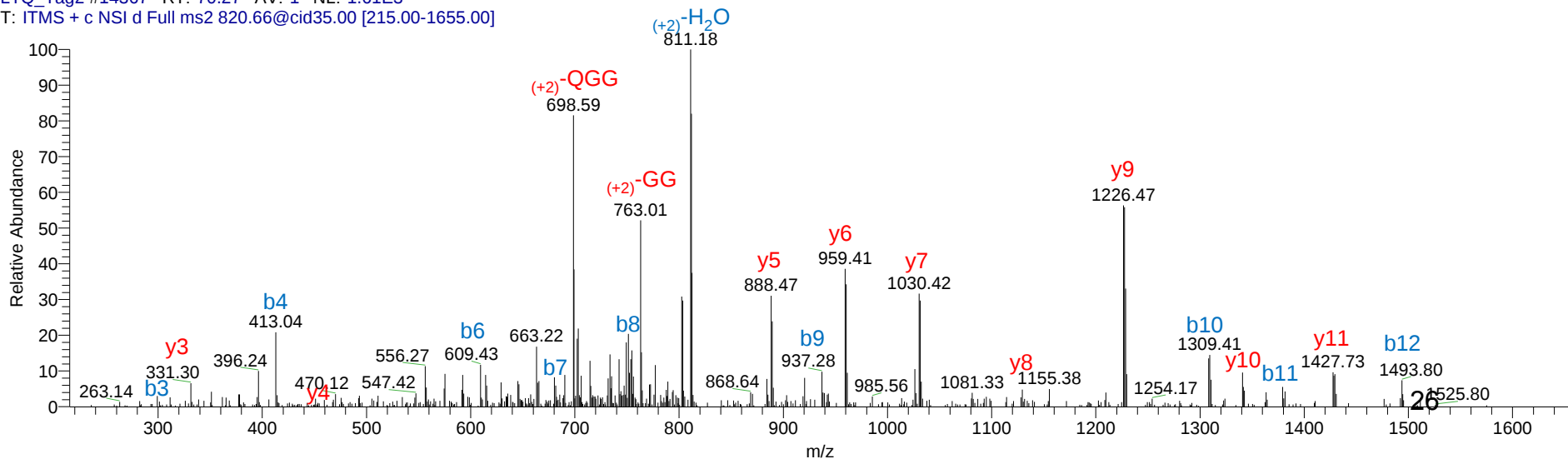
M.PNSNPVAAWK+243ALK.D

Calc MW: 1637.85 Da

Charge = +2



LTQ_Tag2 #14367 RT: 70.27 AV: 1 NL: 1.61E3
T: ITMS + c NSI d Full ms2 820.66@cid35.00 [215.00-1655.00]

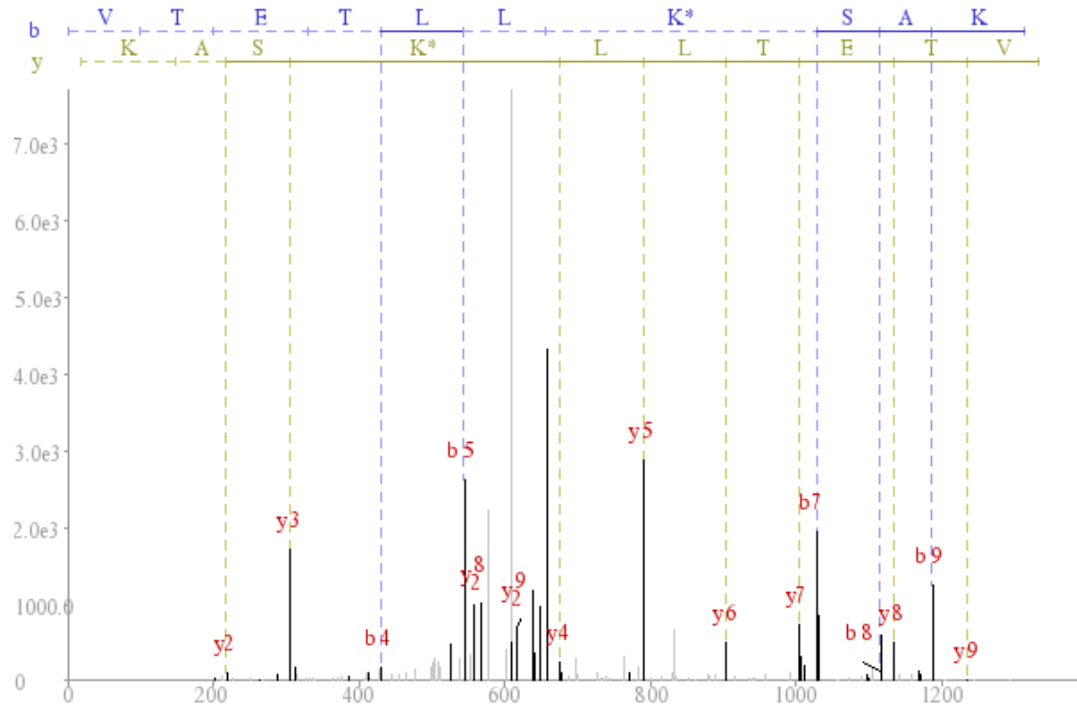
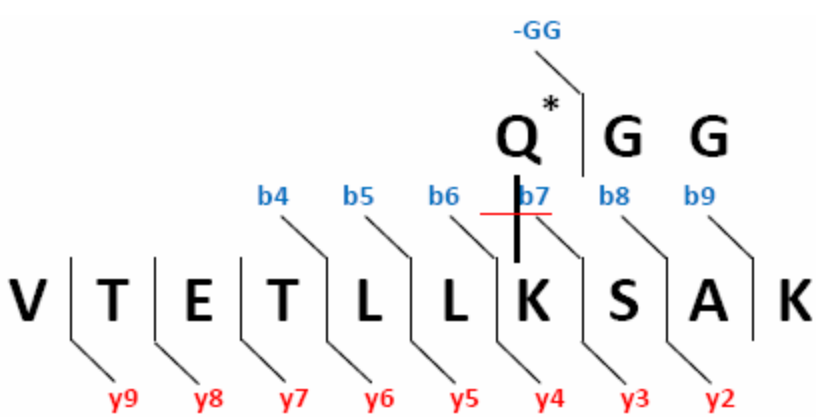


SF17: Chaperonin GroEL

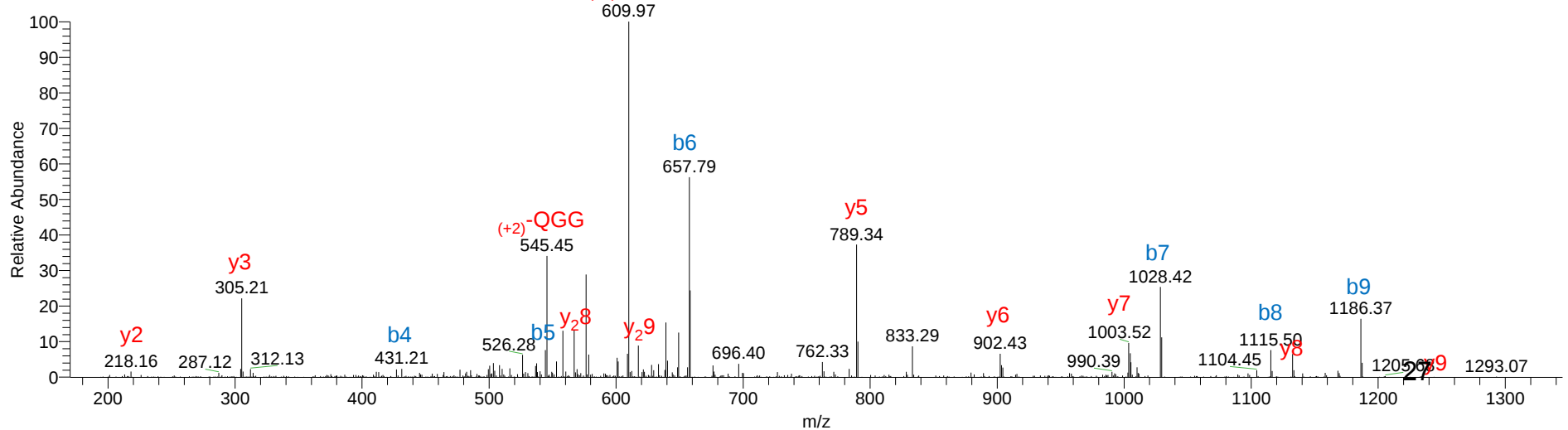
K.VTETLLK+243SAK.E

Calc MW: 1331.73 Da

Charge = +2



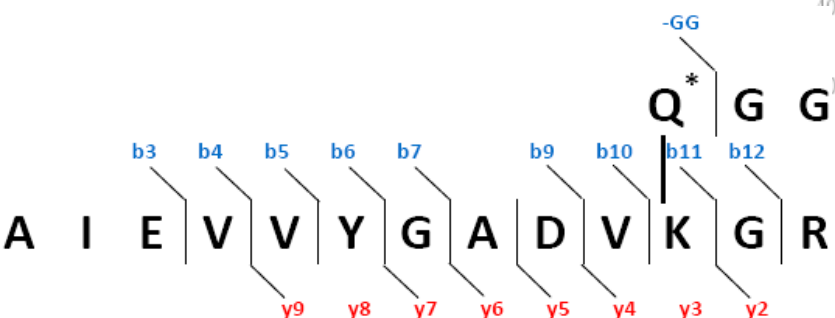
His_1_081107175250 #18414 RT: 63.71 AV: 1 NL: 7.69E3
T: ITMS + c NSI d Full ms2 666.80@cid35.00 [170.00-1345.00]



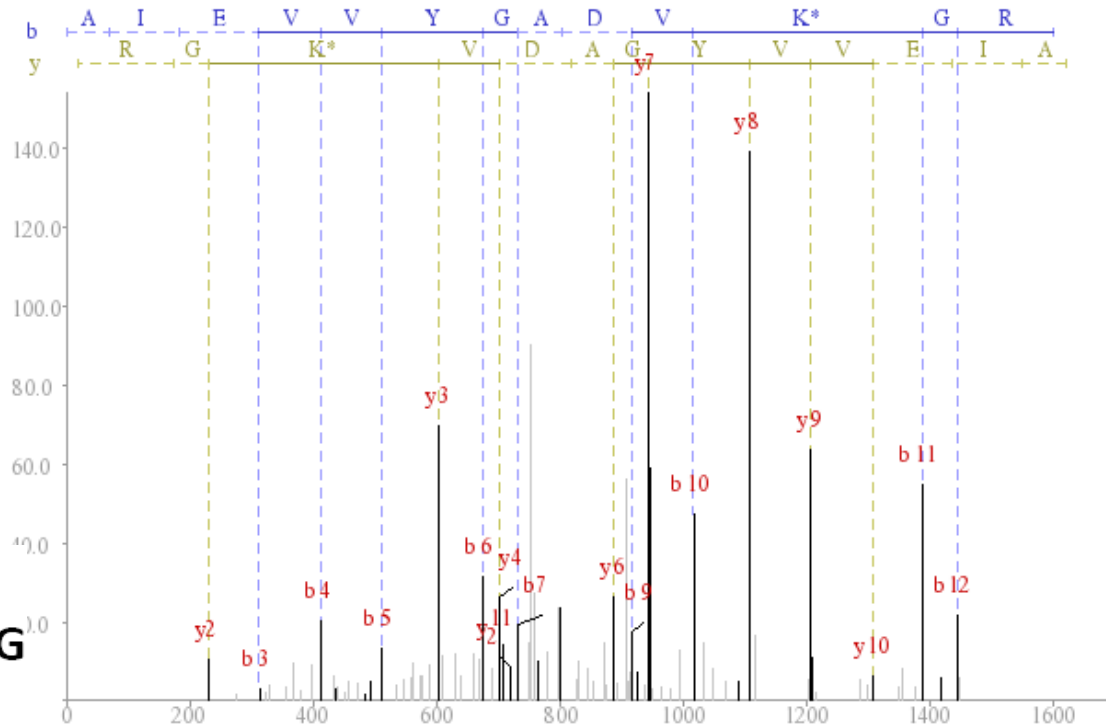
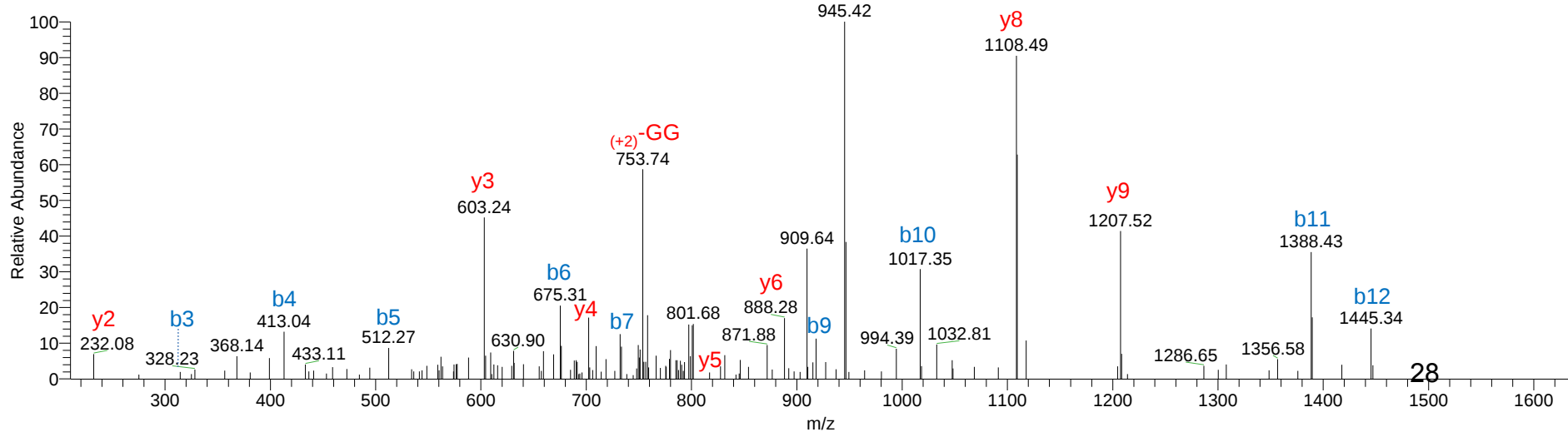
SF18: Conserved hypothetical protein
(MSMEG_4692)

R.AIEVVYGADV K+243GR.G

Calc MW: 1618.83 Da
Charge = +2



LTQ_Tag3 #13677 RT: 69.52 AV: 1 NL: 1.54E2
T: ITMS + c NSI d Full ms2 810.53@cid35.00 [210.00-1635.00]

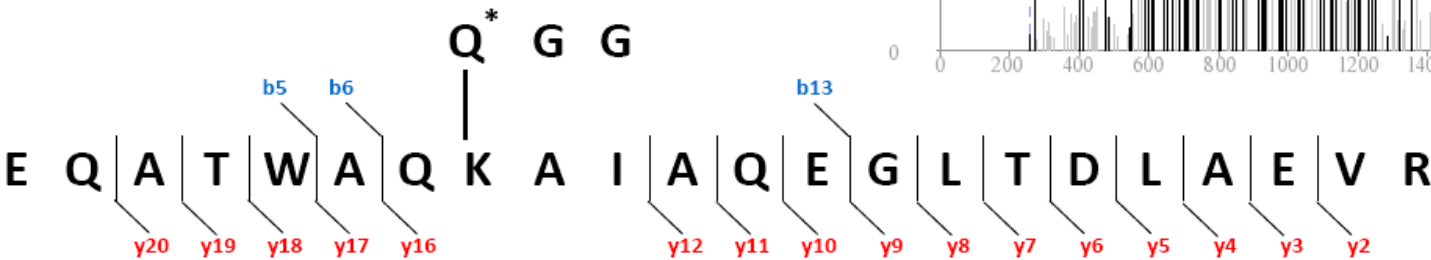


SF19: Cyclopropane-fatty-acyl-phospholipid synthase

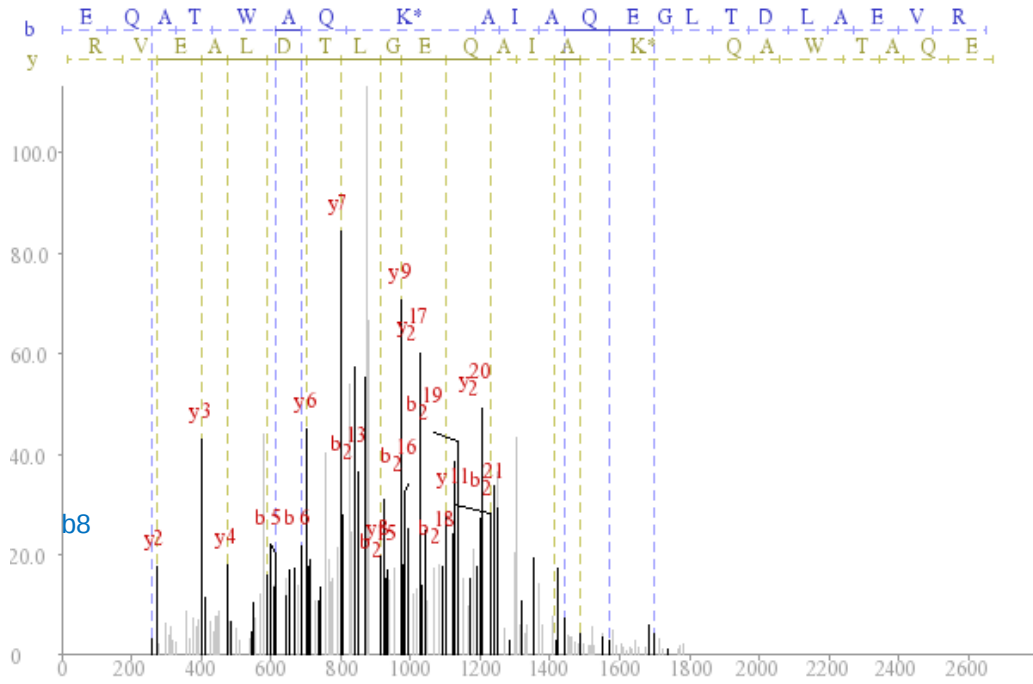
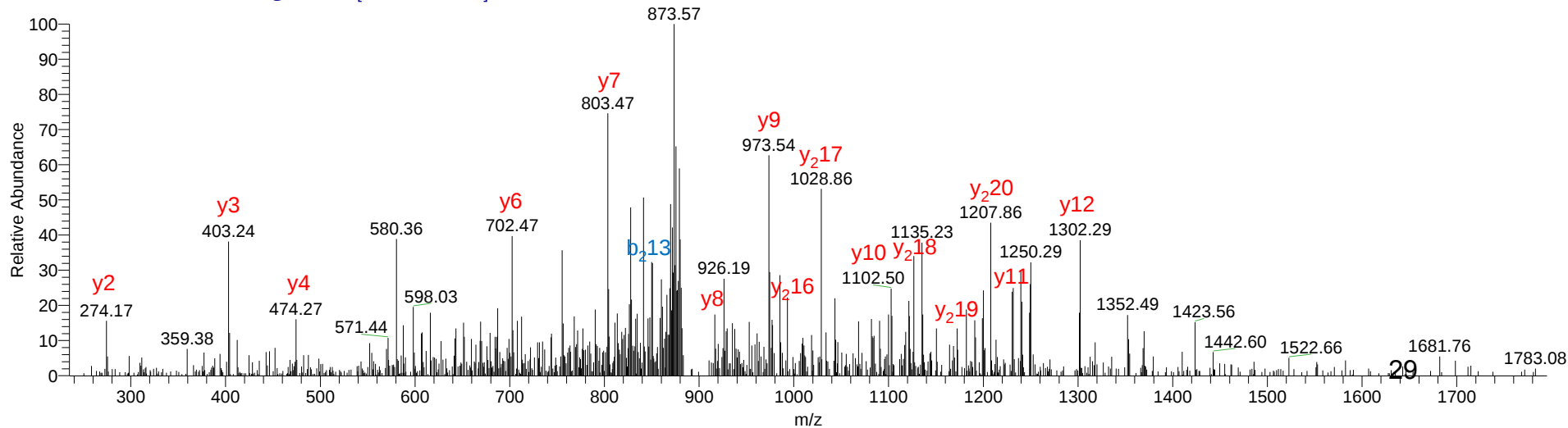
R.EQATWAQK+243AIAQEGLTDLAEV.R.H

Calc MW: 2670.32 Da

Charge = +3



081206_Smeg_Pieter_B_LTQ_2 #6877 RT: 46.86 AV: 1 NL: 1.13E2
T: ITMS + c NSI d Full ms2 891.86@cid35.00 [235.00-1795.00]

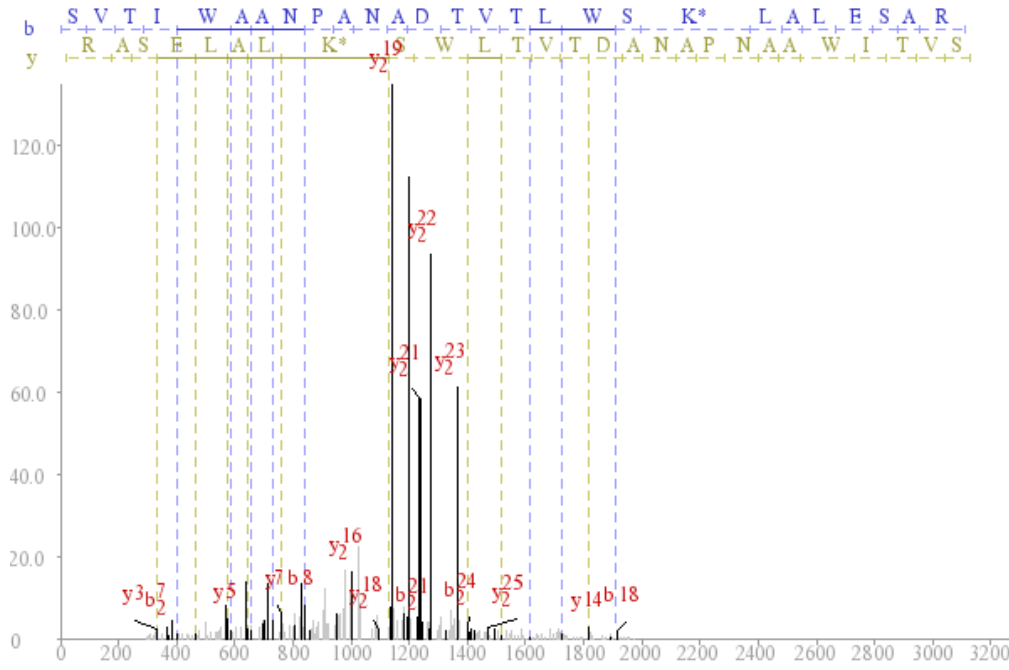


SF20: DNA topoisomerase IV subunit B

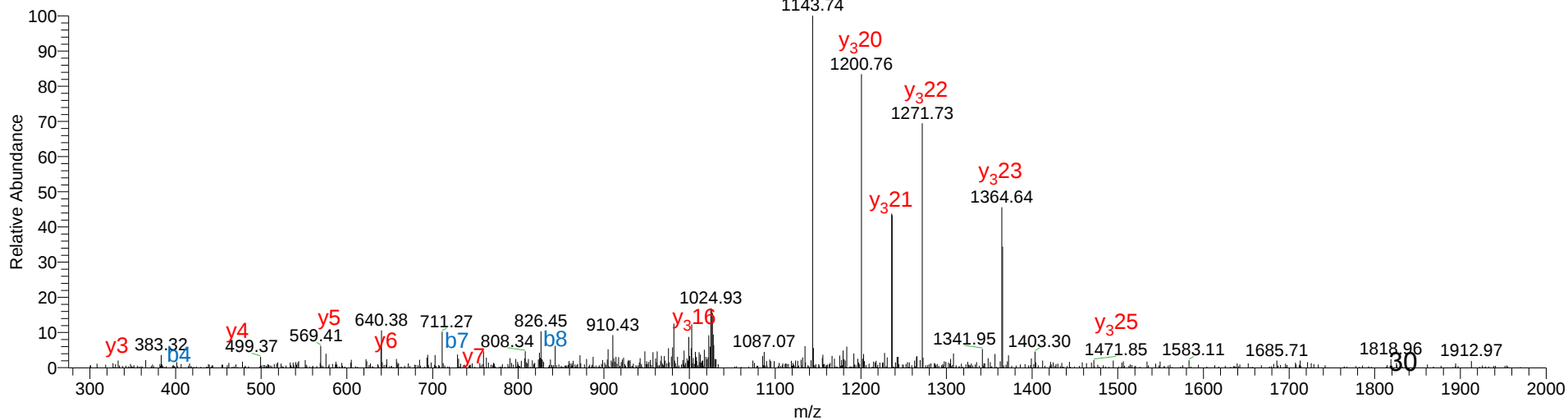
R.SVTIWAANPANADTVTLWSK+243LALESAR.A

Calc MW: 3127.59 Da

Charge = +3



081206_Smeg_Pieter_C_LTQ_Blank #8506 RT: 49.48 AV: 1 NL: 1.34E2
T: ITMS + c NSI d Full ms2 1044.26@cid35.00 [275.00-2000.00]

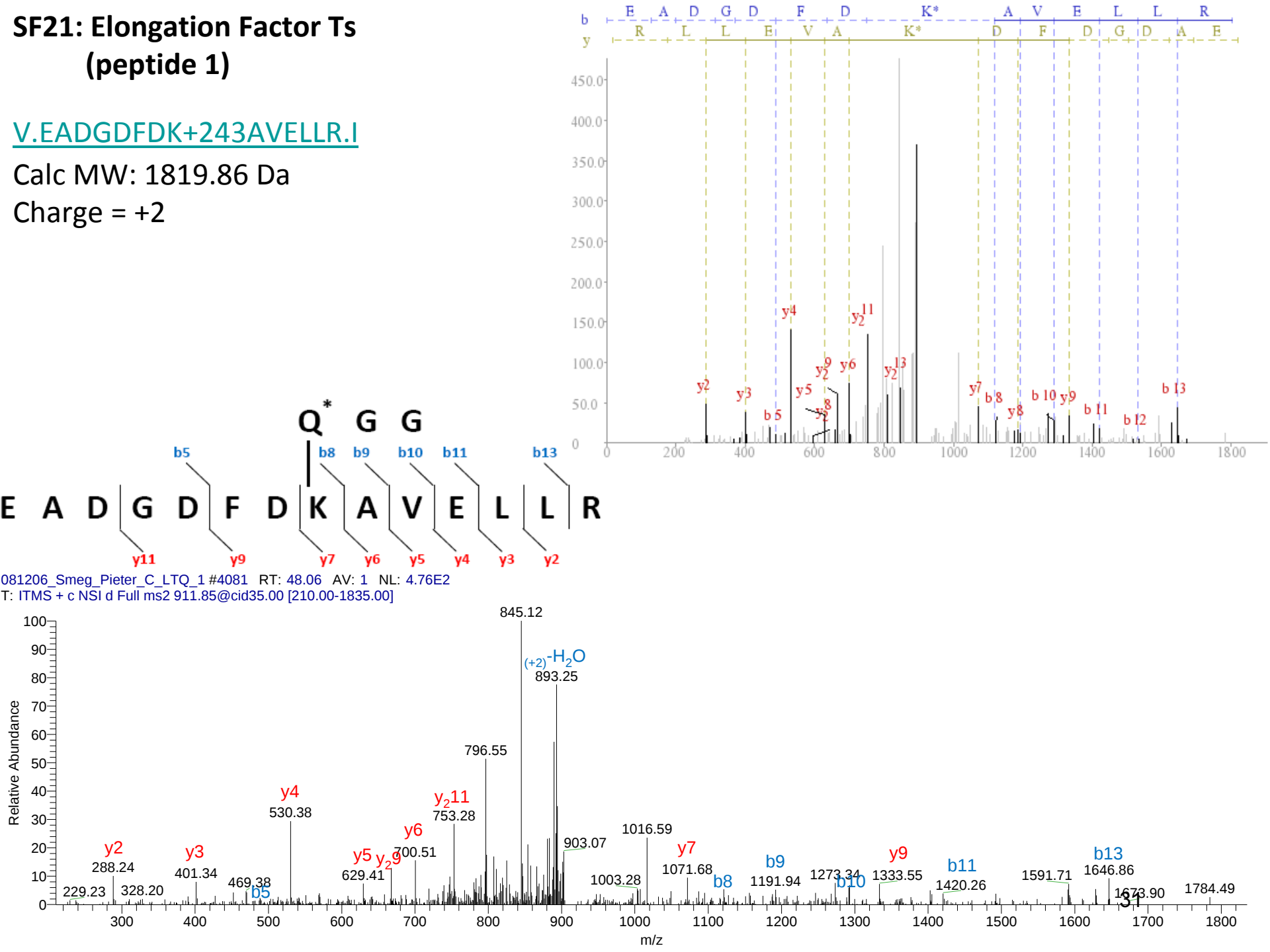


SF21: Elongation Factor Ts
(peptide 1)

V.EADGDFDK+243AVELLR.I

Calc MW: 1819.86 Da

Charge = +2

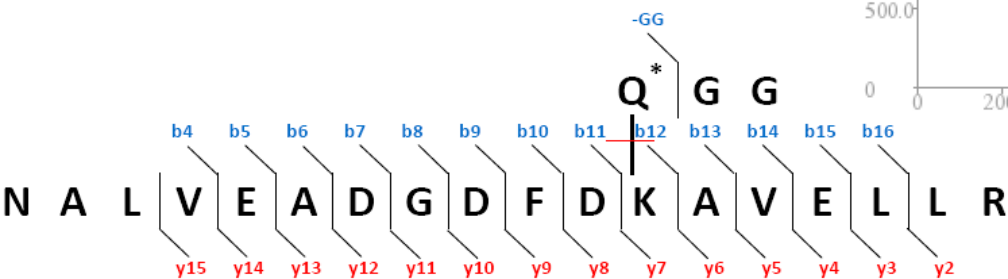


SF22: Elongation factor Ts (peptide 2)

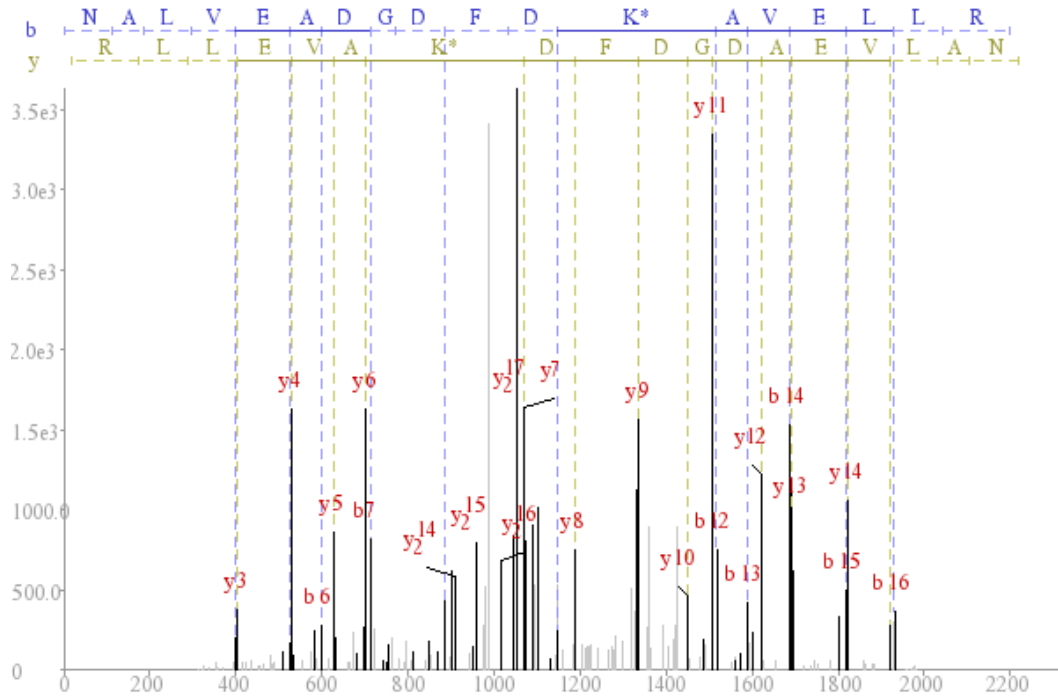
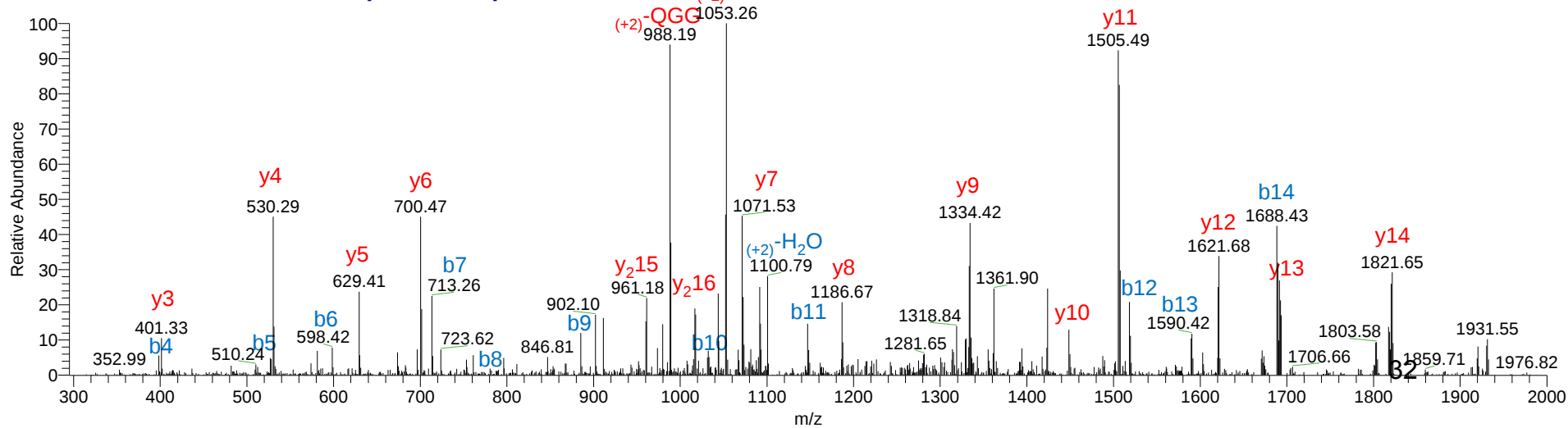
K.NALVEADGDFDK+243AVELLR.I

Calc MW: 2017.09 Da

Charge = +2

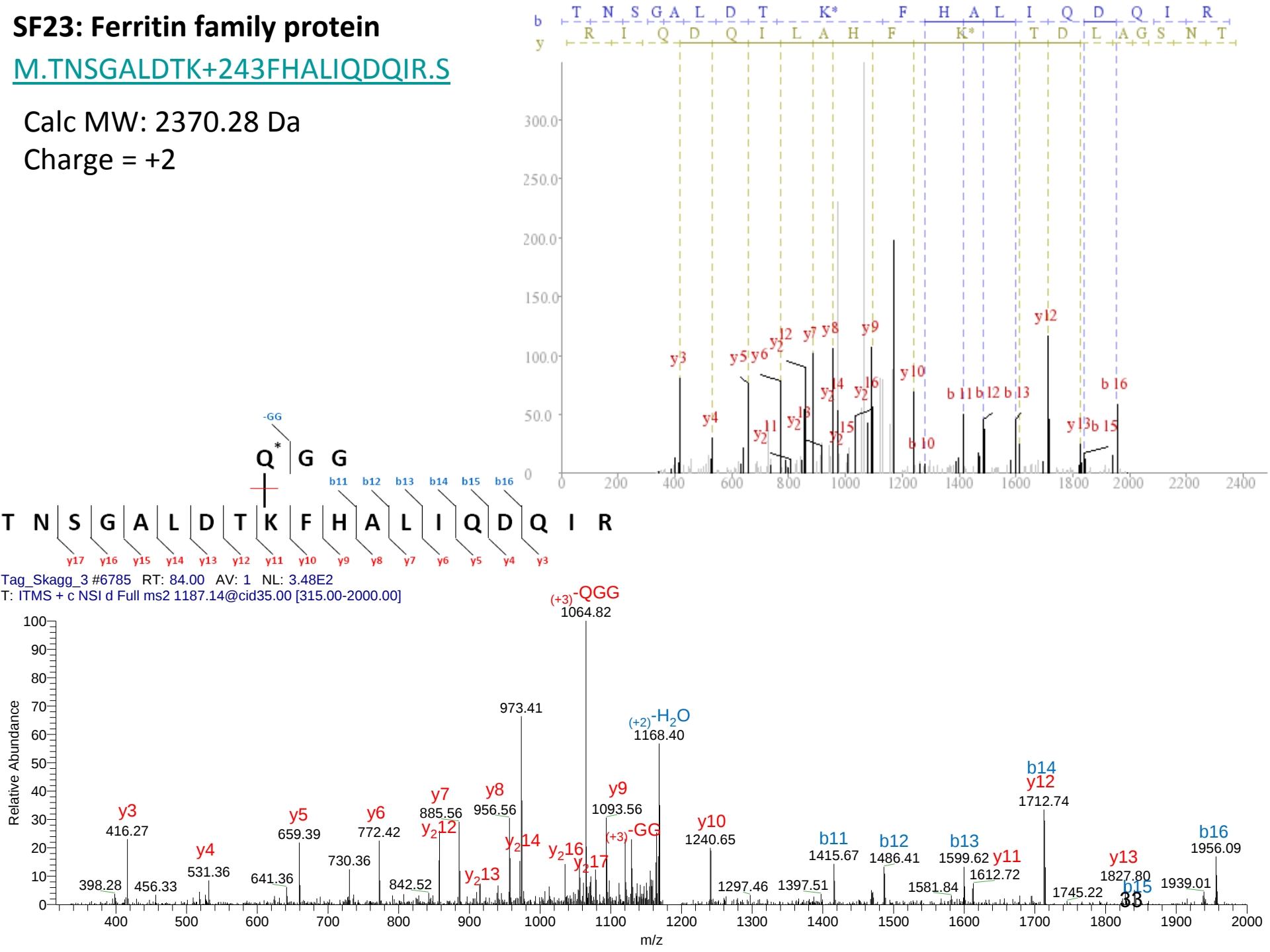


LTQ_Tag2 #18475 RT: 88.34 AV: 1 NL: 3.62E3
T: ITMS + c NSI d Full ms2 1110.25@cid35.00 [295.00-2000.00]



SF23: Ferritin family protein
M.TNSGALDTK+243FHALIQDQIR.S

Calc MW: 2370.28 Da
Charge = +2

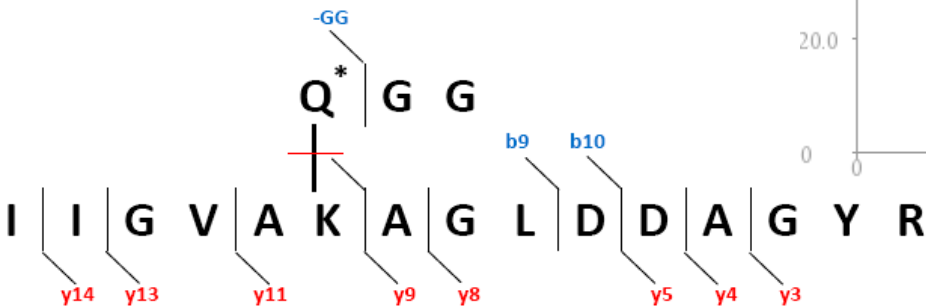


SF24: Glucose-6-phosphate dehydrogenase

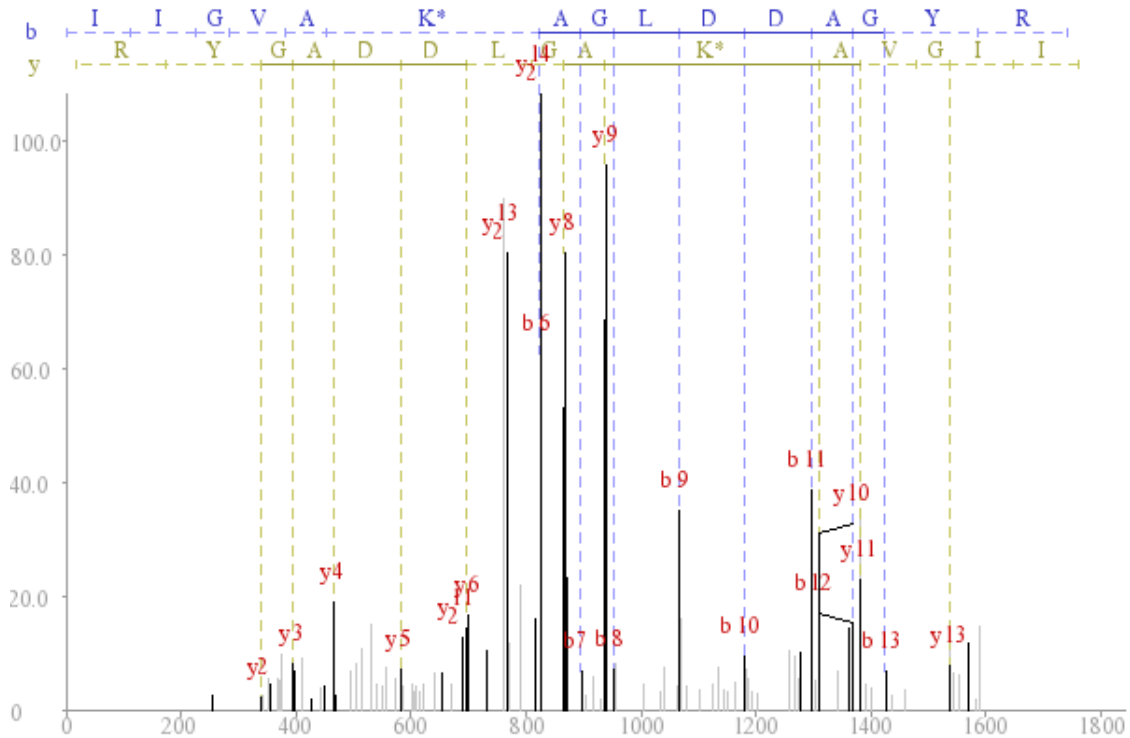
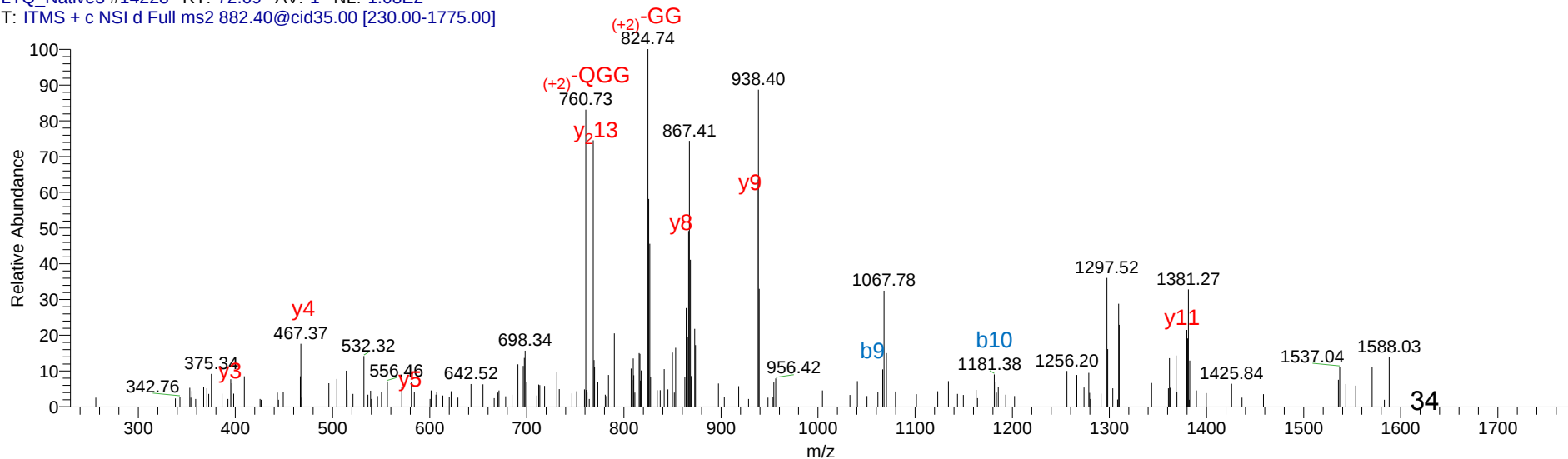
R.IIGVAK+243AGLDDAGYR.N

Calc MW: 1760.91 Da

Charge = +2



LTQ_Native3 #14228 RT: 72.69 AV: 1 NL: 1.08E2
T: ITMS + c NSI d Full ms2 882.40@cid35.00 [230.00-1775.00]

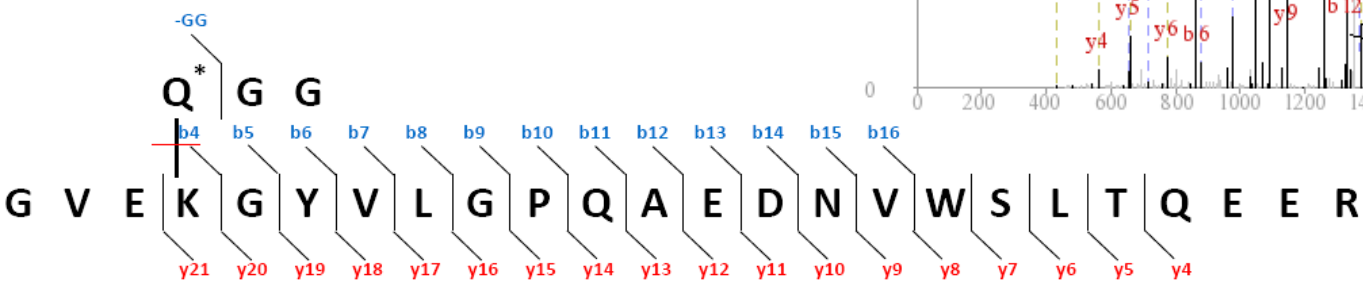
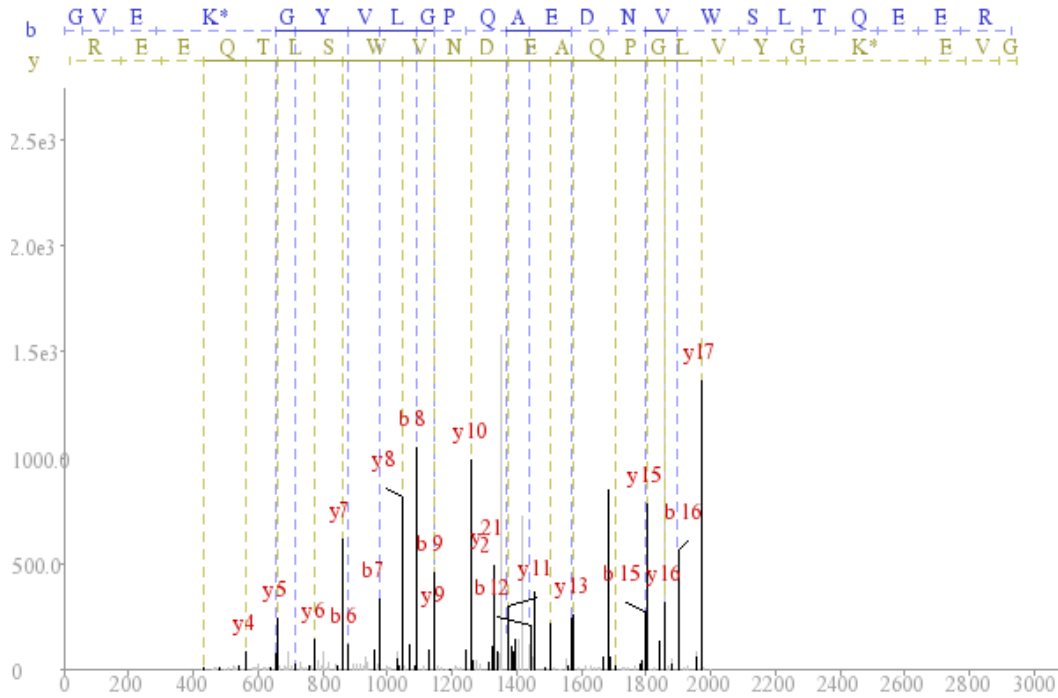


SF25: Glutamine synthetase, type I
(MSMEG_4294)

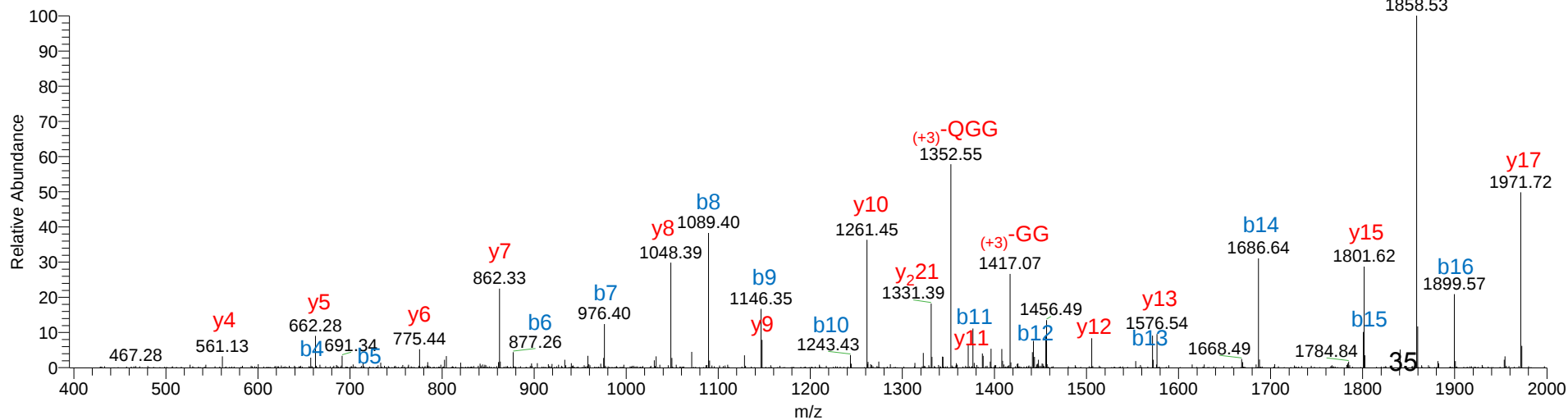
R.GVEK+243GYVLGPQAEDNVWSLTQEER.R

Calc MW: 2947.40 Da

Charge = +2



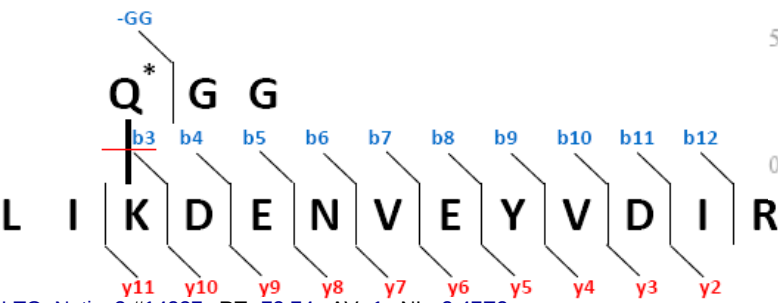
LTQ_Native2 #16701 RT: 81.00 AV: 1 NL: 2.73E3
T: ITMS + c NSI d Full ms2 1474.15@cid35.00 [395.00-2000.00]



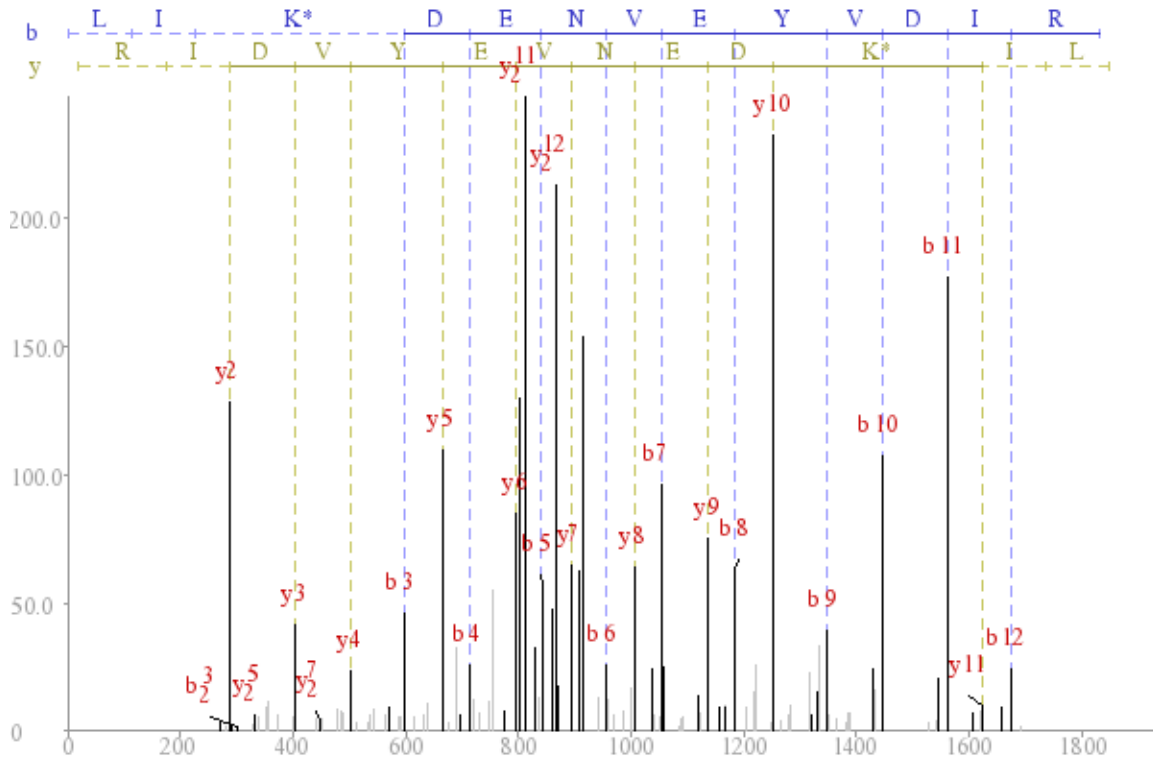
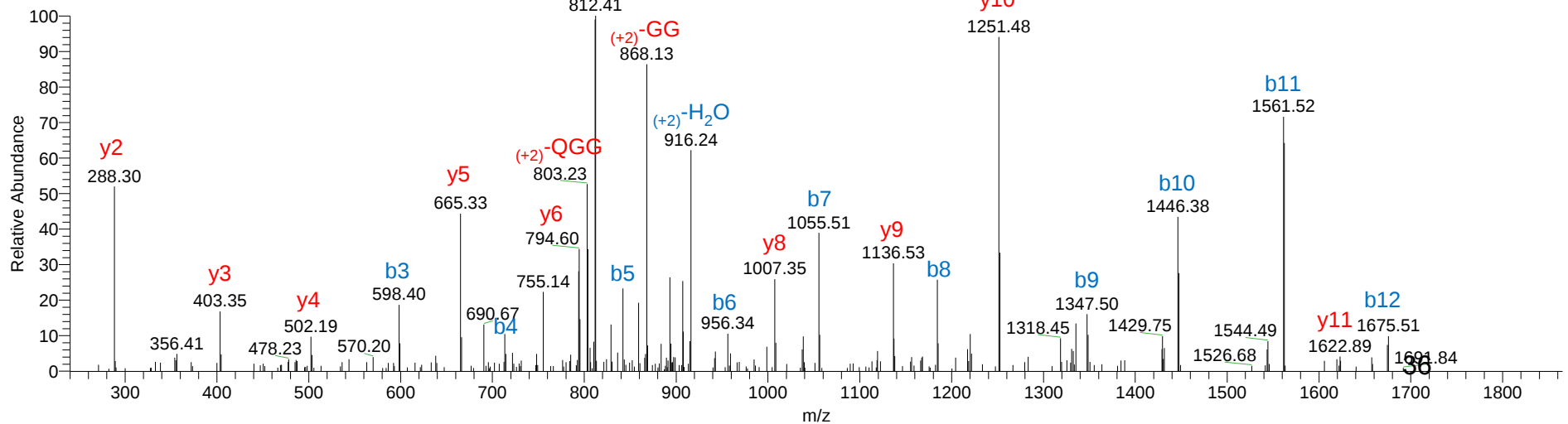
SF26: Glutamine synthetase, type I
(MSMEG_4290)

K.LIK+243DENVEYVDIR.F

Calc MW: 1847.93 Da
Charge = +2



LQ_Native3 #14237 RT: 72.74 AV: 1 NL: 2.47E2
T: ITMS + c NSI d Full ms2 925.05@cid35.00 [240.00-1865.00]

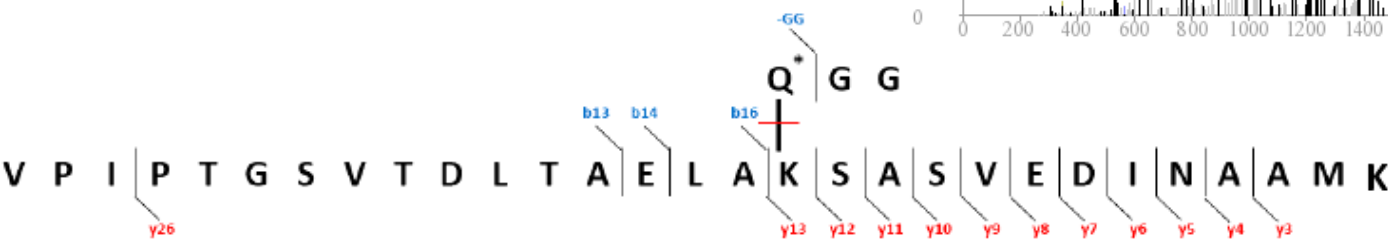
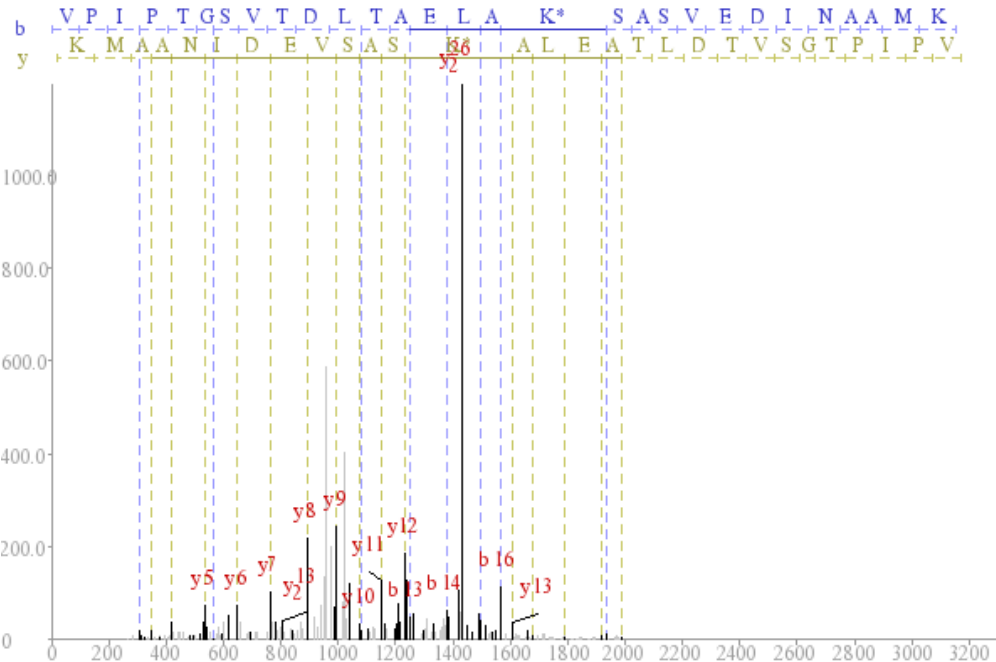


SF27: Glyceraldehyde-3-phosphate dehydrogenase, type I

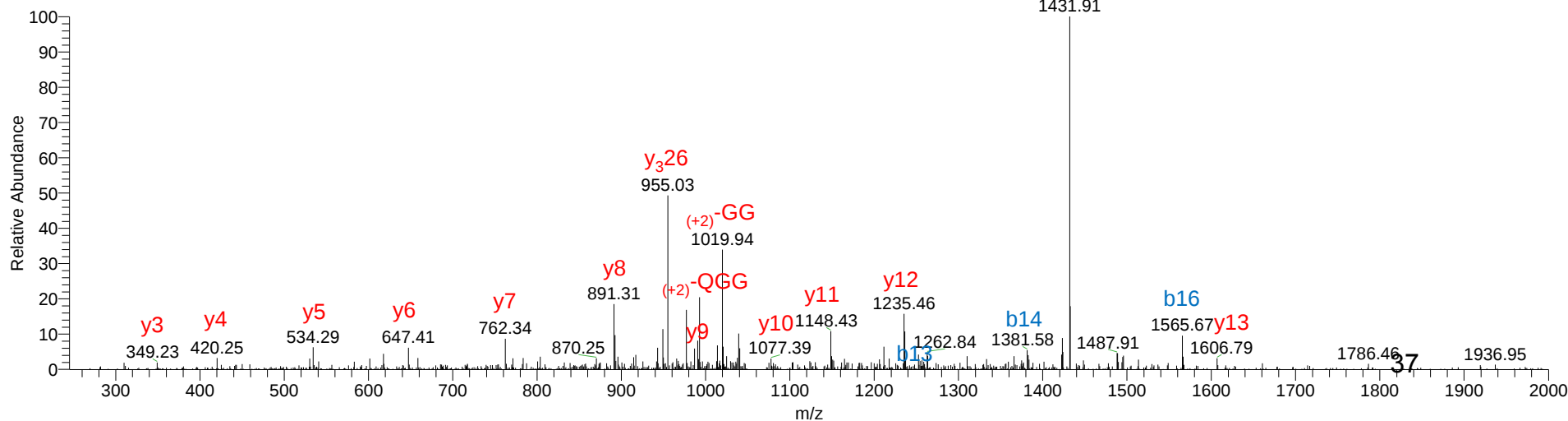
R.VPIPTGSVTDLTAE LAK+243SASVEDINAAMK.A

Calc MW: 3170.60 Da

Charge = +3



081206_Smeg_Pieter_A_LTQ_1#4610 RT: 50.17 AV: 1 NL: 1.19E3
T: ITMS + c NSI d Full ms2 1058.44@cid35.00 [245.00-2000.00]

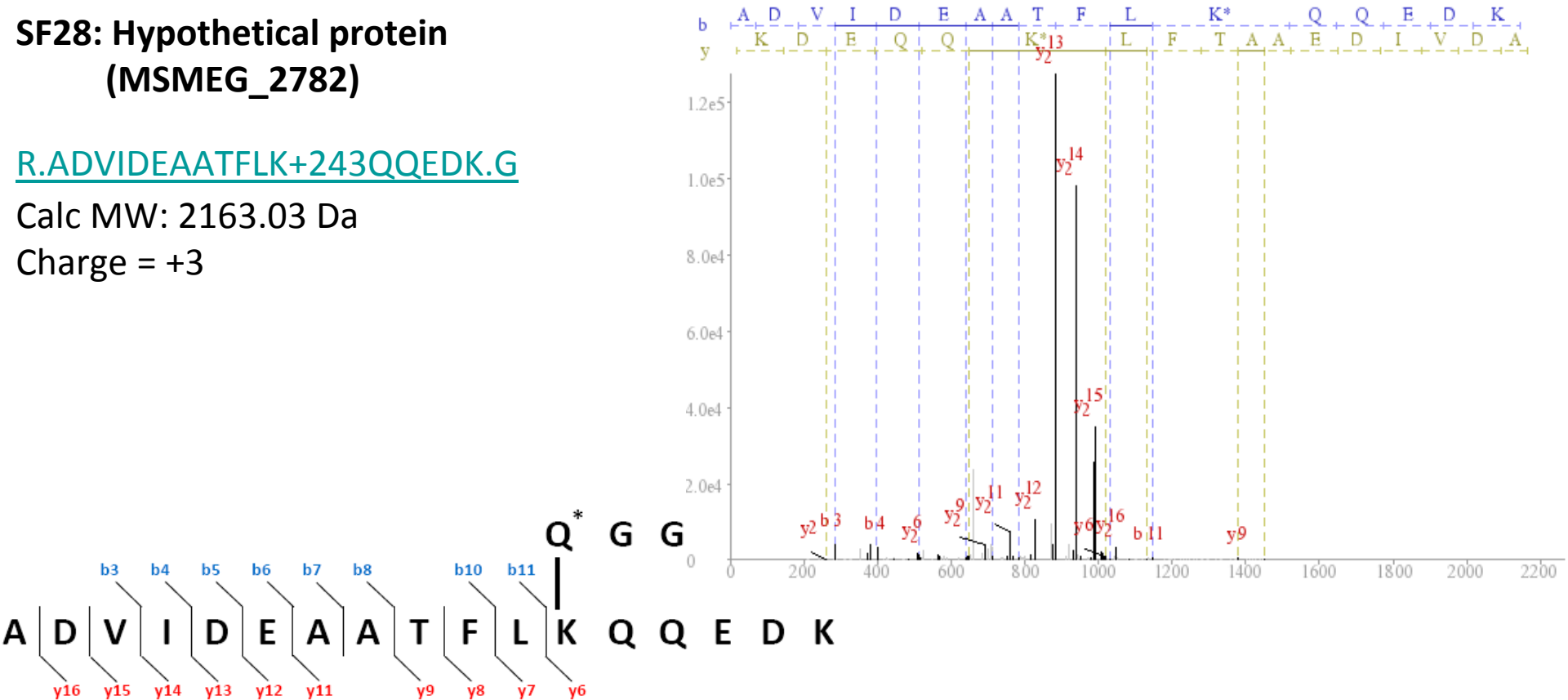


SF28: Hypothetical protein
(MSMEG_2782)

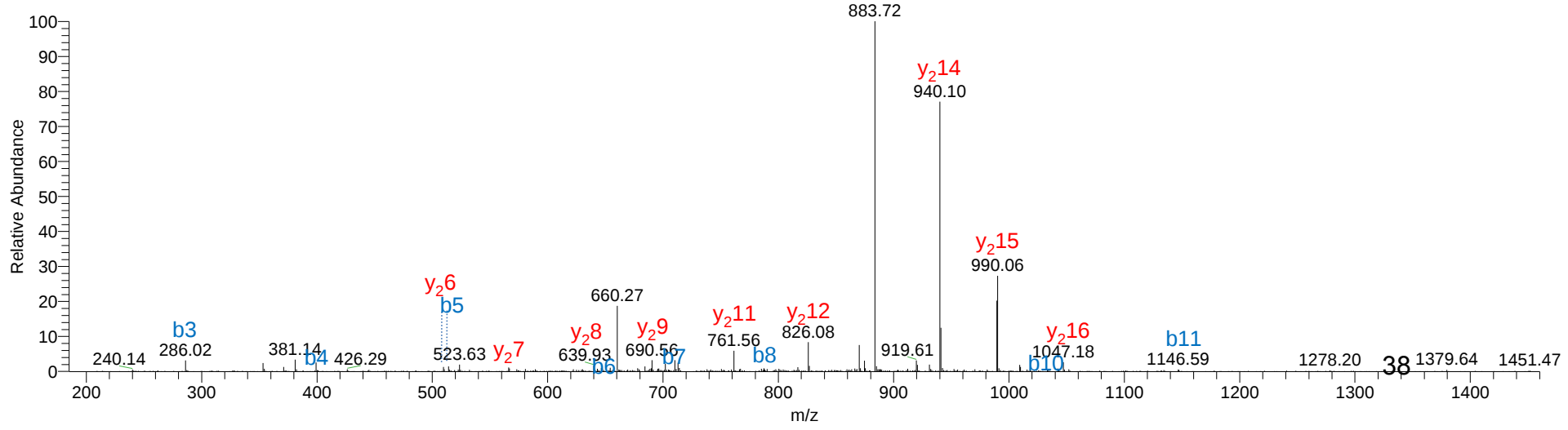
R.ADVIDEAATFLK+243QQEDK.G

Calc MW: 2163.03 Da

Charge = +3



LTQ_Native2 #18008 RT: 86.94 AV: 1 NL: 1.27E5
T: ITMS + c NSI d Full ms2 722.65@cid35.00 [185.00-1460.00]

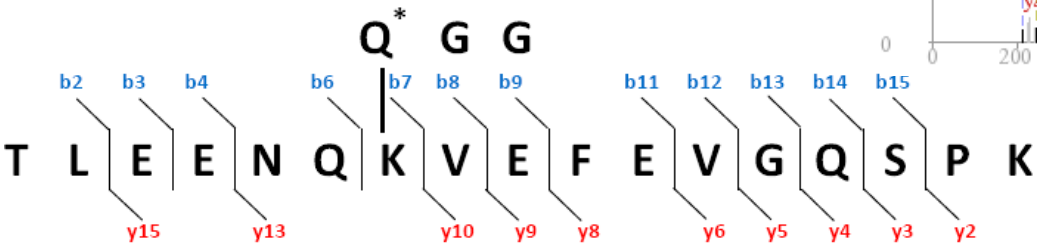


SF29: Hypothetical protein
(MSMEG_6159)

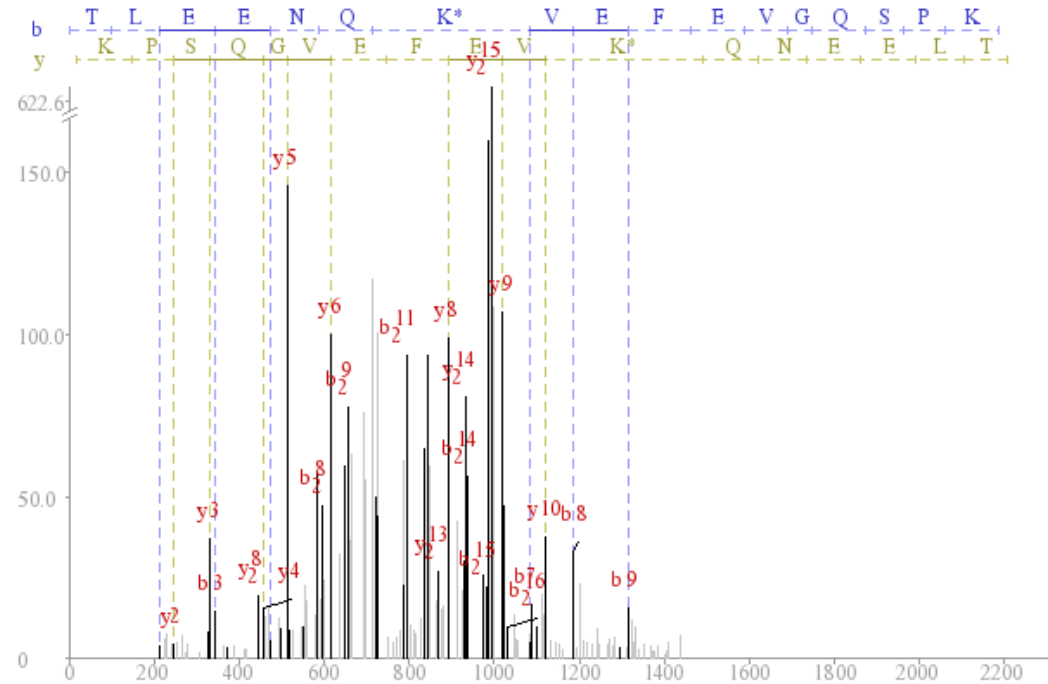
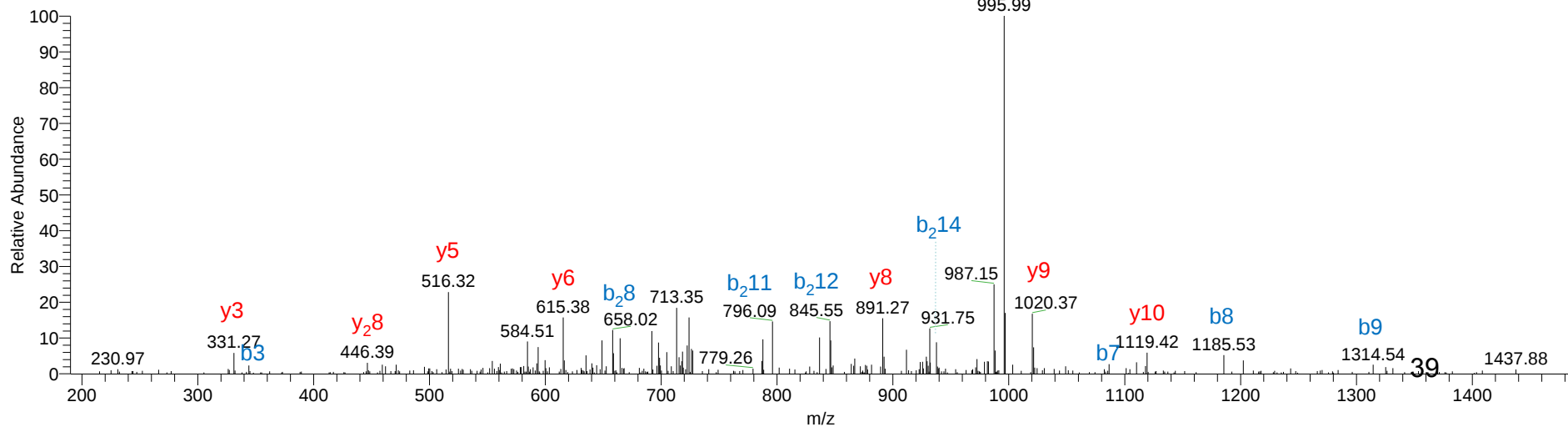
R.TLEENQK+243VEFEVGQSPK.G

Calc MW: 2204.05 Da

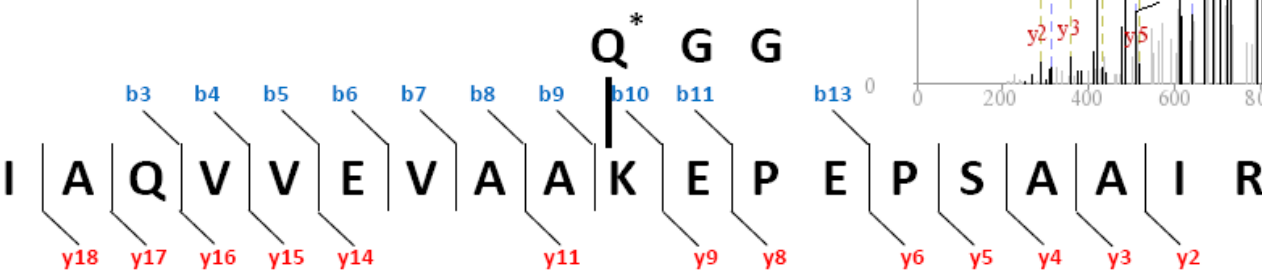
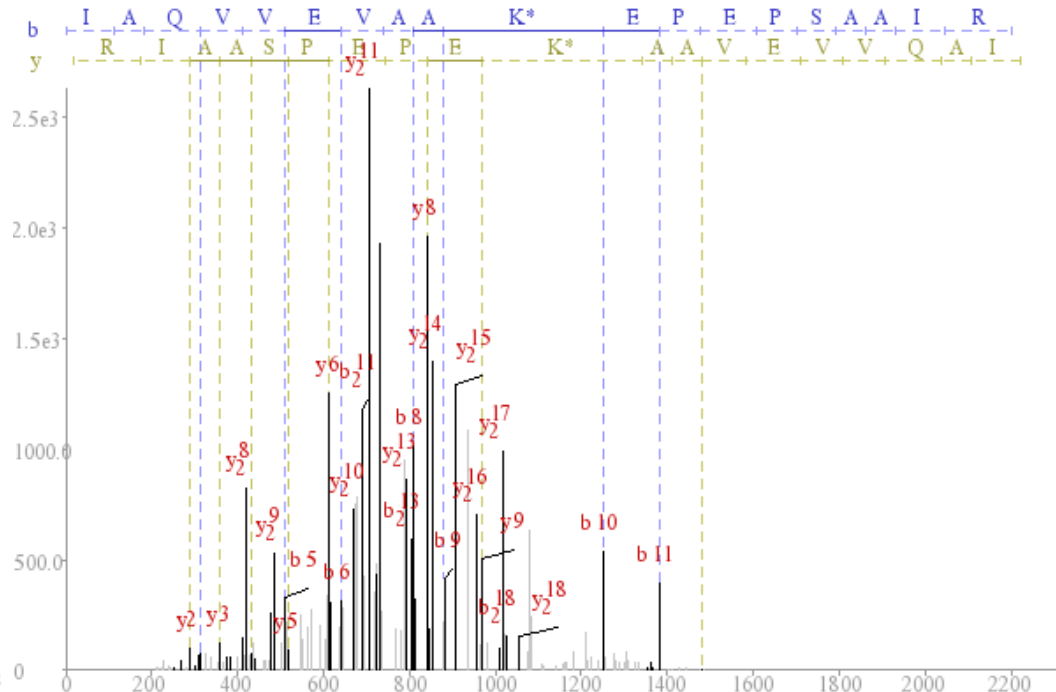
Charge = +3



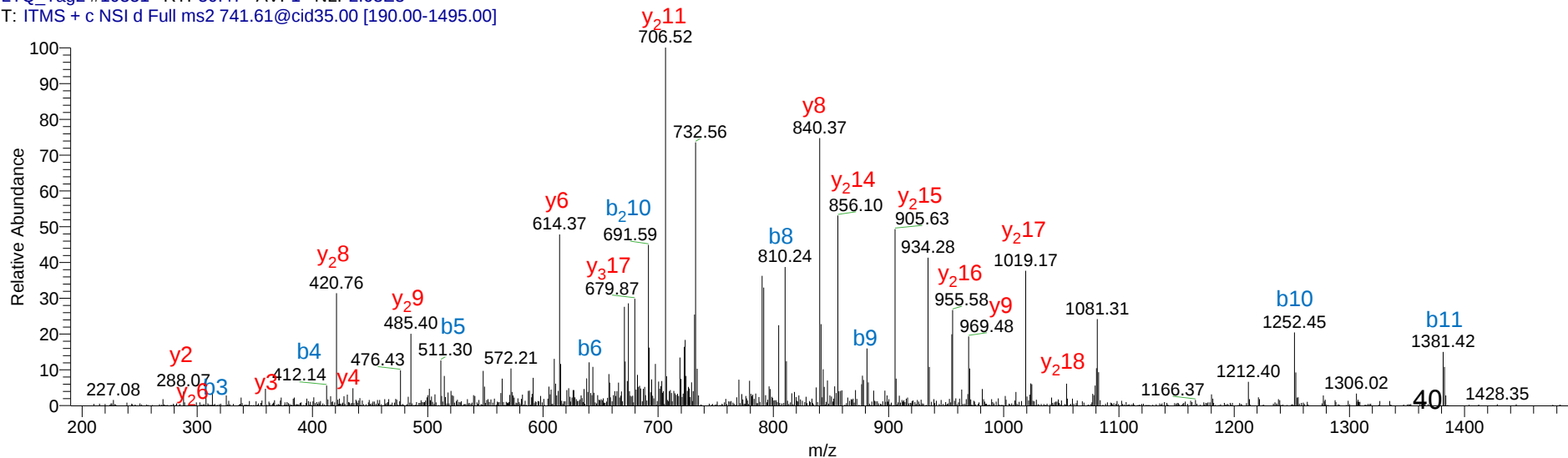
LTQ_Tag3 #13896 RT: 70.62 AV: 1 NL: 6.39E2
T: ITMS + c NSI d Full ms2 736.01@cid35.00 [190.00-1485.00]



Charge = +3



LTQ_Tag2 #10351 RT: 50.47 AV: 1 NL: 2.63E3
T: ITMS + c NSI d Full ms2 741.61@cid35.00 [190.00-1495.00]



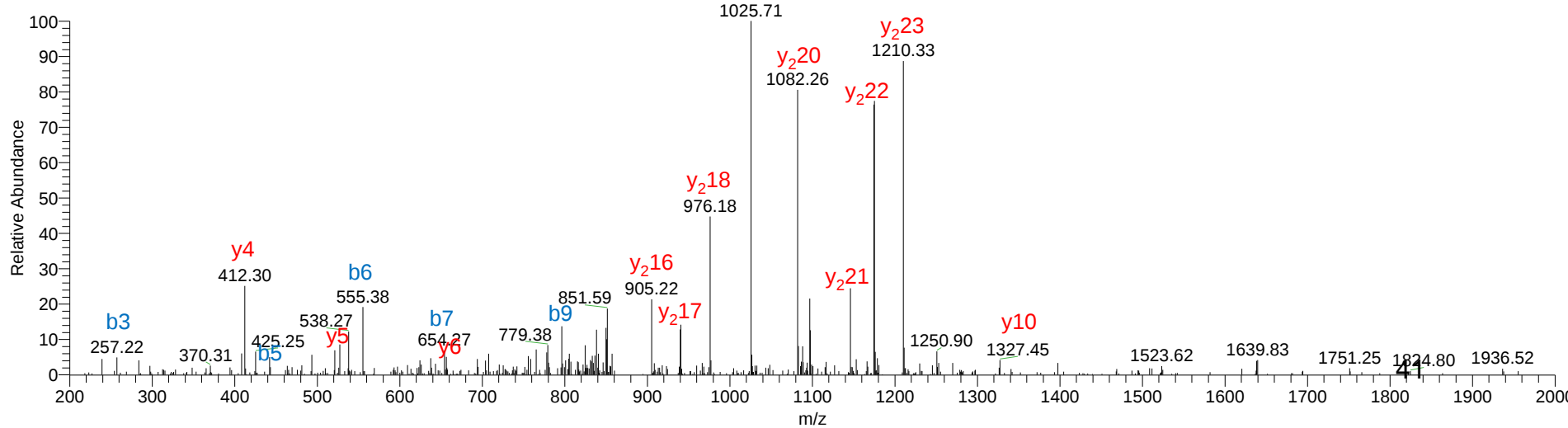
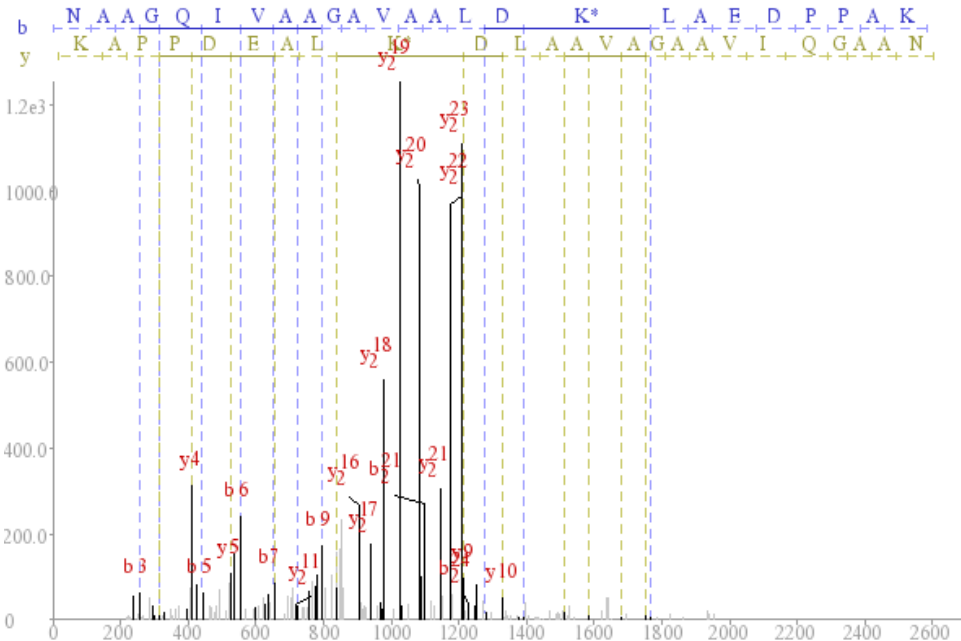
SF31: Malonyl CoA-acyl carrier protein
transacylase

R.NAAGQIVAAGAVAALDK+243LAEDPPAK.A

Calc MW: 2603.36 Da

Charge = +3

-Unable to rule out modification
on terminal lysine



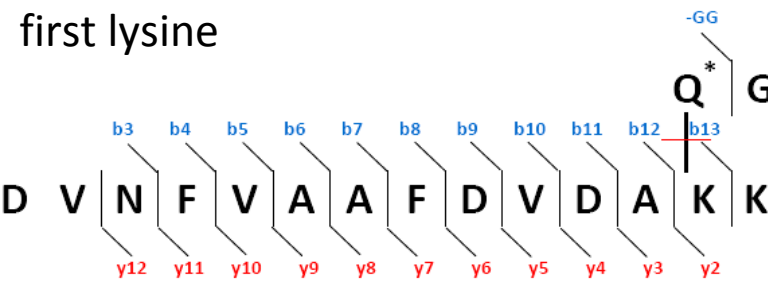
SF32: Myo-inositol-1-phosphate synthase

R.DVNFVAAFDVDAK+243K.V

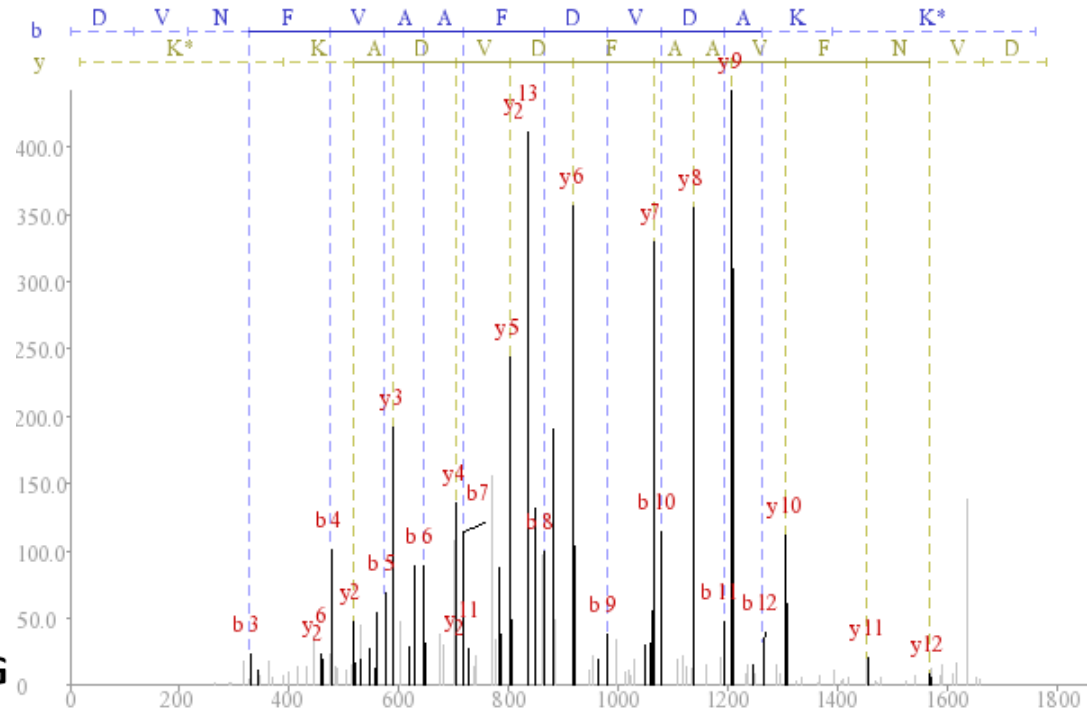
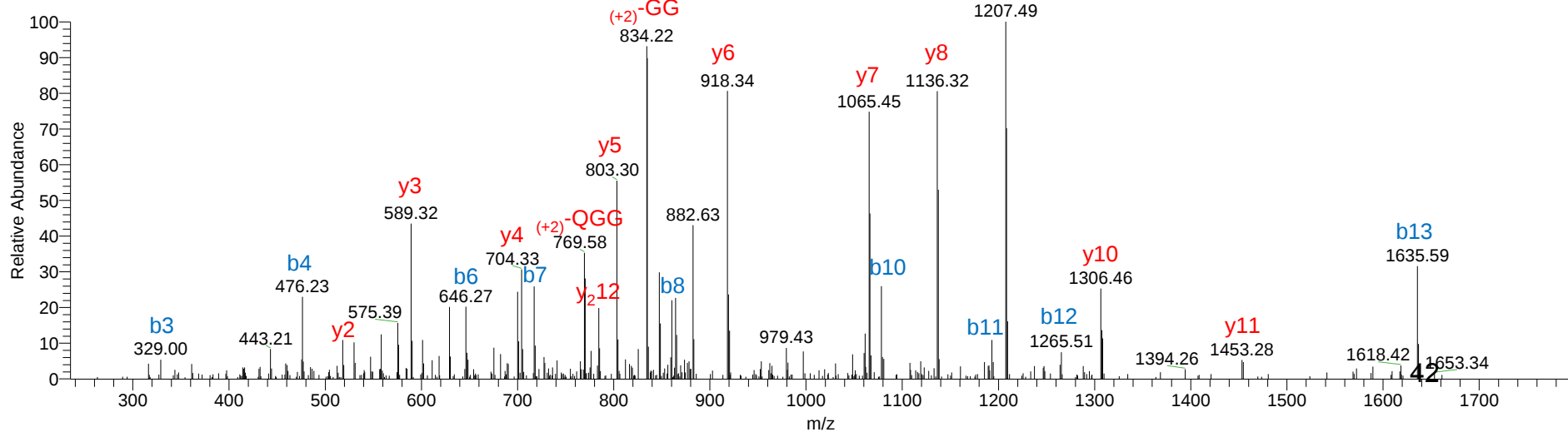
Calc MW: 1780.86 Da

Charge = +2

-Modified b13 ion and no modified y1 ion suggests modification in on the first lysine



LITQ_Tag3 #16289 RT: 82.96 AV: 1 NL: 4.41E2
T: ITMS + c NSI d Full ms2 892.04@cid35.00 [235.00-1795.00]

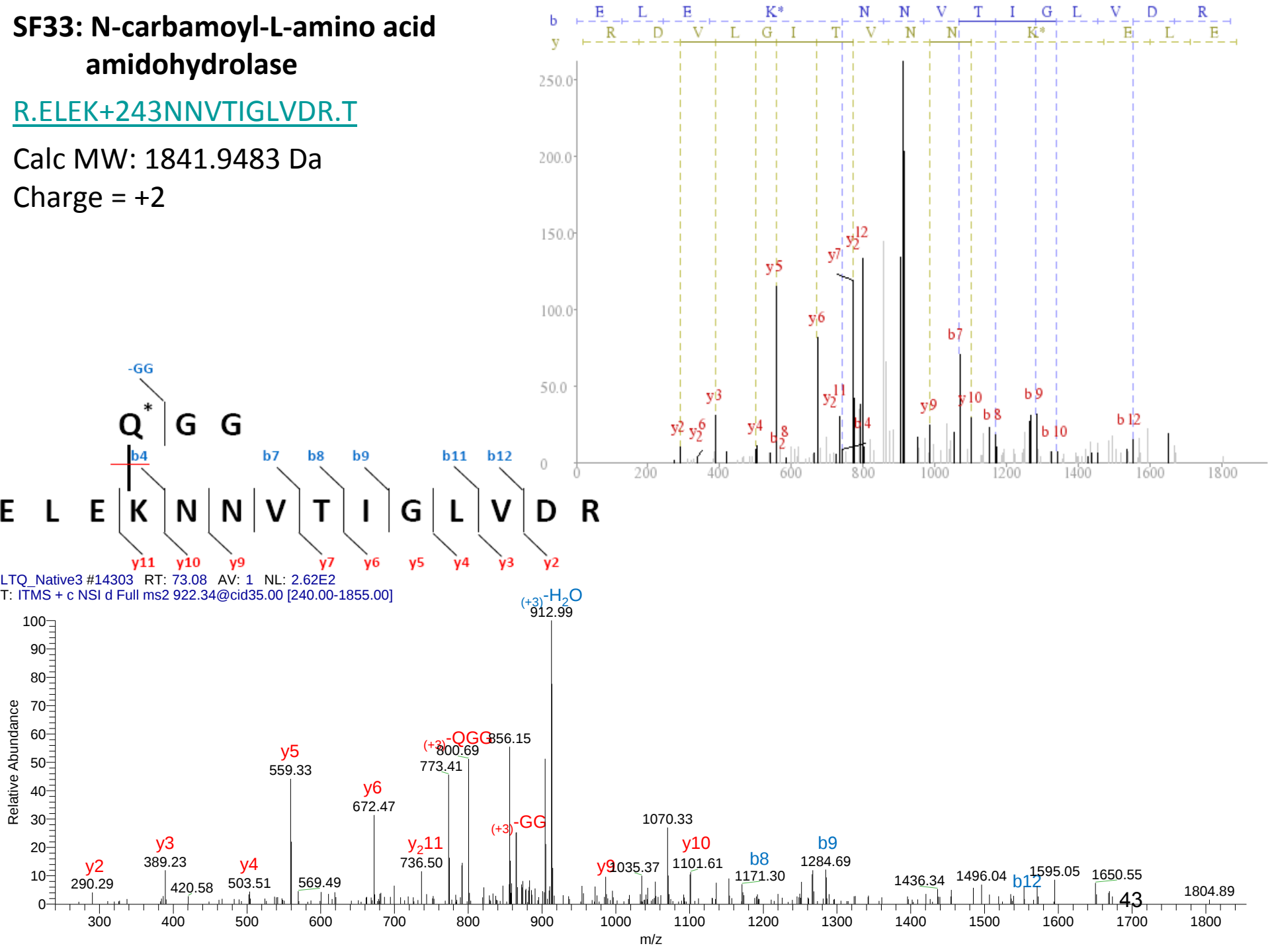


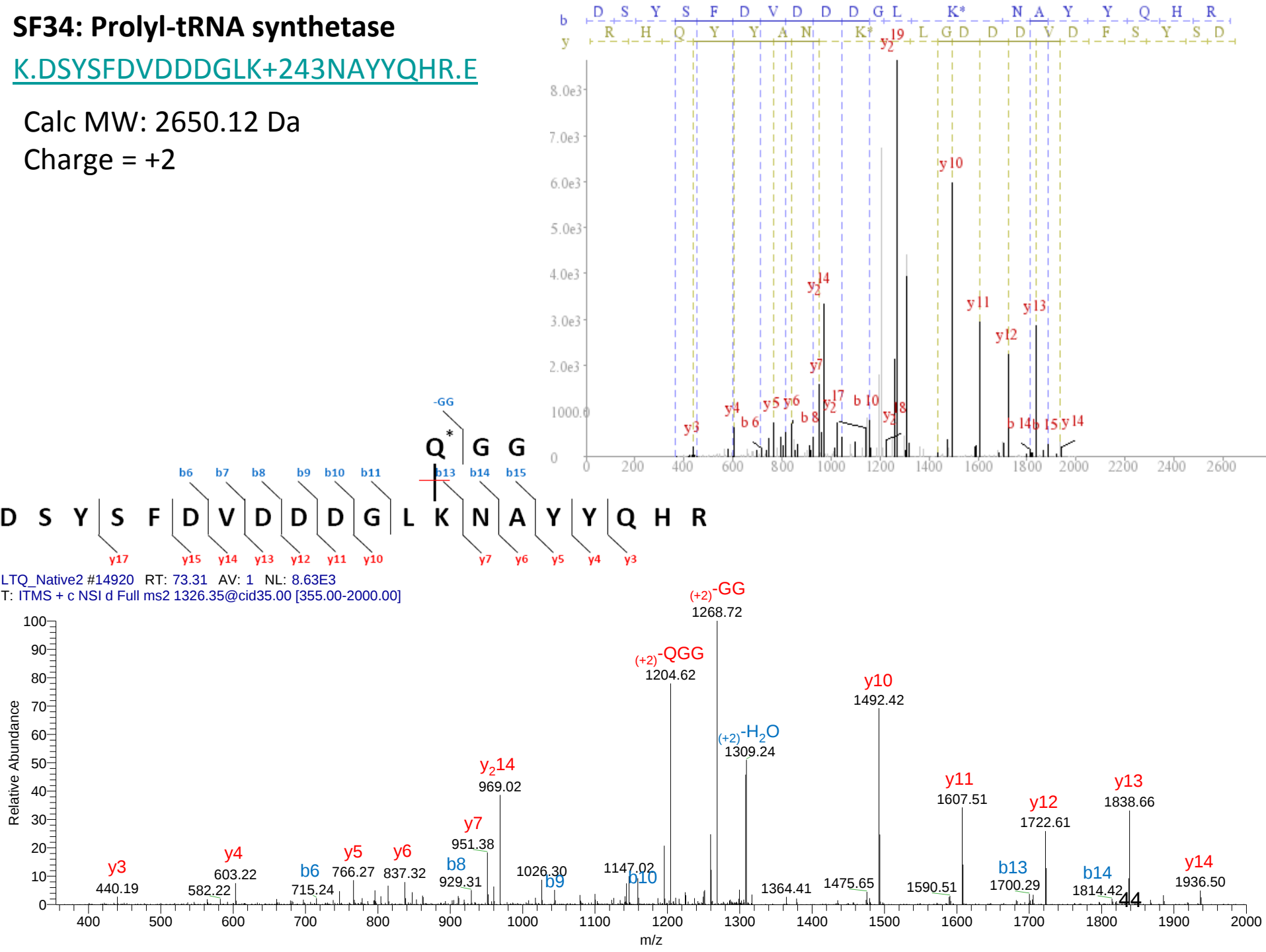
SF33: N-carbamoyl-L-amino acid
amidohydrolase

R.ELEK+243NNVTIGLVDR.T

Calc MW: 1841.9483 Da

Charge = +2



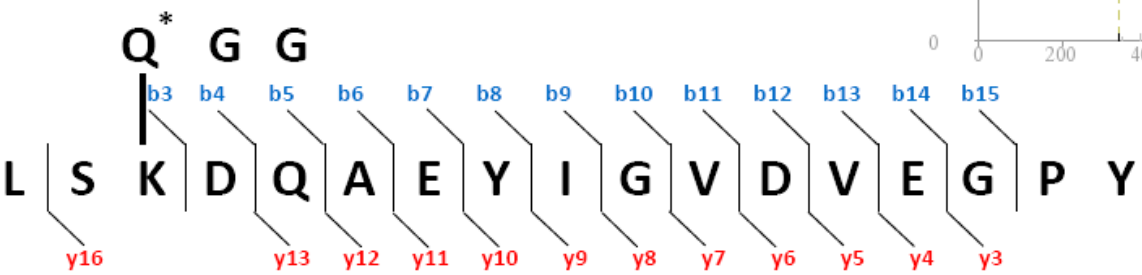


SF35: Putative adenosylhomocysteinase
(peptide 1)

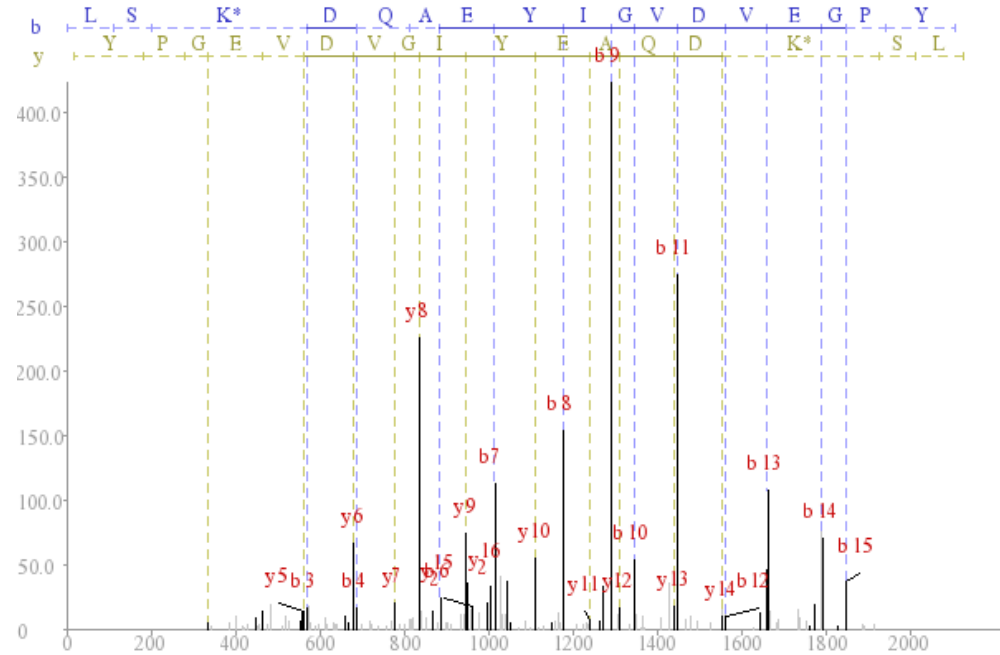
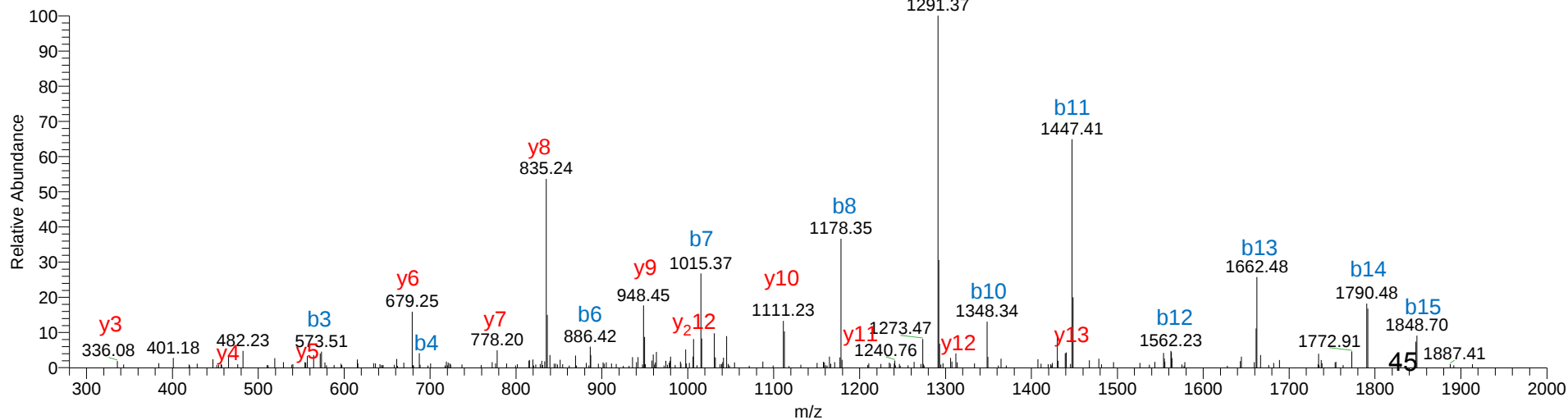
K.LSK+243DQAEYIGVDVEGPY.K

Calc MW: 2125.99 Da

Charge = +2



LTQ_Native3 #15451 RT: 79.05 AV: 1 NL: 4.23E2
T: ITMS + c NSI d Full ms2 1063.57@cid35.00 [280.00-2000.00]

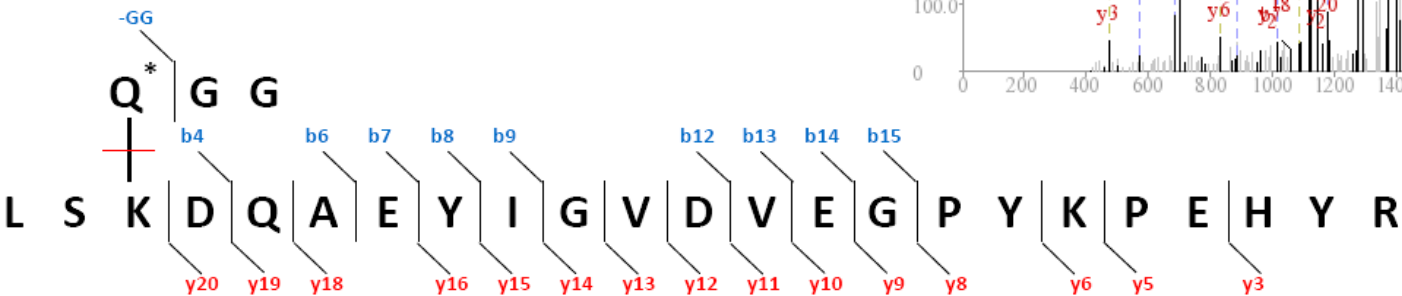


SF36: Putative adenosylhomocysteinase
(peptide 2)

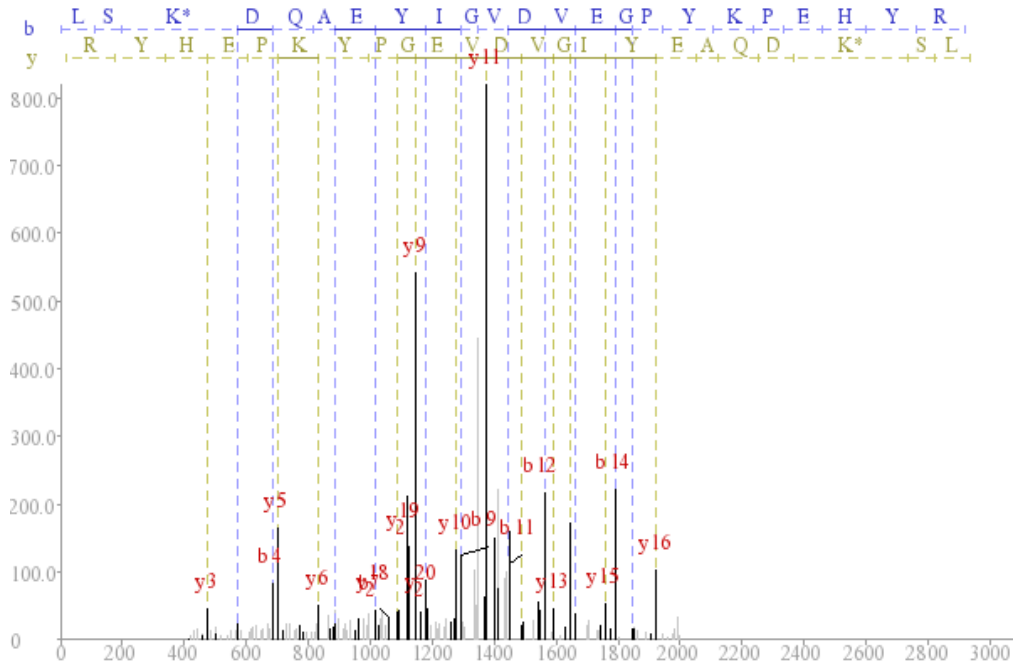
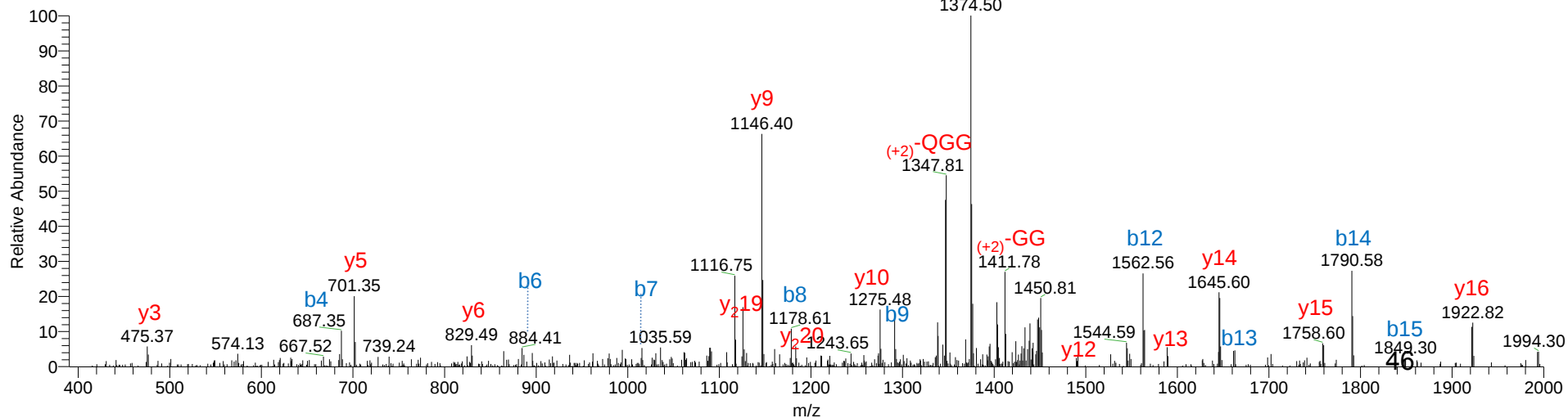
K.LSK+243DQAEYIGVDVEGPKPEHYR.Y

Calc MW: 2936.40 Da

Charge = +2



Native_Skagg_3 #2692 RT: 52.89 AV: 1 NL: 8.17E2
T: ITMS + c NSI d Full ms2 1469.26@cid35.00 [390.00-2000.00]

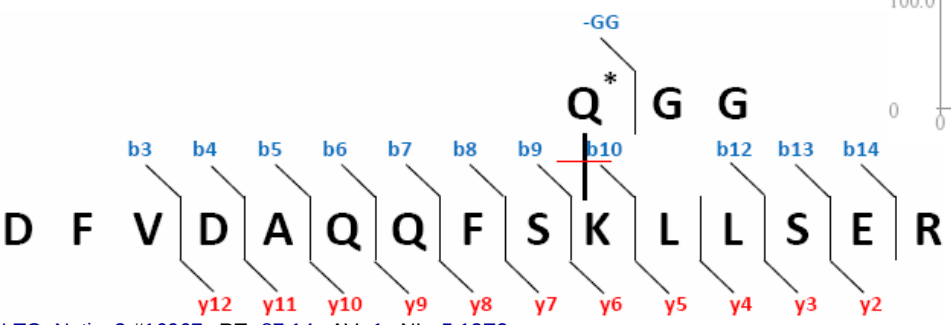
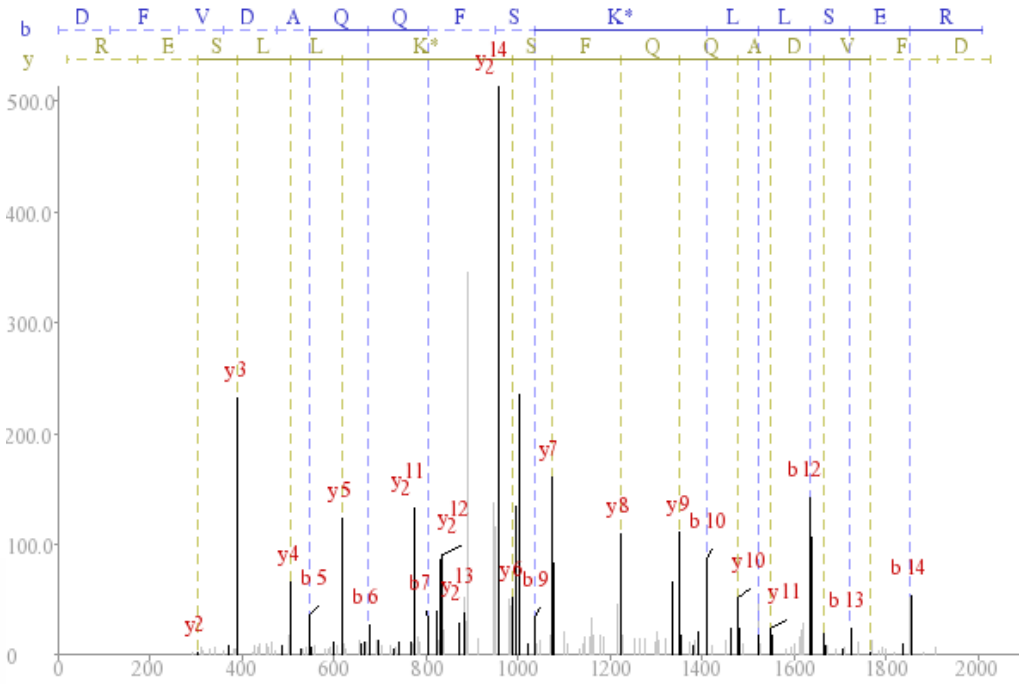


SF37: Putative thiosulfate sulfurtransferase

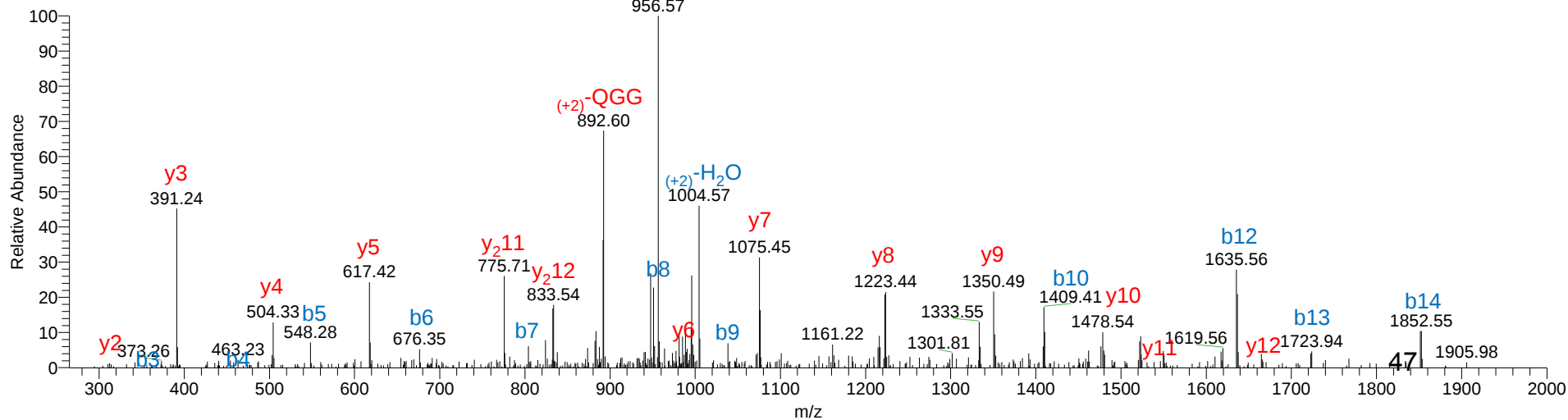
R.DFVDAQFSK+243LLSER.G

Calc MW: 2024.98 Da

Charge = +2



LTQ_Native3 #16967 RT: 87.14 AV: 1 NL: 5.12E2
T: ITMS + c NSI d Full ms2 1014.03@cid35.00 [265.00-2000.00]

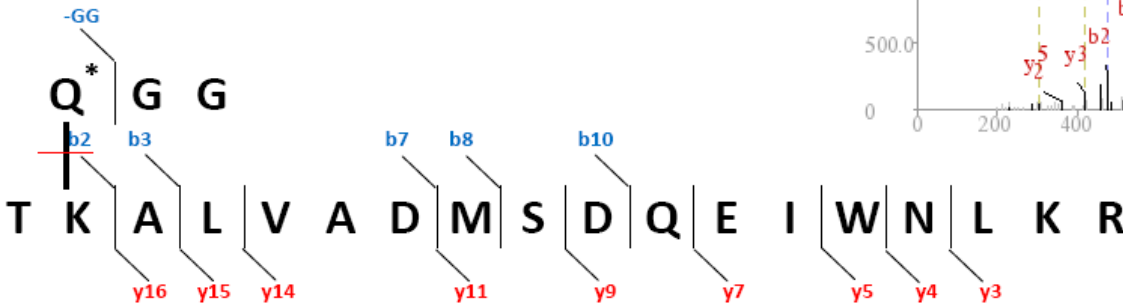
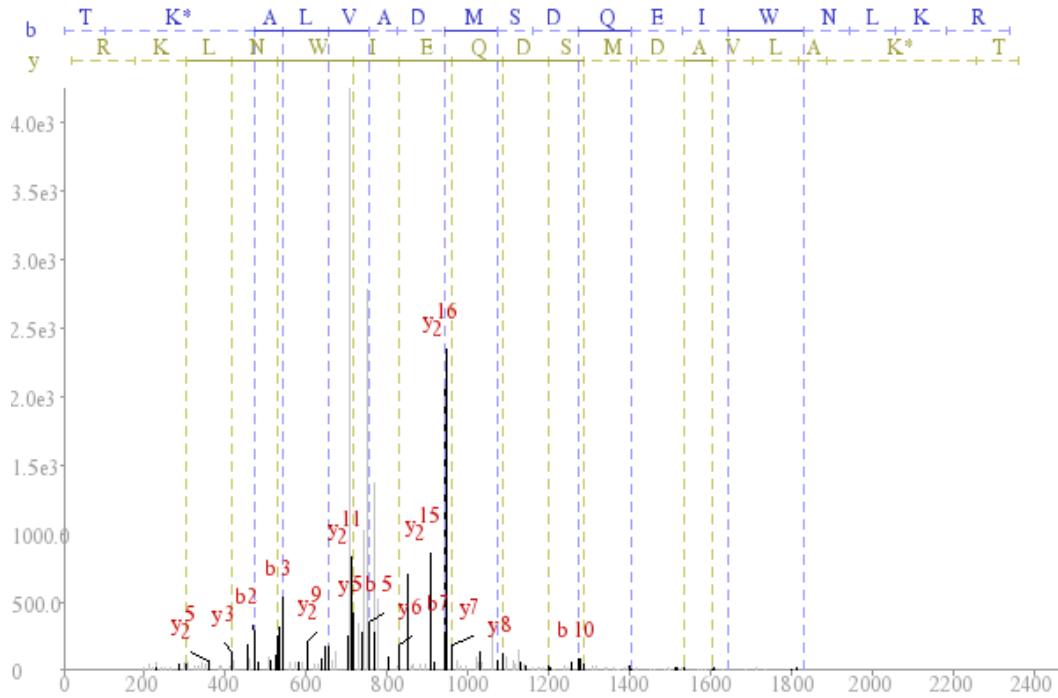


SF38: Pyruvate dehydrogenase
subunit E1 (peptide 1)

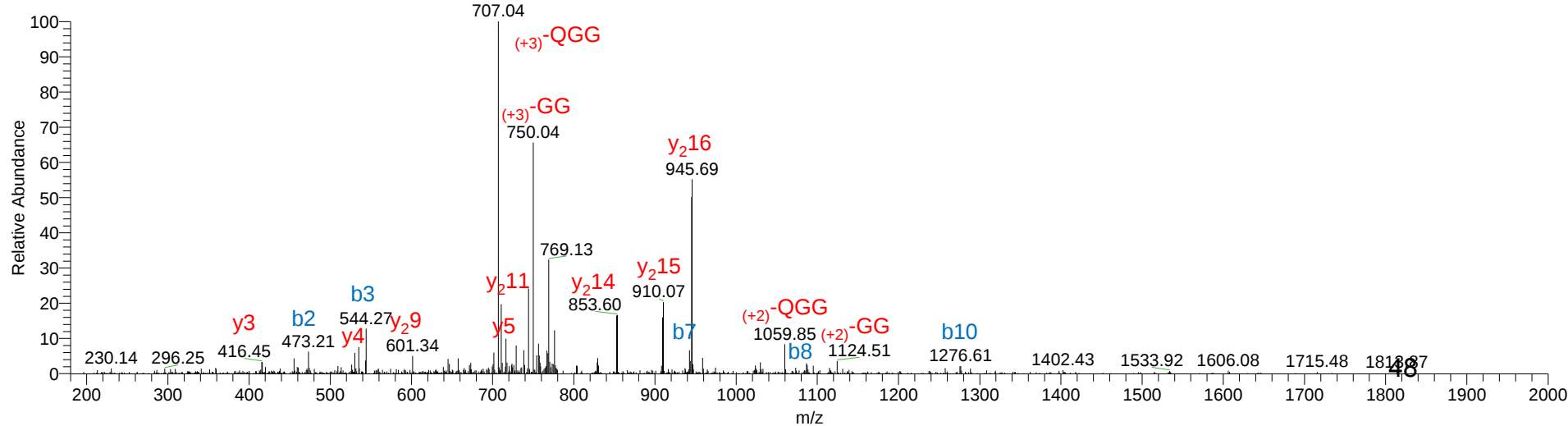
R.TK+243ALVADMSDQEIWNLKR.G

Calc MW: 2360.18 Da

Charge = +3



081206_Smeg_Pieter_A_FT_1 #2723 RT: 43.44 AV: 1 NL: 4.23E3
T: ITMS + c NSI d Full ms2 788.08@cid35.00 [180.00-2000.00]

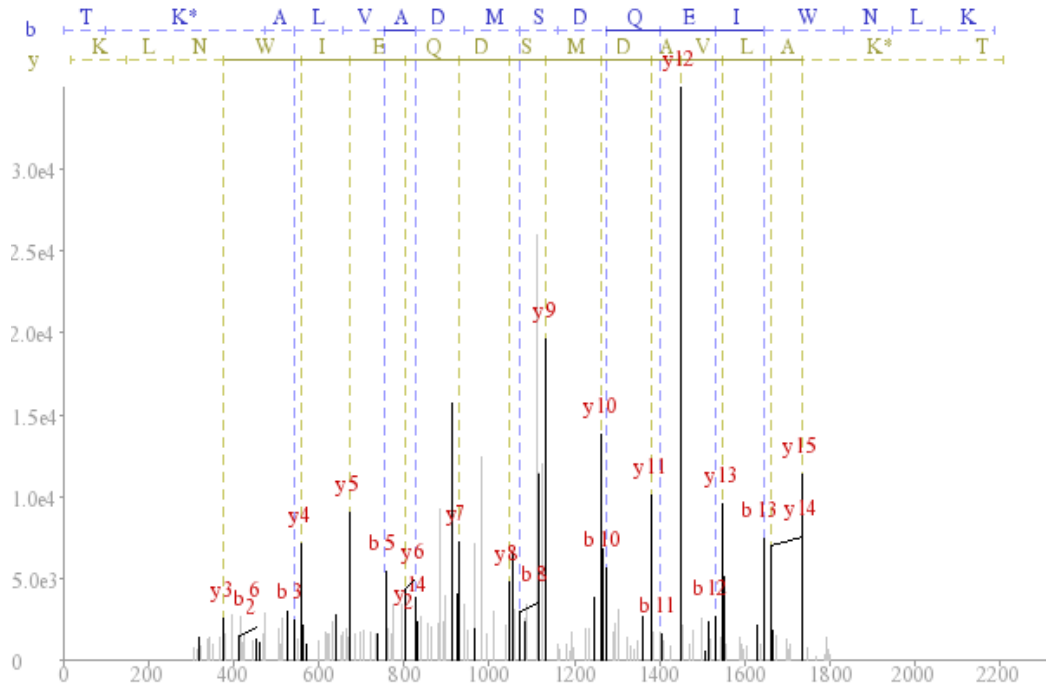
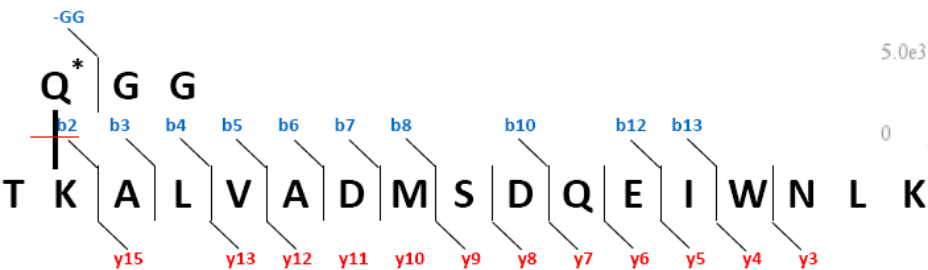


SF39: Pyruvate dehydrogenase subunit E1
(peptide 2)

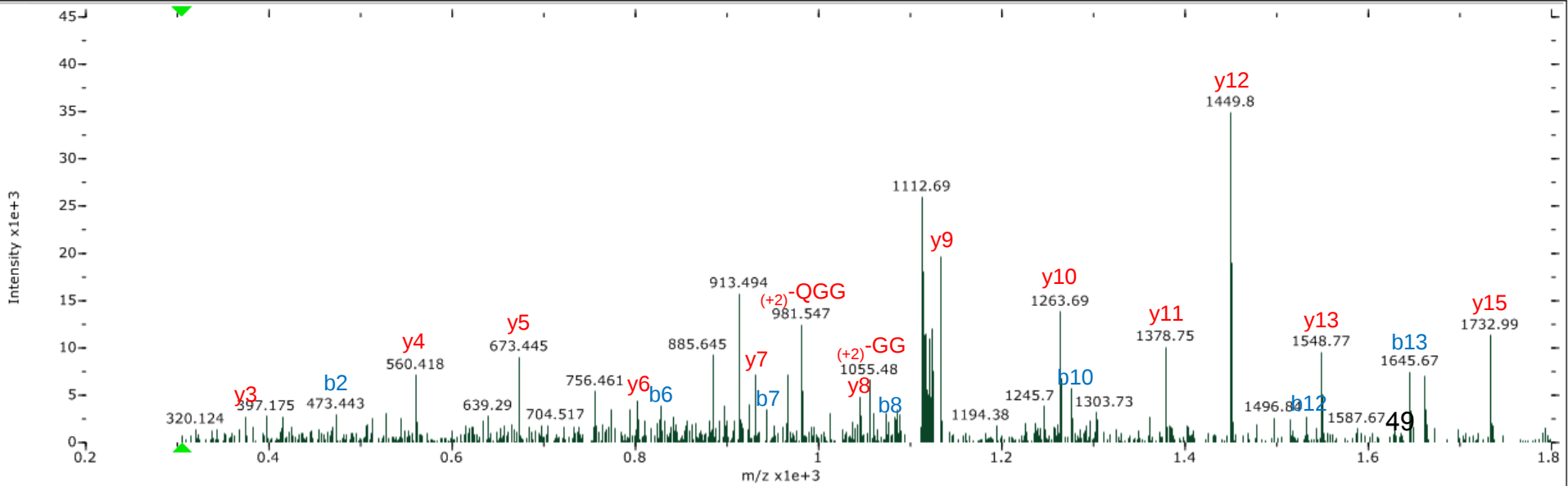
R.TK+243ALVADMSDQEIWNLK.R

Calc MW: 2204.08 Da

Charge = +2



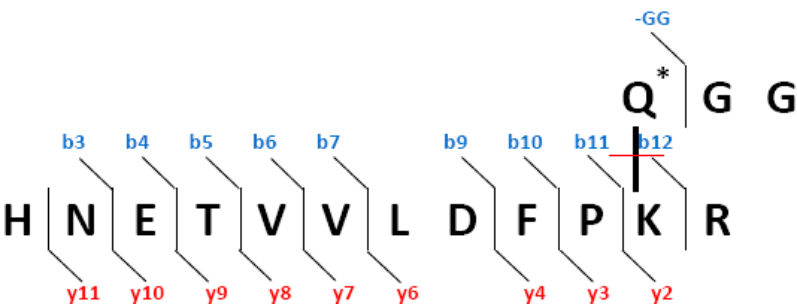
Scan #4700 (1103.95) (t=26.65 BP=34890.64@m/z=1449.80): intensity=761.85@m/z=305.25



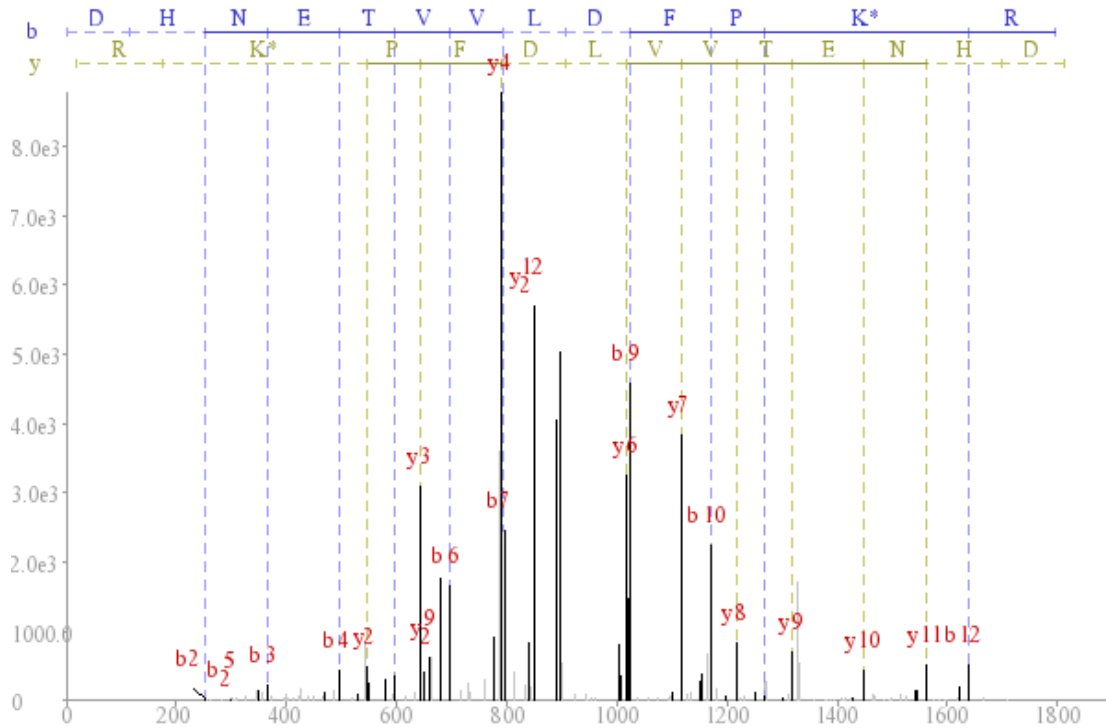
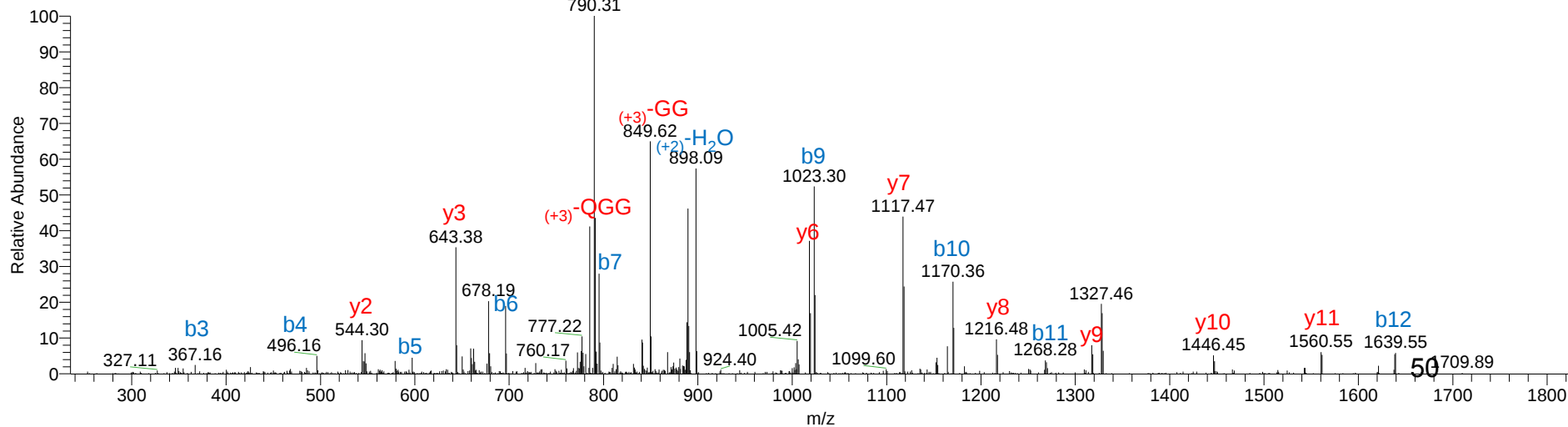
SF40: Rieske 2Fe-2S family protein

R.DHNETVVLDFPK+243R.V

Calc MW: 1811.88 Da
Charge = +2



LTQ_Native2 #9620 RT: 47.81 AV: 1 NL: 8.75E3
T: ITMS + c NSI d Full ms2 907.07@cid35.00 [235.00-1825.00]



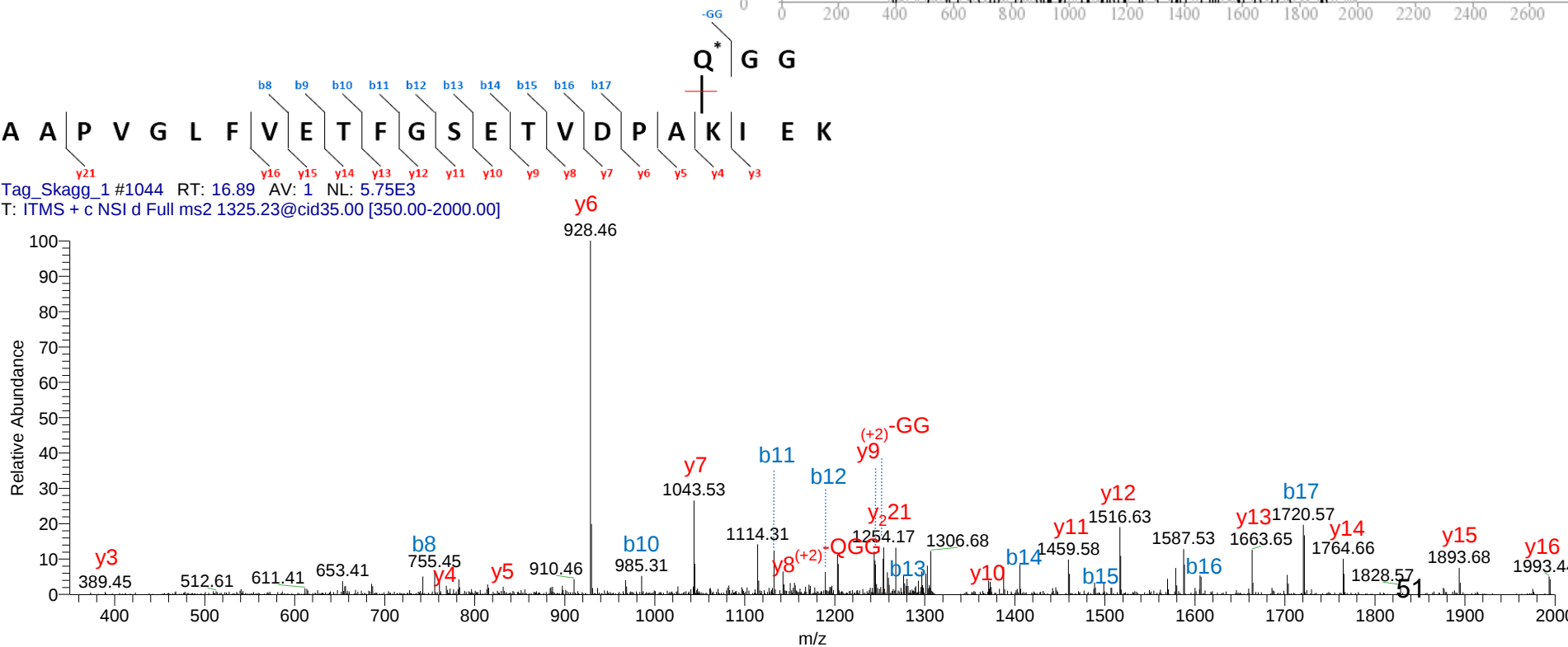
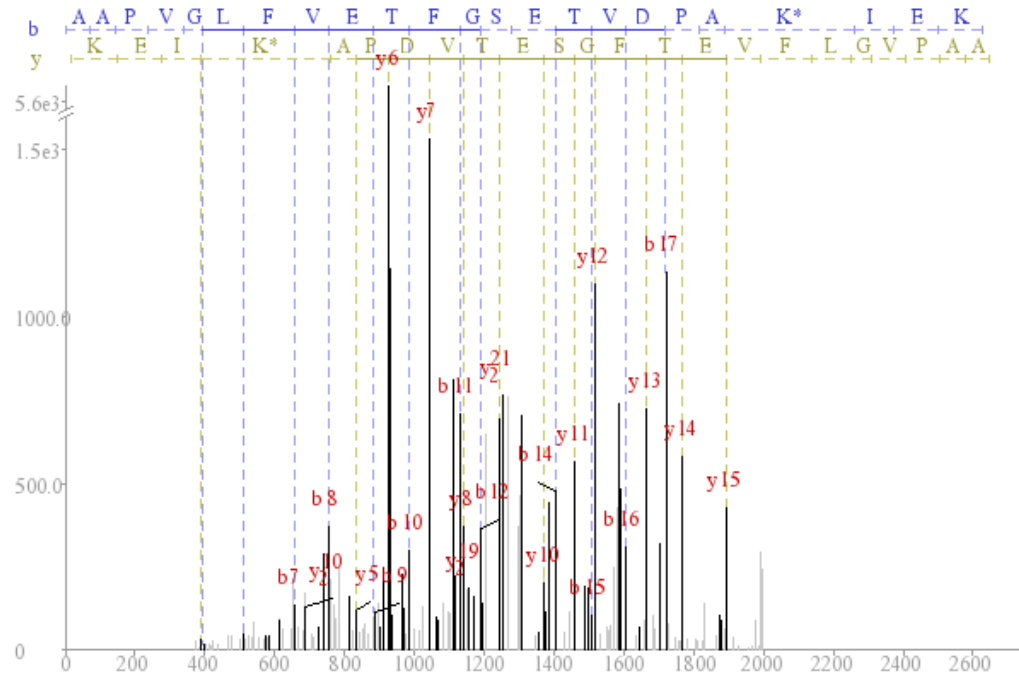
SF41: S-adenosylmethionine synthetase

K.AAPVGLFVETFGSETVDPAK+243IEK.A

Calc MW: 2648.34 Da

Charge = +2

-Unmodified y3 ion
suggests modification is
on the first lysine

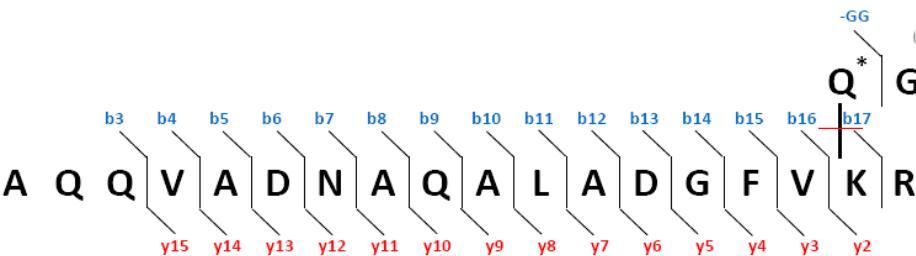


SF42: Serine hydroxymethyltransferase
(peptide 1)

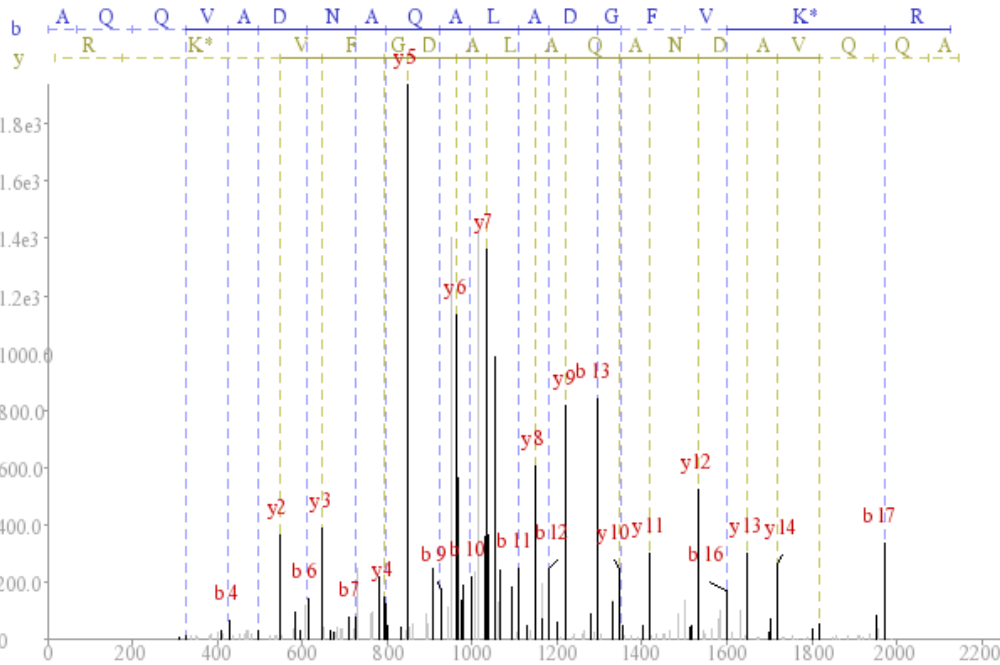
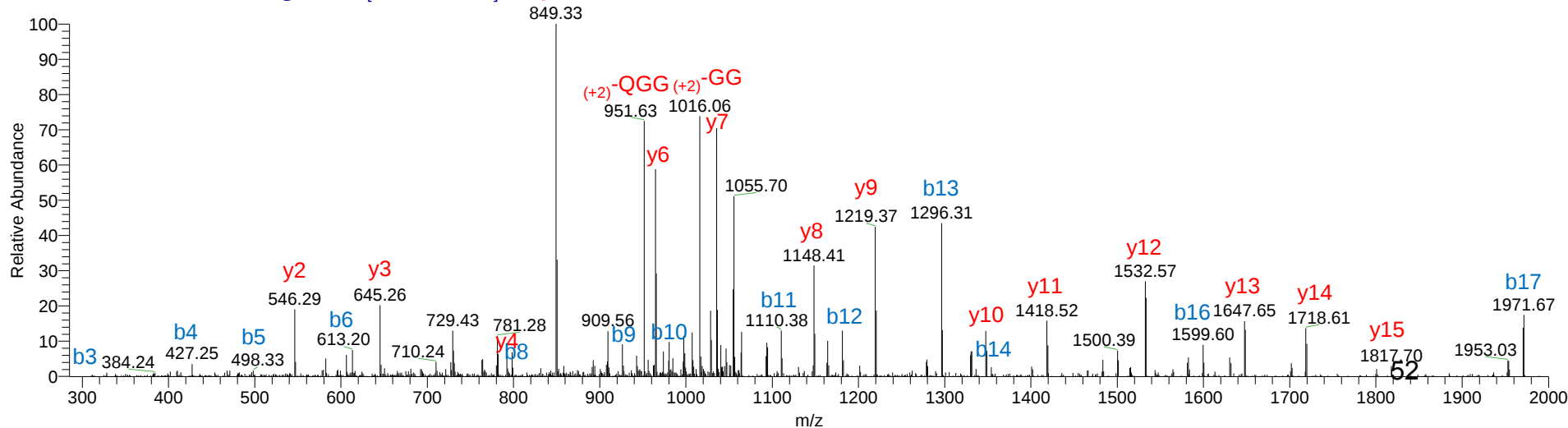
Y.AQQVADNAQALADGFVK+243R.D

Calc MW: 2144.06 Da

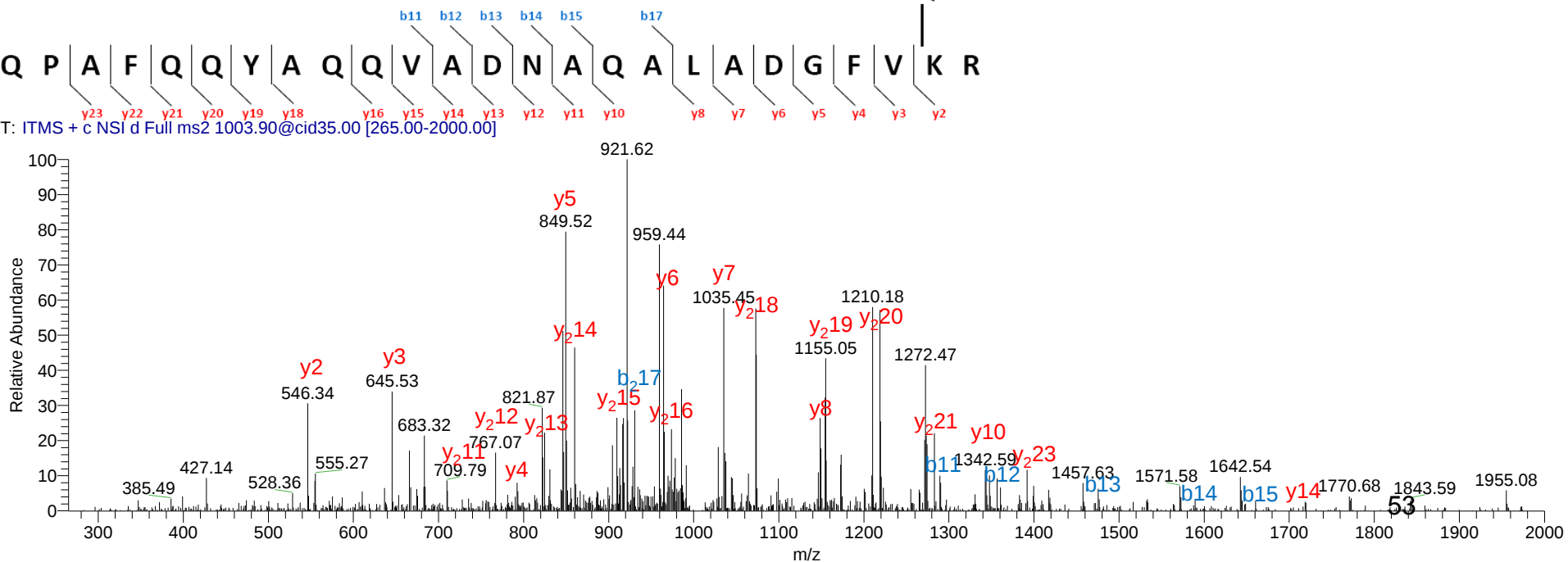
Charge = +3



LITQ_Native2 #14634 RT: 72.06 AV: 1 NL: 1.93E3
T: ITMS + c NSI d Full ms2 1073.12@cid35.00 [285.00-2000.00]



Charge = +3

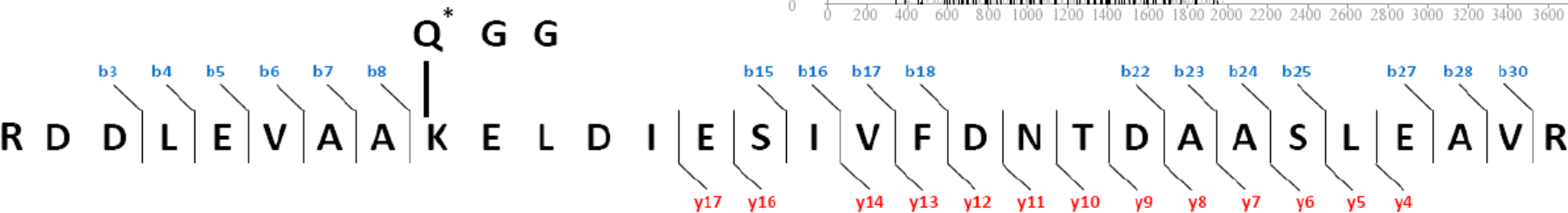
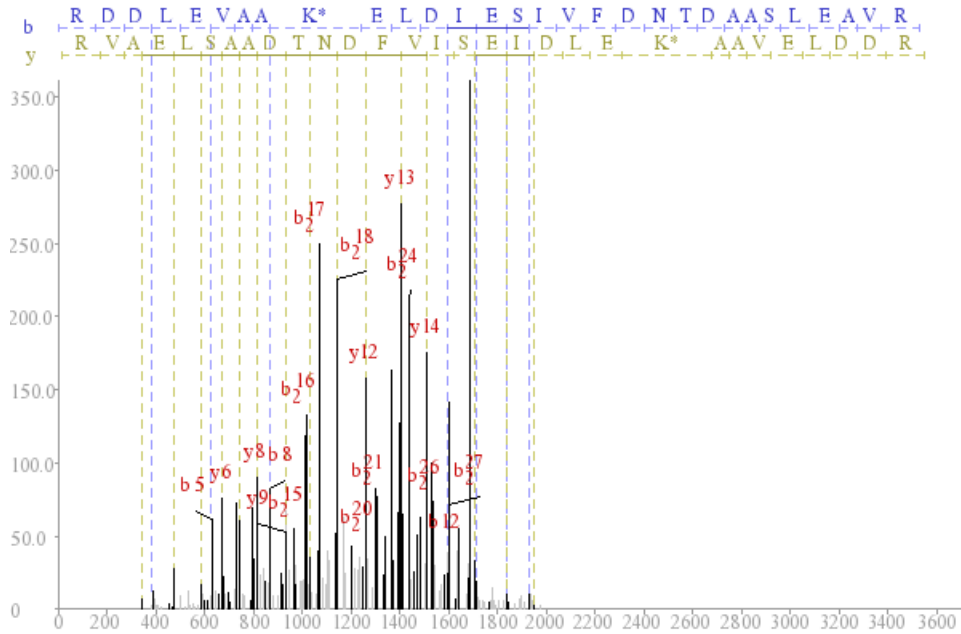


SF44: Short chain dehydrogenase

R.RDDLEVAAK+243ELDIESIVFDNTDAASLEAVR.G

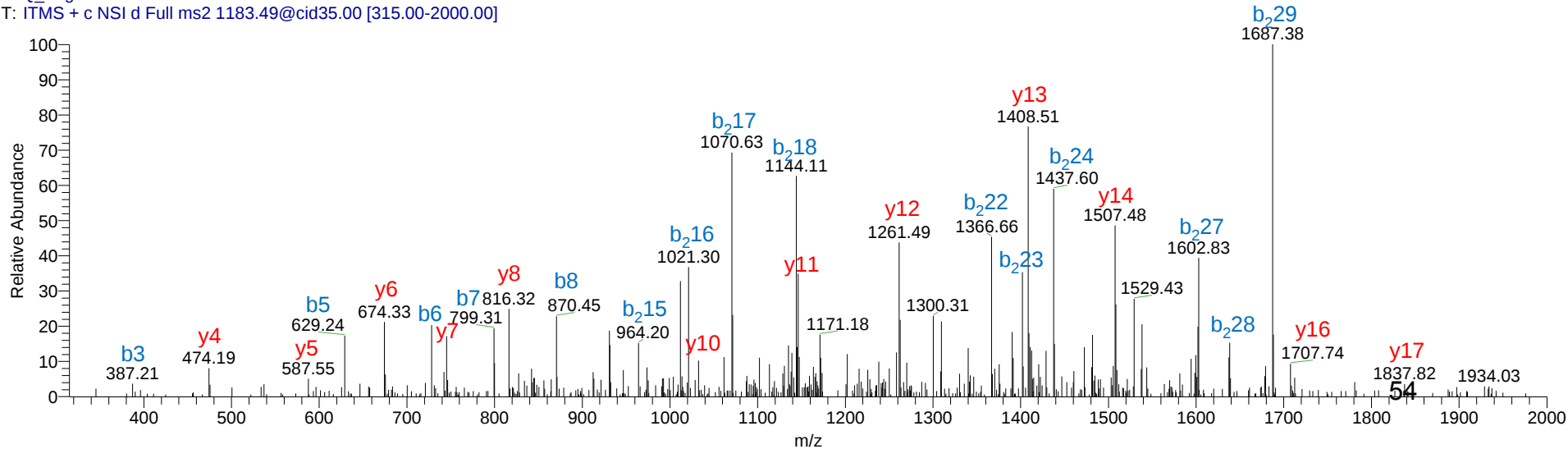
Calc MW: 3546.73 Da

Charge = +3



LTQ_Tag3 #18594 RT: 95.38 AV: 1 NL: 3.60E2

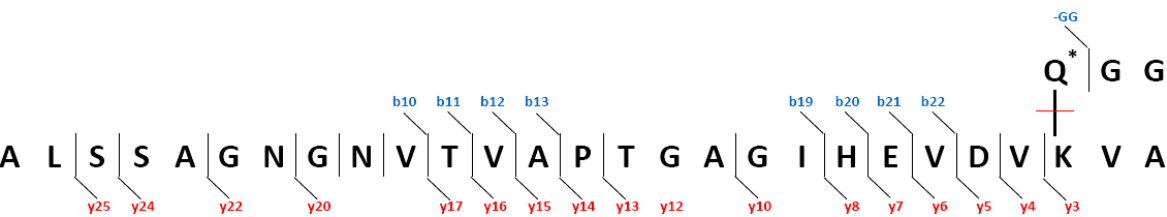
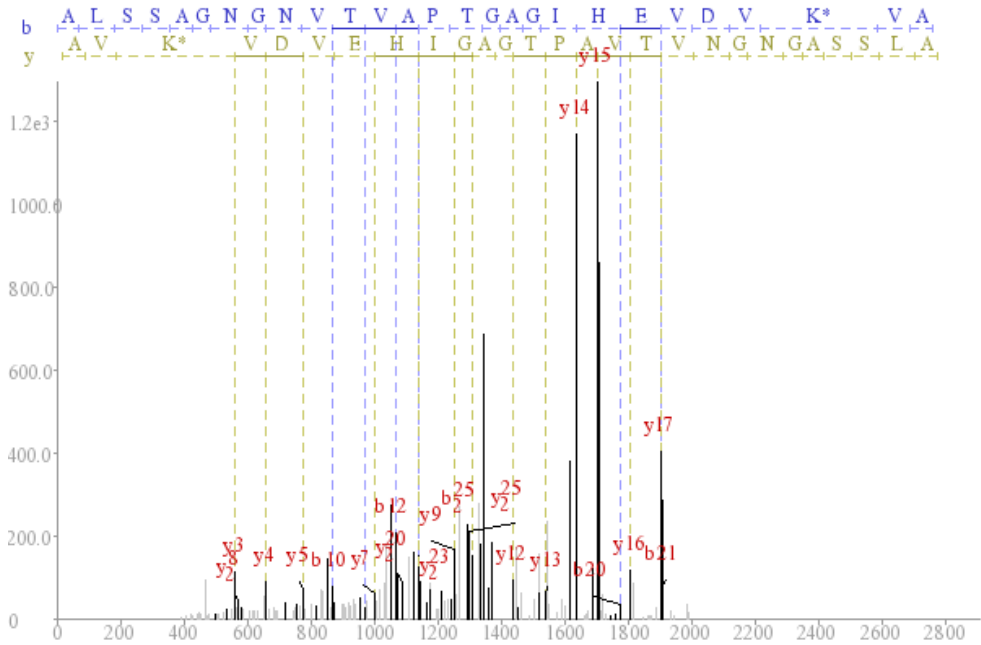
T: ITMS + c NSI d Full ms2 1183.49@cid35.00 [315.00-2000.00]



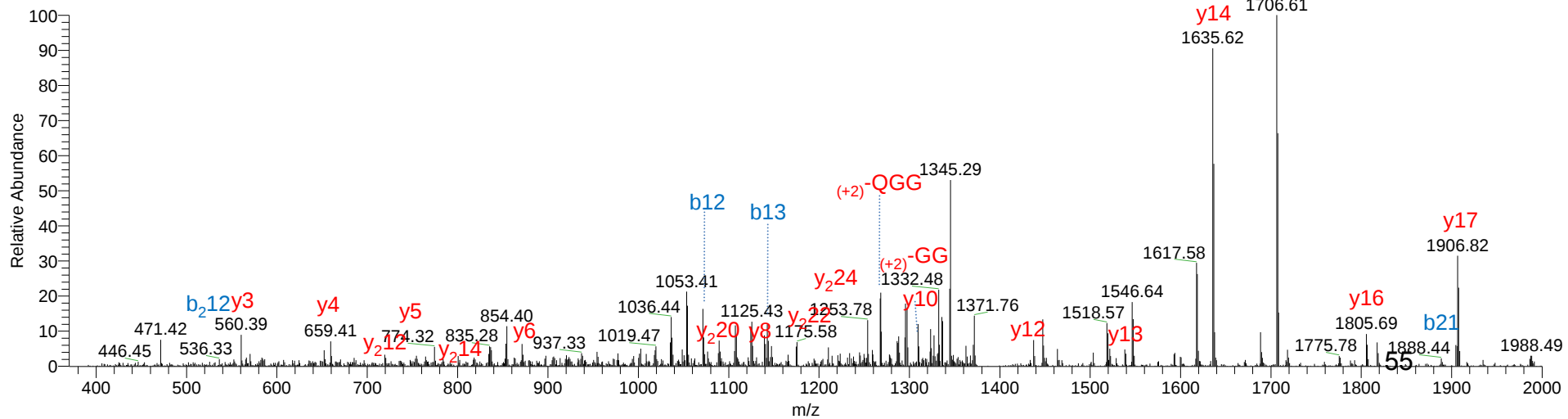
SF45: SseC protein

R.ALSSAGNGNVTVAPTGAGIHEVDVK+243VA

Calc MW: 2776.37 Da
Charge = +2



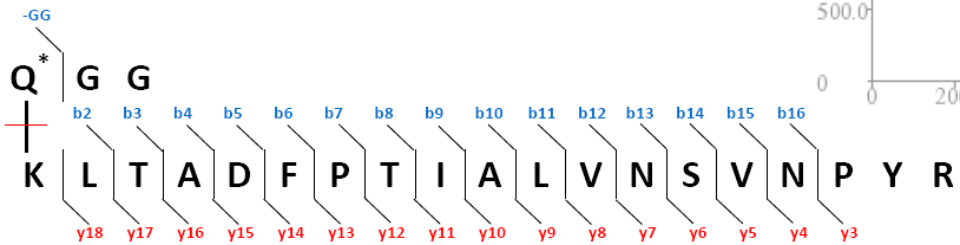
Tag_Skagg_7 #1965 RT: 30.43 AV: 1 NL: 1.29E3
T: ITMS + c NSI d Full ms2 1390.11@cid35.00 [370.00-2000.00]



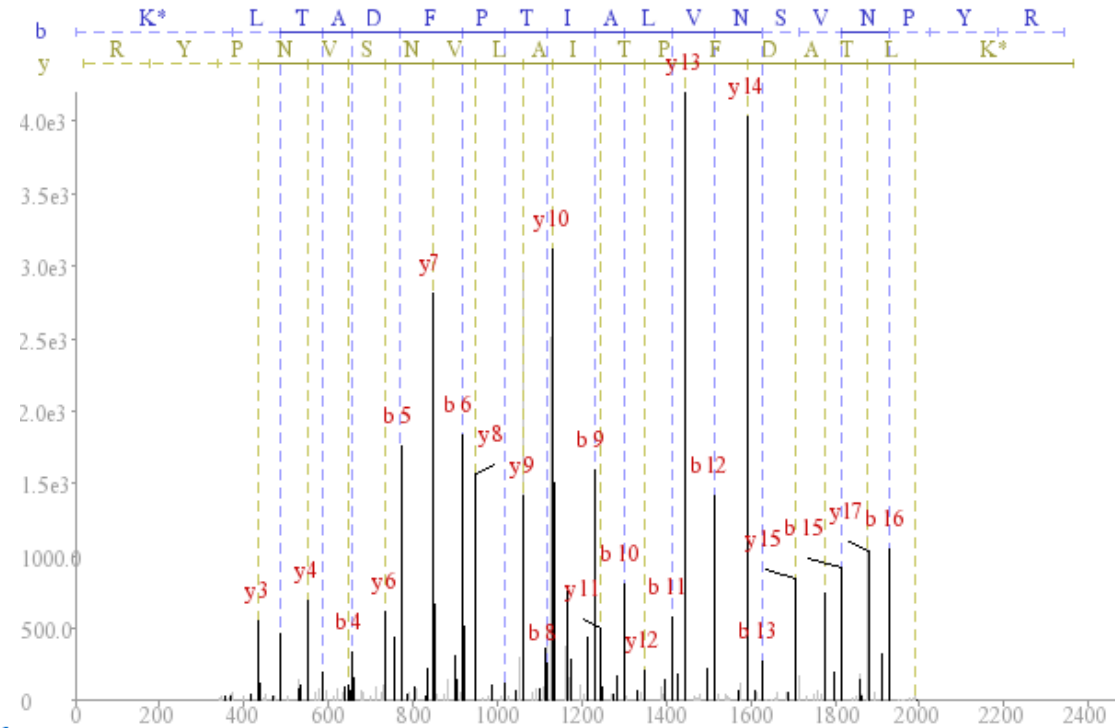
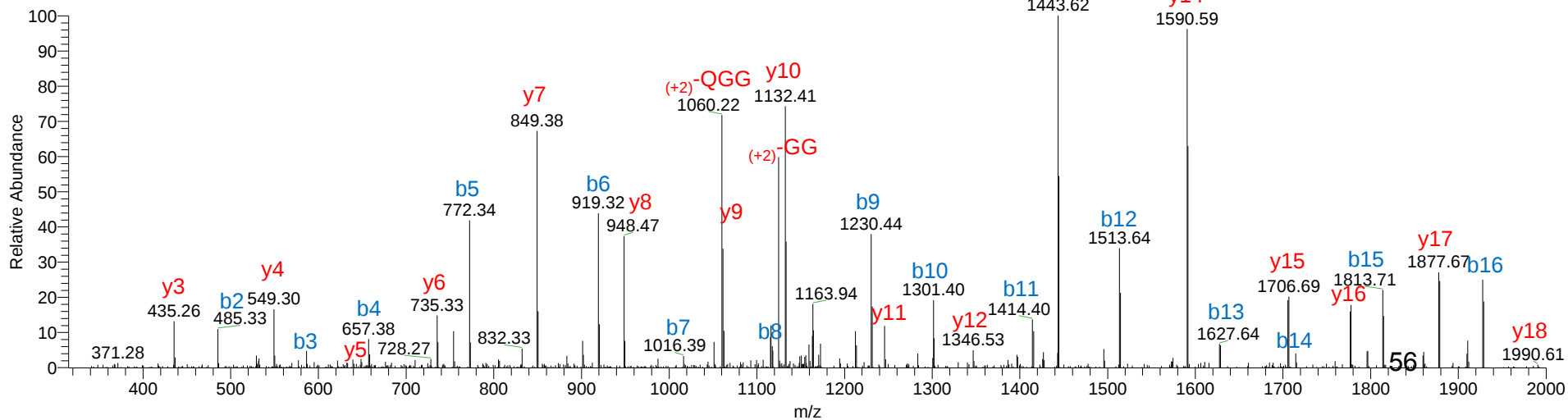
SF46: Threonine synthase

R.K+243LTADFPTIALVNSVNPYR.I

Calc MW: 2361.23 Da
Charge = +2



LTQ_Tag2 #18609 RT: 88.95 AV: 1 NL: 4.19E3
T: ITMS + c NSI d Full ms2 1181.69@cid35.00 [315.00-2000.00]

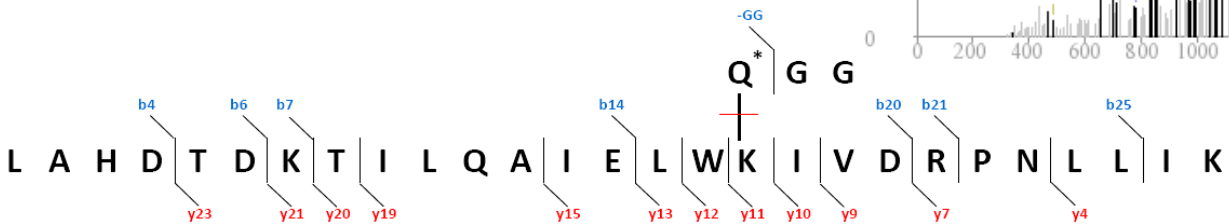
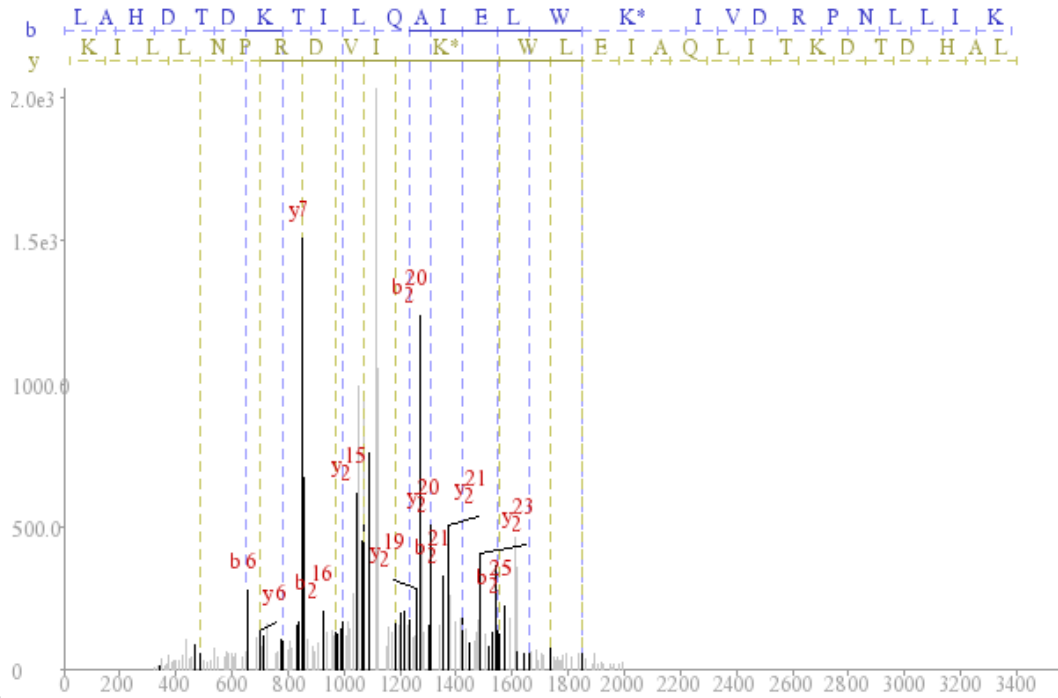


SF47: Transaldolase

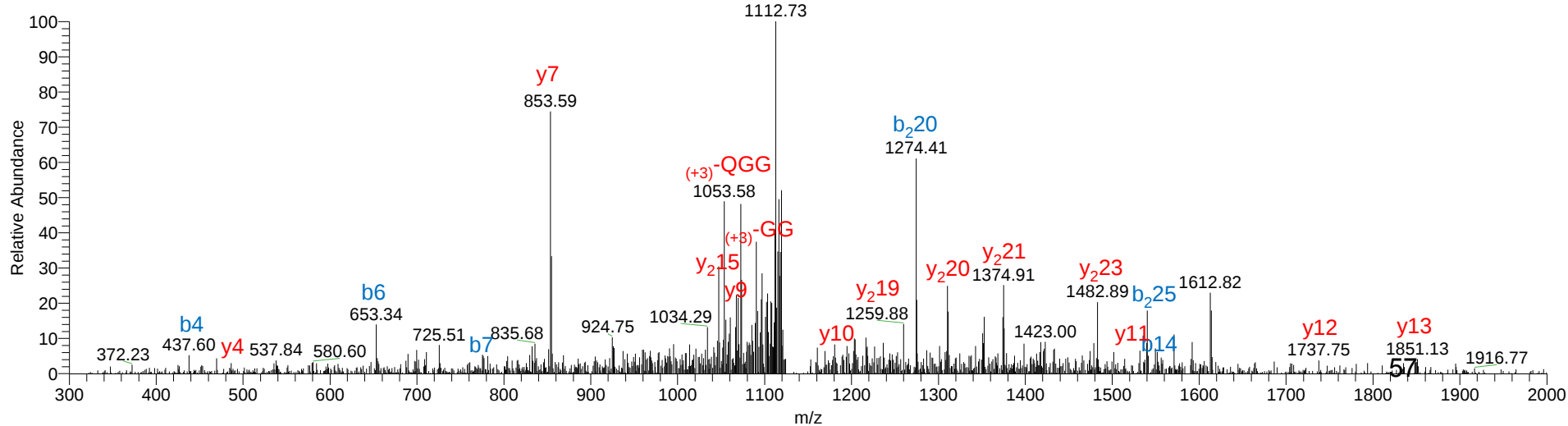
R.LAHDTDKTILQAIELWK+243IVDRPNLLIK.I

Calc MW: 3398.89 Da

Charge = +3



Tag_Skagg_5 #4928 RT: 48.65 AV: 1 NL: 2.03E3
T: ITMS + c NSI d Full ms2 1134.87@cid35.00 [300.00-2000.00]

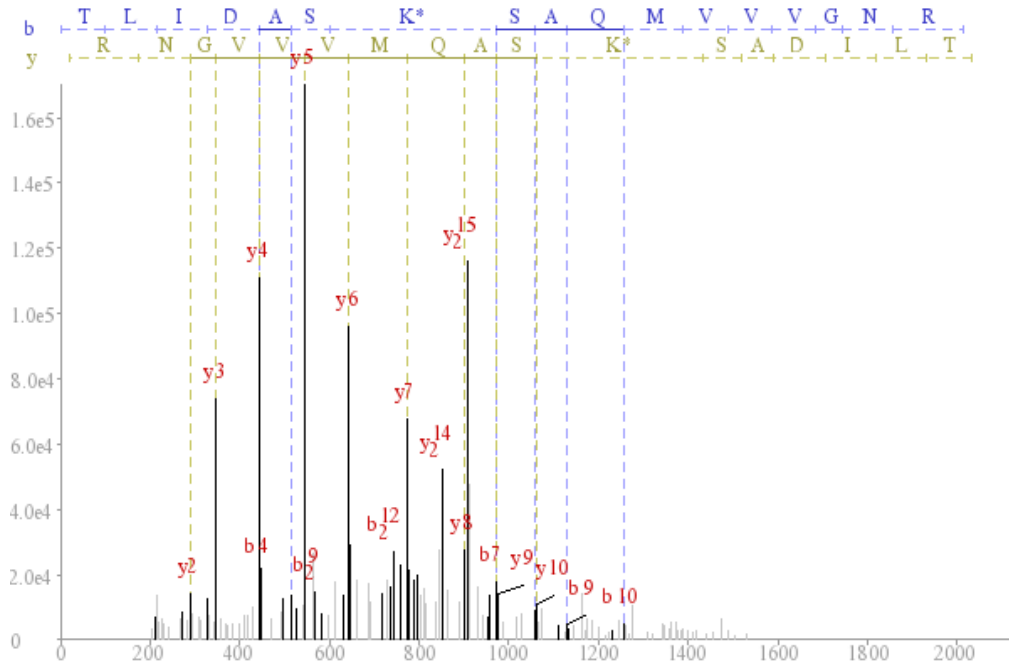
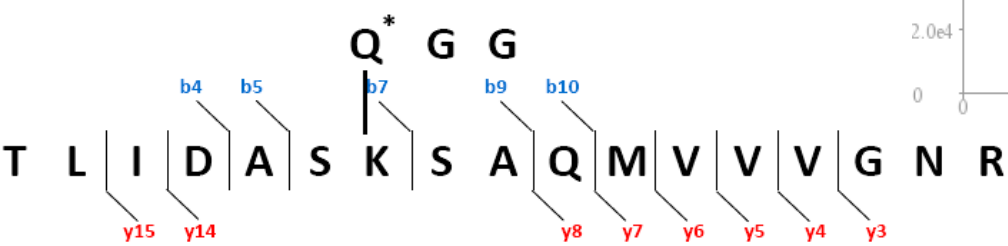


SF48: Universal stress protein family protein

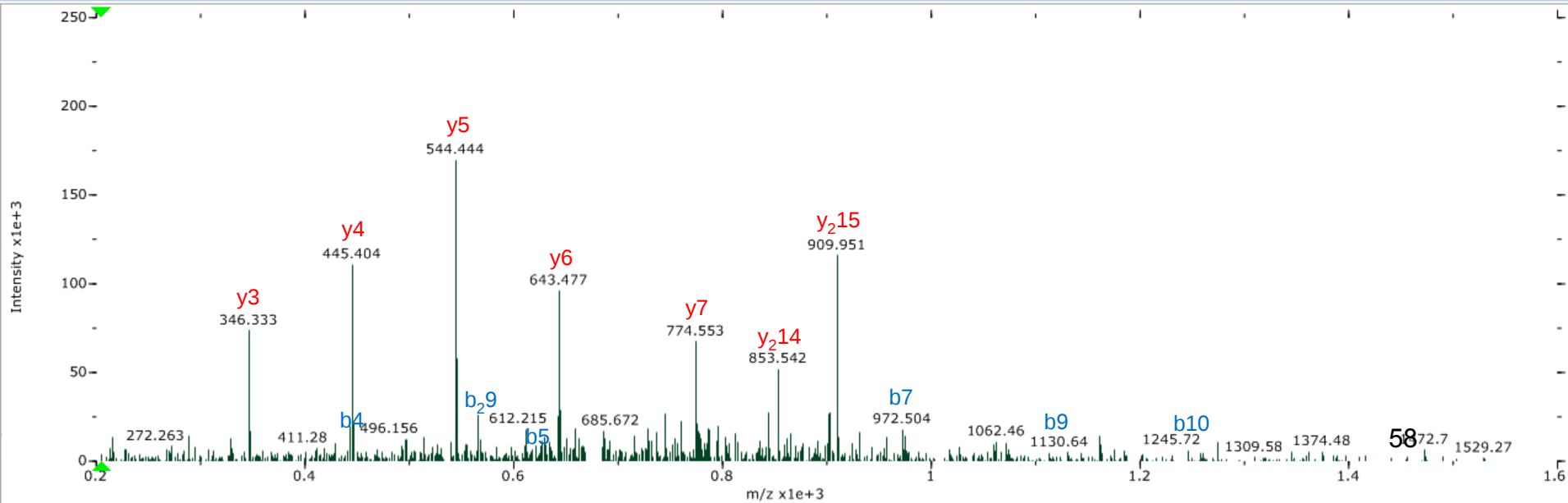
R.TLIDASK+243SAQMVVVGNR.G

Calc MW: 2031.04 Da

Charge = +3



Scan #3606 (678.80) (t=22.36 BP=169777.17@m/z=544.44): intensity=3499.70@m/z=205.12



IV. Sequence Alignments

- a. **Figure SF49: Sequence alignment of all PUPylated proteins centered at site of modification.....60**

This figure shows a ClustalW alignment +/- 10 amino acids around the modified lysine. The result shows no clear target recognition sequence near the site of modification (left) as well as no clear conserved hydrophobic or hydrophilic patches near the site of attachment (right). Alignment was done using the EBI server (www.ebi.ac.uk/Tools/clustalw2/index.html)

- b. **Figure SF50: Full length sequence alignment of all PUPylated proteins using ClustalW algorithm.....61**

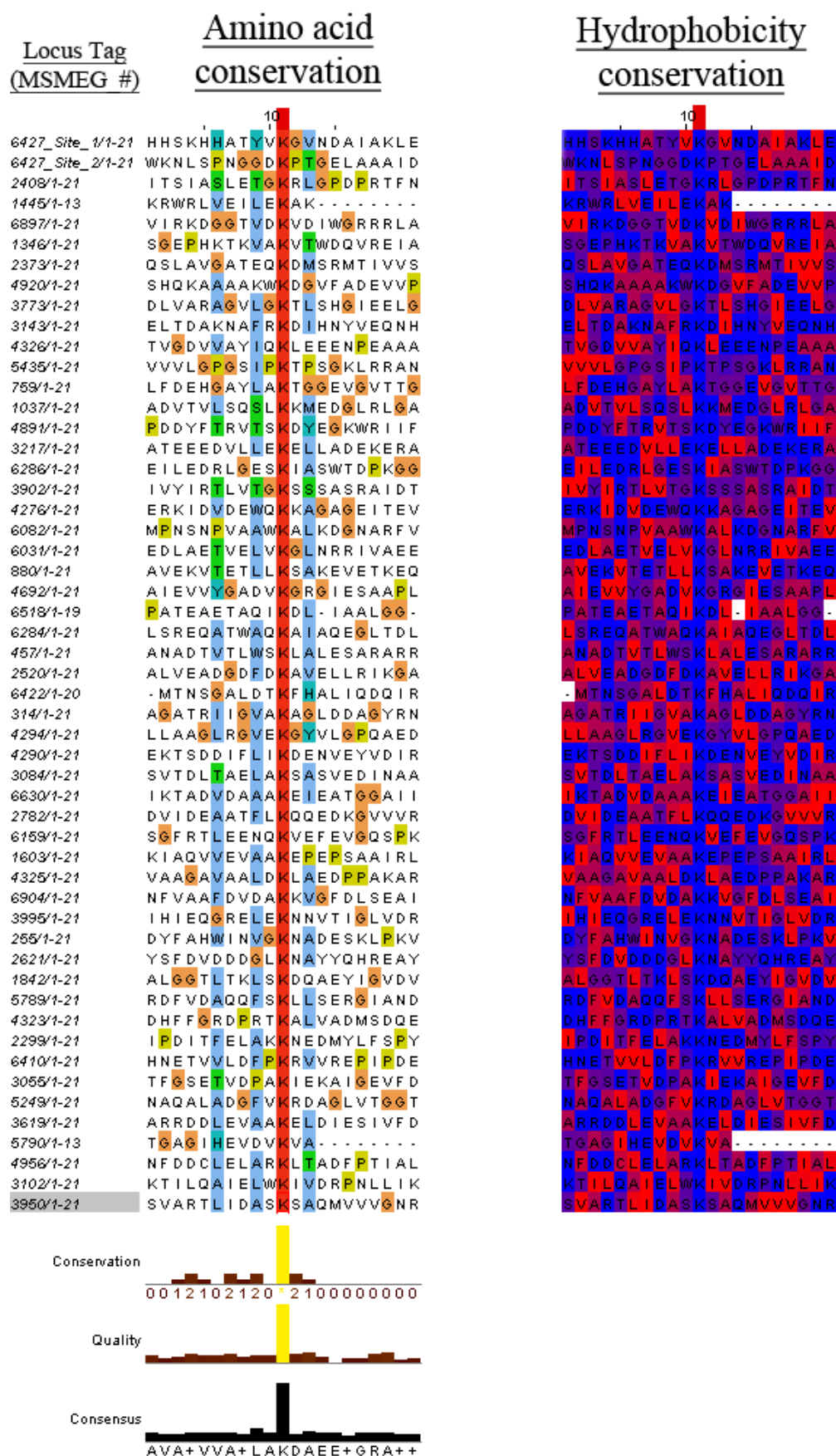
This figure shows a ClustalW alignment of all 52 full length PUPylated proteins in an effort to identify a conserved off site recognition sequence common in all modified proteins. The result shows no clear conservation at any point in the primary sequences.

- c. **Figures SF51-SF102: Individual sequence alignments of PUPylated proteins showing extent of evolutionary sequence conservation at site of modification.....68**

In an attempt to determine if the site of PUPylation was evolutionarily conserved, alignments of individual PUPylated proteins with homologous proteins from neighboring species (from BLAST) was performed. The results show minor conservation at the site of modification for most of the proteins with only a few being highly conserved. Alignment was done using the JGI server (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>)

Each figure for II.c. is oriented with the name of the protein in the upper left corner, the top BLAST hits chosen for alignment in the upper right and the resulting alignment on the bottom with red arrows indicating the site of modification.

SF49: Alignment at site of modification



Locus Tag
(MSMEG #)

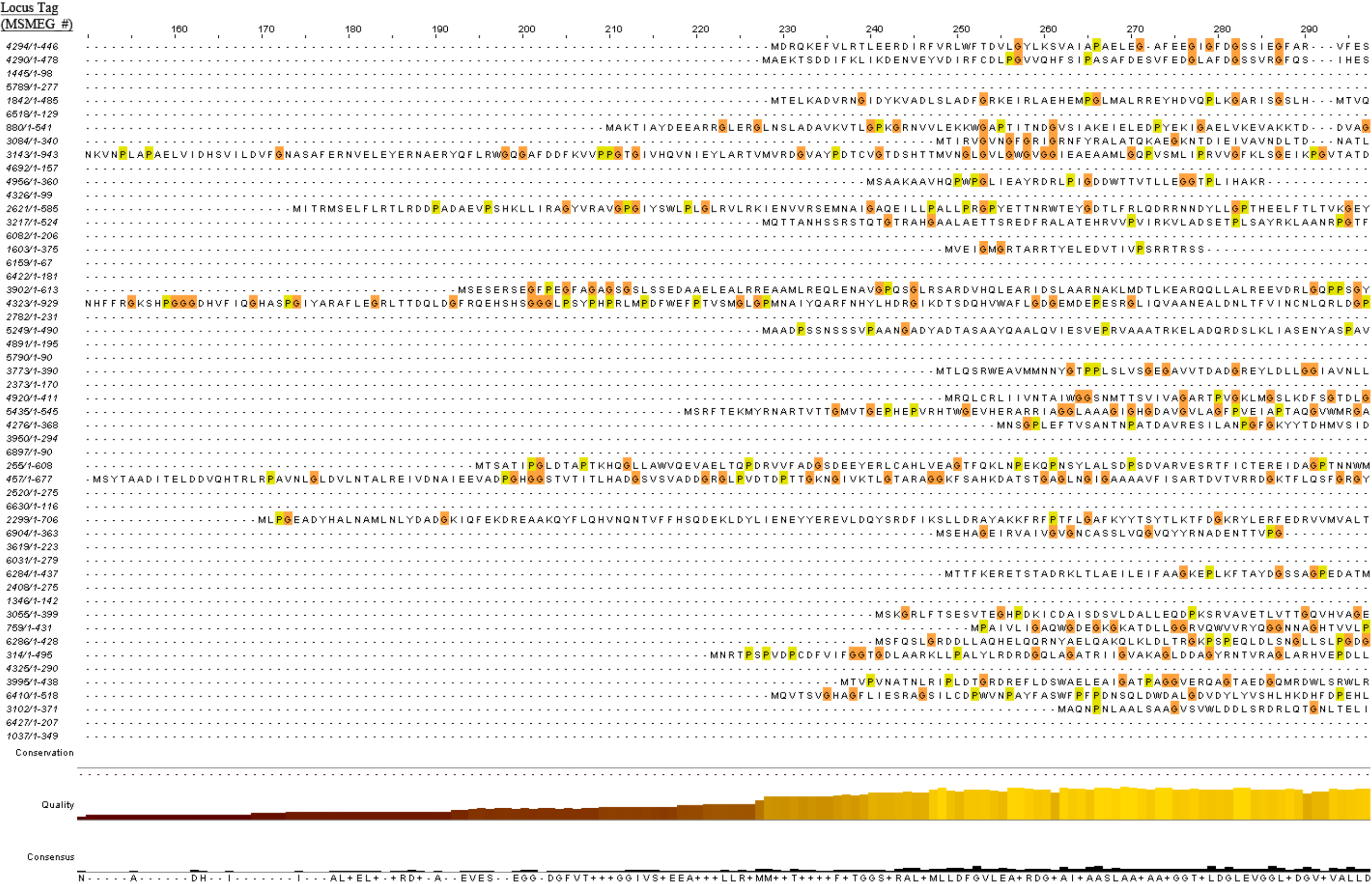
Conservation

Quality

Consensus

Sequence: MSSENTGKSSLSNFGARDTLTVGDQSYE IYRLNAVPGTEKLPYSLKVLAEENLLRTEDGANITKDHIEA IANWD PNAE P SIEIQFT PARVIMQDFTGVPCIVDLATMREAVAALGGDP

SF50: Full Length Sequence Alignment of All PUPylated Proteins Using ClustalW Algorithm (Part 2 of 7)



[illegible]

Consensus

63

Locus Tag
(MSMEG #)

450 460 470 480 490 500 510 520 530 540 550 560 570 580 590

4294/1-446 G I S V E F S H H E G - A P G Q Q E I D L R Y A D A L S M A D N V M T F R Y L V K E V A L A D G V R A S F M P K P F A E H P G S A M H T H M S L F E G D T N A F H - S P D D P L Q L S D V A K S F I A G I L E H A N E I S
4290/1-478 G F T L R G H H E V G T A G Q A E I N Y K F N T L L A A A D D V L L F K Y I I K N T A W Q A G K T M T F M P K P L F G D N G S G M H A H Q S L W K D G Q P L F H - D E S G Y A G L S D I A R H Y I G G I L H H A P S L L
1445/1-98 - - - - - M A D Q K G P - - - - - K Y - - - T P A A E - K P R G - - - - - R R K T A I G - - - - -
5789/1-277 Q Q F S K L L S E R G I A N D D T V I L Y G G N N N W F A A Y A Y W Y F K L Y G - - - H Q D V K L L D G G R K K W E L D A R P L S A E K V E R P Q T S Y T A K E - P D N S I R A F R D E V I A A I G
1842/1-485 V K G V T E E T T T G V L R L Y Q F A A A G E L A F P A I N V N D S V T K S K F D N K Y G T R H S L I D G I N R G T D V L I G G K K V L I C G Y G D V G K G C A E - S L A G Q G A R V S V T E I D P I N A L Q A L M D G F
6518/1-129 - - - - - M Y T A A I D A L - P P V G D A E F P E R A A V L S - - - - -
880/1-541 L E K V I Q S G K P L L I I A E D V E G E A L S T L V N K I R G T F K S V A V K A P G F G D R R K A M L Q D M A I L T G G Q V I S E E V G L S L E T A D V S L L G K A R K V V V T K D E T T I V E G A G D A E A I Q G R
3084/1-340 K G H L D A G A K K V I I S A P A T D E D I T I V L G V N D D K Y D G S Q N I I S N A S C T T N C L G P L A K V L N D E F G I V K G L M T T I H A Y T Q D Q N L Q D G P H K D L R R A R A A A L N I V P
3143/1-943 V E E S F P A S D P V S L S F A D D G A P D M R P S A A N G A T G R P T N P V L V H S E E R G D F V L D H G A V V V A G I T S C T N T S N P S V M L G A A L L A K K A V E K G L T T K P W V K T N M A P G S Q V V T D Y Y N K A G L W P Y L E K L G Y Y L G G Y G C T T C I G N T G P L P E E I S K A I N
4692/1-157 - - - - - M A S G D I A T V A N A E L D L P Y G S A L T S S G R I S A V T E P G E L S V H Y P F F T M D L V - - - - -
4956/1-360 A S A A A Y A A R A G I T C A V L I P Q G K I A M G K L A Q A V M H G A K I I Q V D G N F D D C L E L A R K L T A D F P T I A L V N S V N P Y R I E G Q K T A A F E I V D A L G T A P D V H A L P V G N - - - - -
4326/1-99 - - - - -
2621/1-585 A S L P Q F E G R T V T A A D T L K N V L L K T R E P G G E W E L L A V G V P G D R E V D E K R L G A A L E P A E F A L L D D A D F A A N P F L V K G Y V G P K A L Q D N G V R Y L V D P R V V H G S S W I T G A D A P N - - - - -
3217/1-524 T K A L G Q P L T S A V A T F S R P A P D H R A Q R T M E E Y T E I V D K L V G D I E A G E A F Q V V P S Q R F E M D T A A D P L D V Y R I L R M T N P S P Y M Y L L N I P D A D G G L D F S I V G S S P E A L V T V K D - - - - -
6082/1-206 - - - - - M P N S N P V A A W K A L K D G N A R F V A G Q P L H P S Q G I E R R A S L T Q A Q R P T A V V F G C G D S R V A A E I L F D Q G L G D M F V V R T A G H V I D N - - - - -
1603/1-375 V A A K E P E P S A A I R L L Q Q L H A A P L D P D L L G A A V A R I R E A G V T T A V R V S P Q N A Q A L T P T L V A A G I D L L V I Q G T I I S A E R V A S D G E P L N L K T F I S E L D P V V A G - - - - -
6159/1-67 - - - - -
6422/1-181 - - - - - M T N S G A L D T K F H A L I Q D Q I R S E F T A S Q Q Y I A I A V F F D G A D L P Q L A K H F Y A Q A L E E R N H A M M L V Q Y L L D R D M E V E I P - - - - -
3902/1-613 N D I G G L G R Q I E Q I R D A V E L P E L H K D L Y K E Y S L R P P K G V L L Y G P P G C G K T L I A K A V A N S L A K K M A E V R G D D A R E A K S Y F L N I K G P E L L N K F V G E T E R H I R L I F Q R A R E K A S E G - - - - - T P V I V F F D
4323/1-929 K N F R D V T R V P V S D A Q L E E D P Y L P P Y Y H P G P E A P E I R Y L L E R R R A L G G F V P S R R T K S K P L A L P G S D T Y K A L K K G S G S Q A V A T T M A T V R T F K E L L R D K N I G P R I V P I I P D E A R T F G - - - - - M
2782/1-231 - - - - - M A K L D F A L N S T I R Y L M F S V F A V A P G E L G E D R A D V I D E A A T F L K Q Q E D K G V V V R G L Y D V A G L R A D A D F M I W T H A D N V E A L Q S T Y S D F R R T A L G R - - - - -
5249/1-490 A R E F K P L V L V A G Y S A Y P R R V N F A K M R E I A D E V G A T L M V D M A H F A G L V A G K V F T G D E D P V P H A H V T T T T H K S L R G P R G G M V L A T E E Y A P A V D K G C P M V L G G - - - - -
4891/1-195 - - - - - M A L L T I G D Q F P E Y D L T A V V G G D L S K M D A K Q P D D Y F T R M T S K D Y E G K W R I I F F W P K D F T F V C P T E I A A F G K L N E D F E D R D A K V L G V S - - - - -
5790/1-90 - - - - -
3773/1-390 Q A P F E P L P G N V M H V P Y G D V A A L E A A V D D Q T A A V F L E P I M G E G G V V P P A G Y L V A A R E I T S K H G A L L V L D E V Q T G V R T G A F F A H Q H D G I V P D V V T M A K G L G - - - - -
2373/1-170 - - - - - M S N G T P T H T L S V L V E D K P G V L A R V S S L F S R R G F N I Q S L A V G A T E Q K D M S R M T I V V S V E D S P - - - - -
4920/1-411 V D H L A Y D G L H D V F T D Q P M G A L T E Q R N D V D K F T R A E Q D E Y A A Q S H Q K A A A A W K D G V F A D E V V P V S I P Q R K G D P I E F A E D E G I R A N T T A E S L A G L K P A F R K D G - - - - -
5435/1-545 D F L R D T L L W A K L I D K Y K G T M T A A P N F A Y A L L A K R L R R Q A K P G D F L S T L R F A L S G A E P V E A D V E D F L D A G K P F G L R P E A I L P A Y G M A E T T L A V S F S P C G E G L V V D E V D - - - - -
4276/1-368 G G E E S L Y L R P F V I A T E P G L G V R P S N E Y R Y L I A S P A G A Y F K G G I K P V S V M L S H E Y V R A S P G T G A A K F G G N Y A A S L L A Q A Q A A E M G C D Q V V W L D A I E R R - - - - -
3950/1-294 I A P V V V S W P A G P Y M A T V L E C Q E E N A R H A I E Q A Q K V V A D C L G E T H G L T V Q T E I R K E S V A R T L I D A S K S A Q M V V V G N R G M G A L G R V L L G S T S T S L L H Y A S G P - - - - -
6897/1-90 - - - - -
255/1-608 K L I S P E N K A Y F I A A A F P S A C G K T N L A M L Q P T I E G W R A E T V G D D I A W M R F G K D G R L Y A T N P E F G F F G V A P G T N W S S N P N A M K T I A A G N T V F T N V A K T D D G D V W W E G L E G D P Q - - - - -
457/1-677 A A F C N T V R T P G G G S H L T A A I K G L S E A L A E R A S R M R D L G L A K N E E G P E P Q D F A A V T A L A V D T R A P D V A W D S Q A K T A V S S R S L N L A M A P D V A R S V T I W A A N P A N A D T V T L W S K L A L E S - - - - -
2520/1-275 D K A V E L L R I K G A K D V G K R A E R A T A E G L V A A K D G A L I E L N S E T D F V A K N A E F Q A L A D Q I V A A A V A A K A N D I E T L K A A K T G D T T V E Q A I A D L S A K I G E K L E L R - - - - -
6630/1-116 - - - - -
2299/1-706 S L G V V I P D I T F E L A K K N E D M Y L F S P Y D V E R V Y G V P F A D V S V T E K Y Y E M V D D A R I R K T K I N A R E F F Q T L A K S Q F E S G Y P Y I M F E D T V N R S N P I A G K I T H S N L C S E I L Q V S T P S E F N - - - - - D D L S
6904/1-363 V K V L K E A E V D V L V S Y L P V G S E E A D K F Y A Q C A I D A G V A F V N A L P V F I A S D P W A A K F E D A G V P I V G D D I K S Q V G A T I T H R V M A K L F E D R G

[illegible]

Age Group	Percentage
18-24	18%
25-34	22%
35-44	15%
45-54	12%
55-64	10%
65-74	8%
75-84	5%
85+	3%

VA+V+L+++AG|IAA+++++LLT++G+VAAAAGALRALVVGAEGLDFV++A+++AA+AAAA+ALVVLAGLL+EGA++AALLA+LG+A+A+DVAAIAIGGG+++A+EGDSK...L.....

SF50: Full Length Sequence Alignment of All PUPylated Proteins Using ClustalW Algorithm (Part 5 of 7)

Locus Tag
(MSMEG #)

4294/1-446 --AVTNQWVNSYKRLVHGGEAPTAASWGAANRSALVRVPMYTTHKVSRRVEVRSFDSACNPYLTFVALLAAGLRGVEKG
4290/1-478 --AFTNPTVNSYKRLVPGYEAPINLVYSQRNRSACVRIPITGNNPKAKRLEFRCPDSSGNPYLAFAAMLMAIDGIDKKK
1445/1-98YVVSDDKMQKTIIVVELEDRKS...HPLYG...KIIRTTKKVKAHDENG...AGIGDRVSL
5789/1-277TKNLVDVRSFDEFSGKILAPAHLPQEESQRPFGHIPGAINVFWSKAANEDGTFKSDEELAKLYAEAGLDGEKETIAY
1842/1-485 --DVRTVEEAIGEADIVITATGNKDIITLDHMKAMKNQAILGNIGHFDNEIDMAALERSGAKKINIKPQVDQWIFDDGKS
6518/1-129GLRKLQGSALAEASRSRATPSVIVALSGVRTRYDELMTTAAEGPGATLQRLYVARLRAKLT
880/1-541 --VAQIRAEIENSDDSDYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVALLQSAPS
3084/1-340TSTGAAKAIGLVLPKLGKLDGYALRVPIPTGSVTDLTAEAKSASVEDINAAMKAAAEGLKGLIKYYDAP
3143/1-943 DNDLAVTAVLSGNRNFEGRISPDKVMNYLASPLLVIAYGIAGTMDDFESDPLGQDSEGNDFVLRDIWPSAAEIEETIASSINREMFESYADVFKGDDRWRSLLPT--PEGDTFEWDPASTYVRKAPYFDGMPAEPEPVSDIKGARVLA
4692/1-157VLDDALKYGSRAAKARFAVYIGPLGADTAATAREILANVPTPENAVLLAVSPDQRAIEVVYGAD
4956/1-360AGNITAYWKGYSEYHRDGVSDRLPRMLGTQAAGAAPLVGTGAPVKDRETATAIRIGSPASWNSAVEAQQQSDG
4326/1-99MAATQEEIIAGLAEIIEEVTGIEPSEVTEKSFVDDLDIDSLSMVEIAVQTEDKYGVKIPDED
2621/1-585RHVVGLVAGRDFTPDGTIEAAEVRDGGPSFDDAGVLTSAEGIEIGHIFQLQKRYTDAFSAADVLEDGKPLRLTMGSYG
3217/1-524GRATTHPIAGTRWRGATEEEDVLLKELLADEKERAHMLVDLGRNDLGRVCRPGTIVRVDYSHIERYSHVMHLVST
6082/1-206AVLGSIEYAVTVLKVPLIVVLGHDSCGAVKATLSALDEGEVPSGQFVRDIVERVTPSILLGRKAG
1603/1-375GVLDHRTALHLMRTGAAGVIVGYSSTSGVTTSDDEVLGISVPMATAIADAAAARREYLDDETGGRYVHVLADGDHSS
6159/1-67MPQGTIVKWFNAEKGGFIIAPEDGSADVVFVHYTEIQGSGFRT
6422/1-181GIDPVCNNFTTFRDALALALDQERTVTEQISRLASVARDEGDHLQEQFMQWFLKEQVEEVAAMTT
3902/1-613 EMDSIFRTRGTGVSSDVETTVVQPLLSEIDGVEGLENVIVIGASNREDMIDPAILLRPGRLDVKIKIERPDAAEAQDIFSKYLTEDLPVHADDLTFENG
4323/1-929 DSWFPSLKIYNRNGQLYTSVDSELMMLAYKESEVQQLIHGEGINEAGTSSFTAVGTSYSTHDEPMIPIYIFYSMFGFQRTGDGLWAAADQMARGFVLGATAG
2782/1-231ISDPVMSSVALHRPAEFNKSHIPAFLAGEEPNGYICVYPFVRSYEWYLLPDEERRMLSEHGMAARG
5249/1-490PLSHVMAAKAVALAEARQPAFQQYAAQVADNAQALADGFVKRDAGLVTTGGTDNHIVLLDVTSTGLTGRQAESALLDAG
4891/1-195VDNEFVHFQWRAQHEDLKTLPFPMVSDLKRELTAACGVNLADGVADRATFIVDPNNEVQFVSVT
5790/1-90MPAGVDLEKETVITGRVVDGSGQAVGGAFVRLLDGSDFTAEVVASATGDFRFFAAPG
3773/1-390GGLPIGACLAAGATGLLTPLGLHSGTGGNPNVCTAAGLAVLKTAAEDLVARAGVLGKTLSHGIEELG
2373/1-170LEQITKQLNKLINVIKIVEQEEDNSVSRELALIKVRADATTROQIIIEAVNLFRAKVVDVSTES
4920/1-411TITAGSASQISDGAAGVIVMNKAKAEELGLTWLAEIGAHGVVAGPDSLTQSQPANAIKKAITREG
5435/1-545ADLLAALRRAPATKGNTRKRLASLGLLLNDLEARVVDENGEMVPARGGVIELRGEPTPGYIMGGFVPAQDEYQWYDTG
4276/1-368YVEEMGGMNLFFVFGSGGSARLVTPELSGSLLPGITRDSLLQLATDAGFAVEERKIDVDEWQKKAGAG
3950/1-294VVVVHGGDDQAAHDSRLPVLLGIDGSPASEVATSHAFDEASRRGVDLVALHVMVDVGDIPPIGPTWEEQ
6897/1-90MVILDPTLDERTVAPSLETFNLVIRKDDGTVDKVDIWRRRRLAYEIAKHAEIYAVIDVKAEPAT
255/1-608 -HLIDWKGNDDWTPESEGAHAHPNSRYCTPISQCPTLAPFWDPPQGVPIISAILFGRRKTTVPLITEARDWQHGVFIGATLGSEQ
457/1-677 ---ARARRSAEGAKARARAASKAKGLGTLNLSLPPKLLPSRESGRGSGAELFLCEGDSALGTIKAARDAFQAAFPLKGGKPPNVYGFPLNKARAKDEFDAIERILGCGVRDHCDPELCRYDRILFASDADPDGGNINSSLSIMFLDFYRP
2520/1-275RAAYFDGTVEAYLHKRAADLPAPVGLVVEYQAGDADKQKEAAHAVALQIAALKAKYLTREDVPDIVANE
6630/1-116MSDHEVKMVLSTENLDESIFKYETLGFSLKFRDGAALGAVTLALATPVDHPIPGKVVVG
2299/1-706 YAKVGKDISCNLGSNLIAKAMDSPDFAQTIEVAIRALTAVSDQTHITSVPSIEQGNNDSHAIGLQGMNLHGYLARERIYYGSEEIGDFTNIYFYTVLFLHALRASNKIAIERG
6904/1-363GNMDFLNMLERSRLESKKVSKTQAVTSNLSSALAGKVEDKNVHIGPSDHVAVWLDNRKWAVVRLGFRAGGDVP
3619/1-223NLDAITVSAALTQIVGDHLRSGGSIIVNVLPEAKEGSVEAAIKAAALSDWTAQADDFGTRG
6031/1-279DMIGVNLSGVWKTVKAAPVPHLISGGNGSIIILTSSVGLKAYPHTOHYIAAKHGVVGLMRTFAVELGQHS
6284/1-437LTDLAEVRRHGDIYRDVIESGFDVAVSSIGLTEHIGVHNYPAYFNFLKSKLRTGGLLLNHCIIRPDNRSAPSGGFIDR
2408/1-275FPEHLTTLIEETGIVPAVNQIELHPLLPQQLRQVHAKLGIATEAWSPLGQGSLLADPVIIGIAEQHG
1346/1-142AATESQRGNVIFVEITVYEDRSFTFALKTPFAAKLLLKAAGVQKGSQEPHKTQKVAKVTVWDQVR
3055/1-399LTGRKIIVDTYGGWARHGGGAFSGKDPSSKVDRSAAYAMRWVAKNVVAAGLAERVEVQVAYAIGKAAPVG
759/1-431IGPTRITTVLGIILKAYTTRVSGGPFTELFDHGAYLAKTGGEVGVTTGRARRCGWFDVAVIARYATRVNG
6286/1-428ASADNIAWYLKHAGKKSIGPKVNLRLRHLRFFGDADGVRRQMQRHRELIAPKFALVAEILEDRLGESK
314/1-495VPFYLRGKRMSSQRFSEIVVQFKPVPLPMFPGIEGTSEPNRLIISLQPDFAIRLEMTAKEPGSGGRLRPVS
4325/1-290AVAALDKLAEDPPAKARVRKLATAGAFHTHYMASALDGYAAAAQSVITSEPTATLLSNADGQPVAS
3995/1-438EELHTSCGQLTVLPNSPVVVPREVHMHLDLRSNDDELLAAADAALRRRIAEAEIRAGVKVEHRAHVWPG
6410/1-518WAPSGEPMLLEALRFEIMTQDQICDGIQVPELRLTGRDHNETVVLDFPKRVVREPIDEKFRYGFELP
3102/1-371KAGVANARLAYAAEEVFGSDRFALKKADGARVQRPLWASTGVKNPEYSDTLYVTELVAINTVNTMP
6427/1-207KPTGELAAIIDQFGSFDKFAQQTAAANGLQSGWAVLGYDSLGRLLTFQLYDQGANVPLGIIP
1037/1-349HVAVKLAKAMGADVTVLSQSLKKMEDGLRLGASAYYATSDPETFDKLAGSFDLILNTVSANLDLGAYLE

Conservation

.....0000001011100110102001110010011001011100211102200100010120101000.....0022111110000000.....

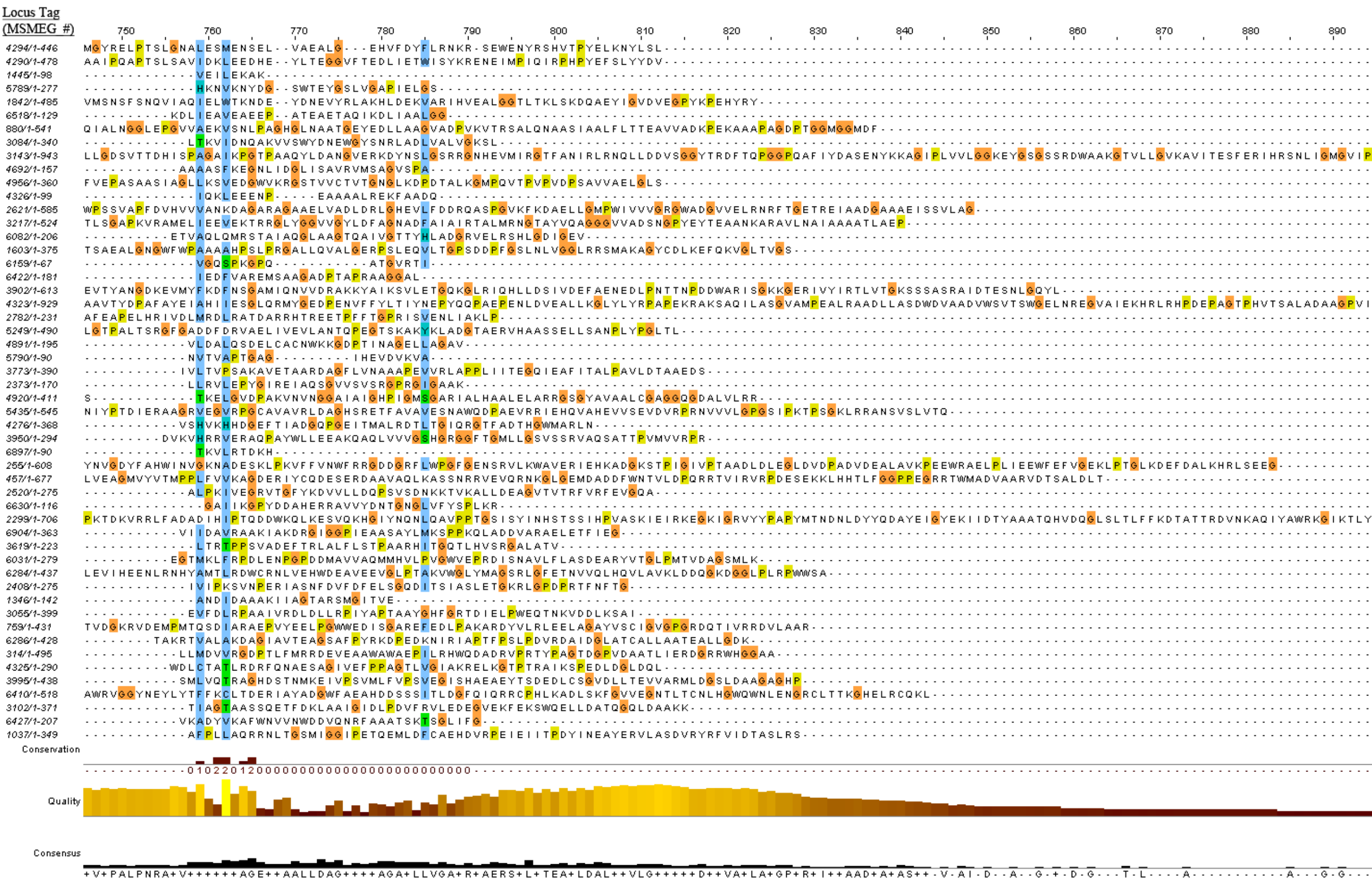
Quality



Consensus

D-DV++R+VAGS+AAV+GGAALAK+LATLA+VP++VLIP+SGASG+K+ELAEALGA++RAGEVLLAAA+LVGAGI+VKAAAGGEEGGP...AR...GDDA-R-L.....E....E.....RALTILELAD+VAA+LLSL+LAE+LA

SF50: Full Length Sequence Alignment of All PUPylated Proteins Using ClustalW Algorithm (Part 6 of 7)



SF50: Full Length Sequence Alignment of All PUPylated Proteins Using ClustalW Algorithm (Part 7 of 7)

Locus Tag (MSMEG #)	900	910	920	930	940	950	960	970	980	
4294/1-446										
4290/1-478										
1445/1-98										
5789/1-277										
1842/1-485										
6518/1-129										
880/1-541										
3084/1-340										
3143/1-943	LQFPAGESAASLKLDGTETDYDIEGIEELNSGKTPKTVHVTATKEDGSKVEFDAVVRIIDTPGEADYYRNGGILQYVLRNMLKSSK									
4692/1-157										
4956/1-360										
4326/1-99										
2621/1-585										
3217/1-524										
6082/1-206										
1603/1-375										
6159/1-67										
6422/1-181										
3902/1-613										
4323/1-929	AVSDWMRAVPEQIRPWVPGTYVTLGTDGFGFSDFRPAARRYFNTDAESVVVAVLQLGLARDGEIDASVAAQAAEQYRIDDVSAAGVSYADTGSA									
2782/1-231										
5249/1-490										
4891/1-195										
5790/1-90										
3773/1-390										
2373/1-170										
4920/1-411										
5435/1-545										
4276/1-368										
3950/1-294										
6897/1-90										
255/1-608										
457/1-677										
2520/1-275										
6630/1-116										
2299/1-706	YIRLRQMALEGTEVEGCVSCML									
6904/1-363										
3619/1-223										
6031/1-279										
6284/1-437										
2408/1-275										
1346/1-142										
3055/1-399										
759/1-431										
6286/1-428										
314/1-495										
4325/1-290										
3995/1-438										
6410/1-518										
3102/1-371										
6427/1-207										
1037/1-349										
Conservation										
Quality										
Consensus	A	G			D	V		G	E	D

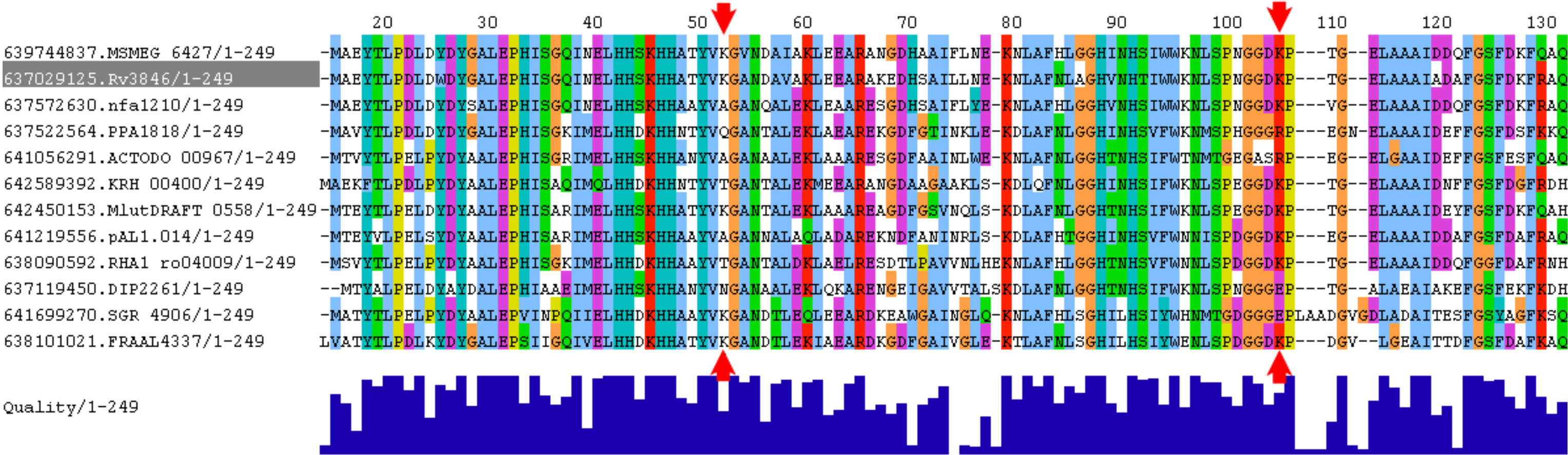
SF51: [Mn] Superoxide Dismutase
(GI: 118469767)

Site 1: HHATYVK+243GVNDIAIAK

Site 2: NLSPNGGDK+243PTGELAAAIDDQFGSFDK

Red arrows indicate PUPylation site on Msm (MSMEG)

Gene Object ID	Locus Tag	Product Name	Genome	Scaffold ID
637029125	Rv3846	SUPEROXIDE DISMUTASE [Fe] SODA	Mycobacterium tuberculosis H37Rv	NC_000962
637119450	DIP2261	manganese superoxide dismutase	Corynebacterium diphtheriae NCTC 13129	NC_002935
637522564	PPA1818	superoxide dismutase [Mn/Fe]	Propionibacterium acnes KPA171202	NC_006085
637572630	nfa1210	putative superoxide dismutase	Nocardia farcinica IFM 10152	NC_006361
638090592	RHA1_ro04009	superoxide dismutase	Rhodococcus sp. RHA1	NC_008268
638101021	FRAAL4337	Superoxide dismutase [Fe-Zn] 1 (FeSOD I)	Frankia alni ACN14a	NC_008278
639744837	MSMEG_6427	[Mn] superoxide dismutase	Mycobacterium smegmatis MC2 155	NC_008596
641056291	ACTODO_00967	hypothetical protein	Actinomyces odontolyticus ATCC 17982	NZ_AAY102000004
641219556	pAL1.014	putative Fe /Mn superoxide dismutase	Arthrobacter nitroguajacolicus Rue61a plasmid pAL1	NC_009453
641699270	SGR_4906	Fe-Zn-superoxide dismutase	Streptomyces griseus griseus NBRC 13350	NC_010572
642450153	MlutDRAFT_0558	Superoxide dismutase	Micrococcus luteus NCTC 2665	NZ_ABLQ01000003
642589392	KRH_00400	superoxide dismutase	Kocuria rhizophila DC2201	NC_010617

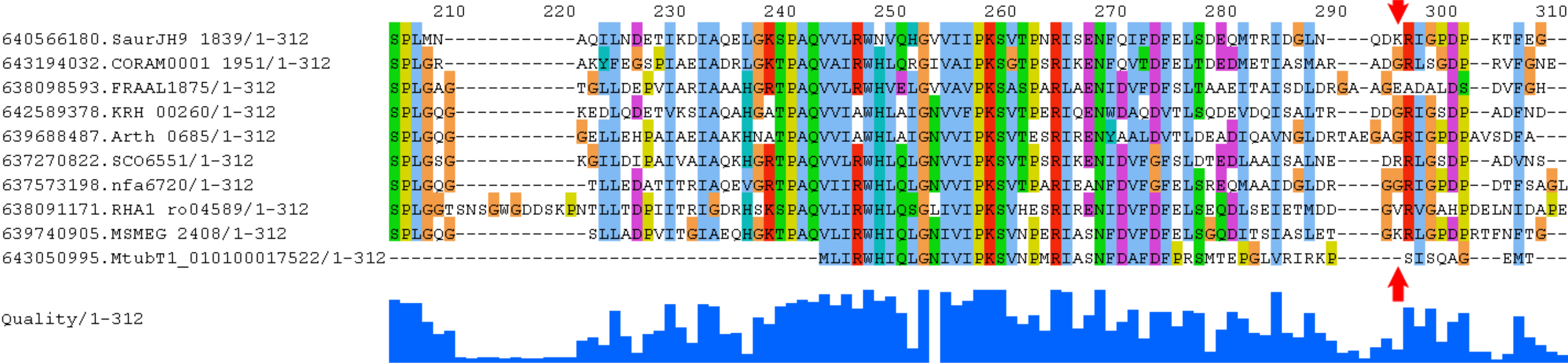


SF52: 2,5-diketo-D-gluconic acid reductase A
(GI: 118173010)

IASNFDVDFELSGQDITSIASLETGK+243R

Red arrows indicate PUPylation site on Msm (MSMEG)

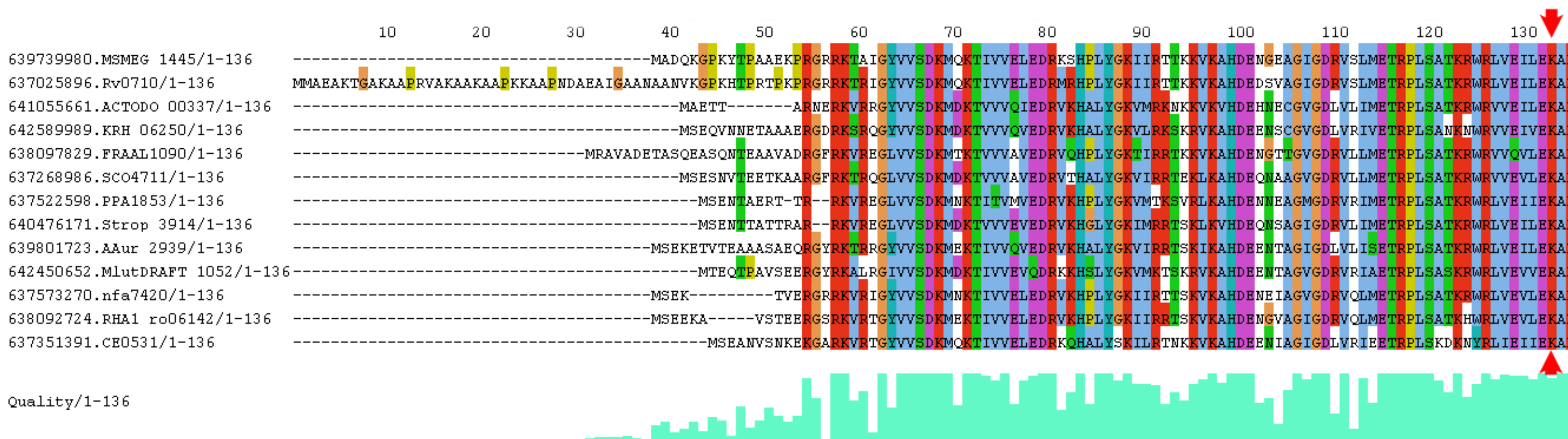
Selection	Gene Object ID	Locus Tag	Product Name	AA Seg. Length	Genome	Batch ¹
☒	639740905	MSMEG_2408	2,5-diketo-D-gluconic acid reductase A (EC:1.1.1.274)	275aa	Mycobacterium smegmatis MC2 155	2
☒	642589378	KRH_00260	2,5-diketo-D-gluconic acid reductase (EC:1.1.1.274)	276aa	Kocuria rhizophila DC2201	2
☒	637270822	SCO6551	oxidoreductase	277aa	Streptomyces coelicolor A3(2)	2
☒	640566180	SaurJH9_1839	2,5-didehydrogluconate reductase (EC:1.1.1.274)	277aa	Staphylococcus aureus aureus JH9	1
☒	638098593	FRAAL1875	2,5-diketo-D-gluconate reductase A (EC:1.1.1.274)	281aa	Frankia alni ACN14a	1
☒	639688487	Arth_0685	2,5-didehydrogluconate reductase (EC:1.1.1.274)	281aa	Arthrobacter sp. FB24	1
☒	637573198	nfa6720	putative oxidoreductase (EC:1.1.1.274)	283aa	Nocardia farcinica IFM 10152	2
☒	643194032	CORAM0001_1951	2,5-diketo-d-gluconic acid reductase a	288aa	Corynebacterium amycolatum SK46	1
☒	638091171	RHA1_ro04589	probable 2,5-diketo-D-gluconate reductase (EC:1.1.1.274)	293aa	Rhodococcus sp. RHA1	2
☒	643050995	MtubT1_010100017522	hypothetical protein	56aa	Mycobacterium tuberculosis T17	2



LVEILEK+243AK*

Quality/1-136

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Bact
☒	641055661	ACTODO_00337	SSU ribosomal protein S17P (IMGterm)	88aa	Actinomyces odontolyticus ATCC 17982	30
☒	639801723	AAur_2939	SSU ribosomal protein S17P (IMGterm)	98aa	Arthrobacter aurescens TC1	30
☒	637351391	CE0531	SSU ribosomal protein S17P (IMGterm)	92aa	Corynebacterium efficiens YS-314	30
☒	638097829	FRAAL1090	SSU ribosomal protein S17P (IMGterm)	105aa	Frankia alni ACN14a	30
☒	642589989	KRH_06250	SSU ribosomal protein S17P (IMGterm)	95aa	Kocuria rhizophila DC2201	30
☒	642450652	MlutDRAFT_1052	SSU ribosomal protein S17P (IMGterm)	93aa	Micrococcus luteus NCTC 2665	30
☒	639739980	MSMEG_1445	SSU ribosomal protein S17P (IMGterm)	98aa	Mycobacterium smegmatis MC2 155	30
☒	637025896	Rv0710	SSU ribosomal protein S17P (IMGterm)	136aa	Mycobacterium tuberculosis H37Rv	30
☒	637573270	nfa7420	SSU ribosomal protein S17P (IMGterm)	89aa	Nocardia farcinica IFM 10152	30
☒	637522598	PPA1853	SSU ribosomal protein S17P (IMGterm)	90aa	Propionibacterium acnes KPA171202	30
☒	638092724	RHA1_ro06142	SSU ribosomal protein S17P (IMGterm)	93aa	Rhodococcus sp. RHA1	30
☒	640476171	Strop_3914	SSU ribosomal protein S17P (IMGterm)	91aa	Salinispora tropica CNB-440	30
☒	637268986	SCO4711	SSU ribosomal protein S17P (IMGterm)	95aa	Streptomyces coelicolor A3(2)	30

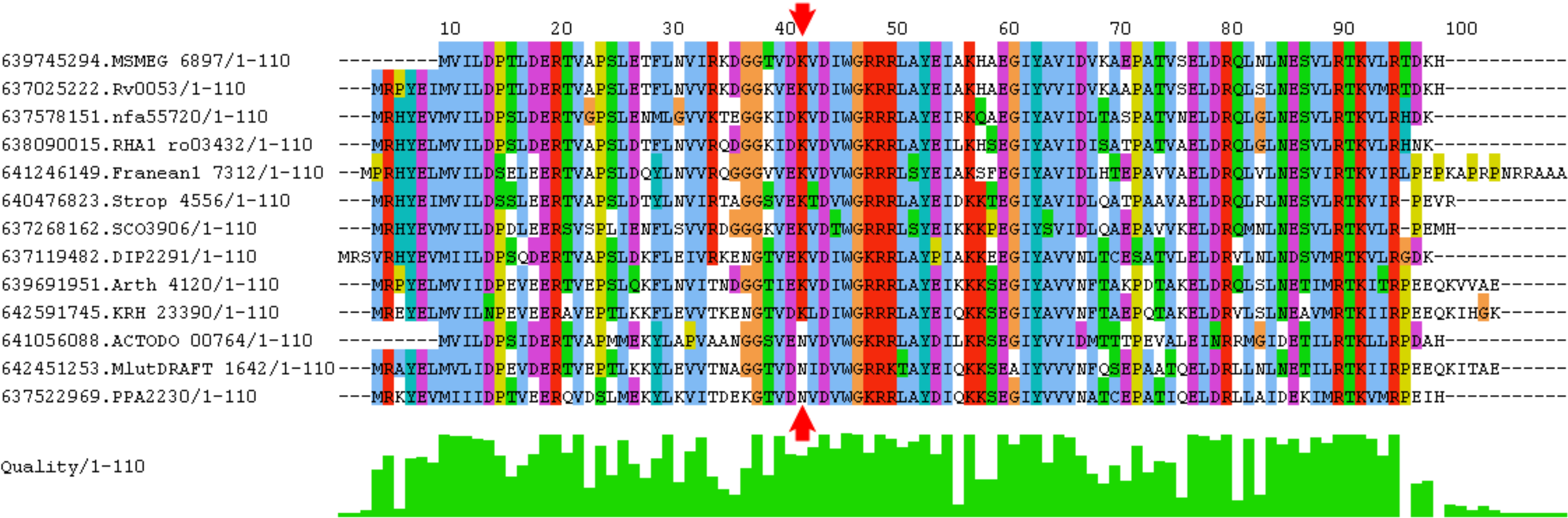


SF54: 30S ribosomal protein S6
(GI: 118171449)

DGGTVDK+243VDIWGR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056088	ACTODO_00764	SSU ribosomal protein S6P (IMGterm)	90aa	Actinomyces odontolyticus ATCC 17982	26
☒	639691951	Arth_4120	SSU ribosomal protein S6P (IMGterm)	101aa	Arthrobacter sp. FB24	26
☒	637119482	DIP2291	SSU ribosomal protein S6P (IMGterm)	98aa	Corynebacterium diphtheriae NCTC 13129	26
☒	641246149	Franean1_7312	SSU ribosomal protein S6P (IMGterm)	108aa	Frankia sp. EAN1pec	26
☒	642591745	KRH_23390	SSU ribosomal protein S6P (IMGterm)	101aa	Kocuria rhizophila DC2201	26
☒	642451253	MlutDRAFT_1642	SSU ribosomal protein S6P (IMGterm)	101aa	Micrococcus luteus NCTC 2665	26
☒	639745294	MSMEG_6897	SSU ribosomal protein S6P (IMGterm)	90aa	Mycobacterium smegmatis MC2 155	26
☒	637025222	Rv0053	SSU ribosomal protein S6P (IMGterm)	96aa	Mycobacterium tuberculosis H37Rv	26
☒	637578151	nfa55720	SSU ribosomal protein S6P (IMGterm)	95aa	Nocardia farcinica IFM 10152	26
☒	637522969	PPA2230	SSU ribosomal protein S6P (IMGterm)	96aa	Propionibacterium acnes KPA171202	26
☒	638090015	RHA1_ro03432	SSU ribosomal protein S6P (IMGterm)	95aa	Rhodococcus sp. RHA1	26
☒	640476823	Strop_4556	SSU ribosomal protein S6P (IMGterm)	96aa	Salinispora tropica CNB-440	26
☒	637268162	SCO3906	SSU ribosomal protein S6P (IMGterm)	96aa	Streptomyces coelicolor A3(2)	26

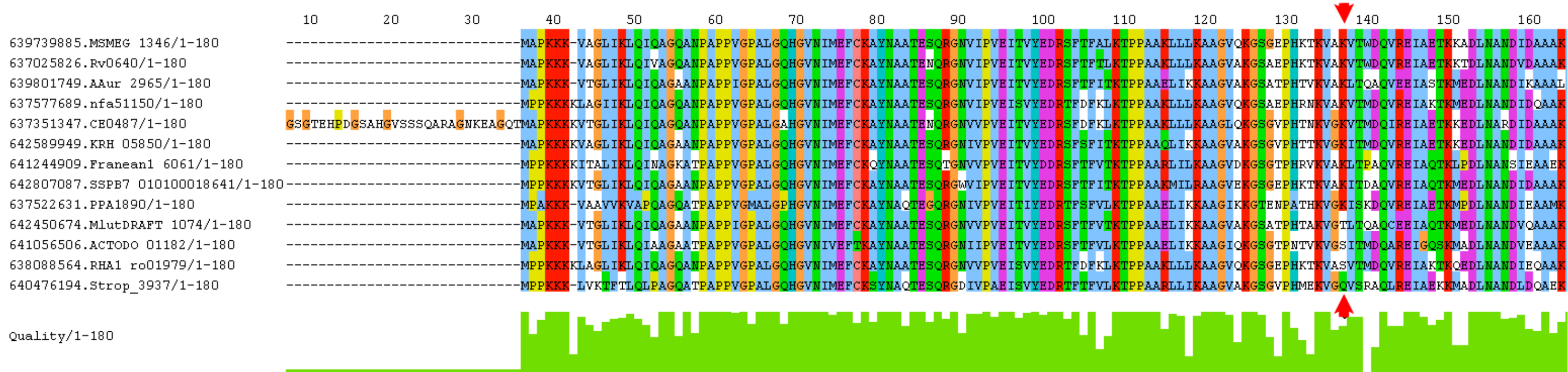


SF55: 50S ribosomal protein L11
(GI: 118473904)

VAK+243VTWDQVR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056506	ACTODO_01182	LSU ribosomal protein L11P (IMGterm)	143aa	Actinomyces odontolyticus ATCC 17982	37
☒	639801749	AAur_2965	LSU ribosomal protein L11P (IMGterm)	143aa	Arthrobacter aurescens TC1	37
☒	637351347	CE0487	LSU ribosomal protein L11P (IMGterm)	180aa	Corynebacterium efficiens YS-314	37
☒	641244909	Franean1_6061	LSU ribosomal protein L11P (IMGterm)	144aa	Frankia sp. EAN1pec	37
☒	642589949	KRH_05850	LSU ribosomal protein L11P (IMGterm)	143aa	Kocuria rhizophila DC2201	37
☒	642450674	MlutDRAFT_1074	LSU ribosomal protein L11P (IMGterm)	143aa	Micrococcus luteus NCTC 2665	37
☒	639739885	MSMEG_1346	LSU ribosomal protein L11P (IMGterm)	142aa	Mycobacterium smegmatis MC2 155	37
☒	637025826	Rv0640	LSU ribosomal protein L11P (IMGterm)	142aa	Mycobacterium tuberculosis H37Rv	37
☒	637577689	nfa51150	LSU ribosomal protein L11P (IMGterm)	144aa	Nocardia farcinica IFM 10152	37
☒	637522631	PPA1890	LSU ribosomal protein L11P (IMGterm)	143aa	Propionibacterium acnes KPA171202	37
☒	638088564	RHA1_ro01979	LSU ribosomal protein L11P (IMGterm)	144aa	Rhodococcus sp. RHA1	37
☒	640476194	Strop_3937	LSU ribosomal protein L11P (IMGterm)	143aa	Salinispora tropica CNB-440	37
☒	642807087	SSPB7_010100018641	LSU ribosomal protein L11P (IMGterm)	144aa	Streptomyces sp. SPB74	37

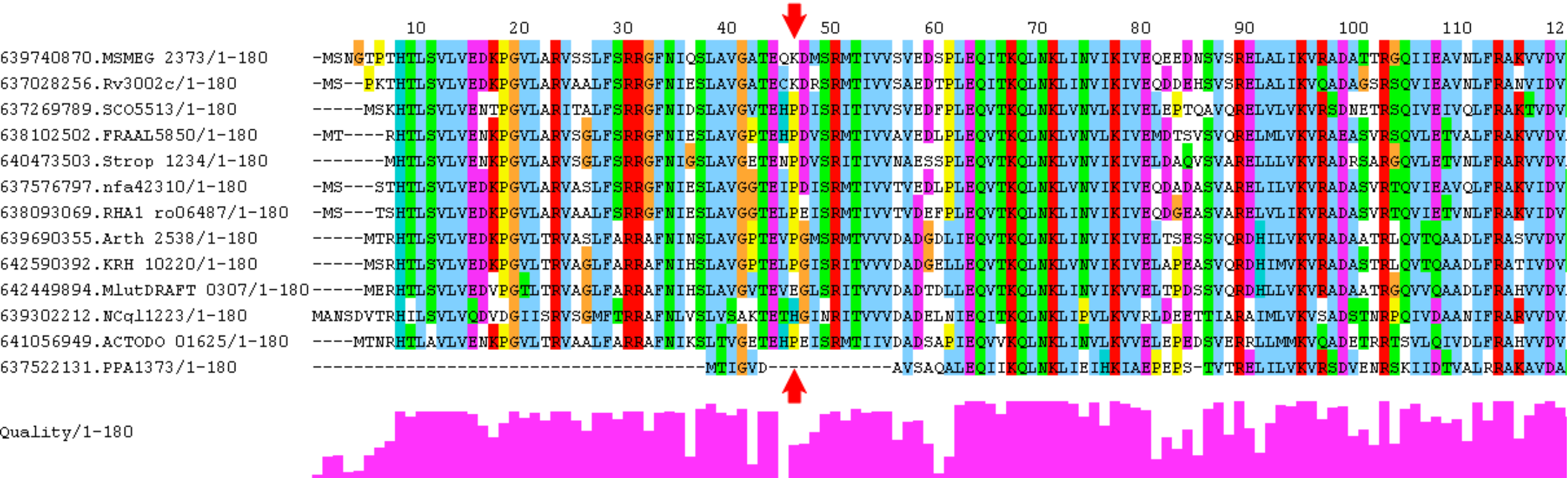


SF56: Acetolactate synthase 3 regulatory subunit
(GI: 118470606)

GFNIQSLAVGATEQK+243DMSR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056949	ACTODO_01625	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	170aa	Actinomyces odontolyticus ATCC 17982	27
☒	639690355	Arth_2538	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	170aa	Arthrobacter sp. FB24	27
☒	639302212	NCgl1223	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	172aa	Corynebacterium glutamicum ATCC 13032 (Kitasato)	27
☒	638102502	FRAAL5850	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	174aa	Frankia alni ACN14a	27
☒	642590392	KRH_10220	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	169aa	Kocuria rhizophila DC2201	27
☒	642449894	MlutDRAFT_0307	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	169aa	Micrococcus luteus NCTC 2665	27
☒	639740870	MSMEG_2373	acetolactate synthase, small subunit (EC:2.2.1.6)	170aa	Mycobacterium smegmatis MC2 155	28
☒	637028256	Rv3002c	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	168aa	Mycobacterium tuberculosis H37Rv	27
☒	637576797	nfa42310	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	167aa	Nocardia farcinica IFM 10152	27
☒	637522131	PPA1373	acetolactate synthase small subunit fragment	81aa	Propionibacterium acnes KPA171202	28
☒	638093069	RHA1_ro06487	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	167aa	Rhodococcus sp. RHA1	27
☒	640473503	Strop_1234	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	169aa	Salinispora tropica CNB-440	27
☒	637269789	SCO5513	acetolactate synthase, small subunit (EC 2.2.1.6) (IMGterm)	174aa	Streptomyces coelicolor A3(2)	27

Red arrows indicate PUPylation site on Msm (MSMEG)

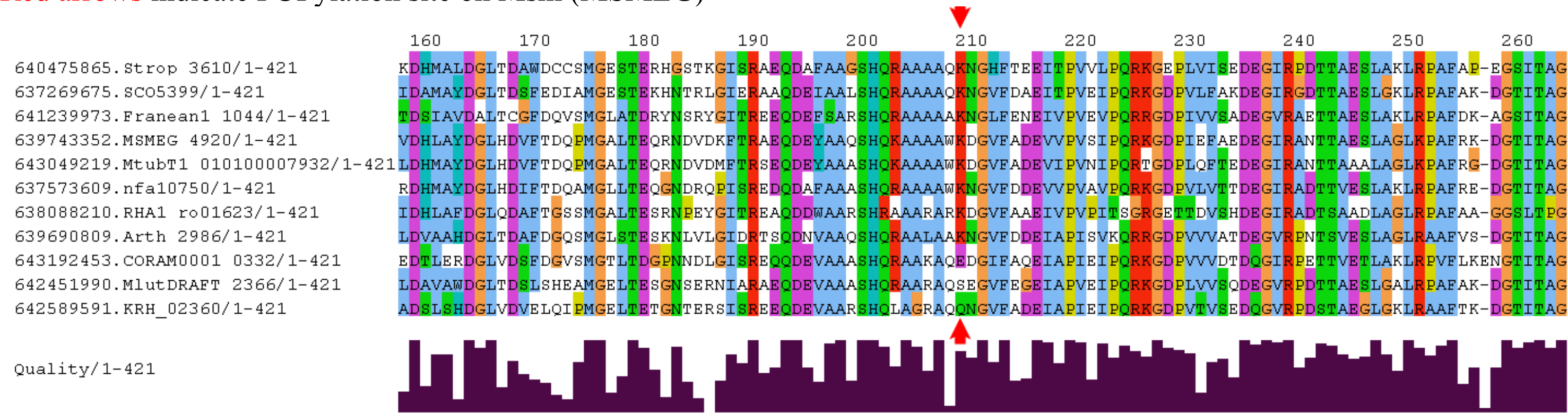


SF57: Acetyl-CoA acetyltransferase
(GI: 118170768)

AAAAWK+243DGVFADEVVPVSIPQR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639690809	Arth_2986	acetyl-CoA acetyltransferases (EC:2.3.1.9)	399aa	Arthrobacter sp. FB24	3
☒	643192453	CORAM0001_0332	acetyl-CoA acetyltransferase (EC:2.3.1.9)	402aa	Corynebacterium amycolatum SK46	3
☒	641239973	Franean1_1044	acetyl-CoA acetyltransferase (EC:2.3.1.9)	394aa	Frankia sp. EAN1pec	3
☒	642589591	KRH_02360	acetyl-CoA acetyltransferase (EC:2.3.1.9)	396aa	Kocuria rhizophila DC2201	3
☒	642451990	MlutDRAFT_2366	acetyl-CoA acetyltransferase (EC:2.3.1.9)	402aa	Micrococcus luteus NCTC 2665	3
☒	639743352	MSMEG_4920	acetyl-CoA acetyltransferase (EC:2.3.1.9)	411aa	Mycobacterium smegmatis MC2 155	3
☒	643049219	MtubT1_010100007932	acetyl-CoA acetyltransferase (EC:2.3.1.9)	389aa	Mycobacterium tuberculosis T17	3
☒	637573609	nfa10750	acetyl-CoA acetyltransferase (EC:2.3.1.9)	392aa	Nocardia farcinica IFM 10152	3
☒	638088210	RHA1_ro01623	acetyl-CoA C-acetyltransferase (EC:2.3.1.9)	397aa	Rhodococcus sp. RHA1	3
☒	640475865	Strop_3610	acetyl-CoA acetyltransferase (EC:2.3.1.9)	399aa	Salinispora tropica CNB-440	3
☒	637269675	SCO5399	acetyl-CoA acetyltransferase (EC:2.3.1.9)	401aa	Streptomyces coelicolor A3(2)	3

Red arrows indicate PUPylation site on Msm (MSMEG)

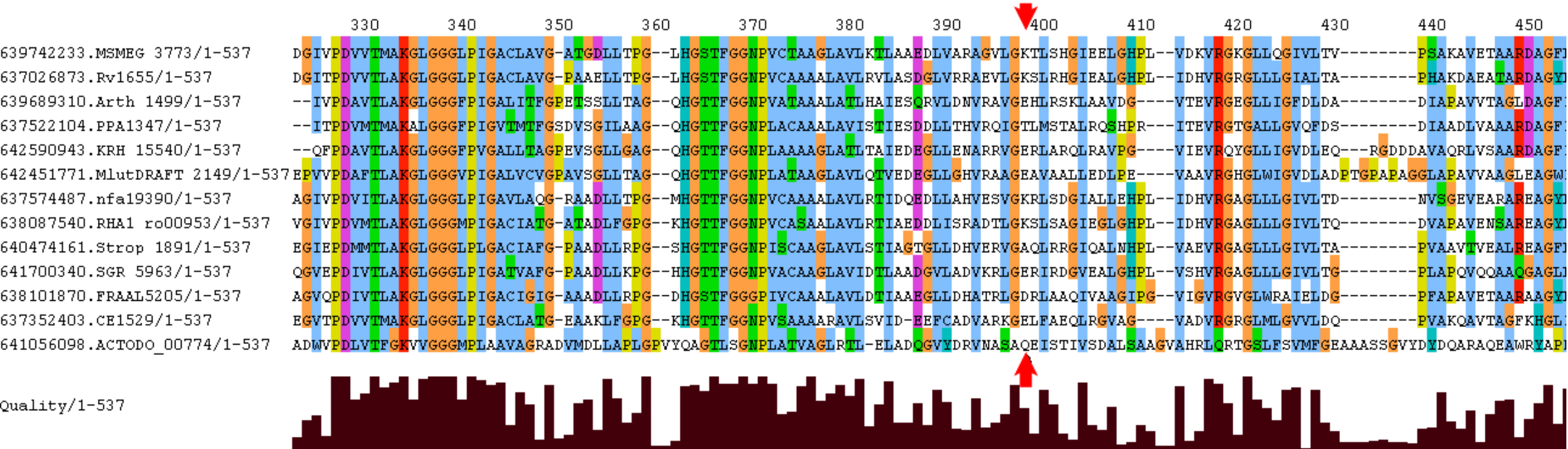


SF58: Acetylornithine aminotransferase
(GI: 118472927)

AGVLGK+243TLSHGIEELGHPLVDK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056098	ACTODO_00774	glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) (IMGterm)	465aa	Actinomyces odontolyticus ATCC 17982	15
☒	639689310	Arth_1499	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	426aa	Arthrobacter sp. FB24	14
☒	637352403	CE1529	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	402aa	Corynebacterium efficiens YS-314	14
☒	638101870	FRAAL5205	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	413aa	Frankia alni ACN14a	14
☒	642590943	KRH_15540	acetylornithine/succinyldiaminopimelate aminotransferase (EC:2.6.1.11, EC:2.6.1.17)	411aa	Kocuria rhizophila DC2201	14
☒	642451771	MlutDRAFT_2149	acetylornithine and succinylornithine aminotransferase (EC:2.6.1.11)	428aa	Micrococcus luteus NCTC 2665	14
☒	639742233	MSMEG_3773	aminotransferase (EC 2.6.1.-) (IMGterm)	390aa	Mycobacterium smegmatis MC2 155	16
☒	637026873	Rv1655	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	400aa	Mycobacterium tuberculosis H37Rv	14
☒	637574487	nfa19390	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	408aa	Nocardia farcinica IFM 10152	14
☒	637522104	PPA1347	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	398aa	Propionibacterium acnes KPA171202	16
☒	638087540	RHA1_ro00953	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	399aa	Rhodococcus sp. RHA1	14
☒	640474161	Strop_1891	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	429aa	Salinispora tropica CNB-440	14
☒	641700340	SGR_5963	acetylornithine aminotransferase (EC 2.6.1.11) (IMGterm)	400aa	Streptomyces griseus griseus NBRC 13350	14

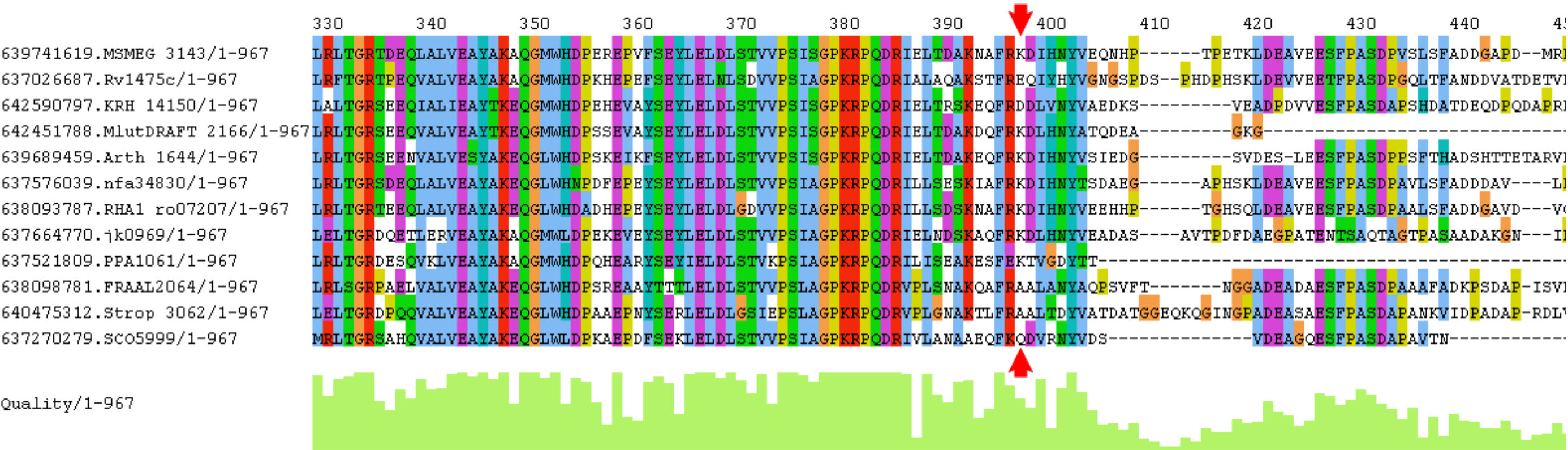


SF59: Aconitate hydratase 1
(GI: 118471676)

K+243DIHNYVEQNHPTPETK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639689459	Arth_1644	aconitase (EC 4.2.1.3) (IMGterm)	936aa	Arthrobacter sp. FB24	39
☒	637664770	jk0969	aconitase (EC 4.2.1.3) (IMGterm)	936aa	Corynebacterium jeikeium K411	39
☒	638098781	FRAAL2064	aconitase (EC 4.2.1.3) (IMGterm)	927aa	Frankia alni ACN14a	39
☒	642590797	KRH_14150	aconitase (EC 4.2.1.3) (IMGterm)	935aa	Kocuria rhizophila DC2201	39
☒	642451788	MlutDRAFT_2166	aconitase (EC 4.2.1.3) (IMGterm)	900aa	Micrococcus luteus NCTC 2665	39
☒	639741619	MSMEG_3143	aconitate hydratase 1 (EC:4.2.1.3)	943aa	Mycobacterium smegmatis MC2 155	39
☒	637026687	Rv1475c	aconitase (EC 4.2.1.3) (IMGterm)	943aa	Mycobacterium tuberculosis H37Rv	39
☒	637576039	nfa34830	aconitase (EC 4.2.1.3) (IMGterm)	933aa	Nocardia farcinica IFM 10152	39
☒	637521809	PPA1061	aconitase (EC 4.2.1.3) (IMGterm)	888aa	Propionibacterium acnes KPA171202	39
☒	638093787	RHA1_ro07207	aconitase (EC 4.2.1.3) (IMGterm)	934aa	Rhodococcus sp. RHA1	39
☒	640475312	Strop_3062	aconitase (EC 4.2.1.3) (IMGterm)	944aa	Salinispora tropica CNB-440	39
☒	637270279	SCO5999	aconitase (EC 4.2.1.3) (IMGterm)	904aa	Streptomyces coelicolor A3(2)	39

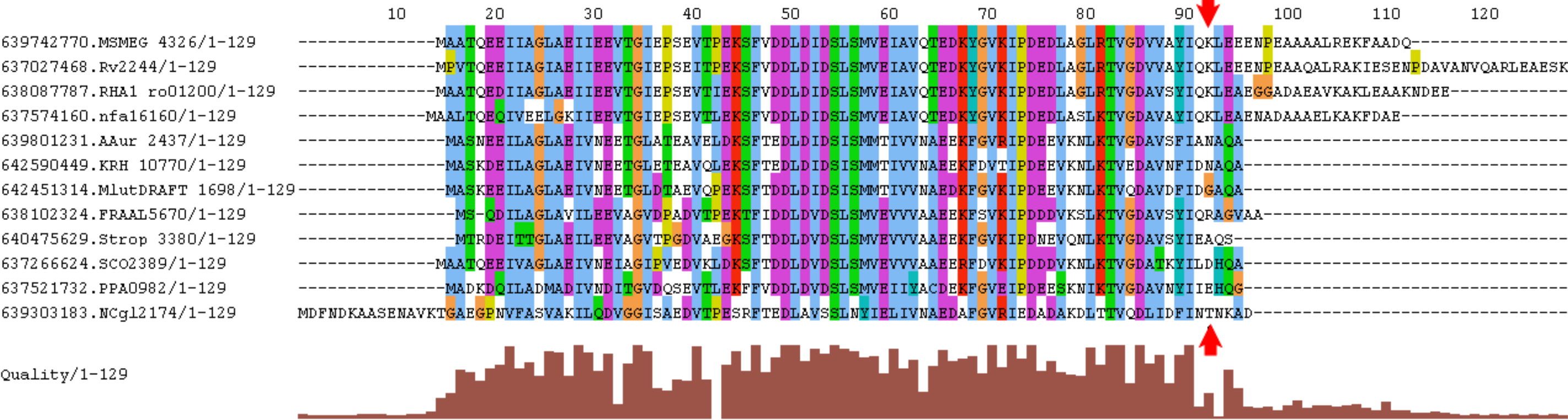


SF60: Acyl carrier protein
(GI: 118471527)

TVGVVAYIQK+243LEEENPEAAAALR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639801231	AAur_2437	putative acyl carrier protein	81aa	Arthrobacter aureus TC1	32
☒	639303183	NCgl2174	acyl carrier protein	97aa	Corynebacterium glutamicum ATCC 13032 (Kitasato)	34
☒	638102324	FRAAL5670	acyl carrier protein (partial match)	81aa	Frankia alni ACN14a	32
☒	642590449	KRH_10770	acyl carrier protein	81aa	Kocuria rhizophila DC2201	32
☒	642451314	MlutDRAFT_1698	phosphopantetheine-binding	81aa	Micrococcus luteus NCTC 2665	32
☒	639742770	MSMEG_4326	acyl carrier protein	99aa	Mycobacterium smegmatis MC2 155	34
☒	637027468	Rv2244	acyl carrier protein	115aa	Mycobacterium tuberculosis H37Rv	32
☒	637574160	nfa16160	acyl carrier protein	99aa	Nocardia farcinica IFM 10152	32
☒	637521732	PPA0982	acyl carrier protein	81aa	Propionibacterium acnes KPA171202	33
☒	638087787	RHA1_ro01200	acyl carrier protein	103aa	Rhodococcus sp. RHA1	32
☒	640475629	Strop_3380	phosphopantetheine-binding	79aa	Salinispora tropica CNB-440	32
☒	637266624	SCO2389	acyl carrier protein	82aa	Streptomyces coelicolor A3(2)	32

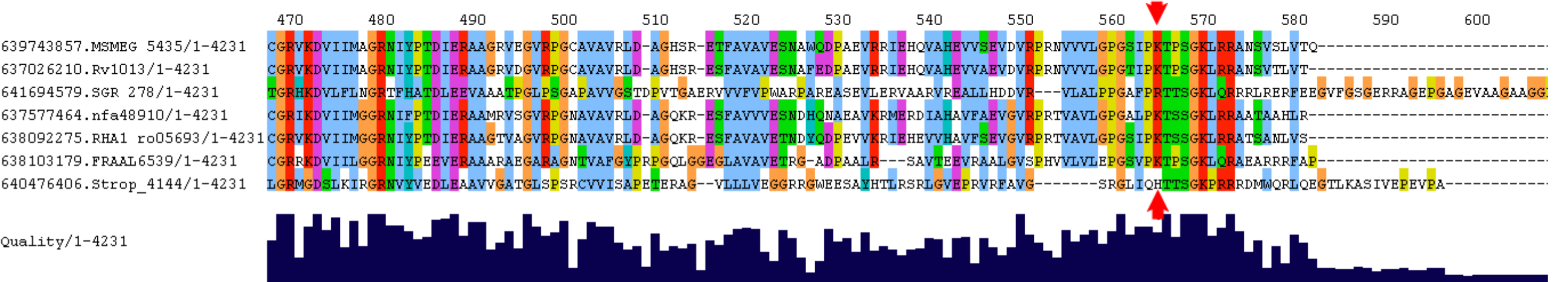


SF61: Acyl-CoA synthetase
(GI: 118470672)

NVVVLGPGSIPK+243TPSGK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	638103179	FRAAL6539	putative fatty-acid--CoA ligase (Acyl-CoA synthetase) (EC:6.2.1.26)	565aa	Frankia alni ACN14a	11
✓	639743857	MSMEG_5435	acyl-CoA synthase	545aa	Mycobacterium smegmatis MC2 155	11
✓	637026210	Rv1013	acyl-CoA synthase	544aa	Mycobacterium tuberculosis H37Rv	11
✓	637577464	nfa48910	acyl-CoA synthase	547aa	Nocardia farcinica IFM 10152	11
✓	638092275	RHA1_ro05693	probable acyl-CoA synthetase (EC:6.2.1.-)	549aa	Rhodococcus sp. RHA1	11
✓	640476406	Strop_4144	AMP-dependent synthetase and ligase	561aa	Salinispora tropica CNB-440	11
✓	641694579	SGR_278	putative NRPS-type-I PKS fusion protein	4185aa	Streptomyces griseus griseus NBRC 13350	11

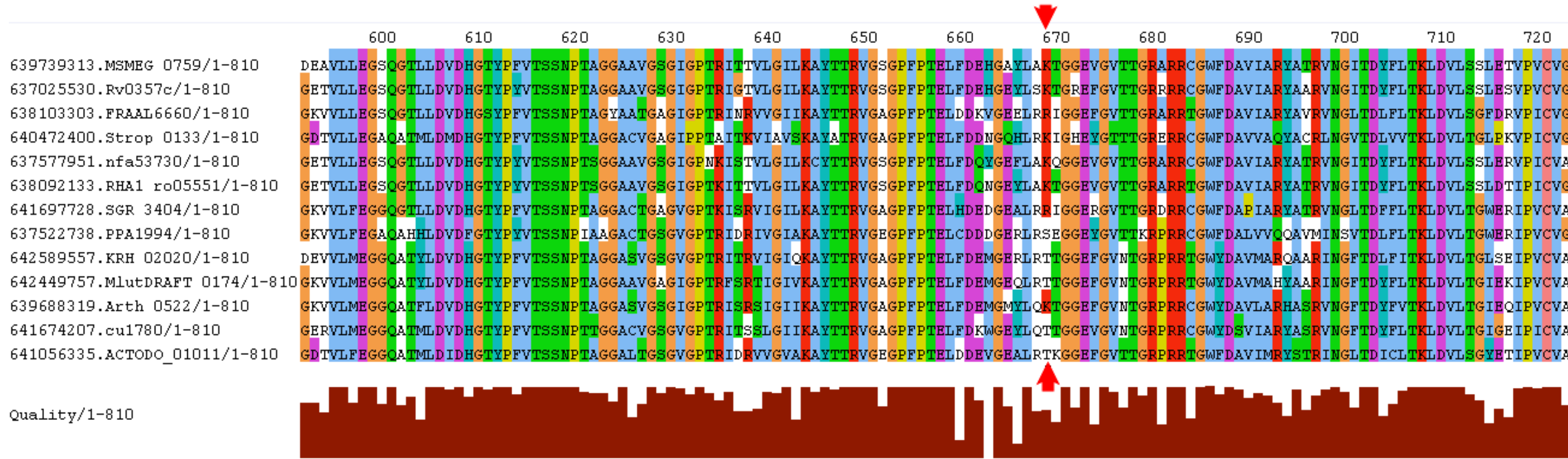


SF62: Adenylosuccinate synthetase
(GI: 166232470)

VGSGPFPTELFDEHGAYLAK+243TGGEVGVTTGR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	641056335	ACTODO_01011	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	426aa	Actinomyces odontolyticus ATCC 17982	12
✓	639688319	Arth_0522	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Arthrobacter sp. FB24	12
✓	641674207	cu1780	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Corynebacterium urealyticum DSM 7109	12
✓	638103303	FRAAL6660	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	427aa	Frankia alni ACN14a	12
✓	642589557	KRH_02020	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Kocuria rhizophila DC2201	12
✓	642449757	MlutDRAFT_0174	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Micrococcus luteus NCTC 2665	12
✓	639739313	MSMEG_0759	adenylosuccinate synthetase (EC:6.3.4.4)	431aa	Mycobacterium smegmatis MC2 155	13
✓	637025530	Rv0357c	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	432aa	Mycobacterium tuberculosis H37Rv	12
✓	637577951	nfa53730	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Nocardia farcinica IFM 10152	12
✓	637522738	PPA1994	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	428aa	Propionibacterium acnes KPA171202	13
✓	638092133	RHA1_ro05551	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	429aa	Rhodococcus sp. RHA1	12
✓	640472400	Strop_0133	Adenylosuccinate synthase (EC:6.3.4.4)	807aa	Salinispora tropica CNB-440	12
✓	641697728	SGR_3404	Adenylosuccinate synthetase (EC 6.3.4.4) (IMGterm)	427aa	Streptomyces griseus griseus NBRC 13350	12

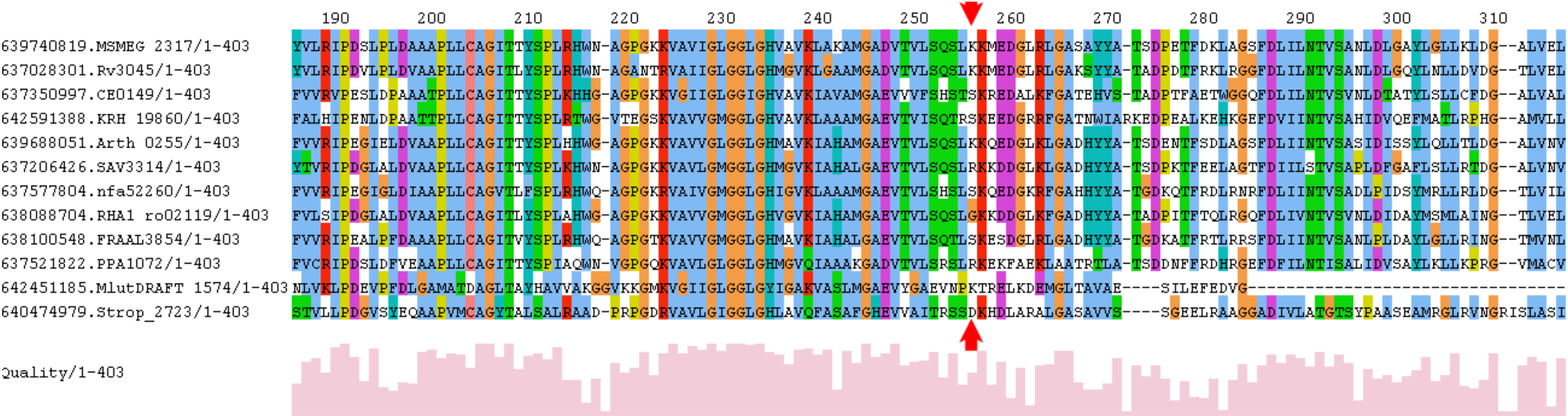


SF63: Alcohol dehydrogenase, zinc-containing (GI: 118468910)

AMGADVTVLSQSLK+243K

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639688051	Arth_0255	Alcohol dehydrogenase GroES domain protein (EC:1.1.1.255)	347aa	Arthrobacter sp. FB24	6
☒	637350997	CE0149	putative alcohol dehydrogenase (EC:1.1.1.255)	387aa	Corynebacterium efficiens YS-314	6
☒	638100548	FRAAL3854	NADP-dependent alcohol dehydrogenase (EC:1.1.1.-, EC:1.1.1.255)	349aa	Frankia alni ACN14a	6
☒	642591388	KRH_19860	NADP-dependent alcohol dehydrogenase (EC:1.1.1.2)	348aa	Kocuria rhizophila DC2201	6
☒	642451185	MlutDRAFT_1574	Alcohol dehydrogenase GroES domain protein	245aa	Micrococcus luteus NCTC 2665	7
☒	639740819	MSMEG_2317	alcohol dehydrogenase, zinc-containing (EC:1.1.1.2, EC:1.1.1.255)	349aa	Mycobacterium smegmatis MC2 155	8
☒	637028301	Rv3045	PROBABLE NADP-DEPENDENT ALCOHOL DEHYDROGENASE ADHC (EC:1.1.1.2, EC:1.1.1.255)	346aa	Mycobacterium tuberculosis H37Rv	6
☒	637577804	nfa52260	putative dehydrogenase	351aa	Nocardia farcinica IFM 10152	6
☒	637521822	PPA1072	Zn-dependent alcohol dehydrogenase (EC:1.1.1.2)	353aa	Propionibacterium acnes KPA171202	6
☒	638088704	RHA1_ro02119	NADP-dependent alcohol dehydrogenase (EC:1.1.1.255)	358aa	Rhodococcus sp. RHA1	6
☒	640474979	Strop_2723	Alcohol dehydrogenase GroES domain protein	338aa	Salinispora tropica CNB-440	8
☒	637206426	SAV3314	NADP-dependent alcohol dehydrogenase (EC:1.1.1.255)	347aa	Streptomyces avermitilis MA-4680	6

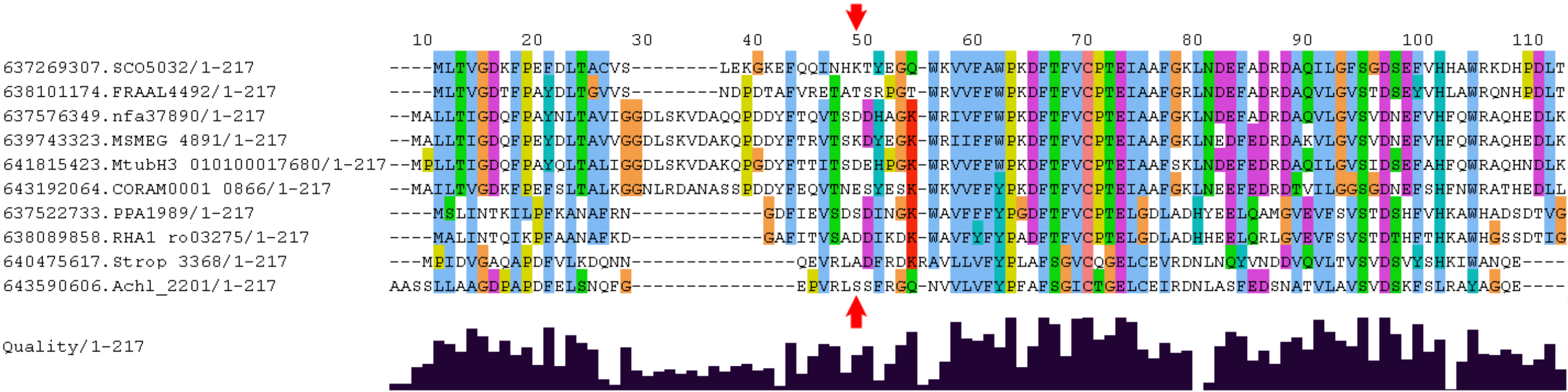


SF64: Alkylhydroperoxide reductase
(GI: 118172773)

VTSK+243DYEGK

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	640475617	Strop_3368	alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen (EC:1.11.1.15)	155aa	Salinispora tropica CNB-440	5
✓	643590606	Achl_2201	alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen (EC:1.11.1.15)	164aa	Arthrobacter chlorophenolicus A6	1
✓	637269307	SCO5032	alkyl hydroperoxide reductase (EC:1.11.1.15)	184aa	Streptomyces coelicolor A3(2)	6
✓	637522733	PPA1989	alkyl hydroperoxide reductase C22 protein (EC:1.11.1.15)	187aa	Propionibacterium acnes KPA171202	5
✓	638089858	RHA1_ro03275	alkyl hydroperoxide reductase (EC:1.11.1.15)	187aa	Rhodococcus sp. RHA1	5
✓	638101174	FRAAL4492	Alkyl hydroperoxide reductase (EC:1.11.1.7)	187aa	Frankia alni ACN14a	3
✓	637576349	nfa37890	putative alkylhydroperoxide reductase (EC:1.11.1.15)	195aa	Nocardia farcinica IFM 10152	4
✓	639743323	MSMEG_4891	alkylhydroperoxide reductase (EC:1.11.1.15)	195aa	Mycobacterium smegmatis MC2 155	6
✓	641815423	MtubH3_010100017680	alkyl hydroperoxide reductase C protein ahpC (EC:1.11.1.15)	195aa	Mycobacterium tuberculosis H37Ra	4
✓	643192064	CORAM0001_0866	alkylhydroperoxide reductase (EC:1.11.1.15)	198aa	Corynebacterium amycolatum SK46	2

Red arrows indicate PUPylation site on Msm (MSMEG)

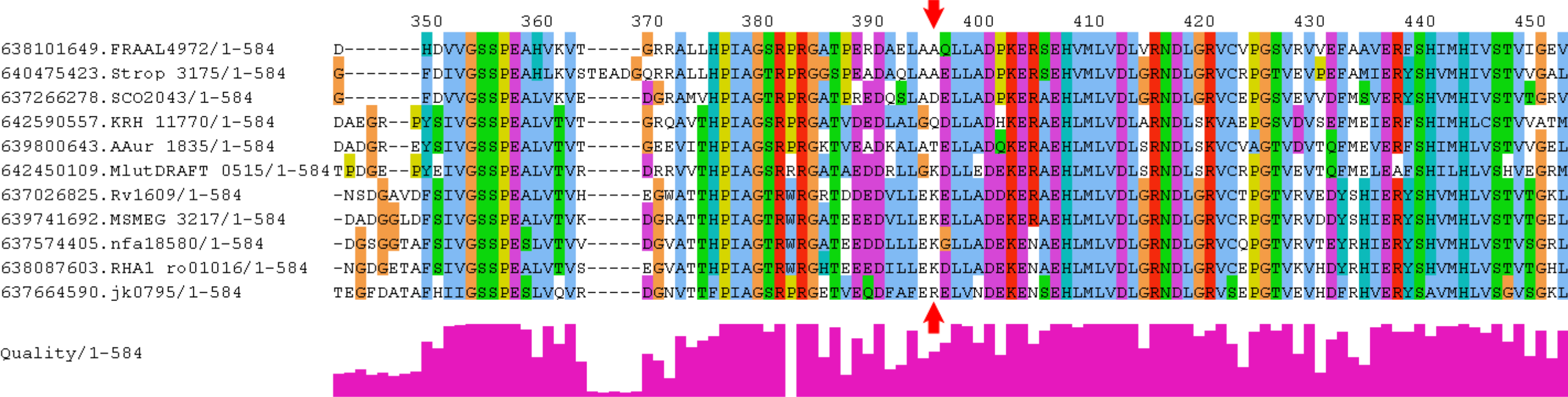


SF65: Anthranilate synthase component I
(GI: 118472702)

GATEEEDVLLEK+243ELLADEKER

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	637266278	SCO2043	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	502aa	Streptomyces coelicolor A3(2)	7
☒	638101649	FRAAL4972	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	512aa	Frankia alni ACN14a	7
☒	637026825	Rv1609	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	516aa	Mycobacterium tuberculosis H37Rv	7
☒	637574405	nfa18580	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	516aa	Nocardia farcinica IFM 10152	7
☒	642450109	MlutDRAFT_0515	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	516aa	Micrococcus luteus NCTC 2665	7
☒	642590557	KRH_11770	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	516aa	Kocuria rhizophila DC2201	7
☒	640475423	Strop_3175	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	517aa	Salinispora tropica CNB-440	7
☒	637664590	jk0795	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	523aa	Corynebacterium jeikeium K411	7
☒	639741692	MSMEG_3217	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	524aa	Mycobacterium smegmatis MC2 155	7
☒	639800643	AAur_1835	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	527aa	Arthrobacter aurescens TC1	7
☒	638087603	RHA1_ro01016	anthranilate synthase, component I (EC 4.1.3.27) (IMGterm)	528aa	Rhodococcus sp. RHA1	7

Red arrows indicate PUPylation site on Msm (MSMEG)

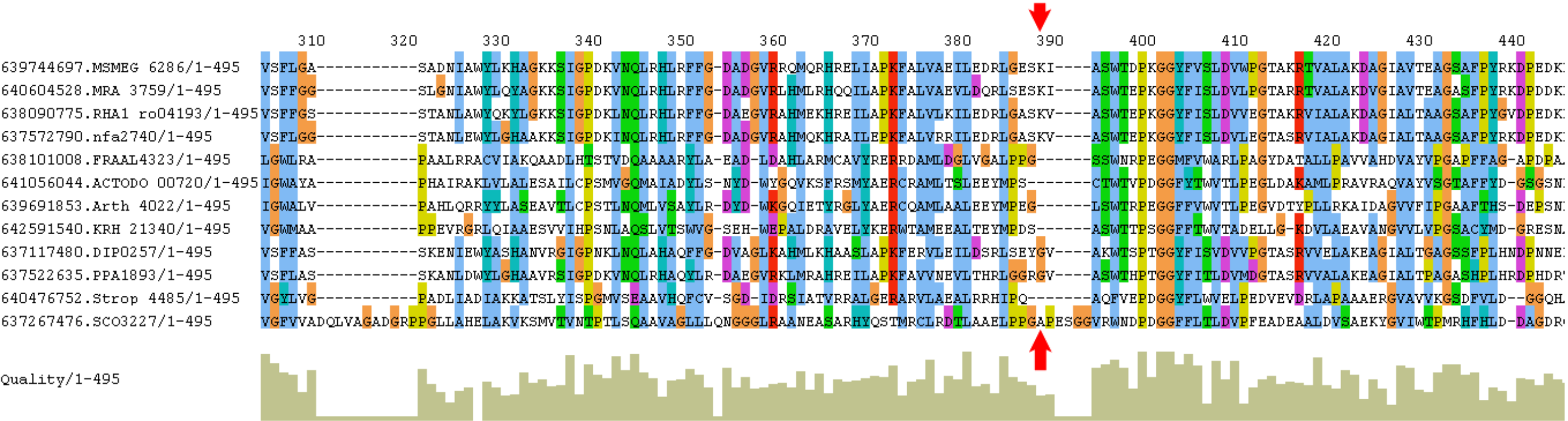


SF66: Aspartate transaminase
(GI: 118468930)

LGESK+243IASWTDPK

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☐	641056044	ACTODO_00720	hypothetical protein	442aa	Actinomyces odontolyticus ATCC 17982	9
☐	639691853	Arth_4022	aminotransferase, class I and II	433aa	Arthrobacter sp. FB24	8
☐	637117480	DIP0257	hypothetical protein	419aa	Corynebacterium diphtheriae NCTC 13129	8
☐	638101008	FRAAL4323	putative aminotransferase (EC:2.6.1.-)	401aa	Frankia alni ACN14a	9
☐	642591540	KRH_21340	aminotransferase (EC:2.6.1.-)	394aa	Kocuria rhizophila DC2201	9
☒	639744697	MSMEG_6286	aspartate transaminase	428aa	Mycobacterium smegmatis MC2 155	10
☐	640604528	MRA_3759	hypothetical protein	435aa	Mycobacterium tuberculosis H37Ra	8
☐	637572790	nfa2740	hypothetical protein	429aa	Nocardia farcinica IFM 10152	8
☐	637522635	PPA1893	aminotransferase (EC:2.6.1.1)	421aa	Propionibacterium acnes KPA171202	8
☐	638090775	RHA1_ro04193	hypothetical protein	432aa	Rhodococcus sp. RHA1	8
☐	640476752	Strop_4485	transcriptional regulator, GntR family (IMGterm)	364aa	Salinispora tropica CNB-440	8
☒	637267476	SCO3227	aminotransferase	444aa	Streptomyces coelicolor A3(2)	10

Red arrows indicate PUPylation site on Msm (MSMEG)

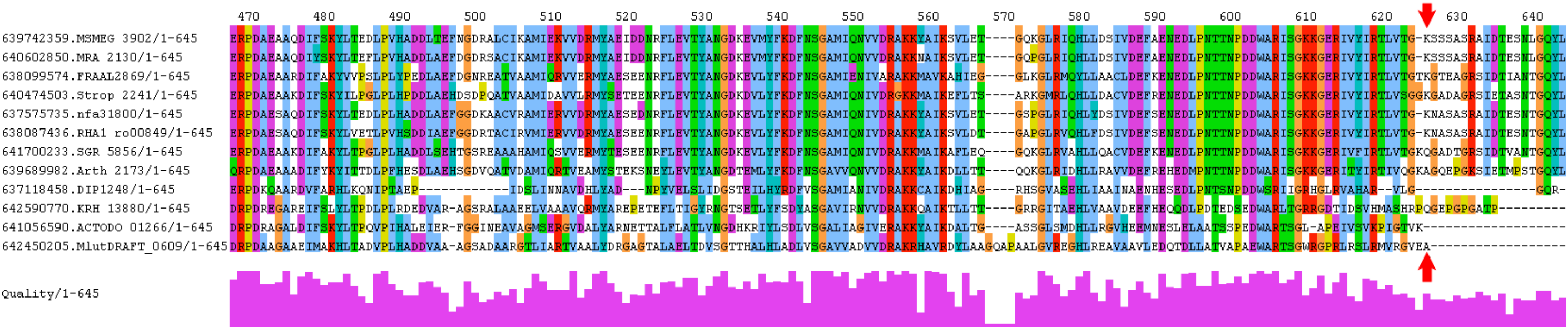


SF67: ATPase, AAA family protein
(GI: 118470801)

TLVTGK+243SSSASR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056590	ACTODO_01266	hypothetical protein	514aa	Actinomyces odontolyticus ATCC 17982	23
☒	639689982	Arth_2173	Vesicle-fusing ATPase (EC:3.6.4.6)	594aa	Arthrobacter sp. FB24	23
☒	637118458	DIP1248	Putative AAA protein family ATPase	509aa	Corynebacterium diphtheriae NCTC 13129	23
☒	638099574	FRAAL2869	H ⁺ -transporting ATP synthase	541aa	Frankia alni ACN14a	23
☒	642590770	KRH_13880	putative ATPase	572aa	Kocuria rhizophila DC2201	23
☒	642450205	MlutDRAFT_0609	AAA ATPase central domain protein	584aa	Micrococcus luteus NCTC 2665	23
☒	639742359	MSMEG_3902	ATPase, AAA family protein	613aa	Mycobacterium smegmatis MC2 155	23
☒	640602850	MRA_2130	AAA family ATPase	609aa	Mycobacterium tuberculosis H37Ra	23
☒	637575735	nfa31800	putative AAA ATPase	586aa	Nocardia farcinica IFM 10152	23
☒	638087436	RHA1_ro00849	ATPase, AAA family (EC:3.6.1.-)	591aa	Rhodococcus sp. RHA1	23
☒	640474503	Strop_2241	AAA ATPase, central domain protein	593aa	Salinispora tropica CNB-440	23
☒	641700233	SGR_5856	putative AAA-family ATPase	588aa	Streptomyces griseus griseus NBRC 13350	23

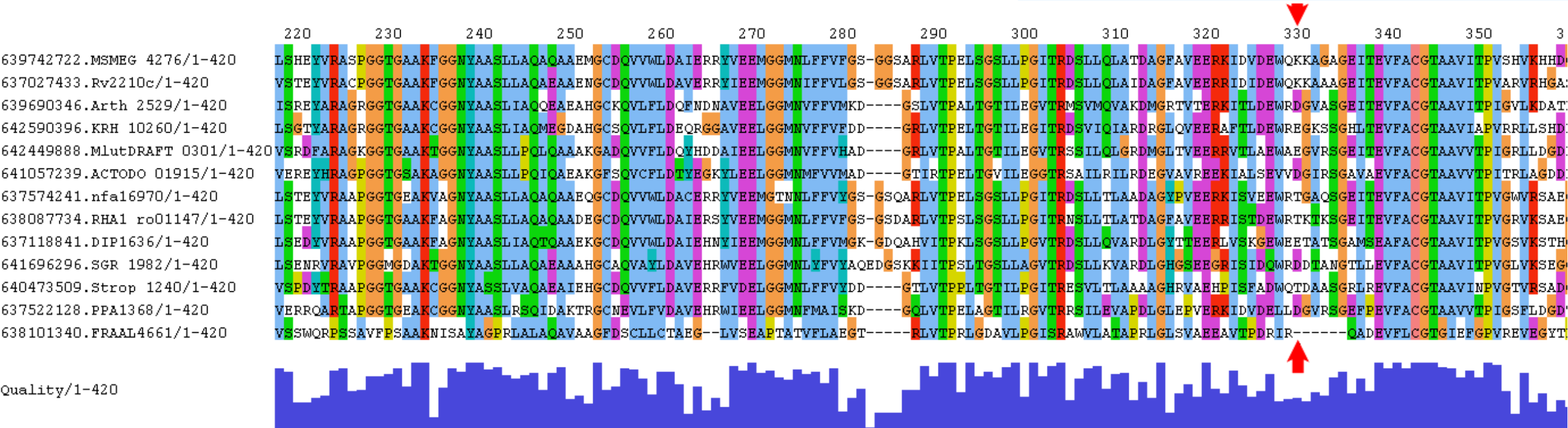
Red arrows indicate PUPylation site on Msm (MSMEG)



SF68: Branched-chain amino acid aminotransferase (GI: 118472496)

KIDVDEWQK+243K

Red arrows indicate PUPylation site on Msm (MSMEG)



Quality/1-420

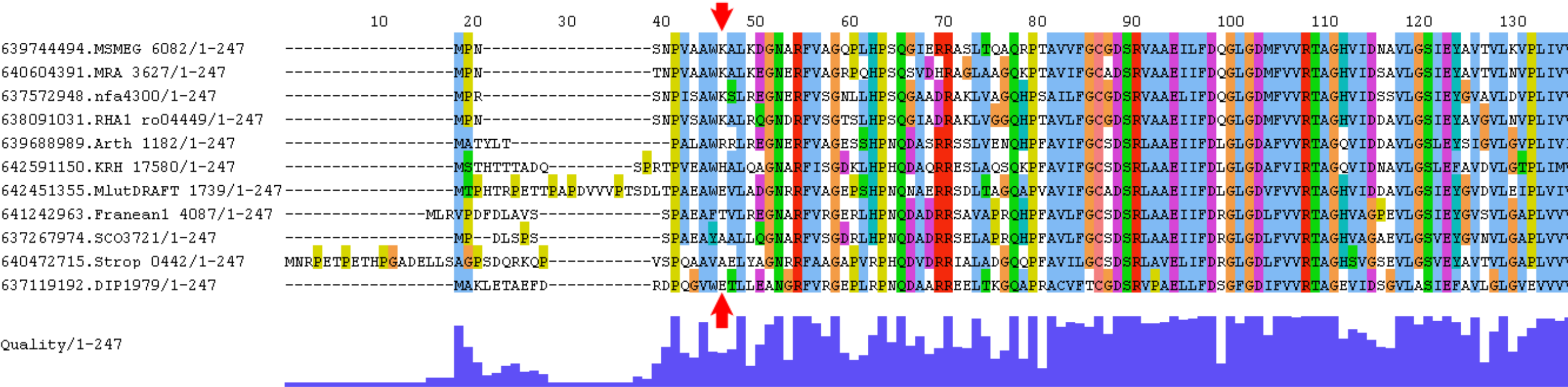
Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641057239	ACTODO_01915	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	382aa	Actinomyces odontolyticus ATCC 17982	10
☒	639690346	Arth_2529	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	370aa	Arthrobacter sp. FB24	10
☒	637118841	DIP1636	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	371aa	Corynebacterium diphtheriae NCTC 13129	10
☒	638101340	FRAAL4661	putative Branched-chain amino acid aminotransferase (EC:2.6.1.42)	342aa	Frankia alni ACN14a	11
☒	642590396	KRH_10260	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	365aa	Kocuria rhizophila DC2201	10
☒	642449888	MlutDRAFT_0301	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	363aa	Micrococcus luteus NCTC 2665	11
☒	639742722	MSMEG_4276	branched-chain amino acid aminotransferase (EC:2.6.1.42)	368aa	Mycobacterium smegmatis MC2 155	11
☒	637027433	Rv2210c	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	368aa	Mycobacterium tuberculosis H37Rv	10
☒	637574241	nfa16970	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	366aa	Nocardia farcinica IFM 10152	10
☒	637522128	PPA1368	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	363aa	Propionibacterium acnes KPA171202	10
☒	638087734	RHA1_ro01147	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	379aa	Rhodococcus sp. RHA1	10
☒	640473509	Strop_1240	branched-chain amino acid aminotransferase (EC:2.6.1.42)	365aa	Salinispora tropica CNB-440	10
☒	641696296	SGR_1982	branched chain amino acid aminotransferase (EC 2.6.1.42) (IMGterm)	367aa	Streptomyces griseus NBRC 13350	10

SF69: Carbonic anhydrase
(GI: 118169435)

*PNSNPVAAWK+243ALK

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	639688989	Arth_1182	carbonic anhydrase	204aa	Arthrobacter sp. FB24	23
✓	637119192	DIP1979	Putative carbonic anhydrase	209aa	Corynebacterium diphtheriae NCTC 13129	23
✓	641242963	Franean1_4087	carbonic anhydrase	216aa	Frankia sp. EAN1pec	24
✓	642591150	KRH_17580	putative carbonic anhydrase (EC:4.2.1.1)	214aa	Kocuria rhizophila DC2201	24
✓	642451355	MlutDRAFT_1739	carbonic anhydrase	223aa	Micrococcus luteus NCTC 2665	24
✓	639744494	MSMEG_6082	carbonic anhydrase	206aa	Mycobacterium smegmatis MC2 155	25
✓	640604391	MRA_3627	carbonic anhydrase	207aa	Mycobacterium tuberculosis H37Ra	24
✓	637572948	nfa4300	hypothetical protein	207aa	Nocardia farcinica IFM 10152	24
✓	638091031	RHA1_ro04449	carbonate dehydratase (EC:4.2.1.1)	208aa	Rhodococcus sp. RHA1	24
✓	640472715	Strop_0442	carbonic anhydrase	236aa	Salinispora tropica CNB-440	24
✓	637267974	SCO3721	carbonic anhydrase	207aa	Streptomyces coelicolor A3(2)	25

Red arrows indicate PUPylation site on Msm (MSMEG)

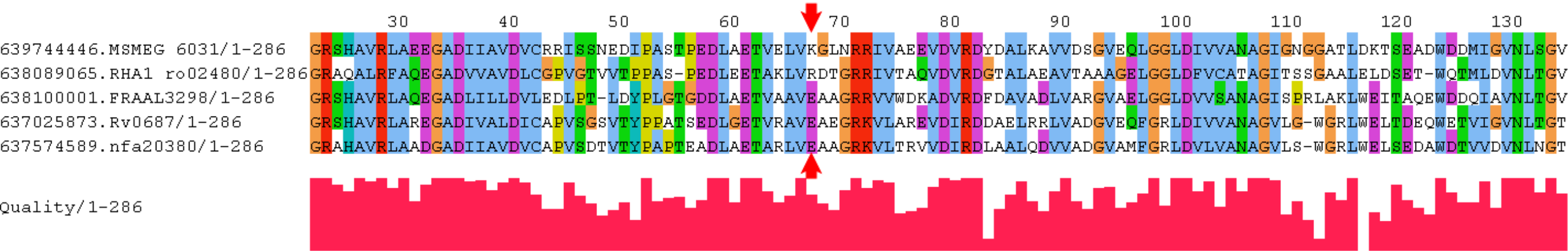


SF70: Carveol dehydrogenase
(GI: 118169959)

ISSNEDIPASTPEDLAETVELVK+243GLNR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	638100001	FRAAL3298	putative carveol dehydrogenase ((+)-trans-carveol dehydrogenase) (CDH)	279aa	Frankia alni ACN14a	9
☒	639744446	MSMEG_6031	carveol dehydrogenase	279aa	Mycobacterium smegmatis MC2 155	10
☒	637025873	Rv0687	short chain dehydrogenase	275aa	Mycobacterium tuberculosis H37Rv	9
☒	637574589	nfa20380	short chain dehydrogenase	280aa	Nocardia farcinica IFM 10152	9
☒	638089065	RHA1_ro02480	carveol dehydrogenase	276aa	Rhodococcus sp. RHA1	9

Red arrows indicate PUPylation site on Msm (MSMEG)



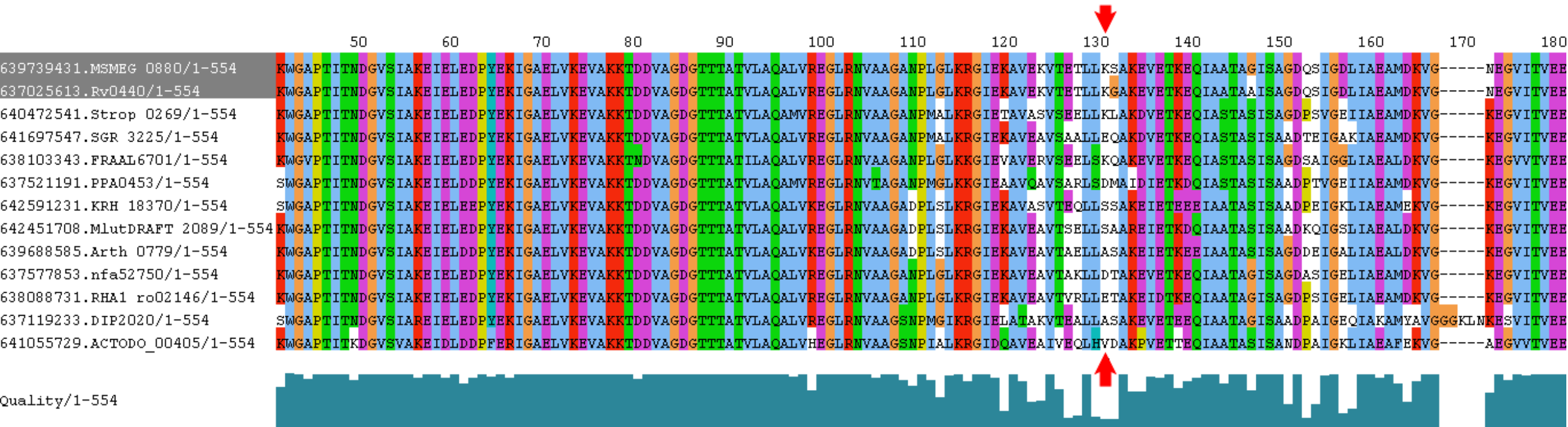
SF71: Chaperonin GroEL

(GI: 118469510)

VTETLLK+243SAK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641055729	ACTODO_00405	hypothetical protein	542aa	Actinomyces odontolyticus ATCC 17982	7
☒	639688585	Arth_0779	chaperonin GroEL	545aa	Arthrobacter sp. FB24	7
☒	637119233	DIP2020	chaperonin GroEL	546aa	Corynebacterium diphtheriae NCTC 13129	7
☒	638103343	FRAAL6701	chaperone Hsp60 (GroEL), part of GroE chaperone system	540aa	Frankia alni ACN14a	7
☒	642591231	KRH_18370	60 kDa chaperonin	544aa	Kocuria rhizophila DC2201	7
☒	642451708	MlutDRAFT_2089	chaperonin GroEL	546aa	Micrococcus luteus NCTC 2665	7
☒	639739431	MSMEG_0880	chaperonin GroL	541aa	Mycobacterium smegmatis MC2 155	7
☒	637025613	Rv0440	chaperonin GroEL	540aa	Mycobacterium tuberculosis H37Rv	7
☒	637577853	nfa52750	chaperonin GroEL	541aa	Nocardia farcinica IFM 10152	7
☒	637521191	PPA0453	chaperonin GroEL	544aa	Propionibacterium acnes KPA171202	7
☒	638088731	RHA1_ro02146	60 kDa chaperonin GroEL	541aa	Rhodococcus sp. RHA1	7
☒	640472541	Strop_0269	chaperonin GroEL	540aa	Salinispora tropica CNB-440	7
☒	641697547	SGR_3225	putative GroEL2	540aa	Streptomyces griseus griseus NBRC 13350	7

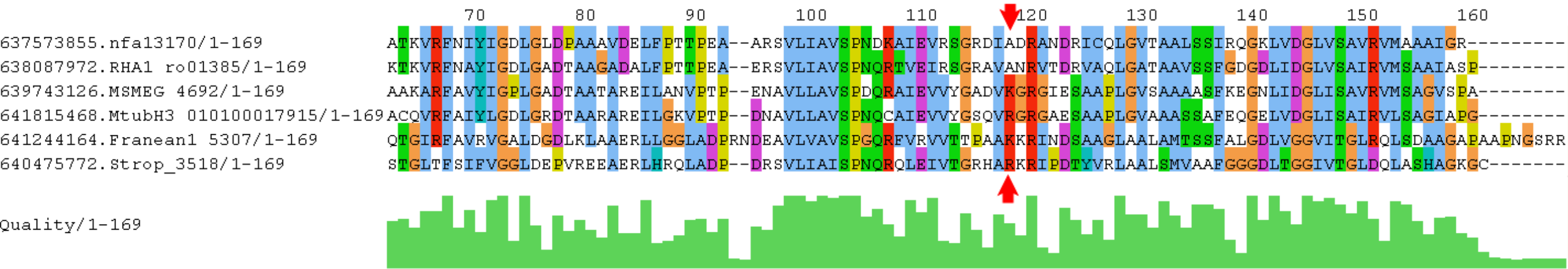


SF72: Conserved hypothetical protein
(GI: 118172453)

AIEVVYGADV K+243GR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641244164	Franean1_5307	hypothetical protein	133aa	Frankia sp. EAN1pec	31
☒	639743126	MSMEG_4692	hypothetical protein	157aa	Mycobacterium smegmatis MC2 155	31
☒	637027705	Rv2468c	hypothetical protein	167aa	Mycobacterium tuberculosis H37Rv	31
☒	637573855	nfa13170	hypothetical protein	154aa	Nocardia farcinica IFM 10152	31
☒	638087972	RHA1_ro01385	hypothetical protein	159aa	Rhodococcus sp. RHA1	31
☒	640475772	Strop_3518	hypothetical protein	132aa	Salinispora tropica CNB-440	31

Red arrows indicate PUPylation site on Msm (MSMEG)



SF73: Conserved hypothetical protein (GI: 118171476)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641579022	MAB_1298c	hypothetical protein	152aa	Mycobacterium abscessus	13
☒	639744928	MSMEG_6518	hypothetical protein	129aa	Mycobacterium smegmatis MC2 155	13
☒	641813108	MtubH3_010100005267	hypothetical alanine rich protein	392aa	Mycobacterium tuberculosis H37Ra	13

KDLIEAVEAEEPATEAETAQIK+243

Red arrows indicate PUPylation site on Msm (MSMEG)

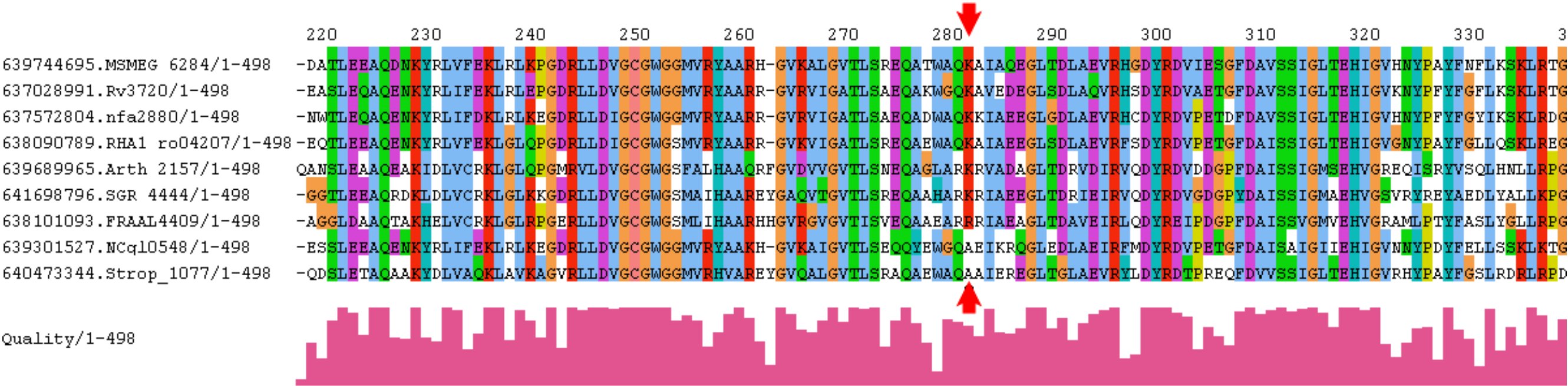


SF74: Cyclopropane-fatty-acyl-phospholipid synthase
(GI: 118173720)

EQATWAQK+243AIAQEGLTDLAIEVR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☐	639689965	Arth_2157	Cyclopropane-fatty-acyl-phospholipid synthase	434aa	Arthrobacter sp. FB24	1
☐	639301527	NCgl0548	putative cyclopropane-fatty-acyl-phospholipid synthase (EC:2.1.1.79)	435aa	Corynebacterium glutamicum ATCC 13032 (Kitsato)	1
☐	638101093	FRAAL4409	Cyclopropane-fatty-acyl-phospholipid synthase (EC:2.1.1.79)	464aa	Frankia alni ACN14a	1
☒	639744695	MSMEG_6284	cyclopropane-fatty-acyl-phospholipid synthase	437aa	Mycobacterium smegmatis MC2 155	2
☐	637028991	Rv3720	POSSIBLE FATTY ACID SYNTHASE (EC:2.1.1.-)	420aa	Mycobacterium tuberculosis H37Rv	1
☒	637572804	nfa2880	putative cyclopropane fatty acid synthase	438aa	Nocardia farcinica IFM 10152	2
☐	638090789	RHA1_ro04207	cyclopropane-fatty-acyl-phospholipid synthase (EC:2.1.1.79)	433aa	Rhodococcus sp. RHA1	1
☐	640473344	Strop_1077	Cyclopropane-fatty-acyl-phospholipid synthase (EC:2.1.1.79)	447aa	Salinispora tropica CNB-440	1
☐	641698796	SGR_4444	putative cyclopropane fatty acid synthase	434aa	Streptomyces griseus griseus NBRC 13350	1

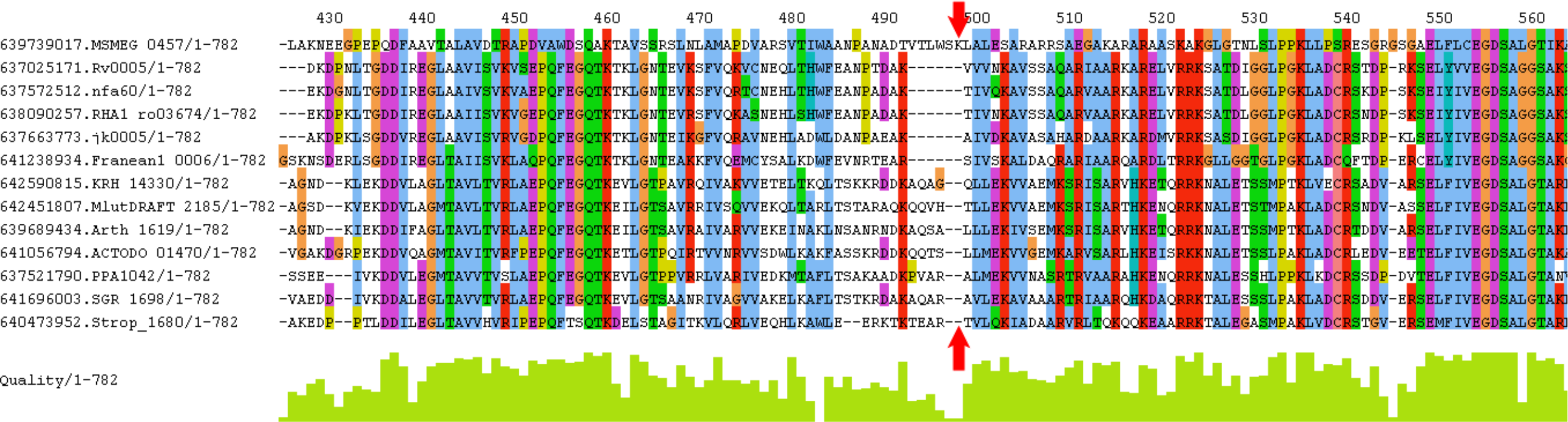


SF75: DNA topoisomerase IV subunit B
(GI: 118468752)

SVTIWAANPANADTVTLWSK+243TGGEVGVTTGR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	641056794	ACTODO_01470	hypothetical protein	706aa	Actinomyces odontolyticus ATCC 17982	3
✓	639689434	Arth_1619	DNA topoisomerase IV subunit B (EC 5.99.1.3) (IMGterm)	702aa	Arthrobacter sp. FB24	3
✓	637663773	jk0005	DNA gyrase subunit B (EC 5.99.1.3) (IMGterm)	674aa	Corynebacterium jeikeium K411	3
✓	641238934	Franean1_0006	DNA gyrase subunit B (EC 5.99.1.3) (IMGterm)	652aa	Frankia sp. EAN1pec	3
✓	642590815	KRH_14330	DNA topoisomerase IV subunit B (EC 5.99.1.3) (IMGterm)	702aa	Kocuria rhizophila DC2201	3
✓	642451807	MlutDRAFT_2185	DNA topoisomerase IV subunit B (EC 5.99.1.3) (IMGterm)	701aa	Micrococcus luteus NCTC 2665	3
✓	639739017	MSMEG_0457	DNA topoisomerase IV subunit B	677aa	Mycobacterium smegmatis MC2 155	5
✓	637025171	Rv0005	DNA gyrase subunit B (EC 5.99.1.3) (IMGterm)	714aa	Mycobacterium tuberculosis H37Rv	3
✓	637572512	nfa60	DNA gyrase subunit B (EC 5.99.1.3) (IMGterm)	696aa	Nocardia farcinica IFM 10152	4
✓	637521790	PPA1042	DNA gyrase subunit B (EC:5.99.1.3)	698aa	Propionibacterium acnes KPA171202	3
✓	638090257	RHA1_ro03674	DNA gyrase subunit B (EC 5.99.1.3) (IMGterm)	696aa	Rhodococcus sp. RHA1	5
✓	640473952	Strop_1680	DNA topoisomerase, type IIA, subunit B, region 2 domain protein	702aa	Salinispora tropica CNB-440	3
✓	641696003	SGR_1698	DNA topoisomerase IV subunit B (EC 5.99.1.3) (IMGterm)	708aa	Streptomyces griseus griseus NBRC 13350	3

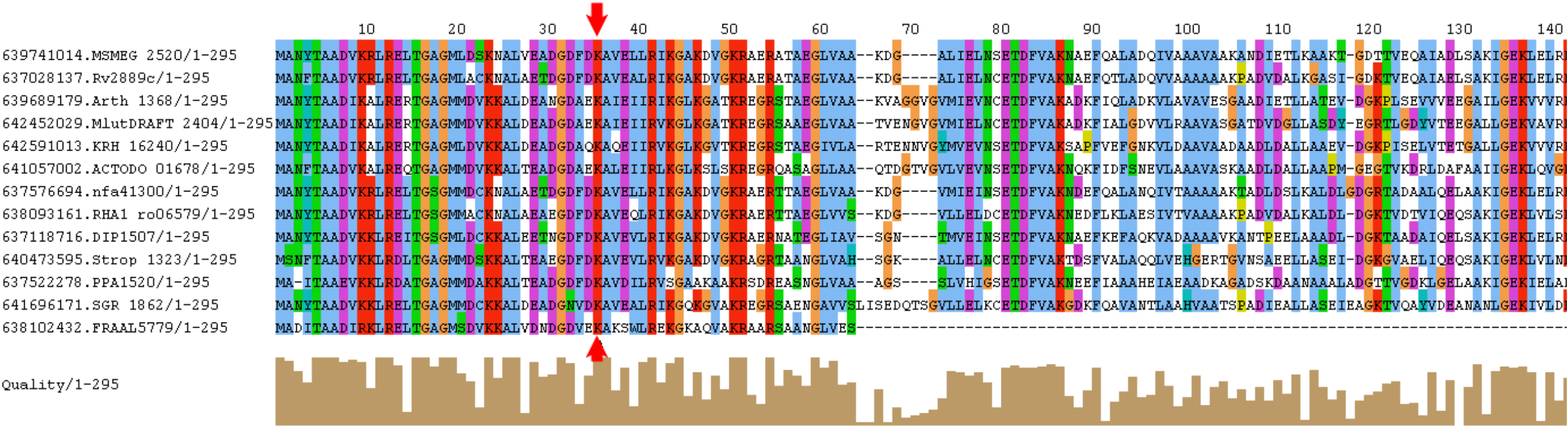


SF76: Elongation factor Ts
(GI: 166221229)

NALVEADGDFDK+243AVELLR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641057002	ACTODO_01678	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Actinomyces odontolyticus ATCC 17982	17
☒	639689179	Arth_1368	translation elongation factor Ts (EF-Ts) (IMGterm)	278aa	Arthrobacter sp. FB24	17
☒	637118716	DIP1507	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Corynebacterium diphtheriae NCTC 13129	17
☒	638102432	FRAAL5779	translation elongation factor Ts (EF-Ts) (IMGterm)	204aa	Frankia alni ACN14a	18
☒	642591013	KRH_16240	translation elongation factor Ts (EF-Ts) (IMGterm)	276aa	Kocuria rhizophila DC2201	19
☒	642452029	MlutDRAFT_2404	translation elongation factor Ts (EF-Ts) (IMGterm)	278aa	Micrococcus luteus NCTC 2665	17
☒	639741014	MSMEG_2520	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Mycobacterium smegmatis MC2 155	19
☒	637028137	Rv2889c	translation elongation factor Ts (EF-Ts) (IMGterm)	271aa	Mycobacterium tuberculosis H37Rv	17
☒	637576694	nfa41300	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Nocardia farcinica IFM 10152	17
☒	637522278	PPA1520	translation elongation factor Ts (EF-Ts) (IMGterm)	270aa	Propionibacterium acnes KPA171202	17
☒	638093161	RHA1_ro06579	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Rhodococcus sp. RHA1	17
☒	640473595	Strop_1323	translation elongation factor Ts (EF-Ts) (IMGterm)	275aa	Salinispora tropica CNB-440	17
☒	641696171	SGR_1862	translation elongation factor Ts (EF-Ts) (IMGterm)	278aa	Streptomyces griseus griseus NBRC 13350	17

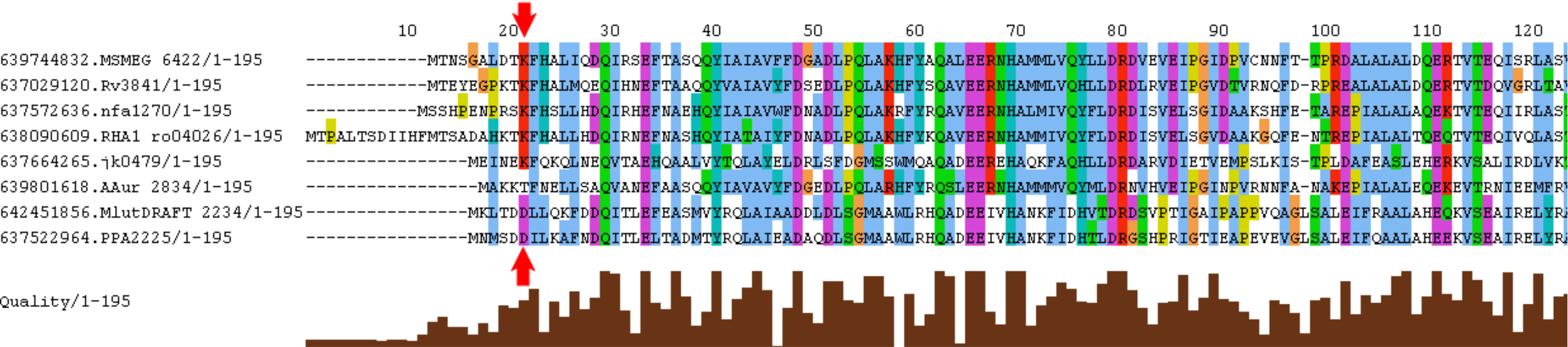


SF77: Ferritin family protein
(GI: 118468959)

TNSGALDTK+243FHALIQDQIR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639801618	AAur_2834	putative ferritin family protein	176aa	Arthrobacter aurescens TC1	9
☒	637664265	jk0479	putative ferritin	162aa	Corynebacterium jeikeium K411	9
☒	642451856	MlutDRAFT_2234	Ferritin Dps family protein	172aa	Micrococcus luteus NCTC 2665	9
☒	639744832	MSMEG_6422	ferritin family protein	181aa	Mycobacterium smegmatis MC2 155	9
☒	637029120	Rv3841	POSSIBLE BACTERIOFERRITIN BFRB	181aa	Mycobacterium tuberculosis H37Rv	9
☒	637572636	nfa1270	putative ferritin family protein	182aa	Nocardia farcinica IFM 10152	9
☒	637522964	PPA2225	ferritin	172aa	Propionibacterium acnes KPA171202	9
☒	638090609	RHA1_ro04026	probable ferritin	193aa	Rhodococcus sp. RHA1	9

Red arrows indicate PUPylation site on Msm (MSMEG)

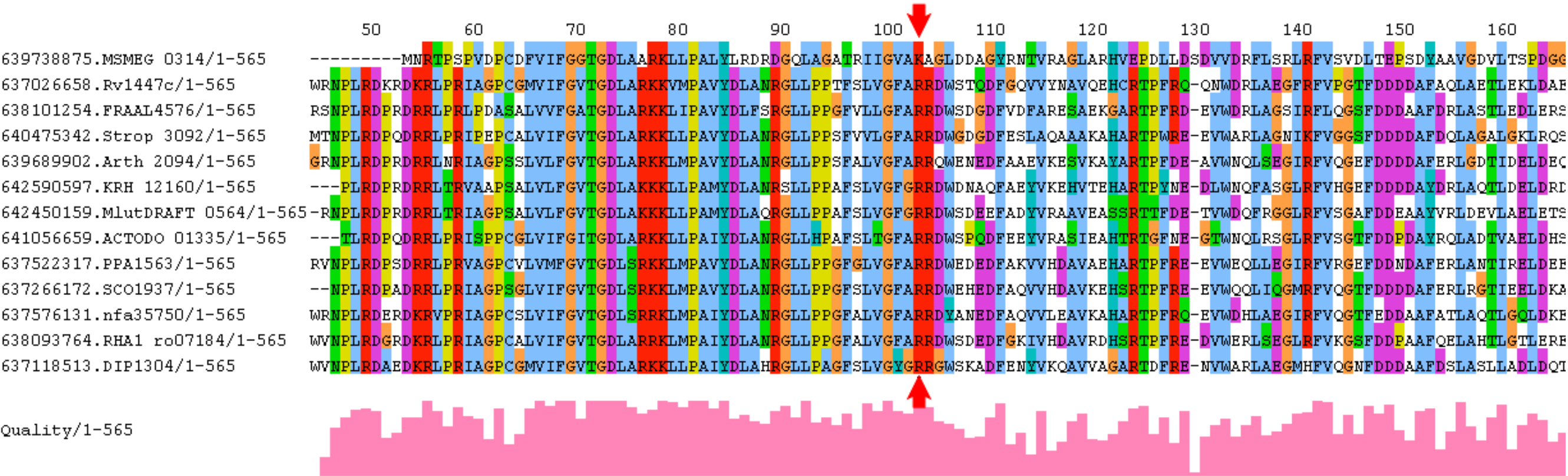


SF78: Glucose-6-phosphate-1-dehydrogenase
(GI: 118171568)

IIGVAK+243AGLDDAGYR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	641056659	ACTODO_01335	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	508aa	Actinomyces odontolyticus ATCC 17982	21
✓	639689902	Arth_2094	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	520aa	Arthrobacter sp. FB24	21
✓	637118513	DIP1304	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	549aa	Corynebacterium diphtheriae NCTC 13129	21
✓	638101254	FRAAL4576	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	510aa	Frankia alni ACN14a	21
✓	642590597	KRH_12160	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	510aa	Kocuria rhizophila DC2201	21
✓	642450159	MlutDRAFT_0564	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	520aa	Micrococcus luteus NCTC 2665	22
✓	639738875	MSMEG_0314	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	495aa	Mycobacterium smegmatis MC2 155	22
✓	637026658	Rv1447c	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	514aa	Mycobacterium tuberculosis H37Rv	22
✓	637576131	nfa35750	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	522aa	Nocardia farcinica IFM 10152	22
✓	637522317	PPA1563	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	524aa	Propionibacterium acnes KPA171202	22
✓	638093764	RHA1_ro07184	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	514aa	Rhodococcus sp. RHA1	22
✓	640475342	Strop_3092	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	505aa	Salinispora tropica CNB-440	22
✓	637266172	SCO1937	glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (IMGterm)	507aa	Streptomyces coelicolor A3(2)	22

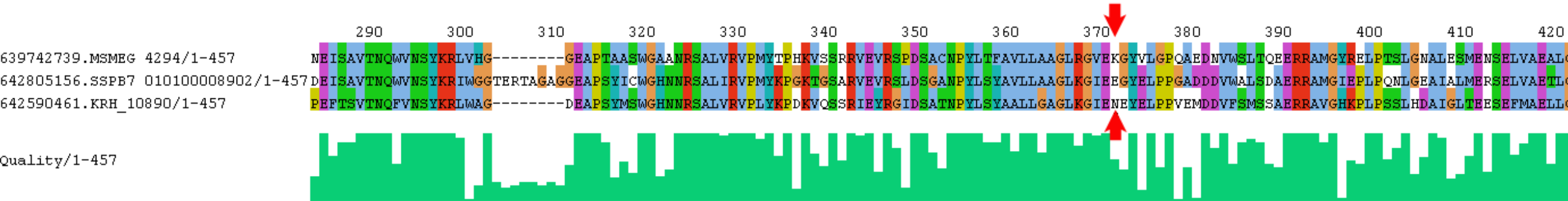


SF79: Glutamine synthetase, type I
(GI: 118474021)

GVEK+243GYVLGPQAEDNVWSLTQEER

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	642590461	KRH_10890	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	447aa	Kocuria rhizophila DC2201	13
☒	639742739	MSMEG_4294	glutamine synthetase, type I (EC:6.3.1.2)	446aa	Mycobacterium smegmatis MC2 155	13
☒	642805156	SSPB7_010100008902	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	453aa	Streptomyces sp. SPB74	13

Red arrows indicate PUPylation site on Msm (MSMEG)

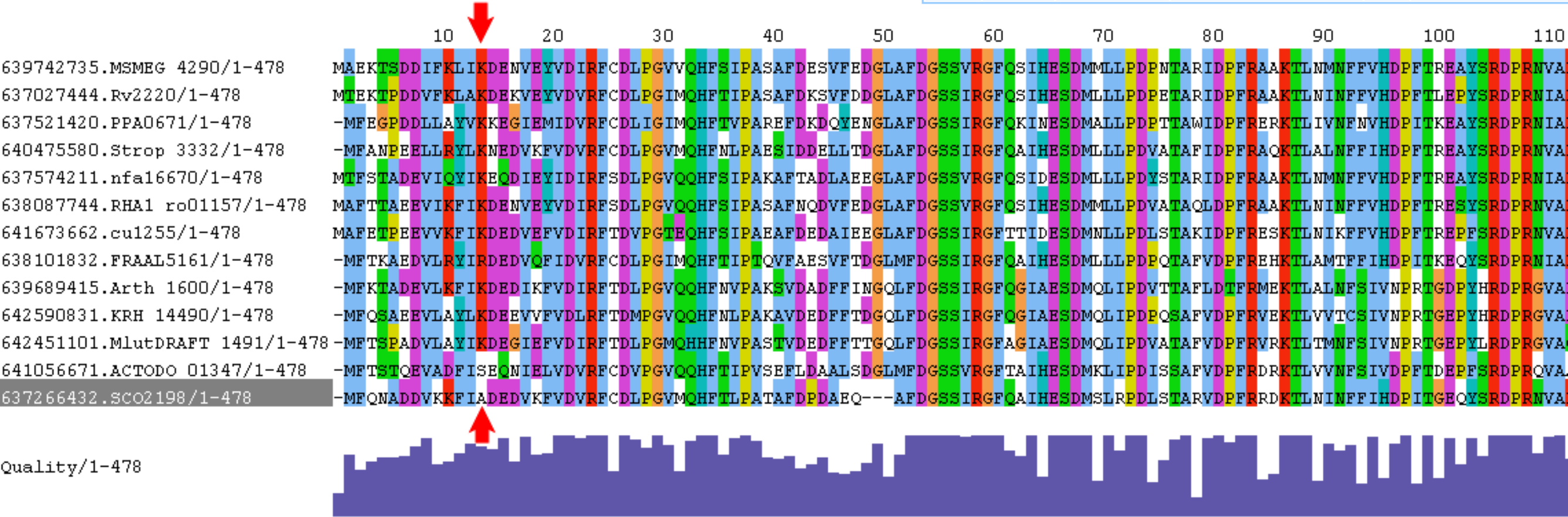


SF80: Glutamine synthetase, type I
(GI: 118169243)

LIK+243DENVEYVDIR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☐	641056671	ACTODO_01347	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Actinomyces odontolyticus ATCC 17982	14
☐	639689415	Arth_1600	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Arthrobacter sp. FB24	15
☐	641673662	cu1255	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	478aa	Corynebacterium urealyticum DSM 7109	14
☐	638101832	FRAAL5161	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Frankia alni ACN14a	14
☐	642590831	KRH_14490	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Kocuria rhizophila DC2201	14
☐	642451101	MlutDRAFT_1491	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Micrococcus luteus NCTC 2665	14
☒	639742735	MSMEG_4290	glutamine synthetase, type I (EC:6.3.1.2)	478aa	Mycobacterium smegmatis MC2 155	16
☐	637027444	Rv2220	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	478aa	Mycobacterium tuberculosis H37Rv	14
☐	637574211	nfa16670	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	478aa	Nocardia farcinica IFM 10152	14
☐	637521420	PPA0671	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	473aa	Propionibacterium acnes KPA171202	14
☐	638087744	RHA1_ro01157	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	478aa	Rhodococcus sp. RHA1	14
☒	640475580	Strop_3332	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	474aa	Salinispora tropica CNB-440	16
☐	637266432	SCO2198	L-glutamine synthetase (EC 6.3.1.2) (IMGterm)	469aa	Streptomyces coelicolor A3(2)	14

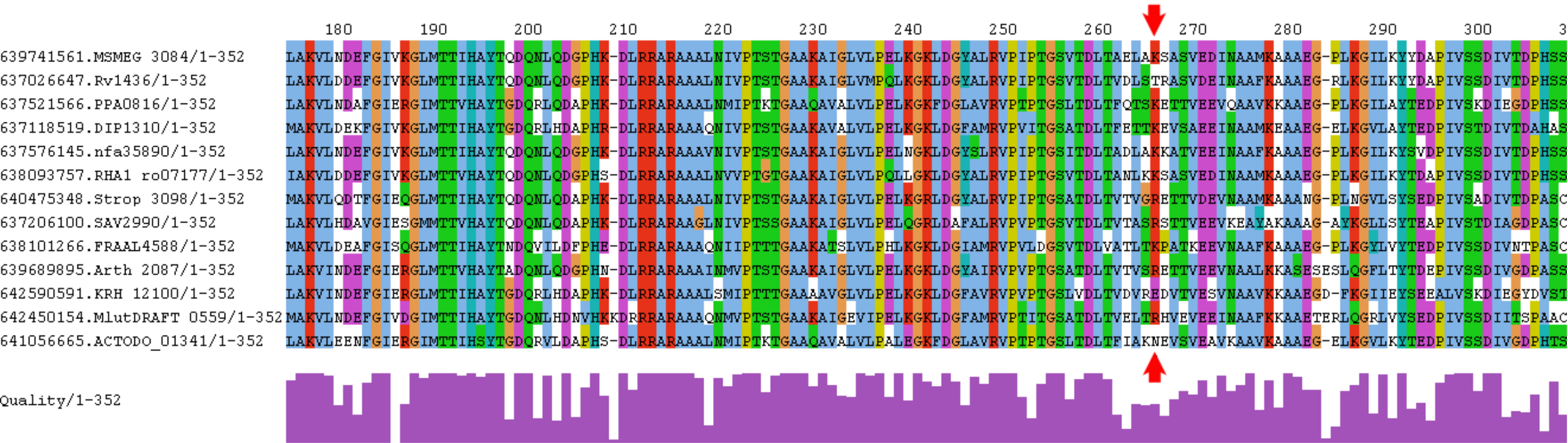


SF81: Glyceraldehyde-3-phosphate dehydrogenase, type I (GI: 118467799)

VPIPTGGSVTDLTAE⁺²⁴³LAKSASVEDINAAMK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	641056665	ACTODO_01341	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	343aa	Actinomyces odontolyticus ATCC 17982	1
✓	639689895	Arth_2087	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	336aa	Arthrobacter sp. FB24	1
✓	637118519	DIP1310	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	334aa	Corynebacterium diphtheriae NCTC 13129	1
✓	638101266	FRAAL4588	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	335aa	Frankia alni ACN14a	1
✓	642590591	KRH_12100	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	335aa	Kocuria rhizophila DC2201	1
✓	642450154	MlutDRAFT_0559	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	336aa	Micrococcus luteus NCTC 2665	2
✓	639741561	MSMEG_3084	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	340aa	Mycobacterium smegmatis MC2 155	2
✓	637026647	Rv1436	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	339aa	Mycobacterium tuberculosis H37Rv	1
✓	637576145	nfa35890	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	339aa	Nocardia farcinica IFM 10152	1
✓	637521566	PPA0816	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	335aa	Propionibacterium acnes KPA171202	1
✓	638093757	RHA1_ro07177	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	339aa	Rhodococcus sp. RHA1	1
✓	640475348	Strop_3098	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	334aa	Salinispora tropica CNB-440	1
✓	637206100	SAV2990	glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (IMGterm)	334aa	Streptomyces avermitilis MA-4680	1

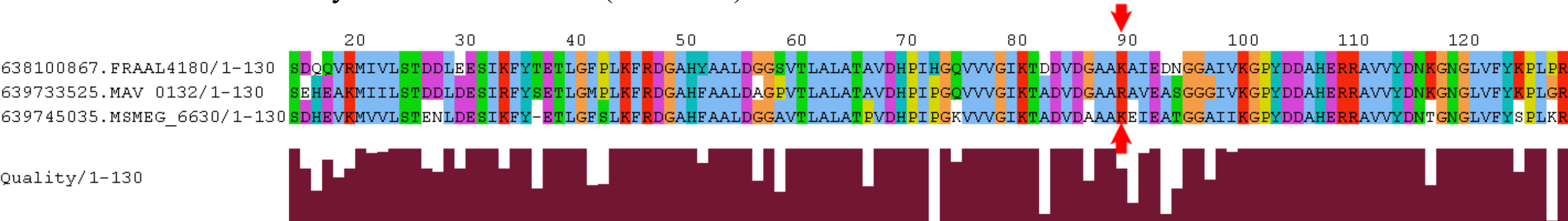


SF82: Glyoxalase family protein
(GI: 118173889)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	638100867	FRAAL4180	hypothetical protein	130aa	Frankia alni ACN14a	14
☒	639733525	MAV_0132	glyoxalase family protein	117aa	Mycobacterium avium 104	14
☒	639745035	MSMEG_6630	glyoxalase family protein	116aa	Mycobacterium smegmatis MC2 155	14

TADVDAAAK+243EIEATGGAIK

Red arrows indicate PUPylation site on Msm (MSMEG)

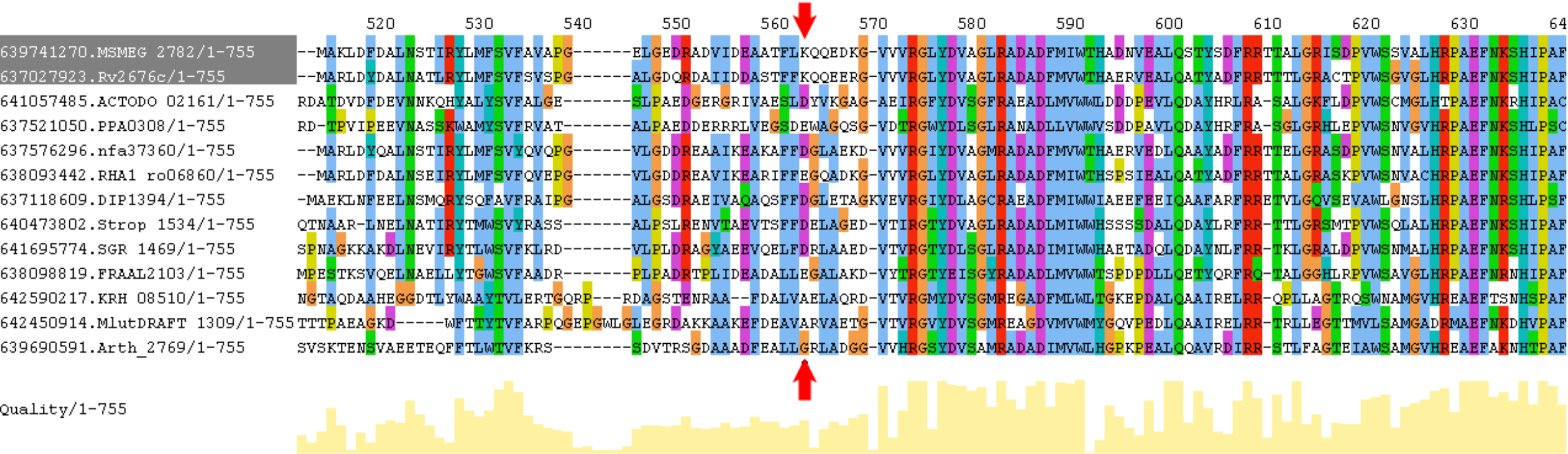


SF83: Hypothetical protein
(GI: 118468878)

ADVIDEAATFLK+243QQEDK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641057485	ACTODO_02161	hypothetical protein	734aa	Actinomyces odontolyticus ATCC 17982	20
☒	639690591	Arth_2769	Chlorite dismutase	237aa	Arthrobacter sp. FB24	20
☒	637118609	DIP1394	hypothetical protein	234aa	Corynebacterium diphtheriae NCTC 13129	20
☒	638098819	FRAAL2103	hypothetical protein	237aa	Frankia alni ACN14a	20
☒	642590217	KRH_08510	hypothetical protein	238aa	Kocuria rhizophila DC2201	20
☒	642450914	MlutDRAFT_1309	Chlorite dismutase	238aa	Micrococcus luteus NCTC 2665	20
☒	639741270	MSMEG_2782	hypothetical protein	231aa	Mycobacterium smegmatis MC2 155	20
☒	637027923	Rv2676c	hypothetical protein	231aa	Mycobacterium tuberculosis H37Rv	20
☒	637576296	nfa37360	hypothetical protein	231aa	Nocardia farcinica IFM 10152	20
☒	637521050	PPA0308	ferrochelatase (EC 4.99.1.1) (IMGterm)	683aa	Propionibacterium acnes KPA171202	20
☒	638093442	RHA1_ro06860	chlorite dismutase	231aa	Rhodococcus sp. RHA1	20
☒	640473802	Strop_1534	Chlorite dismutase	233aa	Salinispora tropica CNB-440	20
☒	641695774	SGR_1469	hypothetical protein	242aa	Streptomyces griseus griseus NBRC 13350	20

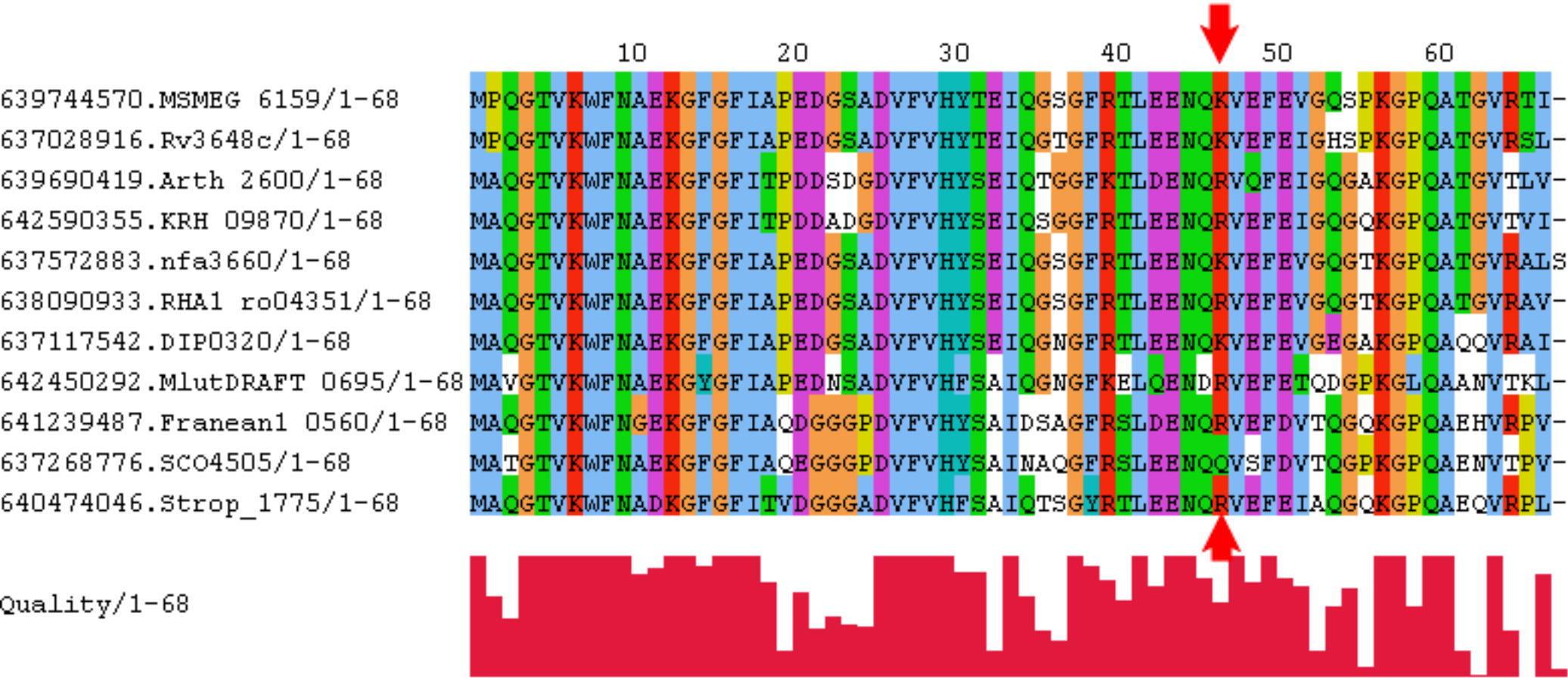


SF84: Hypothetical protein
(GI: 118473705)

TLEENQK+243VEFEVGQSPK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639690419	Arth_2600	cold-shock DNA-binding protein family (IMGterm)	67aa	Arthrobacter sp. FB24	38
☒	637117542	DIP0320	cold-shock DNA-binding protein family (IMGterm)	67aa	Corynebacterium diphtheriae NCTC 13129	38
☒	641239487	Franean1_0560	cold-shock DNA-binding protein family (IMGterm)	67aa	Frankia sp. EAN1pec	38
☒	642590355	KRH_09870	cold shock protein	67aa	Kocuria rhizophila DC2201	38
☒	642450292	MlutDRAFT_0695	cold-shock DNA-binding protein family (IMGterm)	67aa	Micrococcus luteus NCTC 2665	38
☒	639744570	MSMEG_6159	cold-shock DNA-binding protein family (IMGterm)	67aa	Mycobacterium smegmatis MC2 155	38
☒	637028916	Rv3648c	cold-shock DNA-binding protein family (IMGterm)	67aa	Mycobacterium tuberculosis H37Rv	38
☒	637572883	nfa3660	cold-shock DNA-binding protein family (IMGterm)	68aa	Nocardia farcinica IFM 10152	38
☒	638090933	RHA1_ro04351	cold-shock DNA-binding protein family (IMGterm)	67aa	Rhodococcus sp. RHA1	38
☒	640474046	Strop_1775	cold-shock DNA-binding protein family (IMGterm)	67aa	Salinispora tropica CNB-440	38
☒	637268776	SCO4505	cold-shock DNA-binding protein family (IMGterm)	67aa	Streptomyces coelicolor A3(2)	38

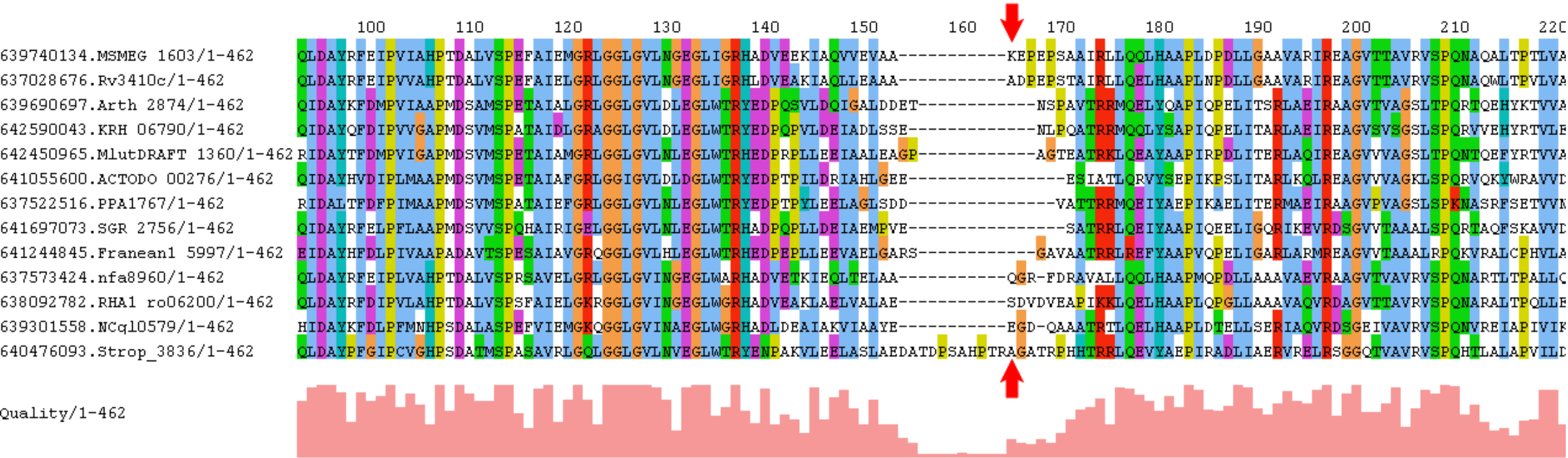


SF85: Inositol-5-monophosphate dehydrogenase
(GI: 118473142)

IAQVVEVAAK+243EPEPSAAIR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641055600	ACTODO_00276	hypothetical protein	374aa	Actinomyces odontolyticus ATCC 17982	35
☒	639690697	Arth_2874	IMP dehydrogenase family protein	378aa	Arthrobacter sp. FB24	35
☒	639301558	NCgl0579	inositol-5-monophosphate dehydrogenase	374aa	Corynebacterium glutamicum ATCC 13032 (Kitasato)	35
☒	641244845	Franean1_5997	IMP dehydrogenase family protein	376aa	Frankia sp. EAN1pec	35
☒	642590043	KRH_06790	IMP dehydrogenase family protein	373aa	Kocuria rhizophila DC2201	35
☒	642450965	MlutDRAFT_1360	IMP dehydrogenase family protein	378aa	Micrococcus luteus NCTC 2665	35
☒	639740134	MSMEG_1603	IMP dehydrogenase family protein	375aa	Mycobacterium smegmatis MC2 155	35
☒	637028676	Rv3410c	inositol-5-monophosphate dehydrogenase (EC:1.1.1.205)	375aa	Mycobacterium tuberculosis H37Rv	35
☒	637573424	nfa8960	inositol-5-monophosphate dehydrogenase	441aa	Nocardia farcinica IFM 10152	35
☒	637522516	PPA1767	inositol-5-monophosphate dehydrogenase	367aa	Propionibacterium acnes KPA171202	35
☒	638092782	RHA1_ro06200	possible IMP dehydrogenase/GMP reductase	379aa	Rhodococcus sp. RHA1	35
☒	640476093	Strop_3836	IMP dehydrogenase family protein	387aa	Salinispora tropica CNB-440	35
☒	641697073	SGR_2756	putative inosine-5'-monophosphate dehydrogenase	374aa	Streptomyces griseus griseus NBRC 13350	35

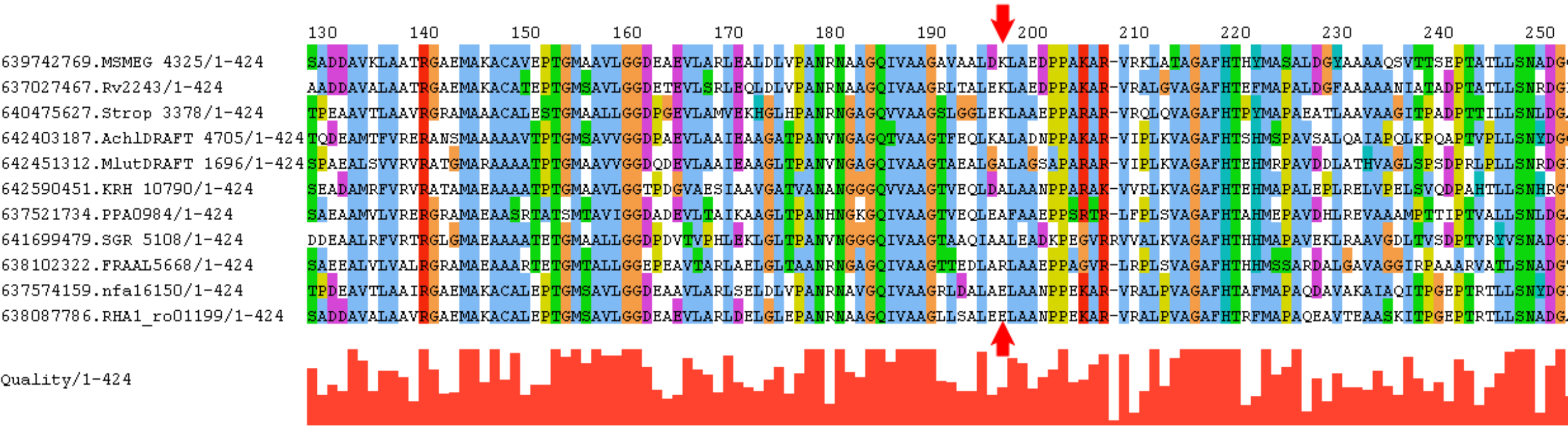


SF86: Malonyl CoA-acyl carrier protein transacylase
(GI: 118471928)

NAAGQIVAAGAVAALDK+243LAEDPPAK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	642403187	Ach1DRAFT_4705	putative (acyl-carrier-protein) S-malonyltransferase	309aa	Arthrobacter chlorophenolicus A6	36
☒	638102322	FRAAL5668	putative malonyl CoA-acyl carrier protein transacylase (MCT) (EC:2.3.1.39)	341aa	Frankia alni ACN14a	36
☒	642590451	KRH_10790	malonyl CoA-acyl carrier protein transacylase (EC:2.3.1.39)	303aa	Kocuria rhizophila DC2201	36
☒	642451312	MlutDRAFT_1696	(Acyl-carrier-protein) S-malonyltransferase (EC:2.3.1.39)	307aa	Micrococcus luteus NCTC 2665	36
☒	639742769	MSMEG_4325	malonyl CoA-acyl carrier protein transacylase (EC:2.3.1.39)	290aa	Mycobacterium smegmatis MC2 155	36
☒	637027467	Rv2243	acyl-carrier-protein S-malonyltransferase (EC:2.3.1.39)	302aa	Mycobacterium tuberculosis H37Rv	36
☒	637574159	nfa16150	acyl-carrier-protein S-malonyltransferase	304aa	Nocardia farcinica IFM 10152	36
☒	637521734	PPA0984	acyl-carrier-protein S-malonyltransferase (EC:2.3.1.39)	333aa	Propionibacterium acnes KPA171202	36
☒	638087786	RHA1_ro01199	[acyl-carrier protein] S-malonyltransferase (EC:2.3.1.39)	305aa	Rhodococcus sp. RHA1	36
☒	640475627	Strop_3378	acyl transferase domain protein	389aa	Salinispora tropica CNB-440	36
☒	641699479	SGR_5108	putative malonyl-CoA:ACP transacylase	305aa	Streptomyces griseus griseus NBRC 13350	36

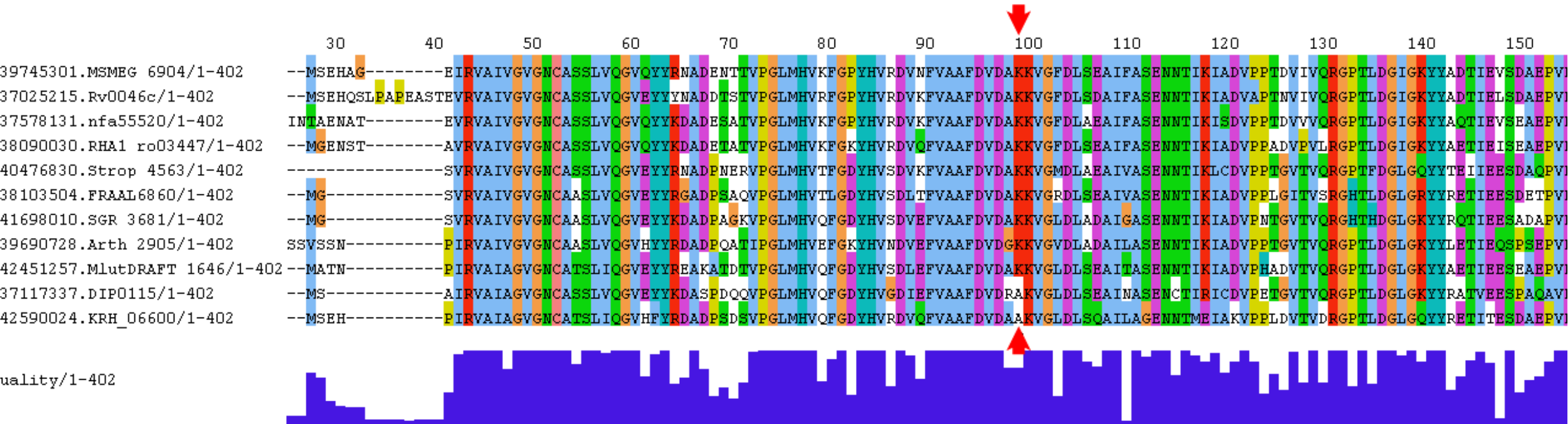


SF87: Myo-inositol-1-phosphate synthase
(GI: 118470644)

DVNFVAAFDVDAK+243K

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639690728	Arth_2905	Myo-inositol-1-phosphate synthase	388aa	Arthrobacter sp. FB24	3
☒	637117337	DIP0115	hypothetical protein	355aa	Corynebacterium diphtheriae NCTC 13129	3
☒	638103504	FRAAL6860	Putative 1L-myo-inositol-1-phosphate synthase	357aa	Frankia alni ACN14a	3
☒	642590024	KRH_06600	myo-inositol-1-phosphate synthase (EC:5.5.1.4)	361aa	Kocuria rhizophila DC2201	3
☒	642451257	MlutDRAFT_1646	inositol 1-phosphate synthase	361aa	Micrococcus luteus NCTC 2665	3
☒	639745301	MSMEG_6904	myo-inositol-1-phosphate synthase	363aa	Mycobacterium smegmatis MC2 155	3
☒	637025215	Rv0046c	MYO-INOSITOL-1-PHOSPHATE SYNTHASE INO1 (Inositol 1-phosphate synthetase) (D-glucose 6-phosphate cyclodolase) (Glucose 6-phosphate cyclase) (Glucocyclodolase) (EC:5.5.1.4)	367aa	Mycobacterium tuberculosis H37Rv	3
☒	637578131	nfa55520	putative 1L-myo-inositol-1-phosphate synthase	364aa	Nocardia farcinica IFM 10152	3
☒	638090030	RHA1_ro03447	possible myo-inositol-1-phosphate synthase	361aa	Rhodococcus sp. RHA1	3
☒	640476830	Strop_4563	Myo-inositol-1-phosphate synthase	359aa	Salinispora tropica CNB-440	3
☒	641698010	SGR_3681	putative 1L-myo-inositol-1-phosphate synthase	360aa	Streptomyces griseus griseus NBRC 13350	3

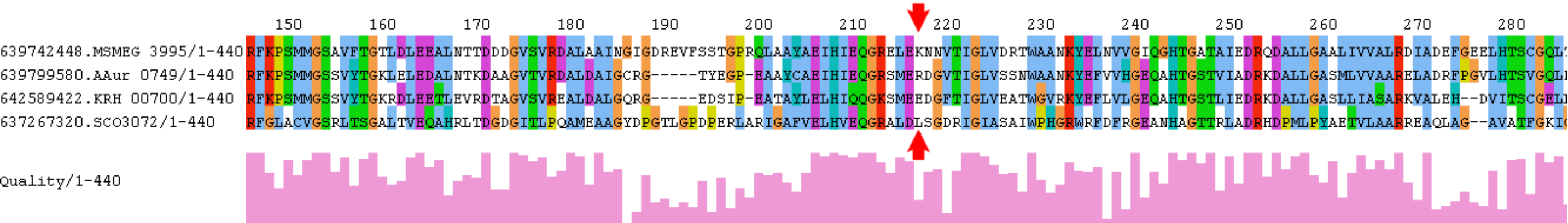


SF88: N-carbamoyl-L-amino acid amidohydrolase
(GI: 118473294)

ELEK+243NNVTIGLVDR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639799580	AAur_0749	putative N-carbamoyl-L-amino acid amidohydrolase	418aa	Arthrobacter aurescens TC1	17
☒	642589422	KRH_00700	putative hydrolase	410aa	Kocuria rhizophila DC2201	17
☒	639742448	MSMEG_3995	N-carbamoyl-L-amino acid amidohydrolase	438aa	Mycobacterium smegmatis MC2 155	18
☒	637267320	SCO3072	N-carbamoyl-L-amino acid amidohydrolase	405aa	Streptomyces coelicolor A3(2)	18

Red arrows indicate PUPylation site on Msm (MSMEG)

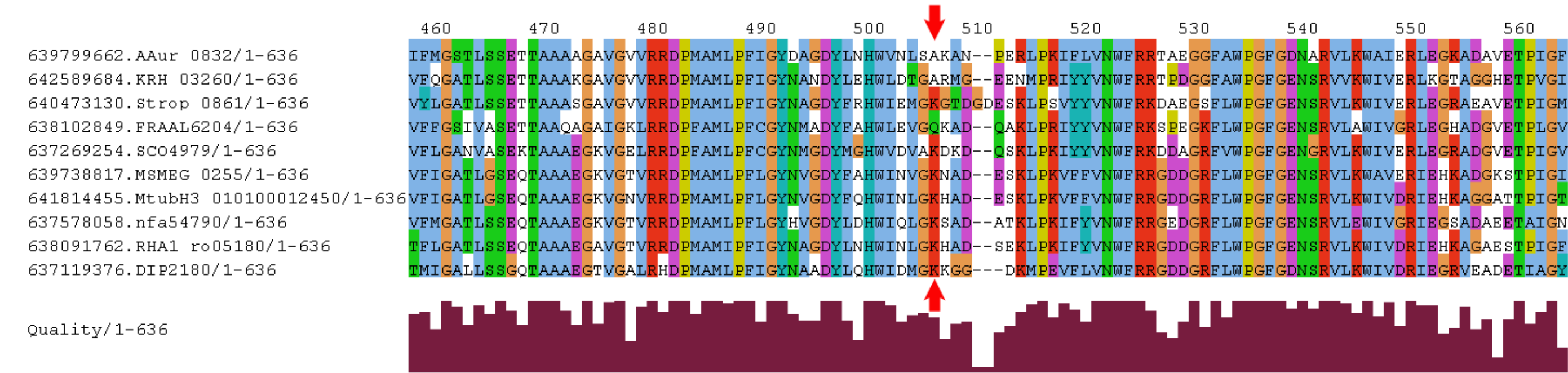


SF89: Phosphoenolpyruvate carboxykinase
(GI: 118173553)

LGYNVGDYFAHWINVVGK+243NADESK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639799662	AAur_0832	phosphoenolpyruvate carboxykinase (EC:4.1.1.32)	608aa	Arthrobacter aureescens TC1	15
☒	637119376	DIP2180	phosphoenolpyruvate carboxykinase (GTP) (EC:4.1.1.32)	612aa	Corynebacterium diphtheriae NCTC 13129	15
☒	638102849	FRAAL6204	Phosphoenolpyruvate carboxykinase [GTP] (PEP carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK) (EC:4.1.1.32)	607aa	Frankia alni ACN14a	15
☒	642589684	KRH_03260	GTP-dependent phosphoenolpyruvate carboxykinase (EC:4.1.1.32)	621aa	Kocuria rhizophila DC2201	15
☒	639738817	MSMEG_0255	phosphoenolpyruvate carboxykinase (EC:4.1.1.32)	608aa	Mycobacterium smegmatis MC2 155	15
☒	641814455	MtubH3_010100012450	phosphoenolpyruvate carboxykinase (EC:4.1.1.32)	606aa	Mycobacterium tuberculosis H37Ra	15
☒	637578058	nfa54790	phosphoenolpyruvate carboxykinase (GTP) (EC:4.1.1.32)	610aa	Nocardia farcinica IFM 10152	15
☒	638091762	RHA1_ro05180	phosphoenolpyruvate carboxykinase (GTP) (EC:4.1.1.32)	609aa	Rhodococcus sp. RHA1	15
☒	640473130	Strop_0861	Phosphoenolpyruvate carboxykinase (GTP) (EC:4.1.1.32)	612aa	Salinispora tropica CNB-440	15
☒	637269254	SCO4979	phosphoenolpyruvate carboxykinase (GTP) (EC:4.1.1.32)	609aa	Streptomyces coelicolor A3(2)	15

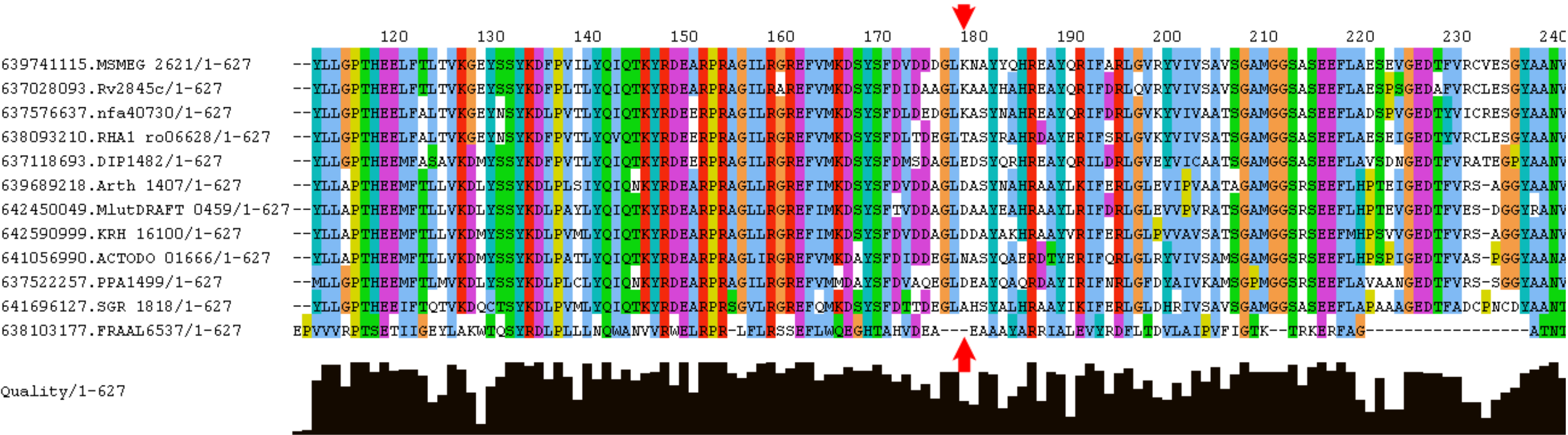


SF90: Prolyl-tRNA synthetase
(GI: 118472270)

DSYSFDVDDDGLK+243NAYYQHR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	641056990	ACTODO_01666	hypothetical protein	602aa	Actinomyces odontolyticus ATCC 17982	6
☐	639689218	Arth_1407	prolyl-tRNA synthetase	603aa	Arthrobacter sp. FB24	4
☐	637118693	DIP1482	prolyl-tRNA synthetase (EC:6.1.1.15)	585aa	Corynebacterium diphtheriae NCTC 13129	4
☐	638103177	FRAAL6537	prolyl-tRNA synthetase (EC:6.1.1.15) (IMGterm)	468aa	Frankia alni ACN14a	5
☐	642590999	KRH_16100	prolyl-tRNA synthetase (EC:6.1.1.15)	606aa	Kocuria rhizophila DC2201	4
☐	642450049	MlutDRAFT_0459	prolyl-tRNA synthetase (EC:6.1.1.15)	608aa	Micrococcus luteus NCTC 2665	4
☒	639741115	MSMEG_2621	prolyl-tRNA synthetase (EC:6.1.1.15)	585aa	Mycobacterium smegmatis MC2 155	6
☐	637028093	Rv2845c	prolyl-tRNA synthetase (EC:6.1.1.15)	582aa	Mycobacterium tuberculosis H37Rv	4
☐	637576637	nfa40730	prolyl-tRNA synthetase	585aa	Nocardia farcinica IFM 10152	4
☐	637522257	PPA1499	prolyl-tRNA synthetase (EC:6.1.1.15)	585aa	Propionibacterium acnes KPA171202	4
☐	638093210	RHA1_ro06628	proline-tRNA ligase (EC:6.1.1.15)	581aa	Rhodococcus sp. RHA1	4
☐	641696127	SGR_1818	putative prolyl-tRNA synthetase	567aa	Streptomyces griseus griseus NBRC 13350	4



SF91: Putative adenosylhomocysteinase
(GI: 116266924)

LSK+243DQAEYIGVDVEGPY

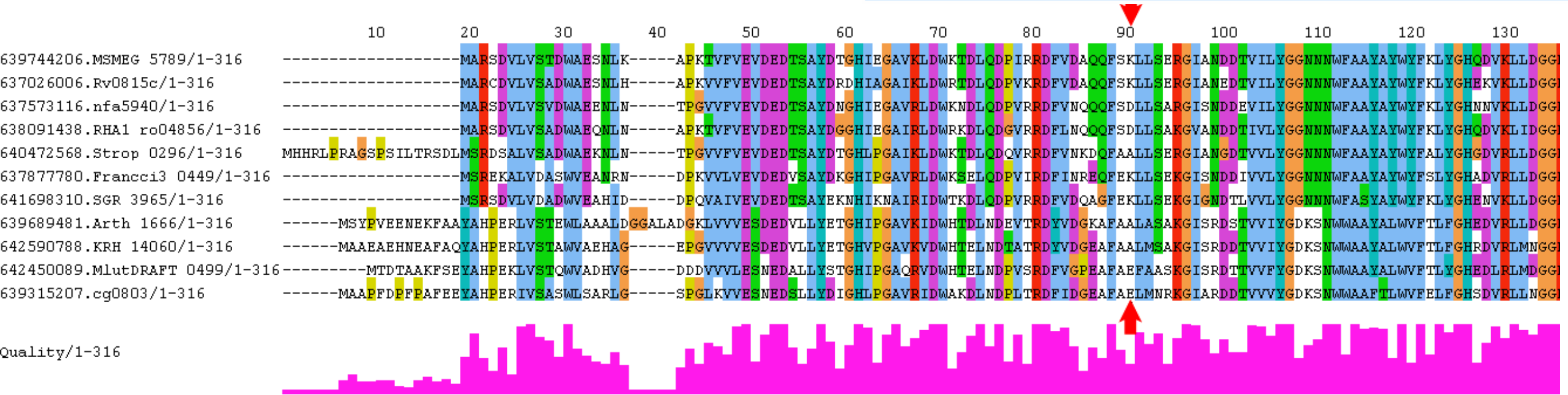
*No homologous proteins found

SF92: Putative thiosulfate sulfurtransferase
(GI: 118472788)

DFVDAQQFSK+243LLSER

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639689481	Arth_1666	Thiosulfate sulfurtransferase (EC:2.8.1.1)	303aa	Arthrobacter sp. FB24	24
☒	639315207	cg0803	THIOSULFATE SULFURTRANSFERASE (EC:2.8.1.1)	301aa	Corynebacterium glutamicum ATCC 13032 (Bielefeld)	24
☒	637877780	Francci3_0449	Rhodanese-like (EC:2.8.1.1)	277aa	Frankia sp. CcI3	24
☒	642590788	KRH_14060	putative thiosulfate sulfurtransferase (EC:2.8.1.1)	302aa	Kocuria rhizophila DC2201	24
☒	642450089	MlutDRAFT_0499	Rhodanese domain protein	297aa	Micrococcus luteus NCTC 2665	24
☒	639744206	MSMEG_5789	putative thiosulfate sulfurtransferase (EC:2.8.1.1)	277aa	Mycobacterium smegmatis MC2 155	24
☒	637026006	Rv0815c	PROBABLE THIOSULFATE SULFURTRANSFERASE CYSA2 (RHODANESE-LIKE PROTEIN) (THIOSULFATE CYANIDE TRANSULFURASE) (THIOSULFATE THIOTRANSFERASE) (EC:2.8.1.1)	277aa	Mycobacterium tuberculosis H37Rv	24
☒	637573116	nfa5940	putative thiosulfate sulfurtransferase (EC:2.8.1.1)	277aa	Nocardia farcinica IFM 10152	24
☒	638091438	RHA1_ro04856	thiosulfate sulfurtransferase (EC:2.8.1.1)	278aa	Rhodococcus sp. RHA1	24
☒	640472568	Strop_0296	Rhodanese domain protein (EC:2.8.1.1)	301aa	Salinispora tropica CNB-440	24
☒	641698310	SGR_3965	putative thiosulfate sulfurtransferase (EC:2.8.1.1)	279aa	Streptomyces griseus griseus NBRC 13350	24

Red arrows indicate PUPylation site on Msm (MSMEG)

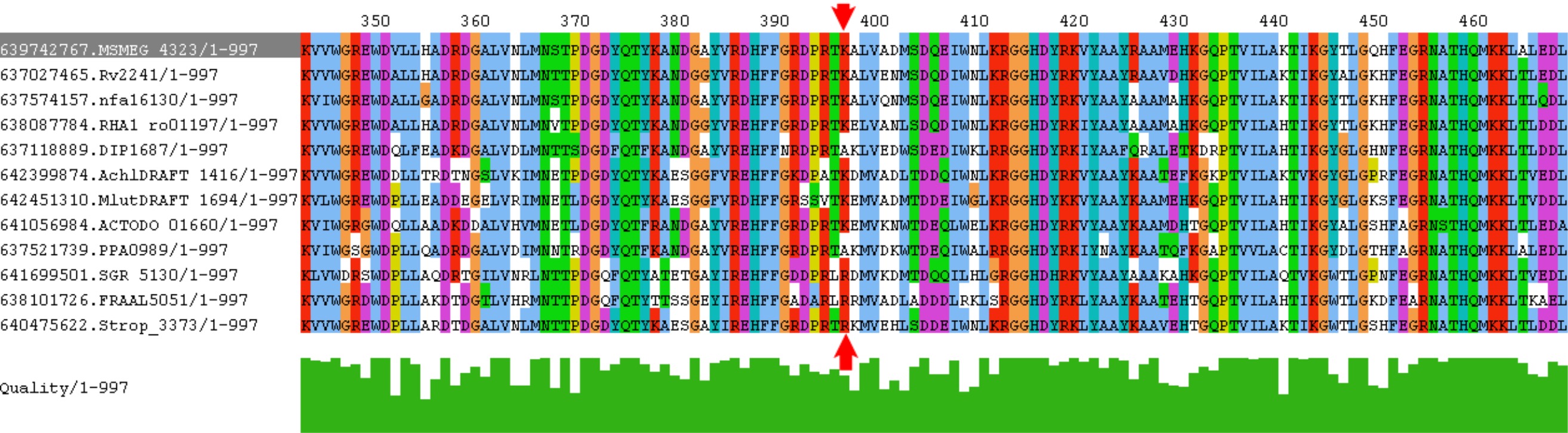


SF93: Pyruvate dehydrogenase subunit E1
(GI: 118470577)

TK+243ALVADMSDQEIWNLK

Red arrows indicate PUPylation site on Msm (MSMEG)

Gene Object ID	Locus Tag	Product Name	Genome	Scaffold ID
637027465	Rv2241	pyruvate dehydrogenase subunit E1	Mycobacterium tuberculosis H37Rv	NC_000962
637118889	DIP1687	pyruvate dehydrogenase subunit E1	Corynebacterium diphtheriae NCTC 13129	NC_002935
637521739	PPA0989	pyruvate dehydrogenase subunit E1	Propionibacterium acnes KPA171202	NC_006085
637574157	nfa16130	pyruvate dehydrogenase subunit E1	Nocardia farcinica IFM 10152	NC_006361
638087784	RHA1_ro01197	pyruvate dehydrogenase E1 component	Rhodococcus sp. RHA1	NC_008268
638101726	FRAAL5051	pyruvate dehydrogenase, decarboxylase subunit, thiamin-binding	Frankia alni ACN14a	NC_008278
639742767	MSMEG_4323	pyruvate dehydrogenase E1 component	Mycobacterium smegmatis MC2 155	NC_008596
640475622	Strop_3373	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	Salinispora tropica CNB-440	NC_009380
641056984	ACTODO_01660	hypothetical protein	Actinomyces odontolyticus ATCC 17982	NZ_AAYI02000004
641699501	SGR_5130	putative pyruvate dehydrogenase E1 component	Streptomyces griseus griseus NBRC 13350	NC_010572
642399874	AchlDRAFT_1416	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	Arthrobacter chlorophenolicus A6	NZ_ABKU01000009
642451310	MlutDRAFT_1694	2-oxo-acid dehydrogenase E1 subunit, homodimeric type	Micrococcus luteus NCTC 2665	NZ_ABLQ01000014

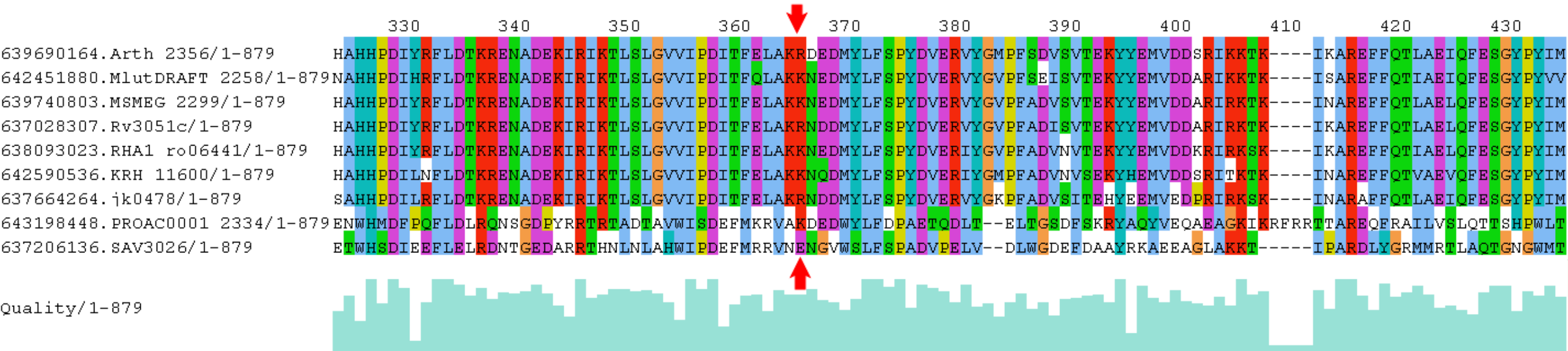


SF94: Ribonucleoside diphosphate reductase, alpha subunit (GI: 118170953)

K+243NEDMYLFSFYDVER

Red arrows indicate PUPylation site on Msm (MSMEG)

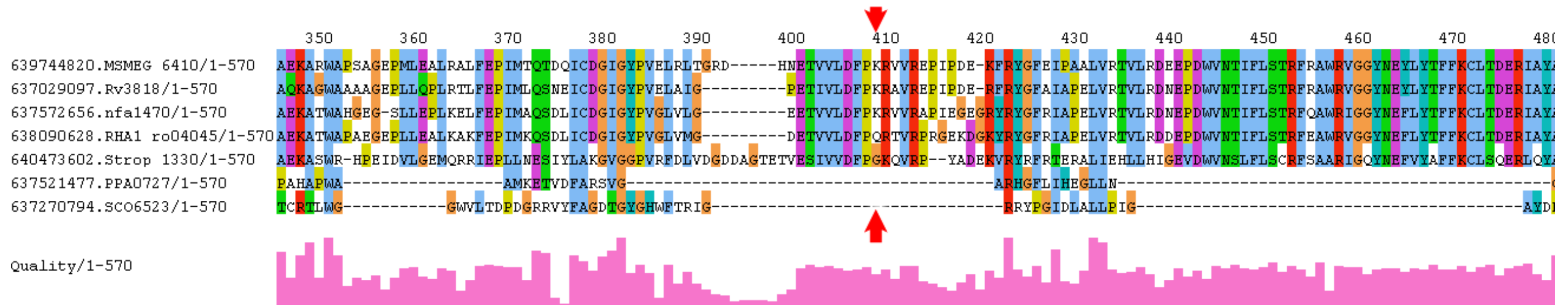
Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639690164	Arth_2356	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	727aa	Arthrobacter sp. FB24	18
☒	637664264	jk0478	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	720aa	Corynebacterium jeikeium K411	18
☒	642590536	KRH_11600	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	719aa	Kocuria rhizophila DC2201	18
☒	642451880	MlutDRAFT_2258	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	725aa	Micrococcus luteus NCTC 2665	18
☒	639740803	MSMEG_2299	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	706aa	Mycobacterium smegmatis MC2 155	18
☒	637028307	Rv3051c	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	693aa	Mycobacterium tuberculosis H37Rv	18
☒	643198448	PROAC0001_2334	ribonucleoside-diphosphate reductase, alpha subunit (EC:1.17.4.1)	856aa	Propionibacterium acnes SK137	18
☒	638093023	RHA1_ro06441	ribonucleoside-diphosphate reductase class Ib alpha subunit (EC 1.17.4.1) (IMGterm)	693aa	Rhodococcus sp. RHA1	18
☒	637206136	SAV3026	ribonucleoside-diphosphate reductase alpha chain (EC:1.17.4.1)	793aa	Streptomyces avermitilis MA-4680	18



DHNETVVLDFPK+243R

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639744820	MSMEG_6410	Rieske 2Fe-2S family protein	518aa	Mycobacterium smegmatis MC2 155	12
☒	637029097	Rv3818	hypothetical protein	516aa	Mycobacterium tuberculosis H37Rv	12
☒	637572656	nfa1470	hypothetical protein	511aa	Nocardia farcinica IFM 10152	12
☒	637521477	PPA0727	putative Zn-dependent hydrolase	215aa	Propionibacterium acnes KPA171202	12
☒	638090628	RHA1_ro04045	hypothetical protein	543aa	Rhodococcus sp. RHA1	12
☒	640473602	Strop_1330	Rieske (2Fe-2S) domain protein	529aa	Salinispora tropica CNB-440	12
☒	637270794	SCO6523	hypothetical protein	346aa	Streptomyces coelicolor A3(2)	12

Red arrows indicate PUPylation site on Msm (MSMEG)

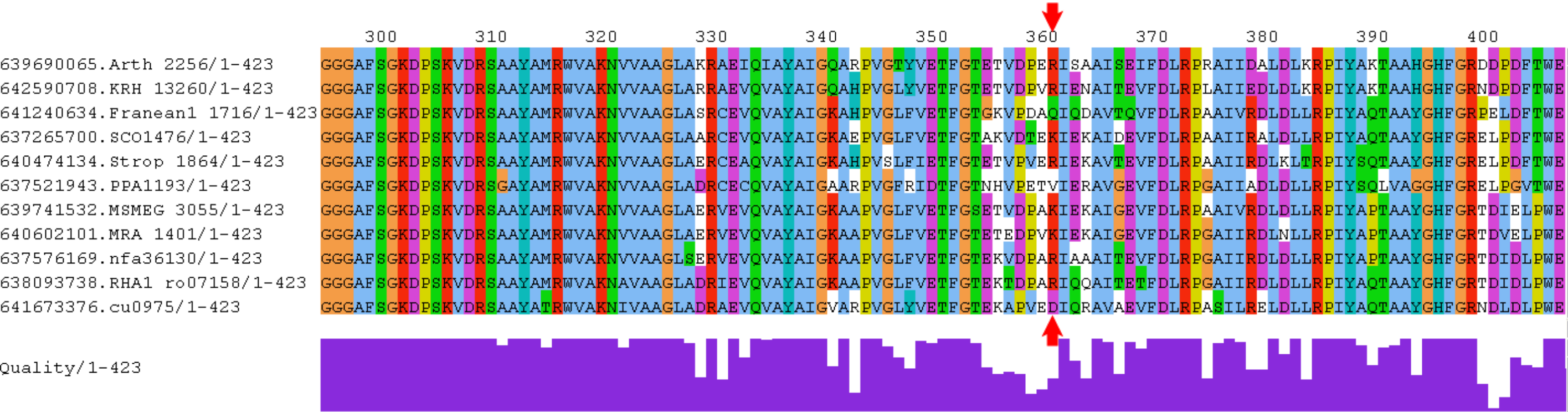


SF96: S-adenosylmethionine synthetase
(GI: 118169294)

AAPVGLFVETFGSETVDPK+243IEK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	639690065	Arth_2256	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	411aa	Arthrobacter sp. FB24	19
✓	641673376	cu0975	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	401aa	Corynebacterium urealyticum DSM 7109	19
✓	641240634	Franean1_1716	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	398aa	Frankia sp. EAN1pec	19
✓	642590708	KRH_13260	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	401aa	Kocuria rhizophila DC2201	19
✓	639741532	MSMEG_3055	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	399aa	Mycobacterium smegmatis MC2 155	19
✓	640602101	MRA_1401	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	403aa	Mycobacterium tuberculosis H37Ra	19
✓	637576169	nfa36130	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	404aa	Nocardia farcinica IFM 10152	19
✓	637521943	PPA1193	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	397aa	Propionibacterium acnes KPA171202	19
✓	638093738	RHA1_ro07158	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	404aa	Rhodococcus sp. RHA1	19
✓	640474134	Strop_1864	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	399aa	Salinispora tropica CNB-440	19
✓	637265700	SCO1476	methionine adenosyltransferase (EC 2.5.1.6) (IMGterm)	402aa	Streptomyces coelicolor A3(2)	19

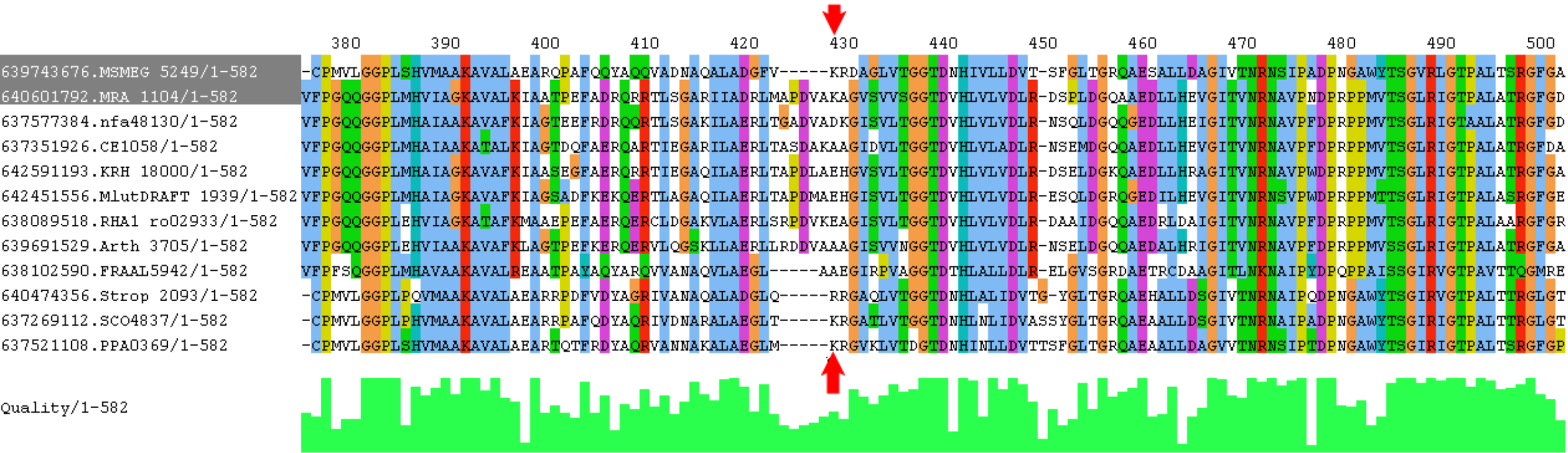


SF97: Serine hydroxymethyltransferase
(GI: 118468818)

AQQVADNAQALADGFVK+243R

Red arrows indicate PUPylation site on Msm (MSMEG)

Gene Object ID	Locus Tag	Product Name	Genome	Scaffold ID
637269112	SCO4837	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Streptomyces coelicolor A3(2)	NC_003888
637351926	CE1058	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Corynebacterium efficiens YS-314	NC_004369
637521108	PPA0369	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Propionibacterium acnes KPA171202	NC_006085
637577384	nfa48130	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Nocardia farcinica IFM 10152	NC_006361
638089518	RHA1_ro02933	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Rhodococcus sp. RHA1	NC_008268
638102590	FRAAL5942	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Frankia alni ACN14a	NC_008278
639691529	Arth_3705	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Arthrobacter sp. FB24	NC_008541
639743676	MSMEG_5249	serine hydroxymethyltransferase	Mycobacterium smegmatis MC2 155	NC_008596
640474356	Strop_2093	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Salinispora tropica CNB-440	NC_009380
640601792	MRA_1104	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Mycobacterium tuberculosis H37Ra	NC_009525
642451556	MlutDRAFT_1939	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Micrococcus luteus NCTC 2665	NZ_ABLQ01000018
642591193	KRH_18000	serine hydroxymethyltransferase (EC 2.1.2.1) (IMGterm)	Kocuria rhizophila DC2201	NC_010617

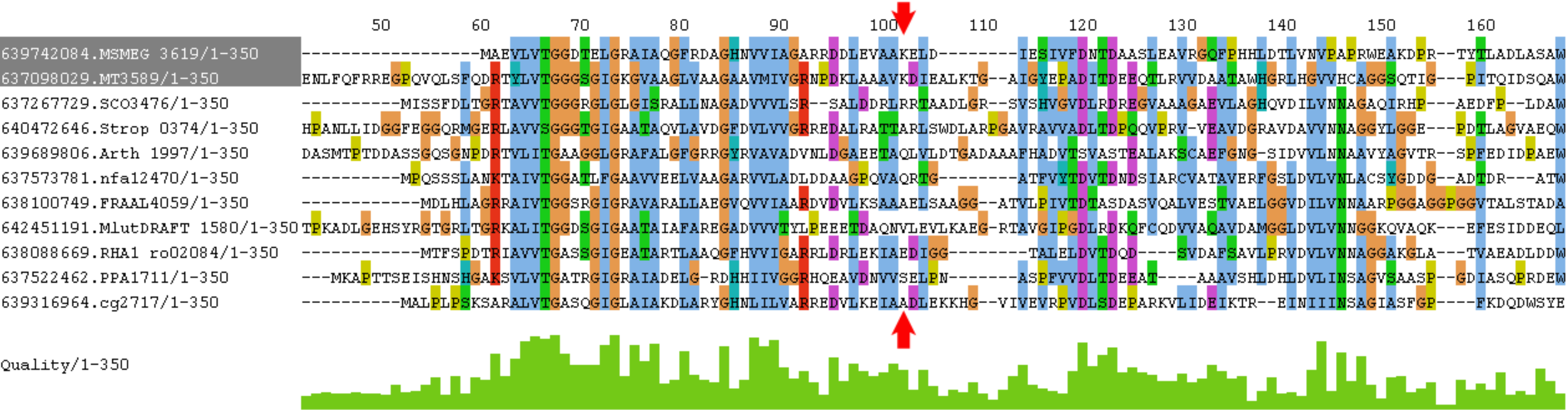


SF98: Short chain dehydrogenase
(GI: 118473731)

RDDLEVAAK+243ELDIESIVFDNTDAASLEAVR

Red arrows indicate PUPylation site on Msm (MSMEG)

Gene Object ID	Locus Tag	Product Name	Genome	Scaffold ID
637098029	MT3589	short chain dehydrogenase	Mycobacterium tuberculosis CDC1551	NC_002755
637267729	SCO3476	short-chain dehydrogenase	Streptomyces coelicolor A3(2)	NC_003888
637522462	PPA1711	short chain dehydrogenase	Propionibacterium acnes KPA171202	NC_006085
637573781	nfa12470	short chain dehydrogenase	Nocardia farcinica IFM 10152	NC_006361
638088669	RHA1_ro02084	short chain dehydrogenase	Rhodococcus sp. RHA1	NC_008268
638100749	FRAAL4059	putative short-chain dehydrogenase	Frankia alni ACN14a	NC_008278
639316964	cg2717	Short-chain dehydrogenase	Corynebacterium glutamicum ATCC 13032 (Bielefeld)	NC_006958
639689806	Arth_1997	short-chain dehydrogenase/reductase SDR	Arthrobacter sp. FB24	NC_008541
639742084	MSMEG_3619	short chain dehydrogenase	Mycobacterium smegmatis MC2 155	NC_008596
640472646	Strop_0374	short-chain dehydrogenase/reductase SDR	Salinispora tropica CNB-440	NC_009380
642451191	MlutDRAFT_1580	short-chain dehydrogenase/reductase SDR	Micrococcus luteus NCTC 2665	NZ_ABLQ01000013

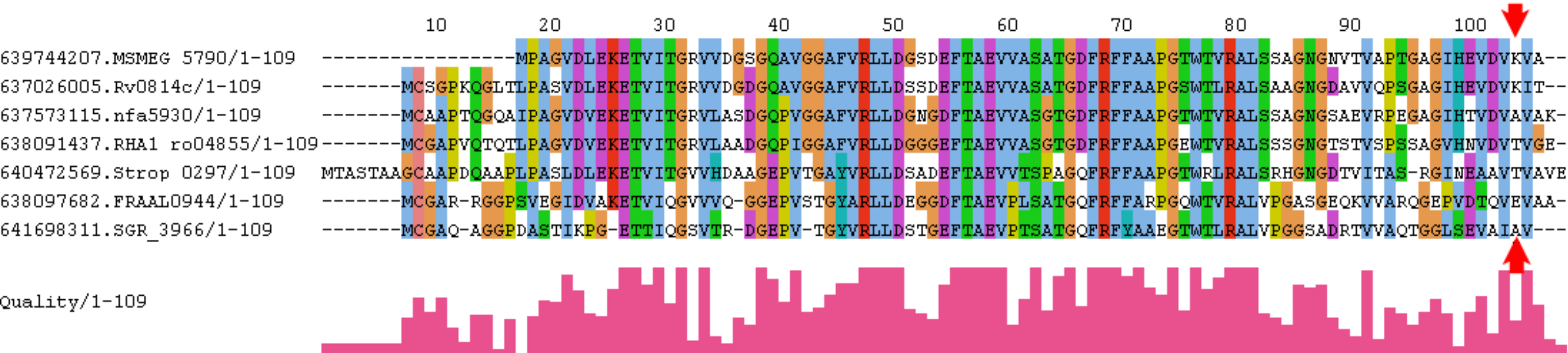


SF99: SseC protein
(GI: 118468699)

ALSSAGNGNVTVAPTGAGIHEVDVK+243VA*

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	638097682	FRAAL0944	hypothetical protein	99aa	Frankia alni ACN14a	19
☒	639744207	MSMEG_5790	SseC protein	90aa	Mycobacterium smegmatis MC2 155	19
☒	637026005	Rv0814c	CONSERVED HYPOTHETICAL PROTEIN SSEC2	100aa	Mycobacterium tuberculosis H37Rv	19
☒	637573115	nfa5930	hypothetical protein	101aa	Nocardia farcinica IFM 10152	19
☒	638091437	RHA1_ro04855	hypothetical protein	101aa	Rhodococcus sp. RHA1	19
☒	640472569	Strop_0297	protein of unknown function DUF1416	108aa	Salinispora tropica CNB-440	19
☒	641698311	SGR_3966	hypothetical protein	95aa	Streptomyces griseus griseus NBRC 13350	19

Red arrows indicate PUPylation site on Msm (MSMEG)

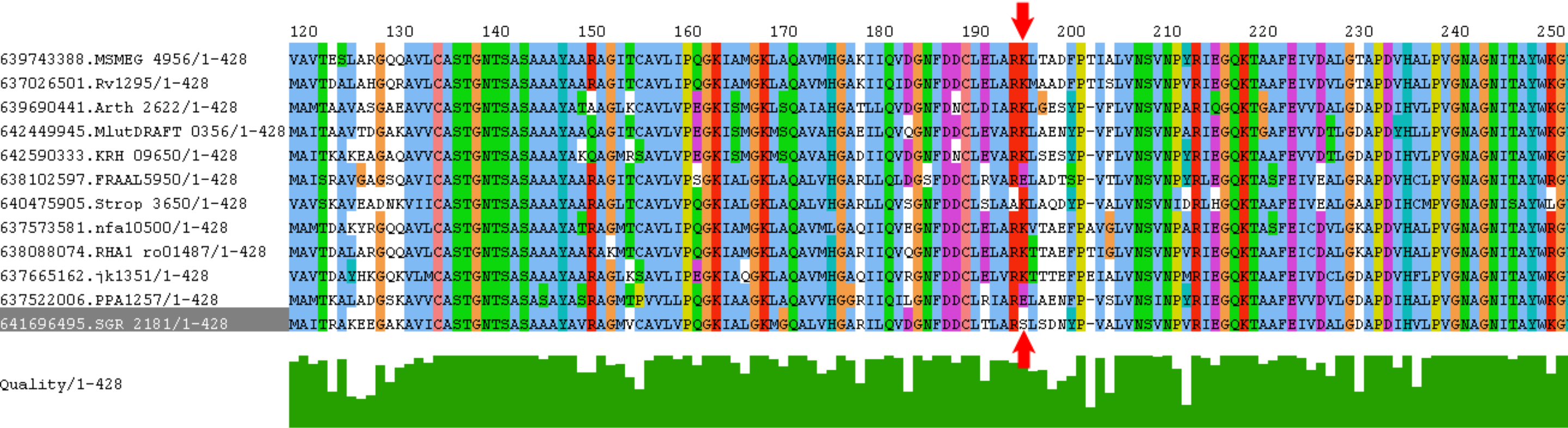


SF100: Threonine synthase
(GI: 118468890)

K+243LTADFPTIALVNSVNPYR

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639690441	Arth_2622	L-threonine synthase (EC 4.2.3.1) (IMGterm)	368aa	Arthrobacter sp. FB24	22
☒	637665162	jk1351	hypothetical protein (EC:4.2.3.1)	357aa	Corynebacterium jeikeium K411	22
☒	638102597	FRAAL5950	L-threonine synthase (EC 4.2.3.1) (IMGterm)	405aa	Frankia alni ACN14a	22
☒	642590333	KRH_09650	L-threonine synthase (EC 4.2.3.1) (IMGterm)	369aa	Kocuria rhizophila DC2201	22
☒	642449945	MlutDRAFT_0356	L-threonine synthase (EC 4.2.3.1) (IMGterm)	370aa	Micrococcus luteus NCTC 2665	22
☒	639743388	MSMEG_4956	threonine synthase (EC:4.2.3.1)	360aa	Mycobacterium smegmatis MC2 155	22
☒	637026501	Rv1295	L-threonine synthase (EC 4.2.3.1) (IMGterm)	360aa	Mycobacterium tuberculosis H37Rv	22
☒	637573581	nfa10500	L-threonine synthase (EC 4.2.3.1) (IMGterm)	364aa	Nocardia farcinica IFM 10152	22
☒	637522006	PPA1257	L-threonine synthase (EC 4.2.3.1) (IMGterm)	356aa	Propionibacterium acnes KPA171202	22
☒	638088074	RHA1_ro01487	L-threonine synthase (EC 4.2.3.1) (IMGterm)	360aa	Rhodococcus sp. RHA1	22
☒	640475905	Strop_3650	L-threonine synthase (EC 4.2.3.1) (IMGterm)	349aa	Salinispora tropica CNB-440	22
☒	641696495	SGR_2181	L-threonine synthase (EC 4.2.3.1) (IMGterm)	356aa	Streptomyces griseus griseus NBRC 13350	22

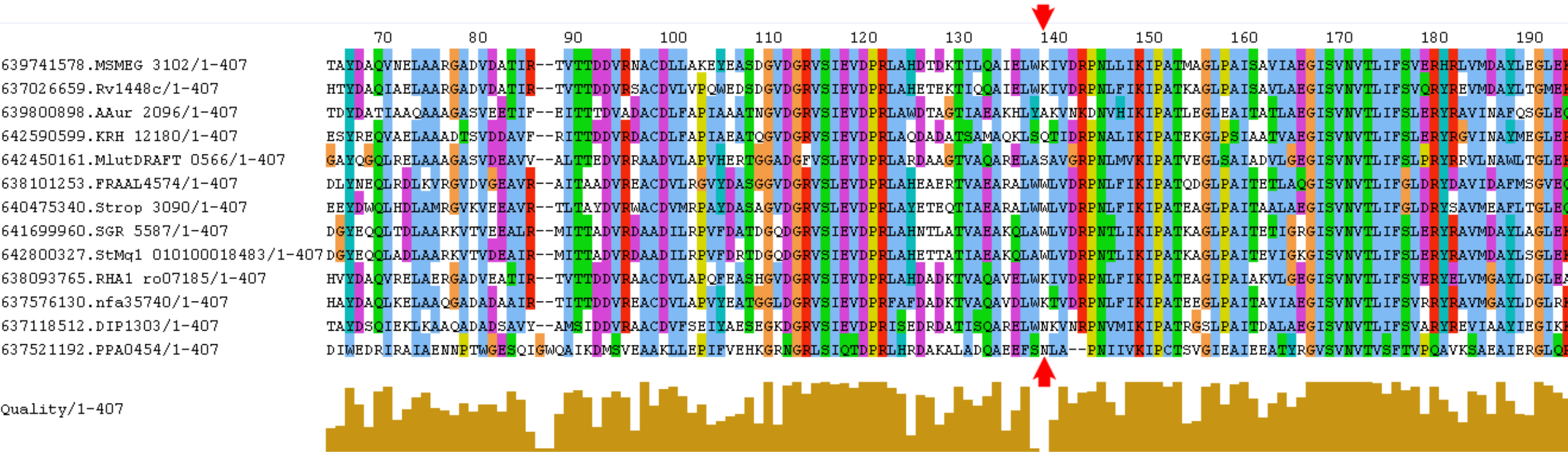


SF101: Transaldolase
(GI: 118171487)

LAHDTDKTILQAIELWK+243IVDRPNLLIK

Red arrows indicate PUPylation site on Msm (MSMEG)

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
☒	639800898	AAur_2096	transaldolase (EC 2.2.1.2) (IMGterm)	372aa	Arthrobacter aureescens TC1	20
☒	637118512	DIP1303	transaldolase (EC 2.2.1.2) (IMGterm)	360aa	Corynebacterium diphtheriae NCTC 13129	20
☒	638101253	FRAAL4574	transaldolase (EC 2.2.1.2) (IMGterm)	371aa	Frankia alni ACN14a	20
☒	642590599	KRH_12180	transaldolase (EC 2.2.1.2) (IMGterm)	368aa	Kocuria rhizophila DC2201	20
☒	642450161	MlutDRAFT_0566	transaldolase (EC 2.2.1.2) (IMGterm)	372aa	Micrococcus luteus NCTC 2665	20
☒	639741578	MSMEG_3102	transaldolase (EC 2.2.1.2) (IMGterm)	371aa	Mycobacterium smegmatis MC2 155	21
☒	637026659	Rv1448c	transaldolase (EC 2.2.1.2) (IMGterm)	373aa	Mycobacterium tuberculosis H37Rv	20
☒	637576130	nfa35740	transaldolase (EC 2.2.1.2) (IMGterm)	380aa	Nocardia farcinica IFM 10152	20
☒	637521192	PPA0454	transaldolase (EC 2.2.1.2) (IMGterm)	358aa	Propionibacterium acnes KPA171202	21
☒	638093765	RHA1_ro07185	transaldolase (EC 2.2.1.2) (IMGterm)	374aa	Rhodococcus sp. RHA1	20
☒	640475340	Strop_3090	transaldolase (EC 2.2.1.2) (IMGterm)	392aa	Salinispora tropica CNB-440	20
☒	641699960	SGR_5587	transaldolase (EC 2.2.1.2) (IMGterm)	375aa	Streptomyces griseus griseus NBRC 13350	20
☒	642800327	StMg1_010100018483	transaldolase (EC 2.2.1.2) (IMGterm)	372aa	Streptomyces sp. Mg1	20

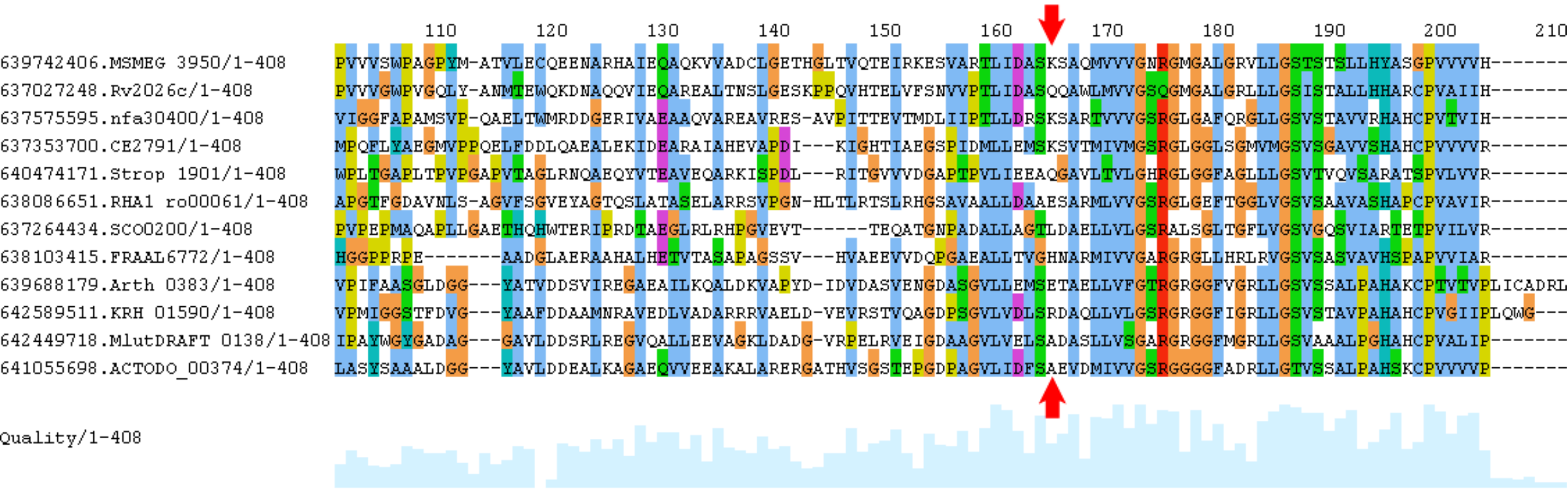


SF102: Universal stress protein family protein
(GI: 118471745)

TLIDASK+243SAQMVVVGNR

Selection	Gene Object ID	Locus Tag	Product Name	AA Seq. Length	Genome	Batch ¹
✓	641055698	ACTODO_00374	hypothetical protein	324aa	Actinomyces odontolyticus ATCC 17982	29
✓	639688179	Arth_0383	UspA domain protein	343aa	Arthrobacter sp. FB24	29
✓	637353700	CE2791	hypothetical protein	318aa	Corynebacterium efficiens YS-314	29
✓	638103415	FRAAL6772	hypothetical protein	268aa	Frankia alni ACN14a	29
✓	642589511	KRH_01590	hypothetical protein	315aa	Kocuria rhizophila DC2201	29
✓	642449718	MlutDRAFT_0138	UspA domain protein	324aa	Micrococcus luteus NCTC 2665	29
✓	639742406	MSMEG_3950	universal stress protein family protein	294aa	Mycobacterium smegmatis MC2 155	29
✓	637027248	Rv2026c	hypothetical protein	294aa	Mycobacterium tuberculosis H37Rv	29
✓	637575595	nfa30400	hypothetical protein	345aa	Nocardia farcinica IFM 10152	29
✓	638086651	RHA1_ro00061	possible universal stress protein	295aa	Rhodococcus sp. RHA1	29
✓	640474171	Strop_1901	UspA domain protein	295aa	Salinispora tropica CNB-440	29
✓	637264434	SCO0200	hypothetical protein	301aa	Streptomyces coelicolor A3(2)	29

Red arrows indicate PUPylation site on Msm (MSMEG)



V. Structural Annotations

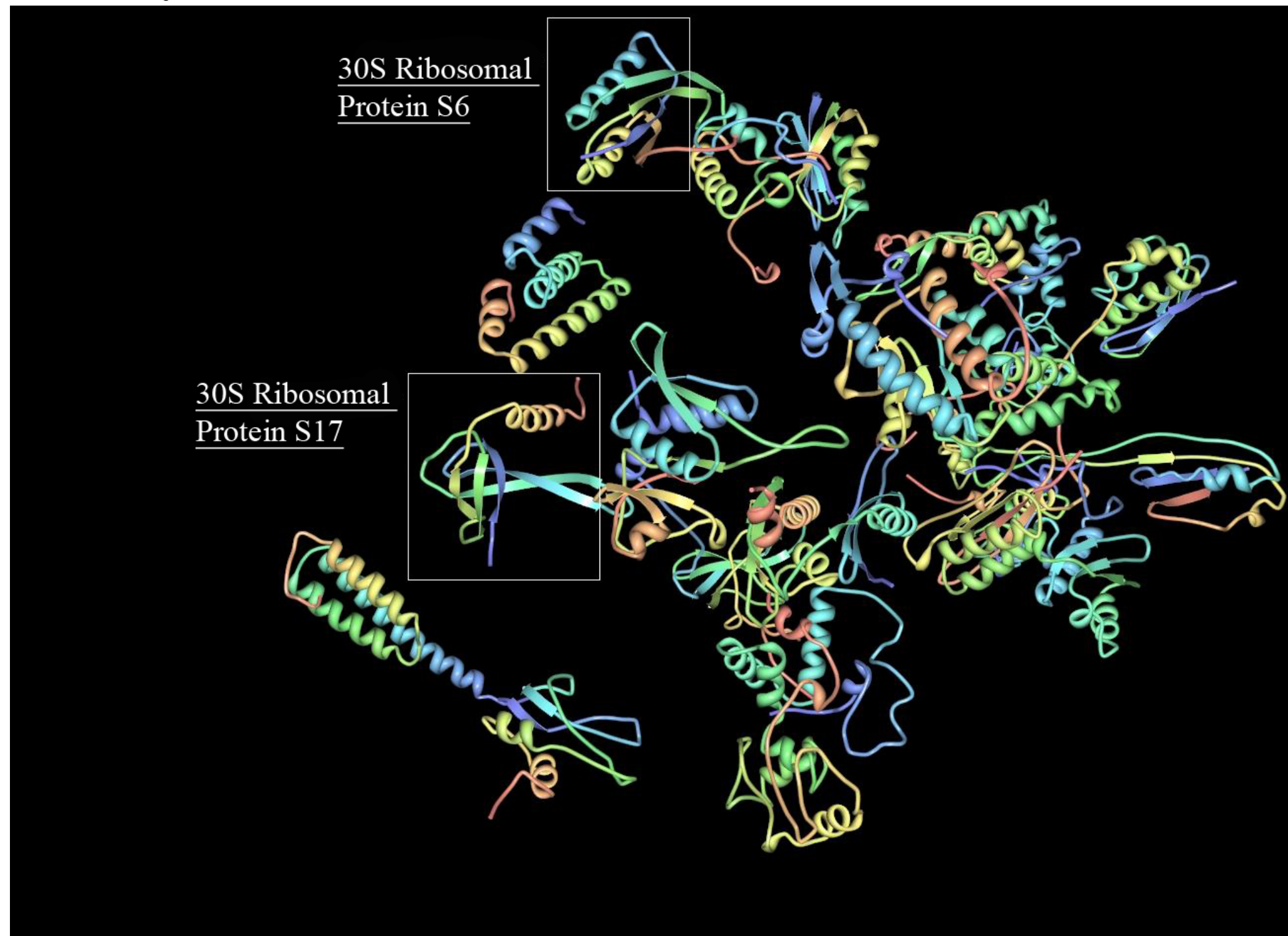
a. Figures SF103-SF123: Indication of site of modification on crystal structures of homologous proteins.....121

Using NCBI pBLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and the Protein Data Bank (www.pdb.org), crystal structures from homologous proteins were found for some of the PUPylated proteins. The location of the modified lysine (or homologous residue determined from sequence alignment) is indicated by the white arrows in each figure. Overall, there is a general trend for the modified lysine to be located on the peripheral parts of the protein and on peripheral subunits of multi-subunit protein complexes (mainly the ribosome). The modified lysine was found to be located on either an alpha helix or unstructured region of the protein with modification on a beta sheet not seen.

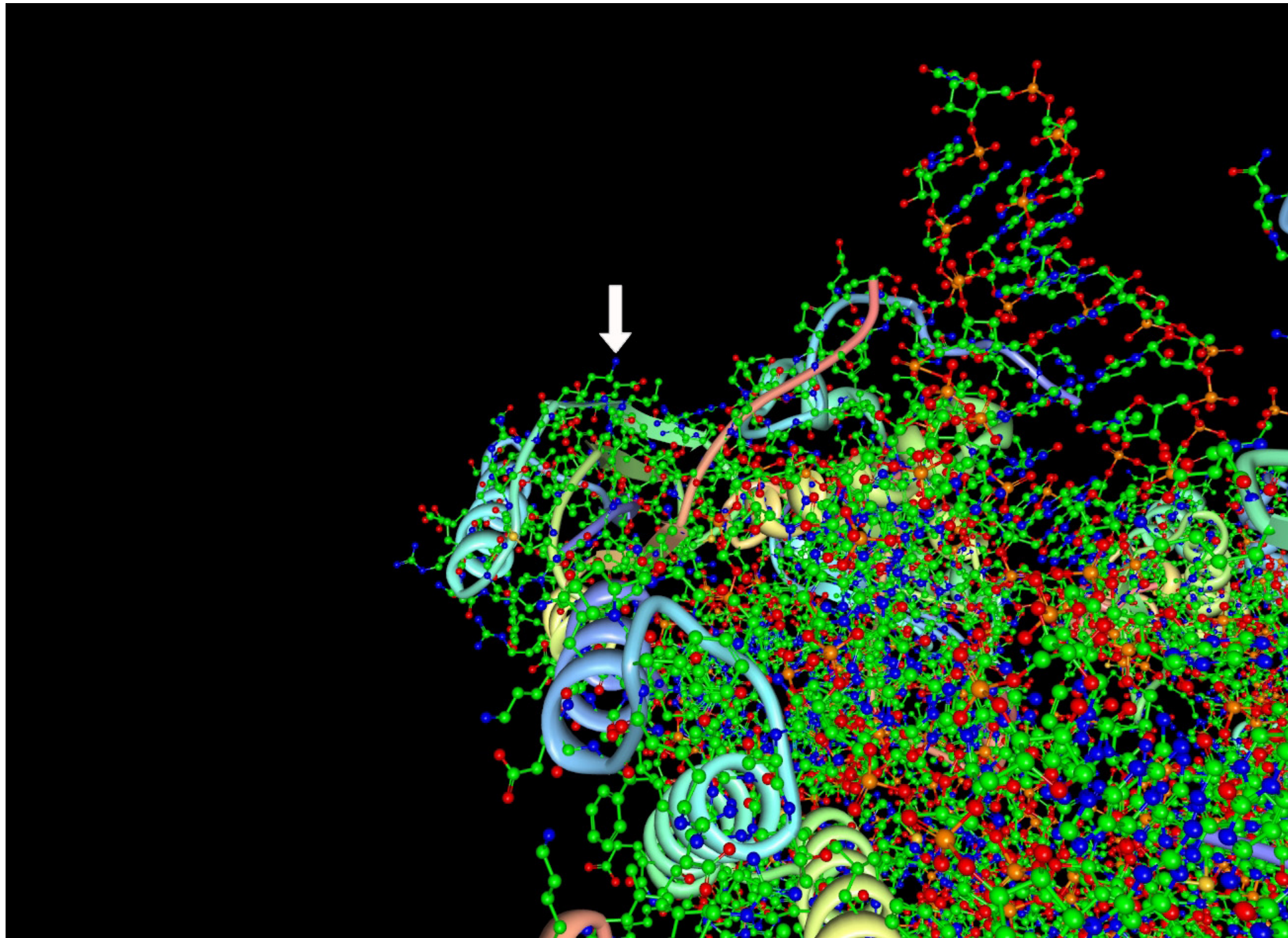
It is well known that proteins are dynamic and that a crystal structure is simply a snapshot of one of the many conformations that a protein can occupy. This being said, the location of the modified lysine may not always be as exposed as the following figures show, but it does indicate that at certain times the following proteins occupy a conformation allowing for access to the target lysine.

All screenshots taken from Protein Workshop, which is available from the PDB website.

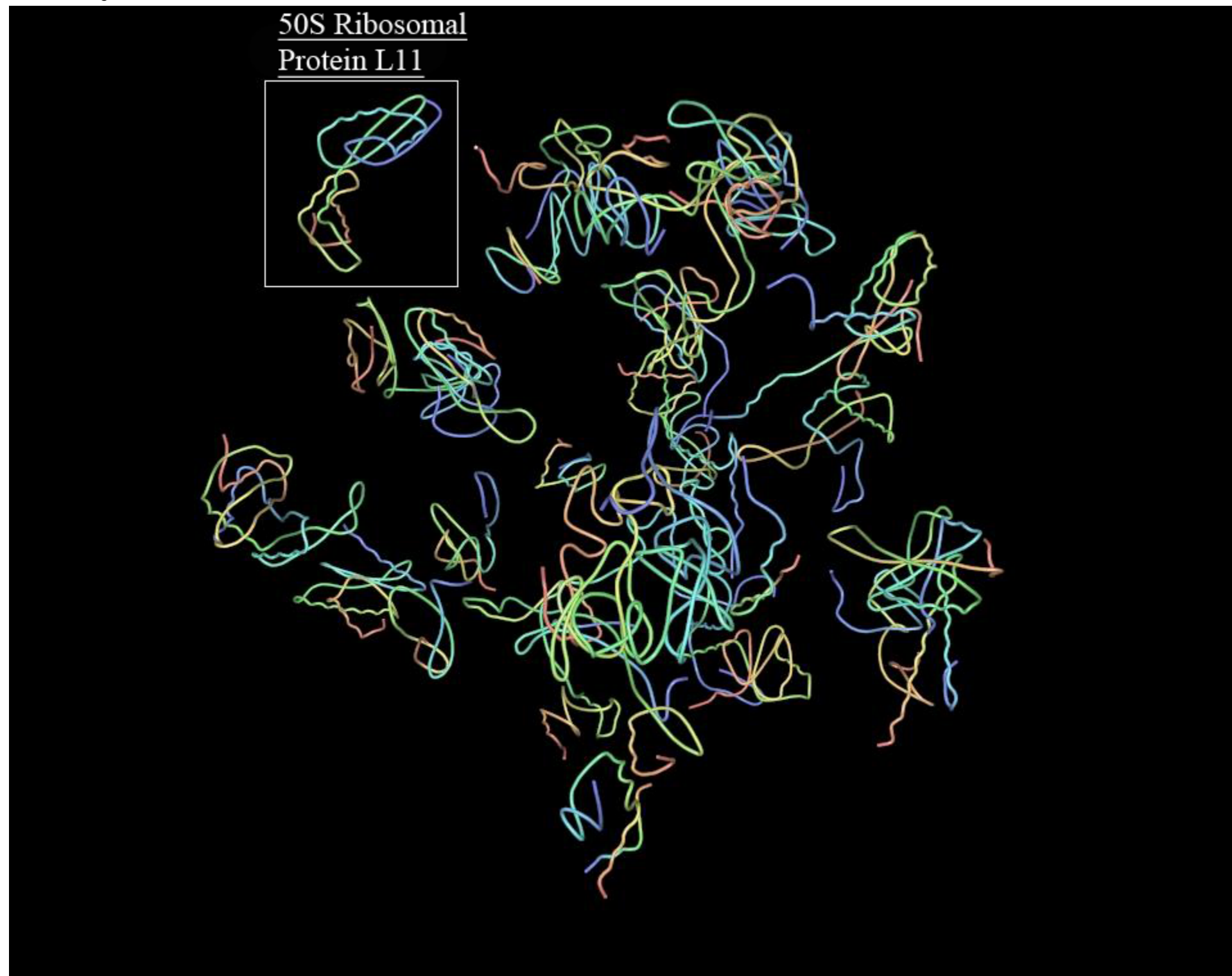
SF103: Global view of 30S ribosome from *E. coli* (PDB: 2GYB) showing location of PUPylated subunits



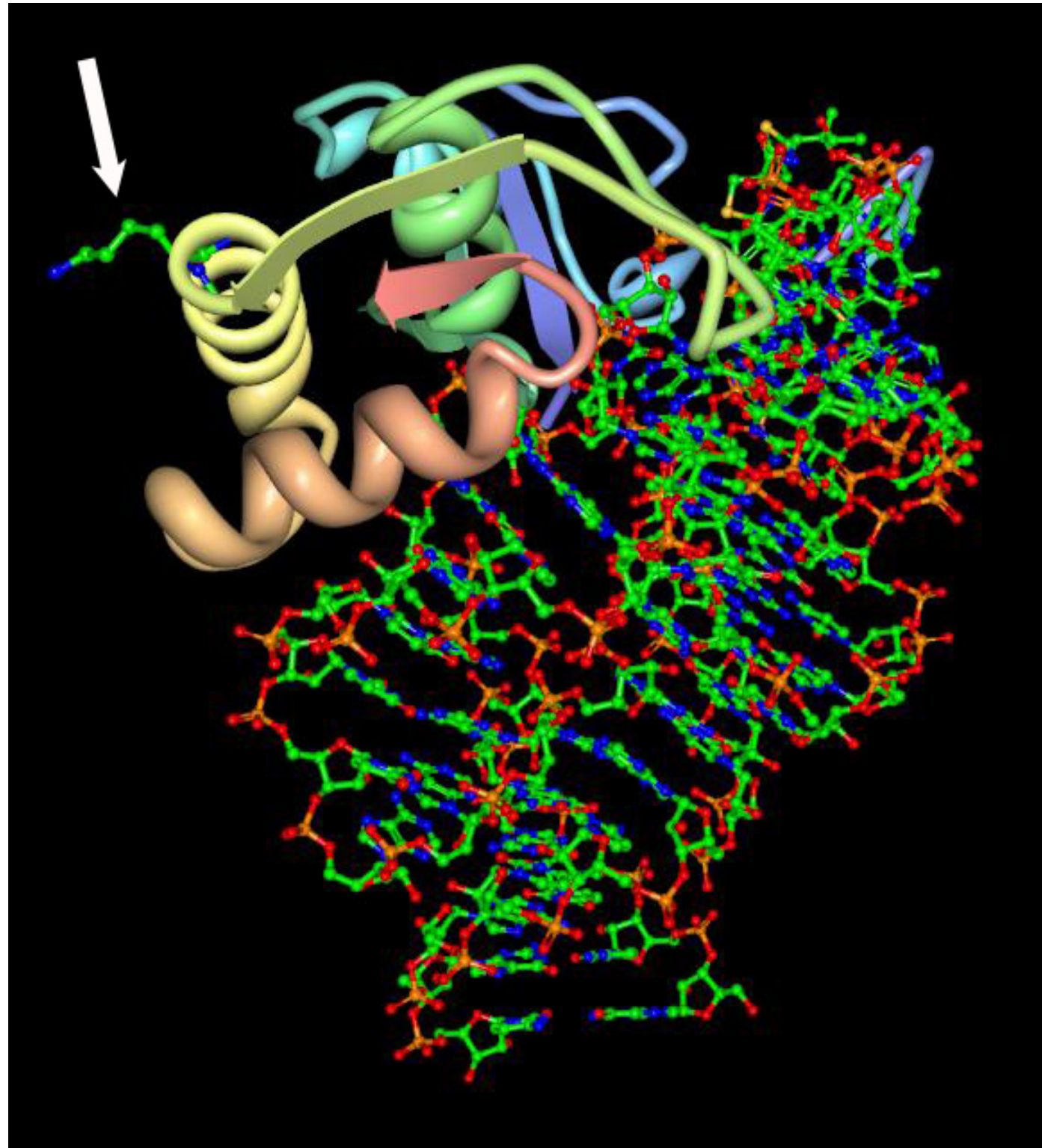
SF104: Homologous modification site on 30S ribosomal subunit S6 from *E. coli* (PDB: 2GYB)



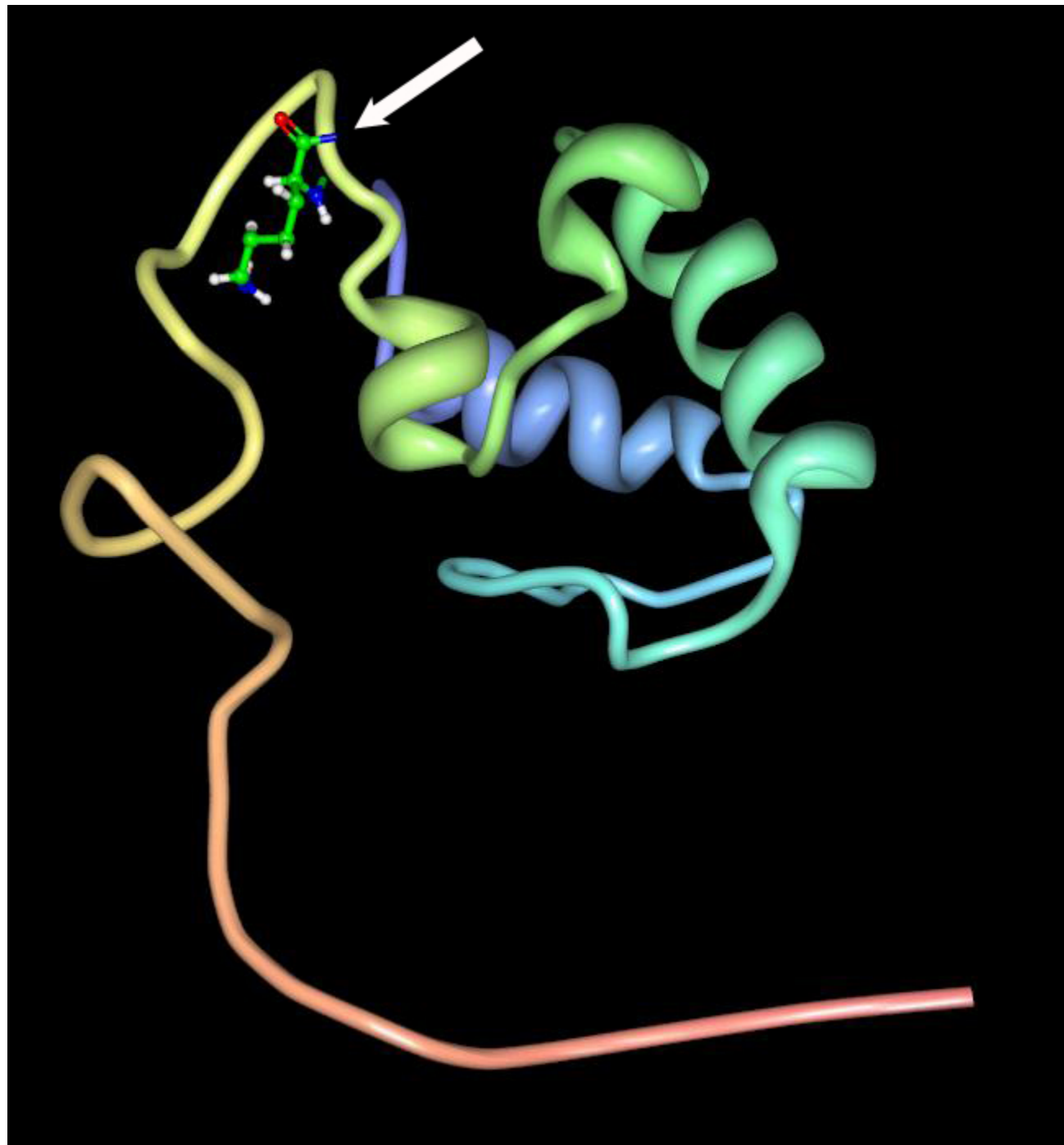
SF105: Global view of 50S ribosome from *E. coli* (PDB: 2RDO) showing location of PUPylated subunit



SF106: Homologous modification site on 50S ribosomal subunit L11 from *Streptomyces* (PDB: 1OLN)



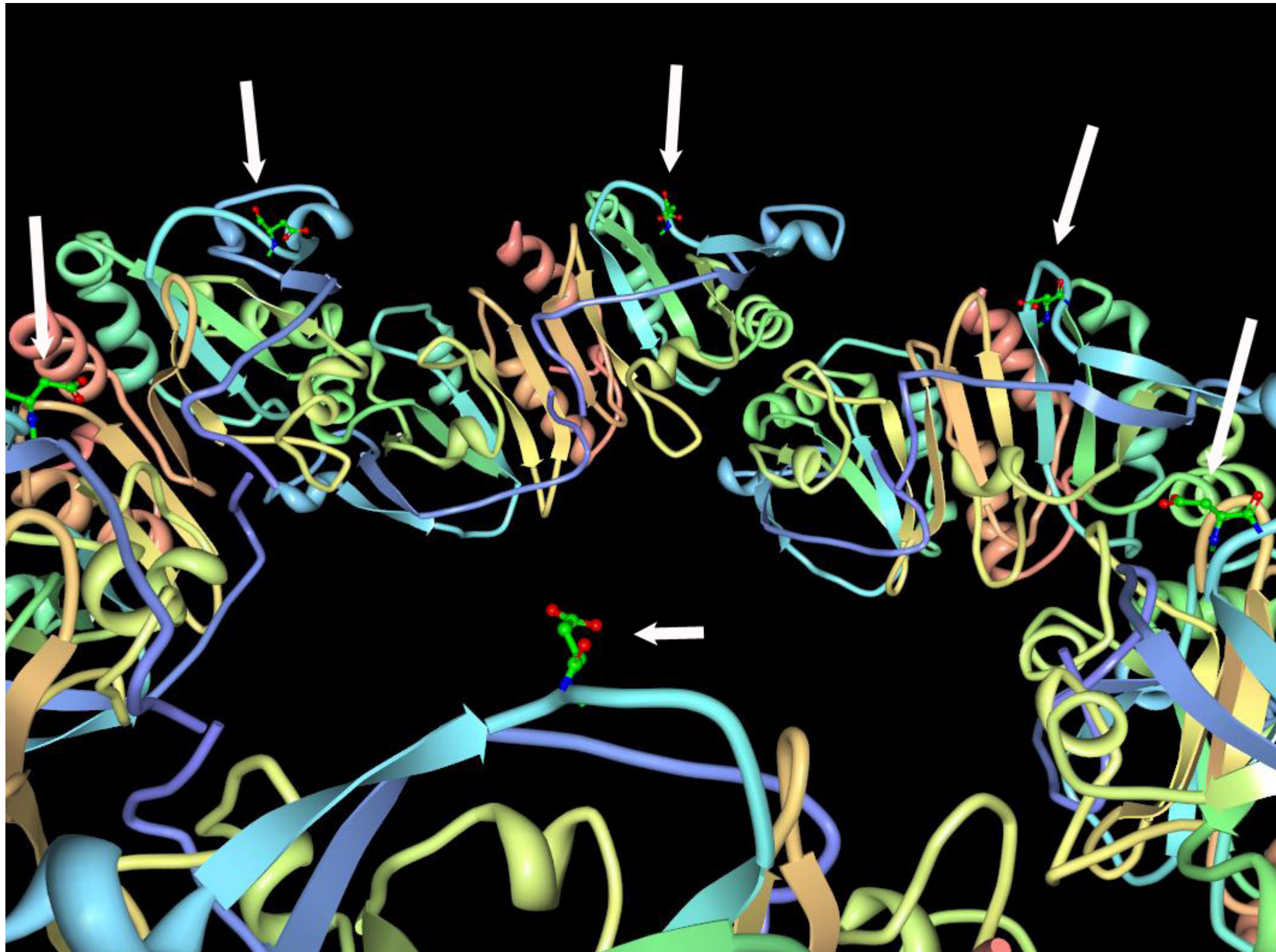
SF107: Homologous modification site on acyl carrier protein (ACP) from *Mycobacterium tuberculosis* (PDB: 1KLP)



SF108: Homologous modification site on adenylylsuccinate synthetase from *E.coli* (PDB: 1QF4)



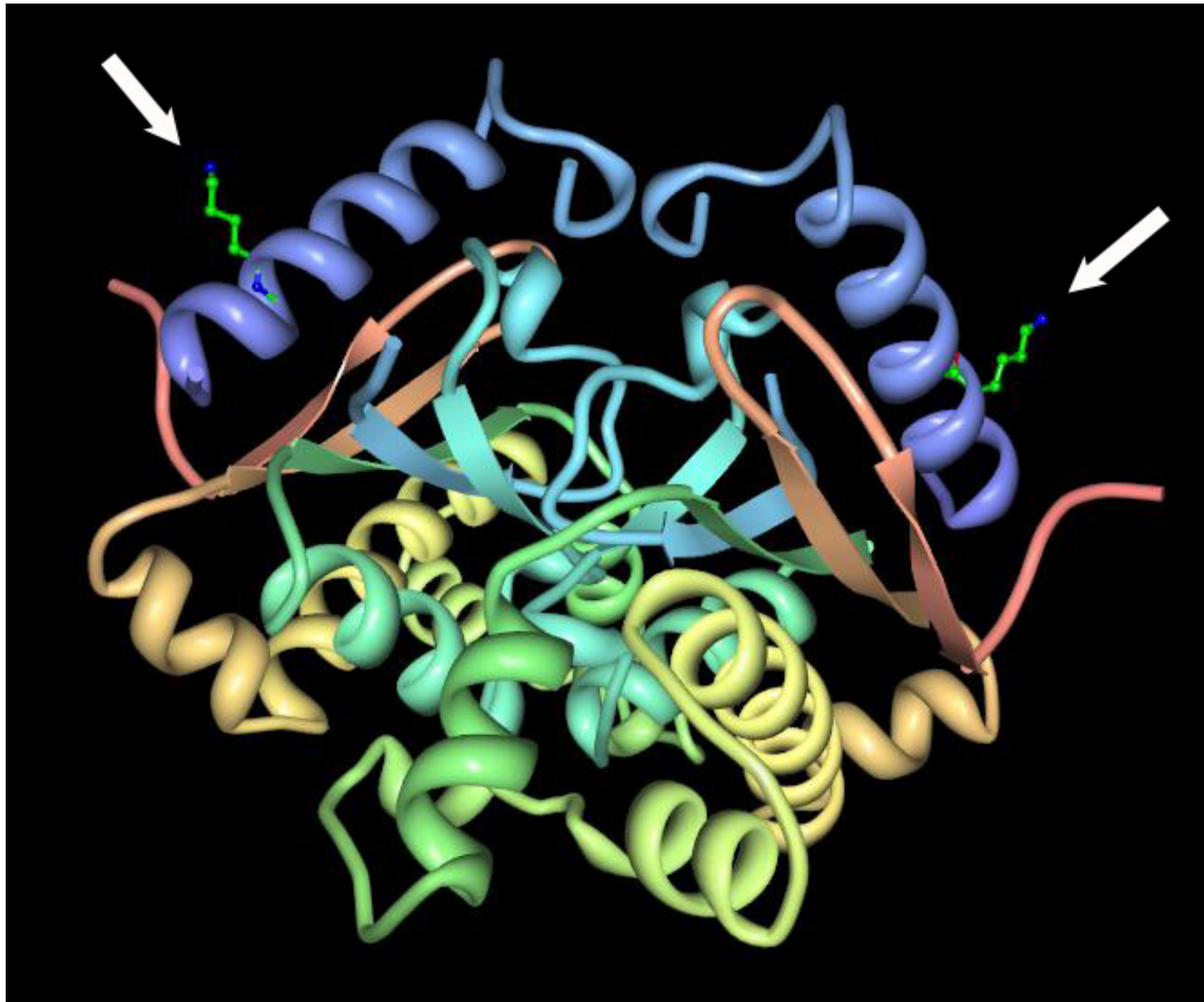
SF109: Homologous modification site on alkylperoxide reductase (hexamer) from *Mycobacterium tuberculosis* (PDB: 2BMX)



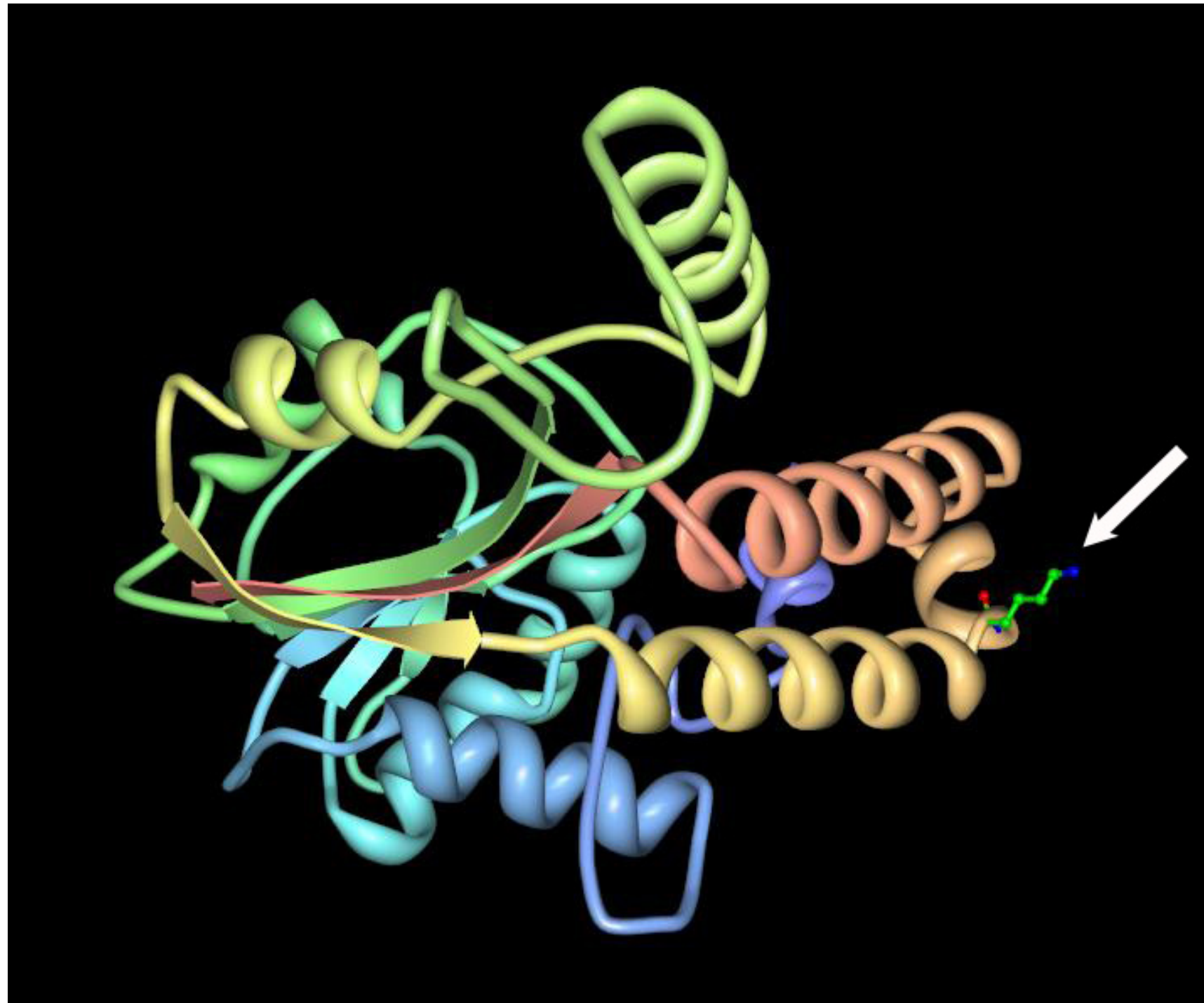
SF110: Homologous modification site on aspartate transaminase from *Corneibacterium* (PDB: 3D6K)



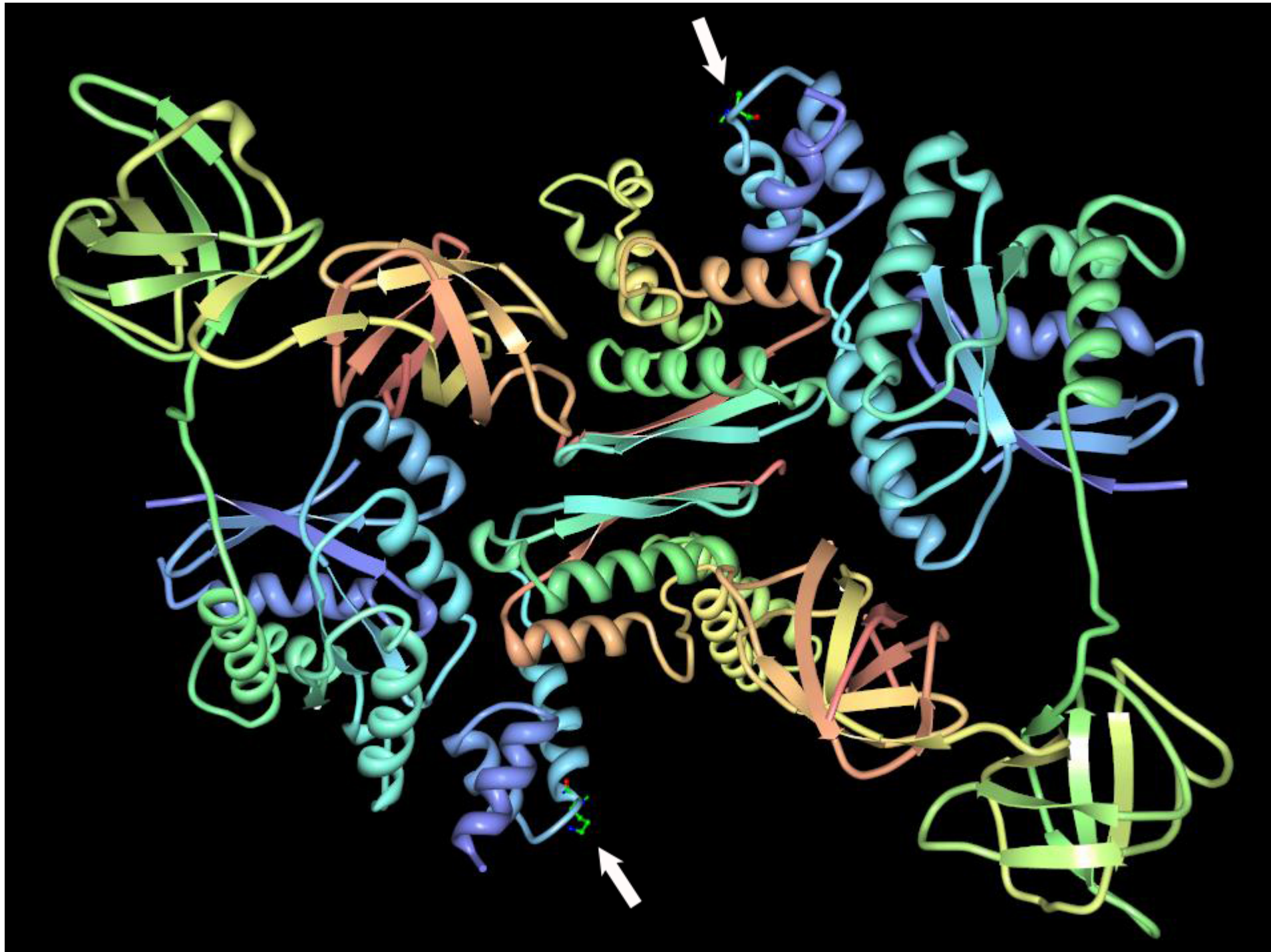
SF111: Homologous modification site on carbonic anhydrase (dimmer) from *Mycobacterium tuberculosis* (PDB: 1YM3)



SF112: Homologous modification site on cyclopropane-fatty-acyl-phospholipid synthase from *Mycobacterium tuberculosis* (PDB: 1KP9)



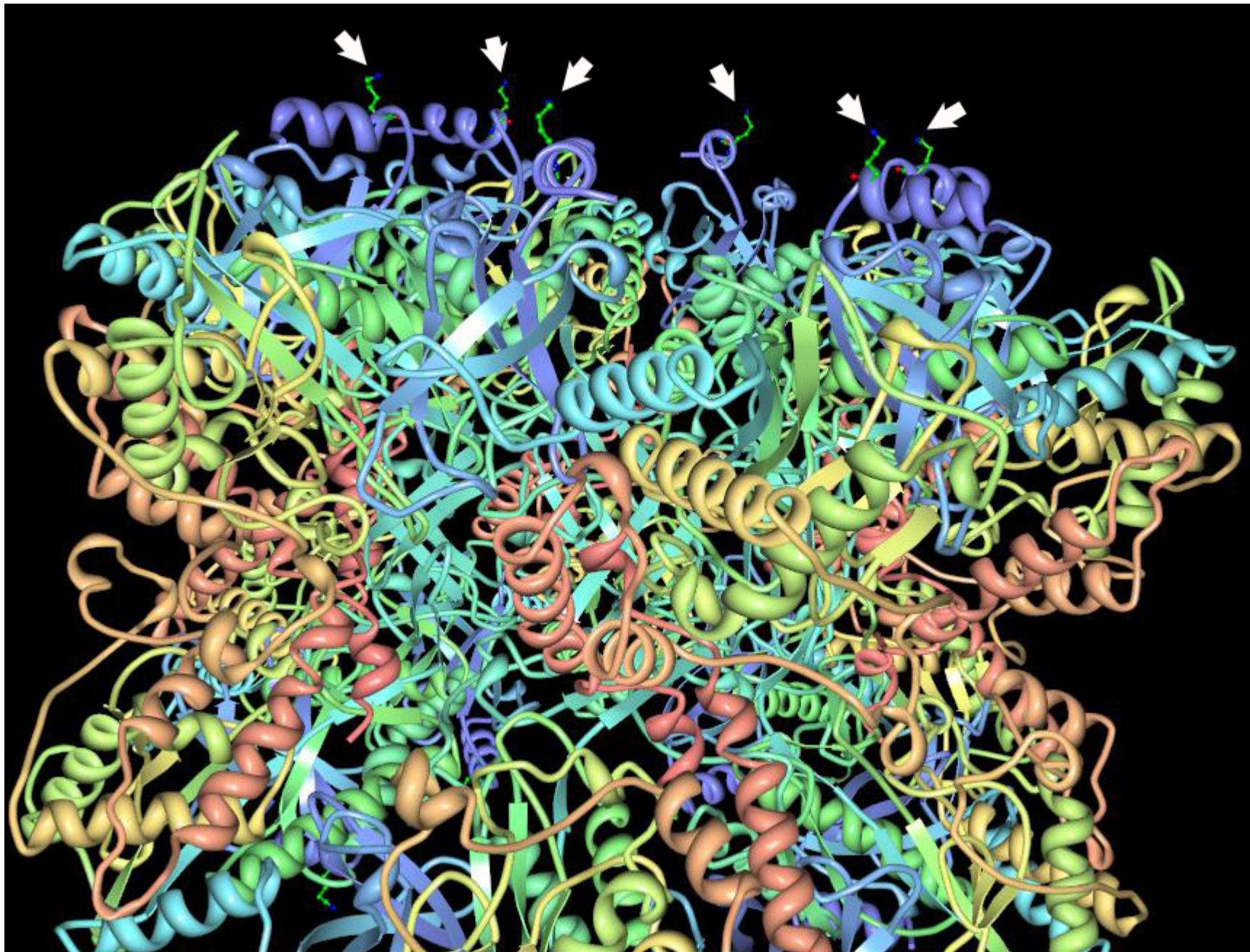
SF113: Homologous modification site on elongation factor Ts (dimer) from *Thermophilus* (PDB: 1AIP)



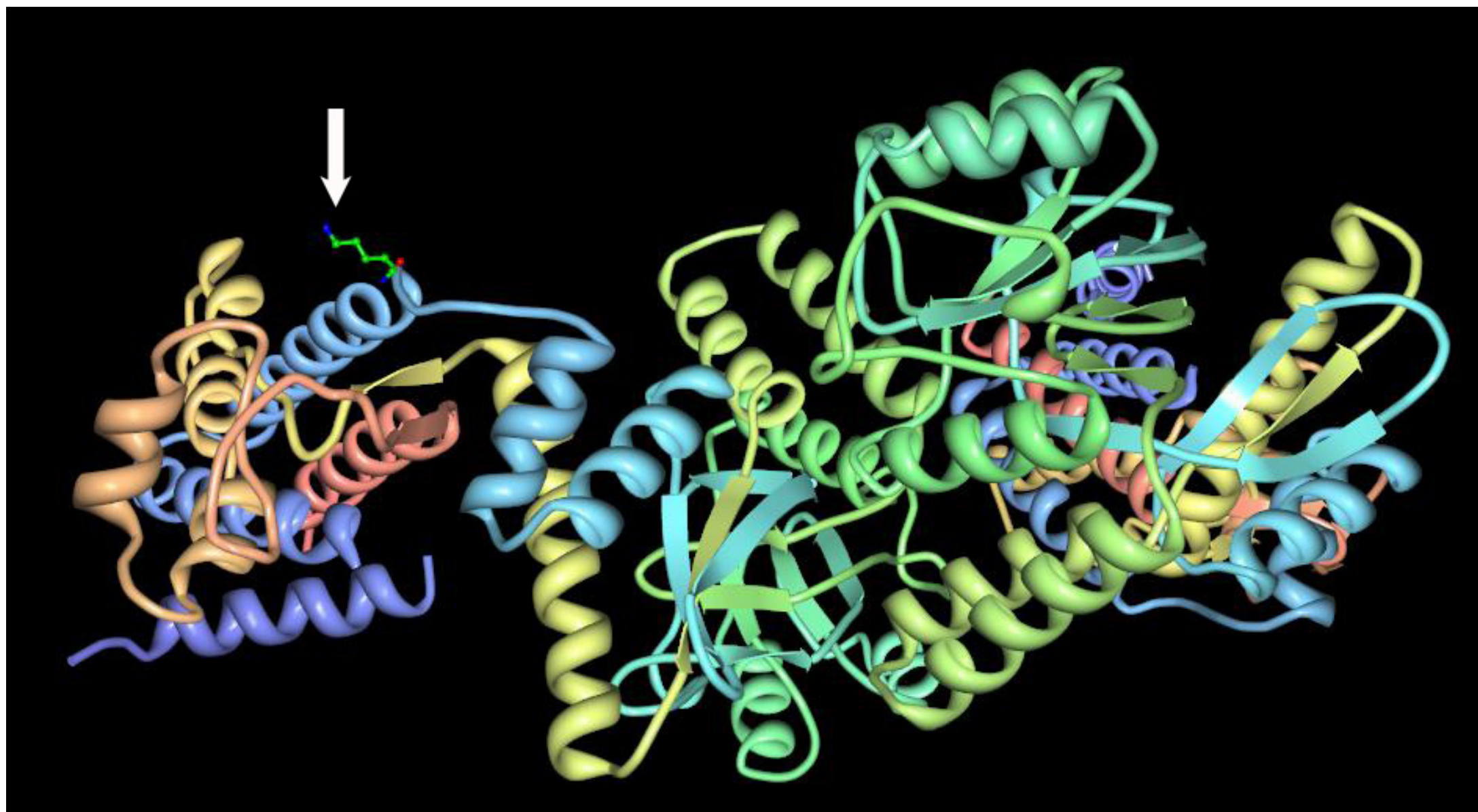
SF114: Homologous modification site on glyceraldehyde-3-phosphate dehydrogenase, type I, from *Bacillus* (PDB: 2GD1)



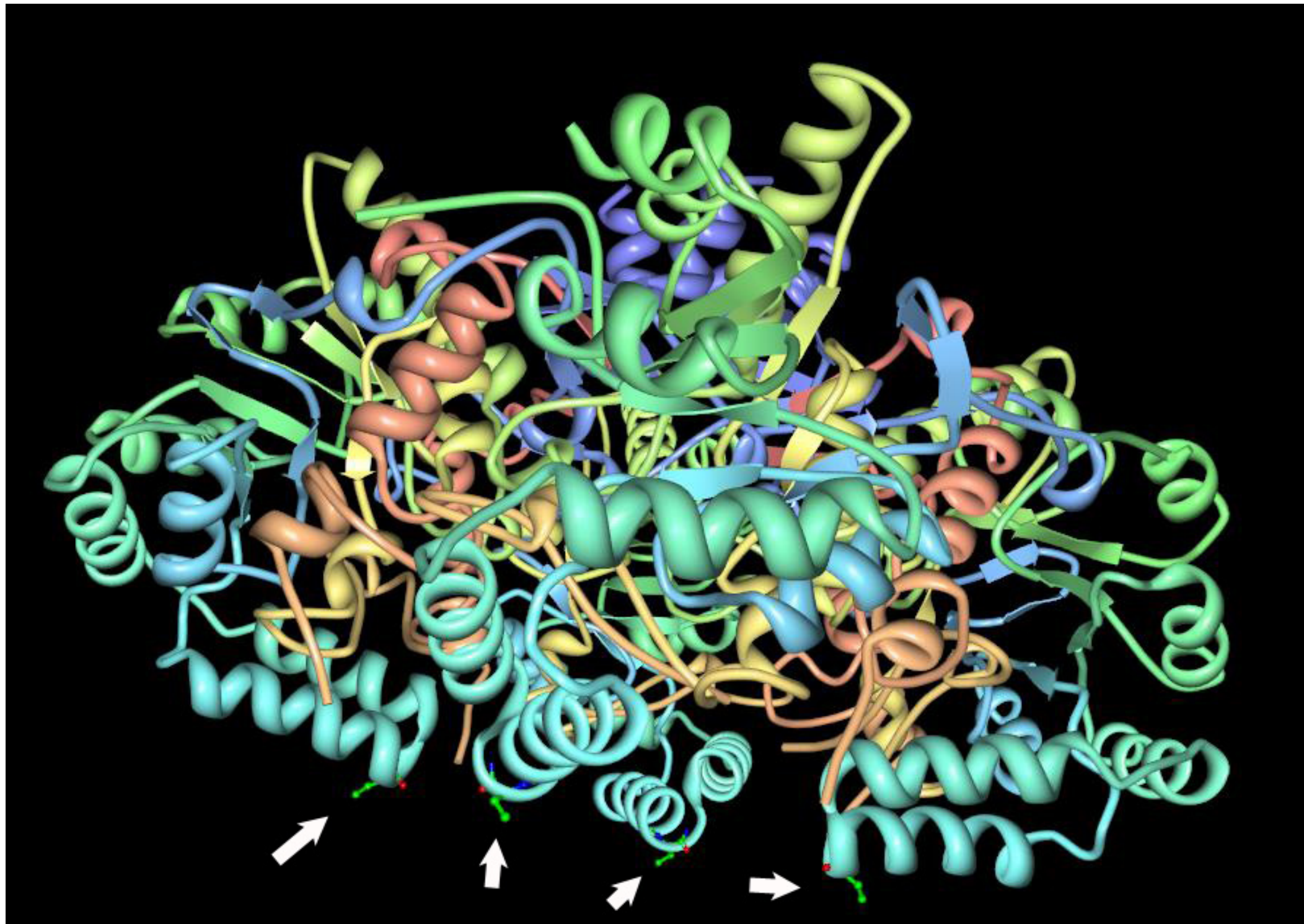
SF115: Homologous modification site on glutamine synthetase, type 1, from *Mycobacterium tuberculosis* (PDB: 2BVC)



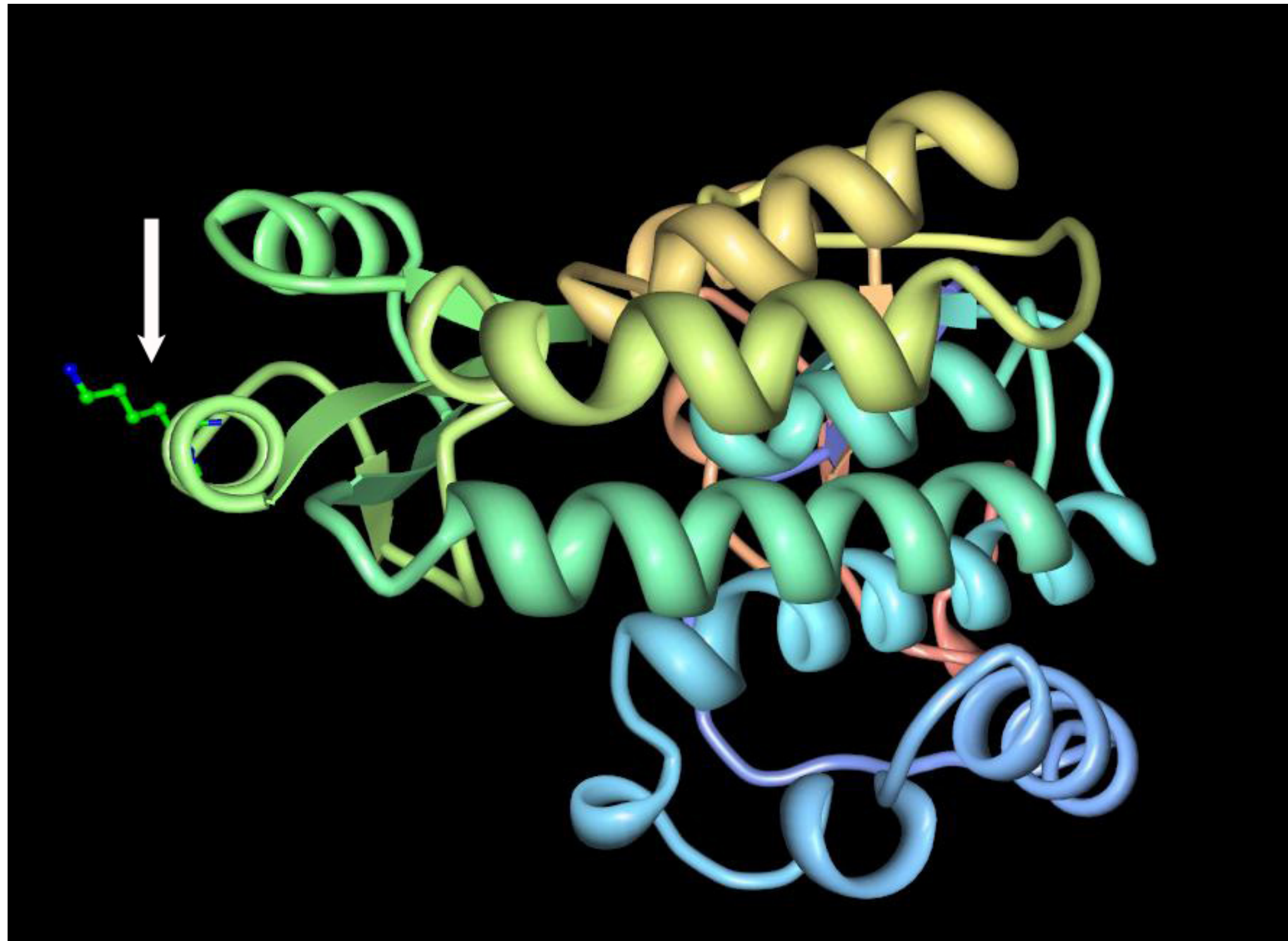
**SF116: Homologous modification site on GroEL from *Mycobacterium tuberculosis*
(PDB: 1SJP)**



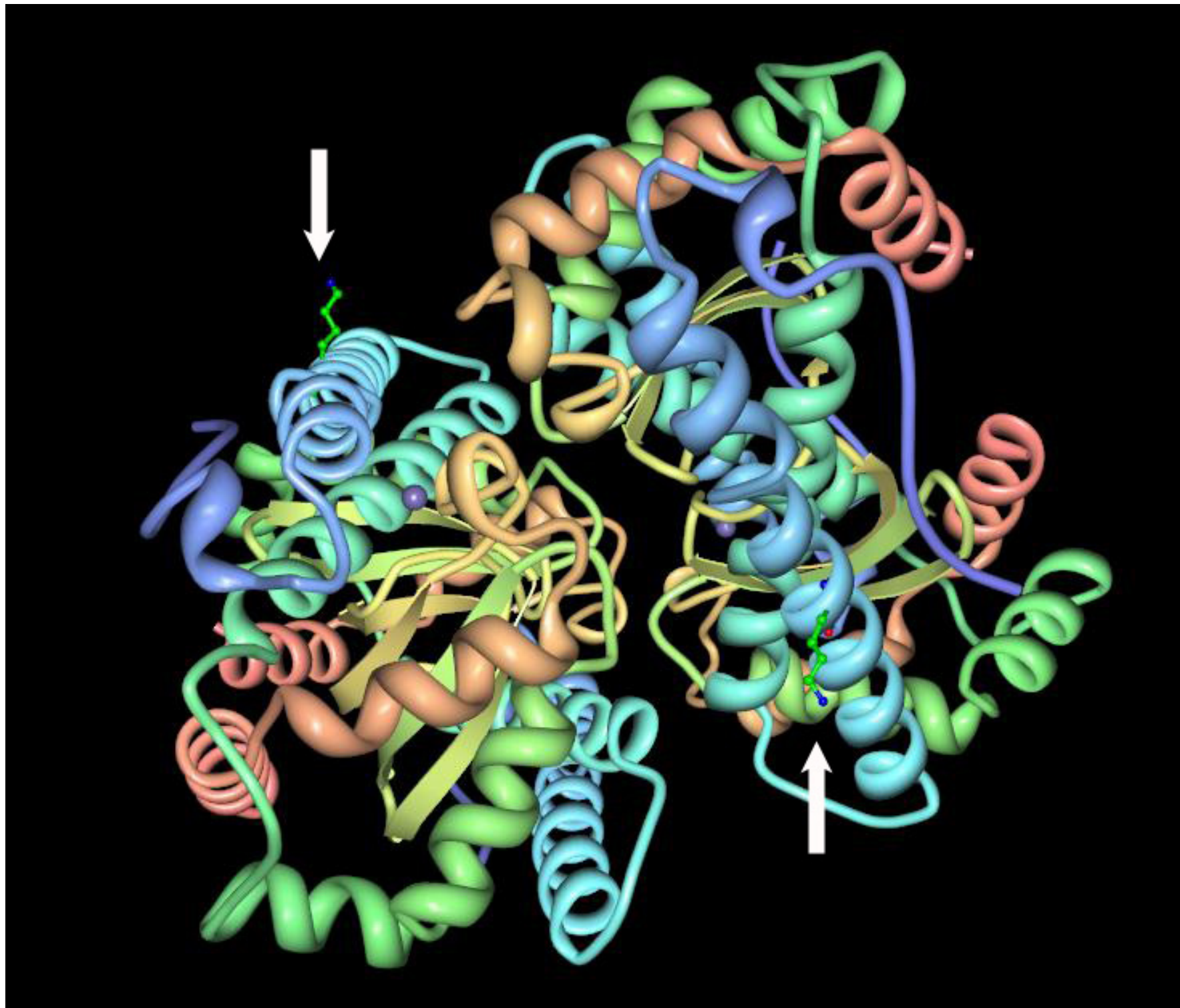
SF117: Homologous modification site on inositol-5-monophosphate dehydrogenase (tetramer) from *Cornaybacterium* (PDB: 2QR6)



SF118: Homologous modification site on malonyl CoA-acyl carrier protein transacylase from *Mycobacterium tuberculosis* (PDB: 2QC3)



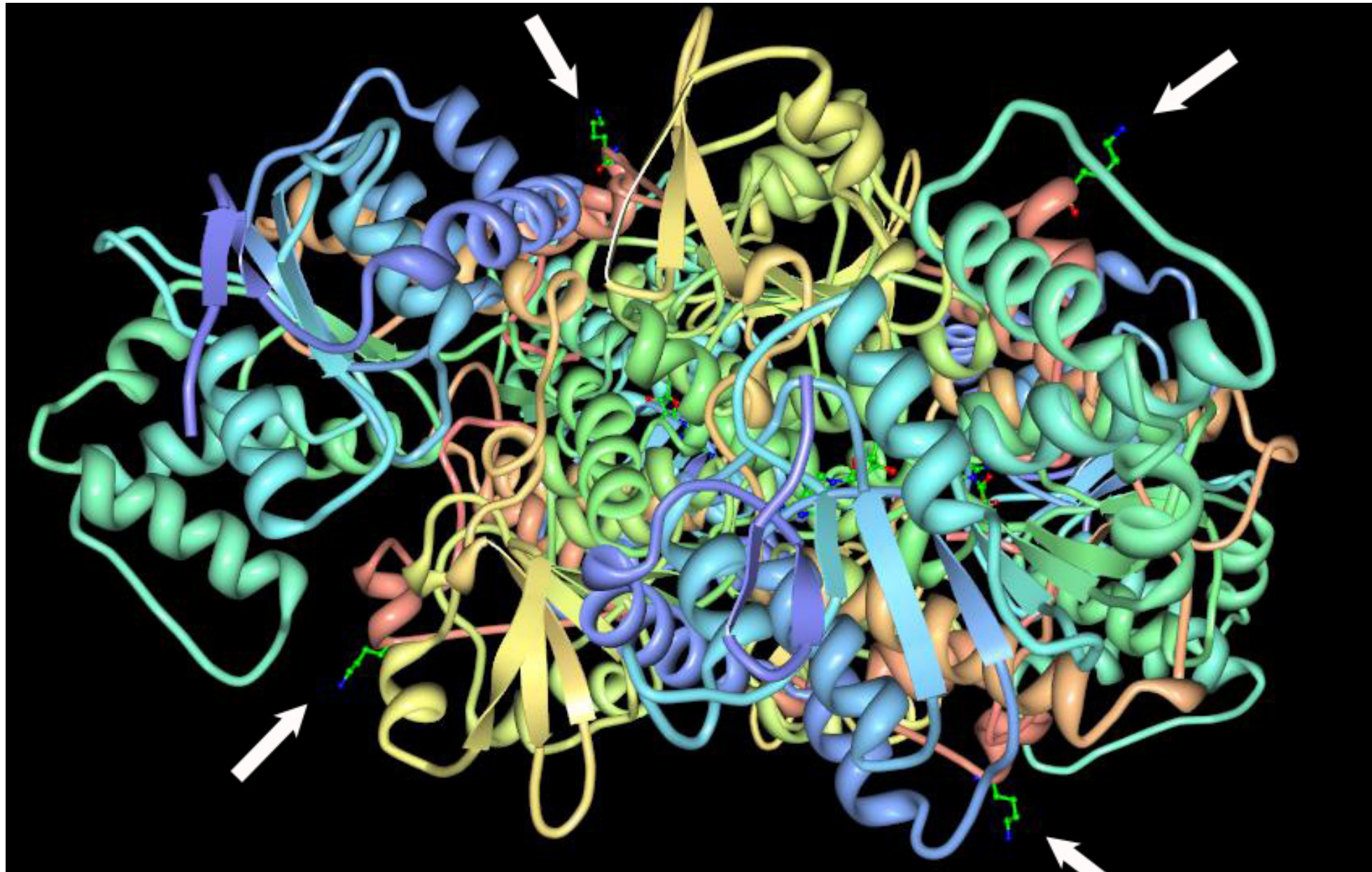
SF119: Homologous modification site on [Mn] superoxide dismutase from *Mycobacterium tuberculosis* (PDB: 1GN3)



SF120: Homologous modification site on myo-inositol-1-phosphate synthase from *Mycobacterium tuberculosis* (PDB: 1GR0)



SF121: Homologous modification site on putative adenosylhomocysteinase (tetramer) from *Mycobacterium tuberculosis* (PDB: 3CE6)



SF122: Homologous modification site on threonine synthase from *Mycobacterium tuberculosis* (PDB: 2D1F)



SF123: Homologous modification site on universal stress protein (dimer) from *Mycobacterium tuberculosis* (PDB: 2JAX)



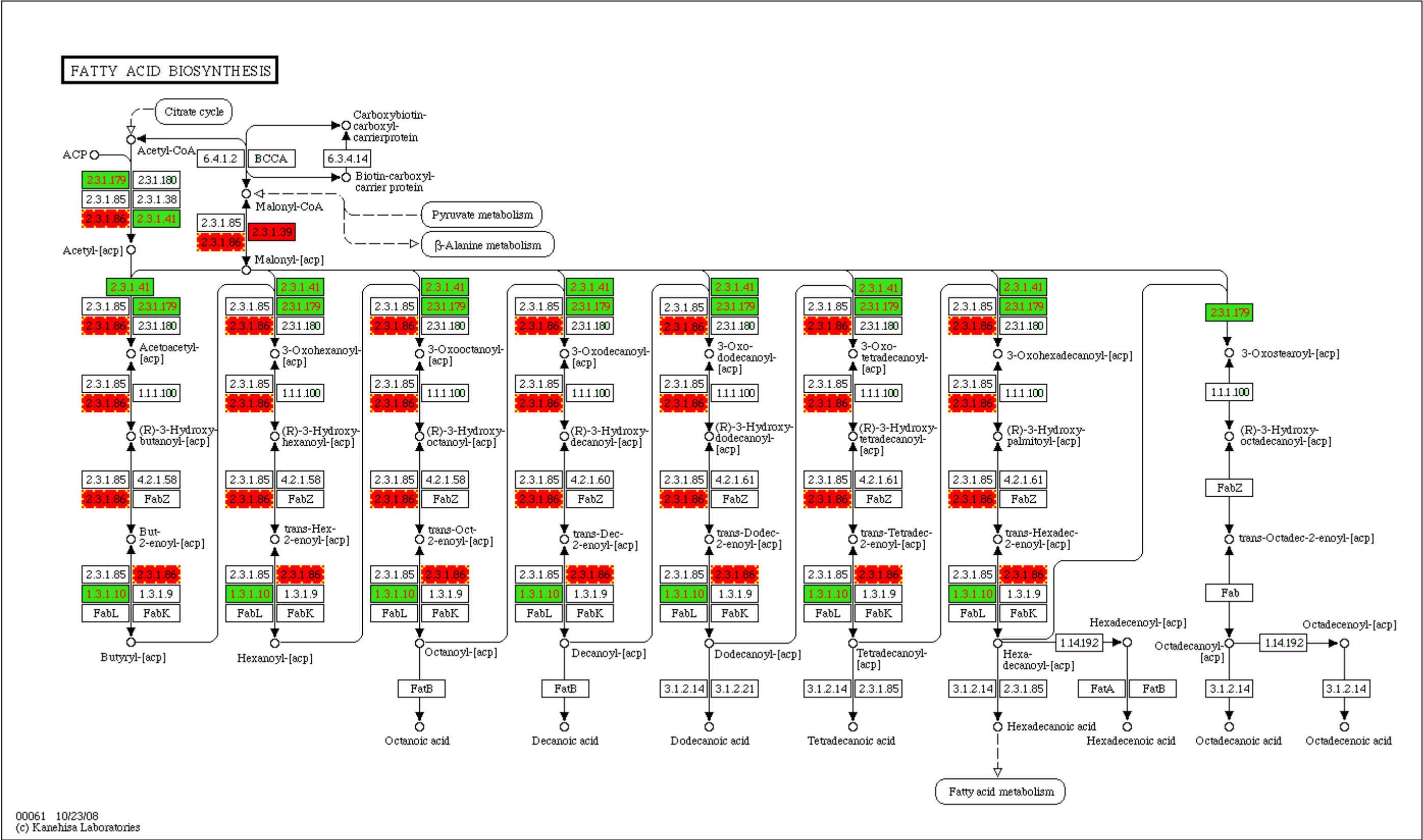
VI. Pathway Annotations

a. Figures SF124-SF136: Annotated pathway maps showing location of PUPylated and candidate proteins in pathway.....143

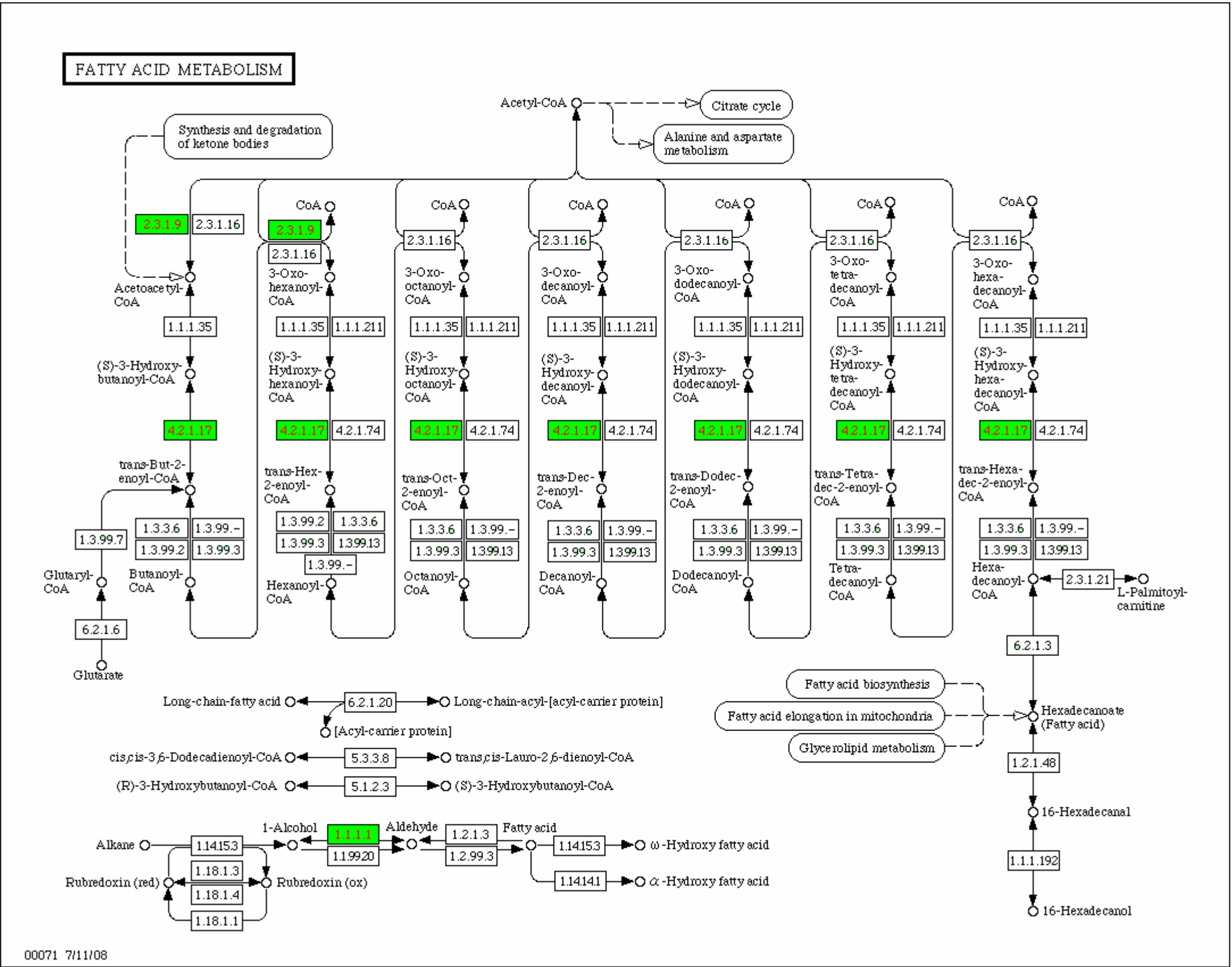
Enzymes were found for 103 PUPylated and candidate proteins, which were then plotted on pathway maps from *Mycobacterium smegmatis* strain MC²-155 using the Kegg Kyoto database (www.genome.jp/kegg/). The resulting figures show possible regulation by PUP in many metabolic pathways such as fatty acid biosynthesis, amino acid biosynthesis and glycolysis.

Enzymes in the following figures were color coded as follows: **red** indicates a confirmed PUPylated protein, **green** indicates a candidate PUPylated protein, **blue** indicates a candidate protein complex and any enzyme with a **red** / **yellow** dashed border indicates that it is one of the common nitrosylated proteins (See section II).

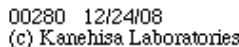
SF124: Presence of PUPylated proteins in fatty acid biosynthesis



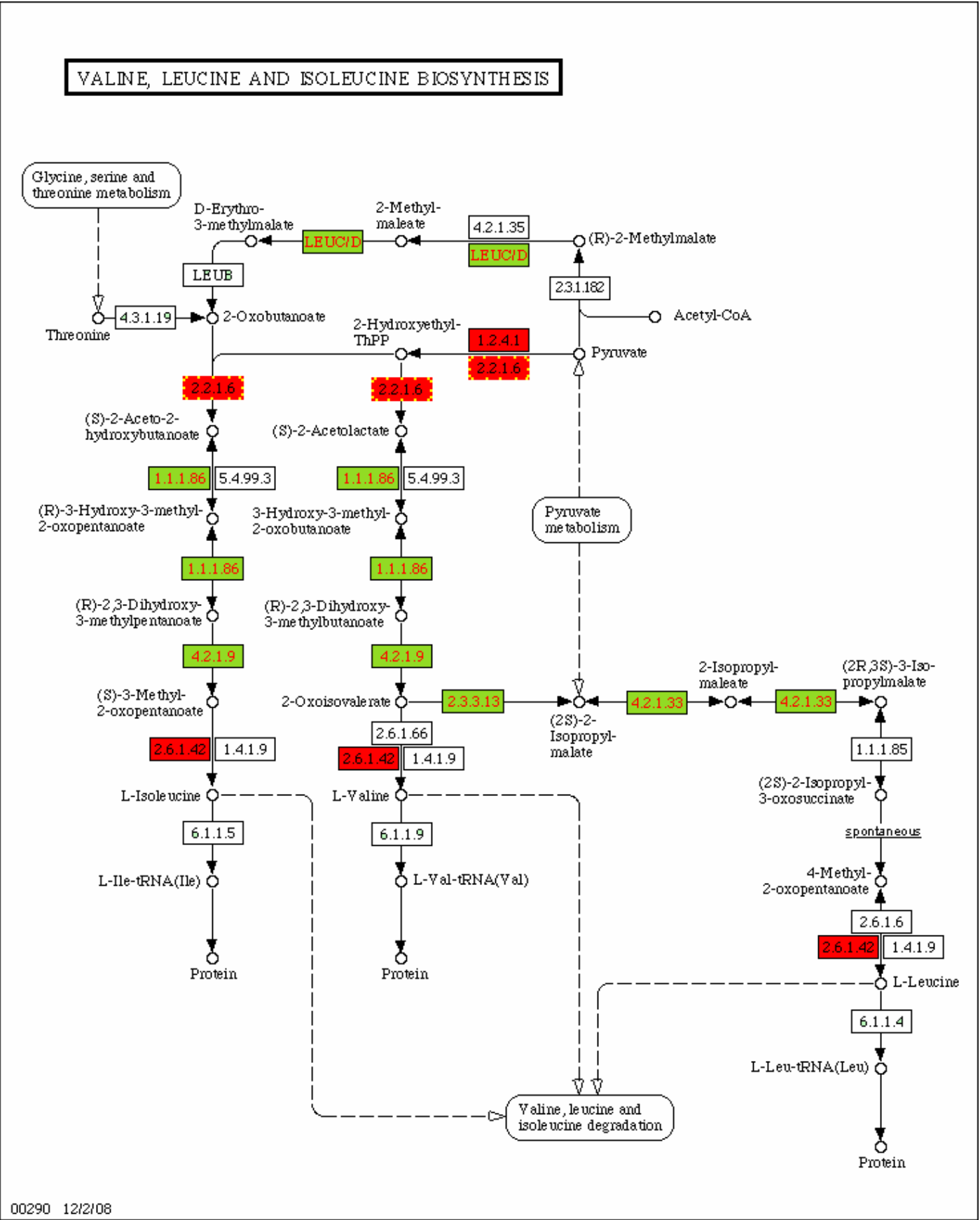
SF125: Presence of PUPylated proteins in fatty acid metabolism



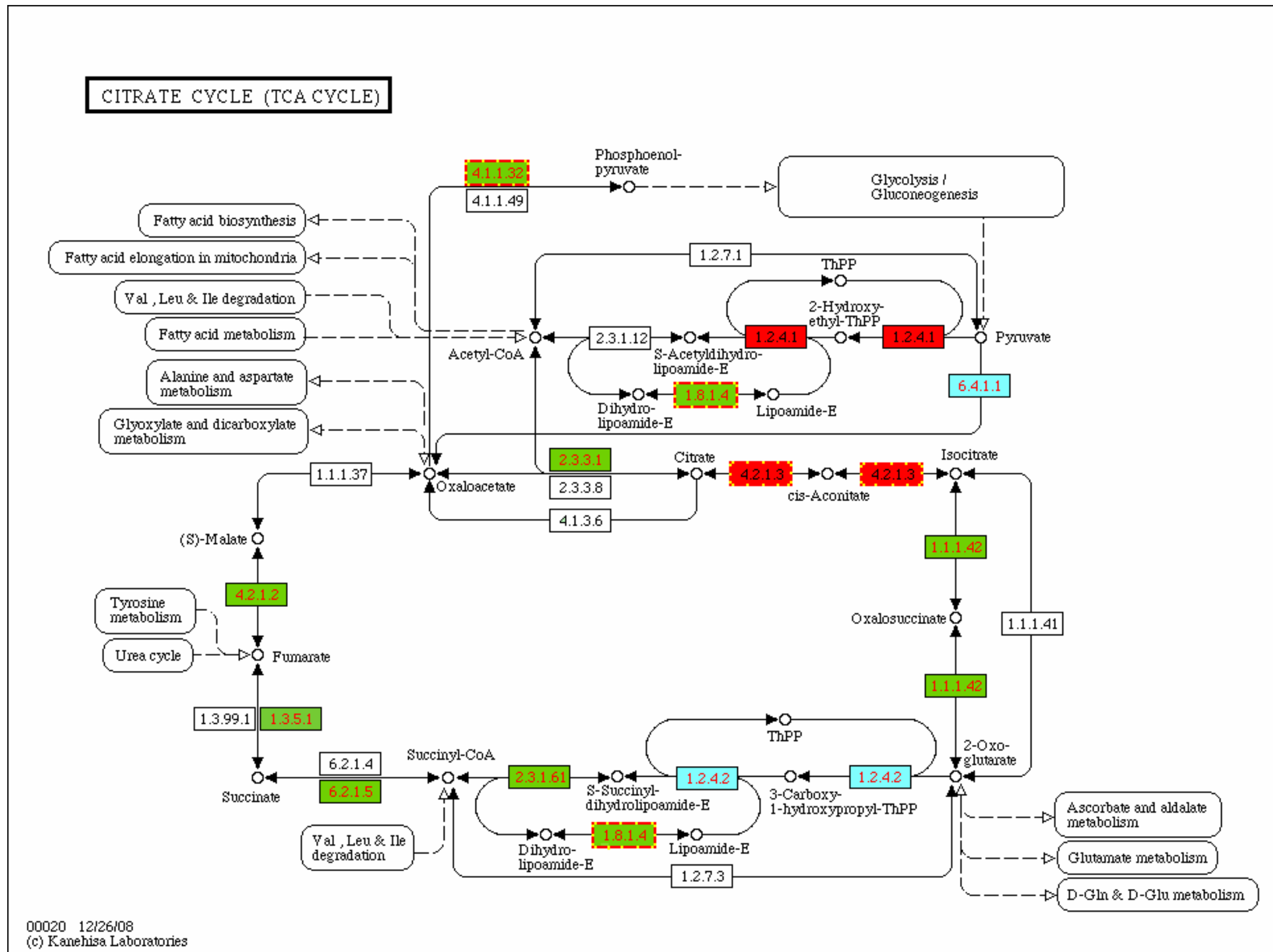
SF126: Presence of PUPylated proteins in valine, leucine and isoleucine degradation



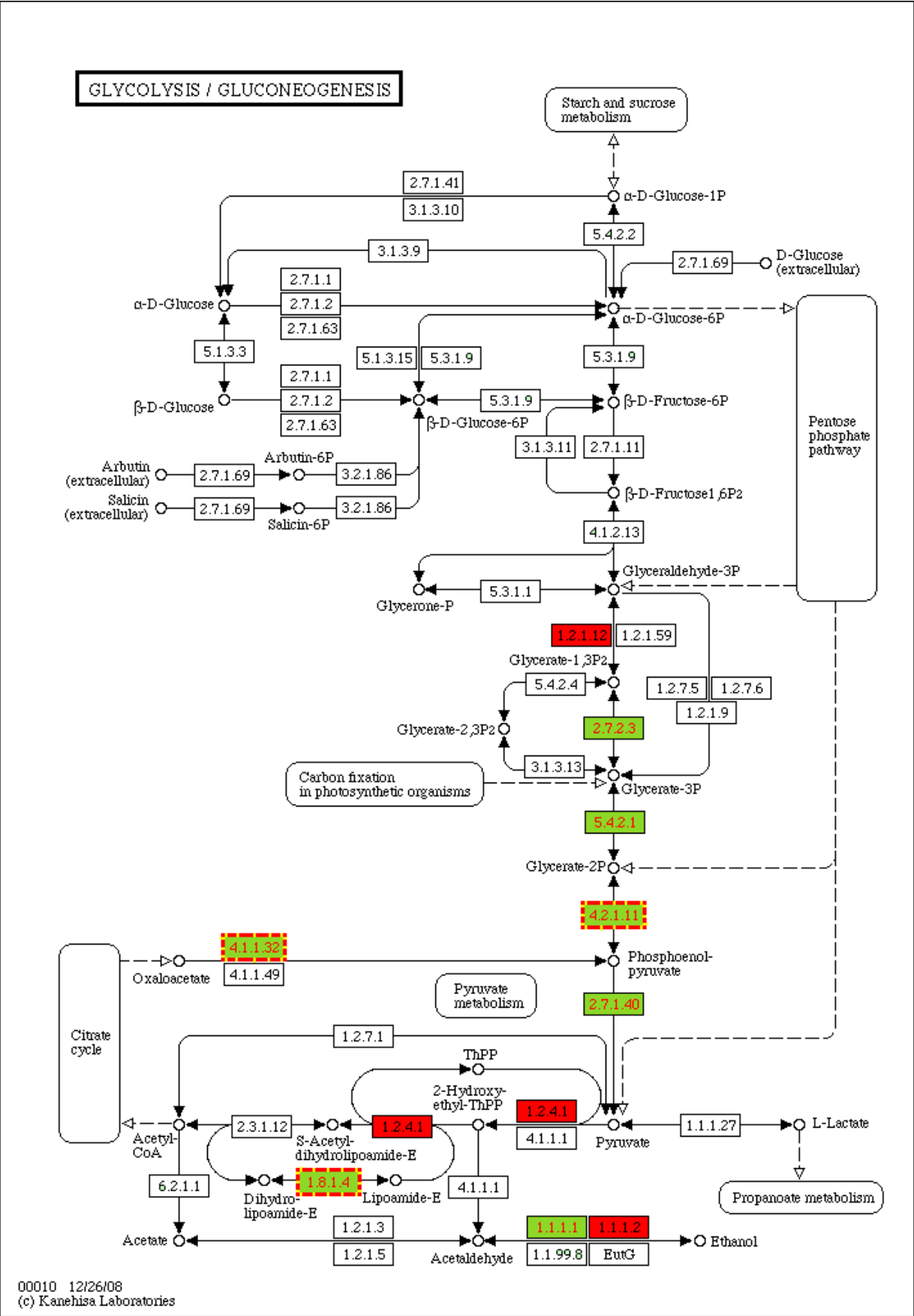
SF127: Presence of PUPylated proteins in valine, leucine and isoleucine biosynthesis



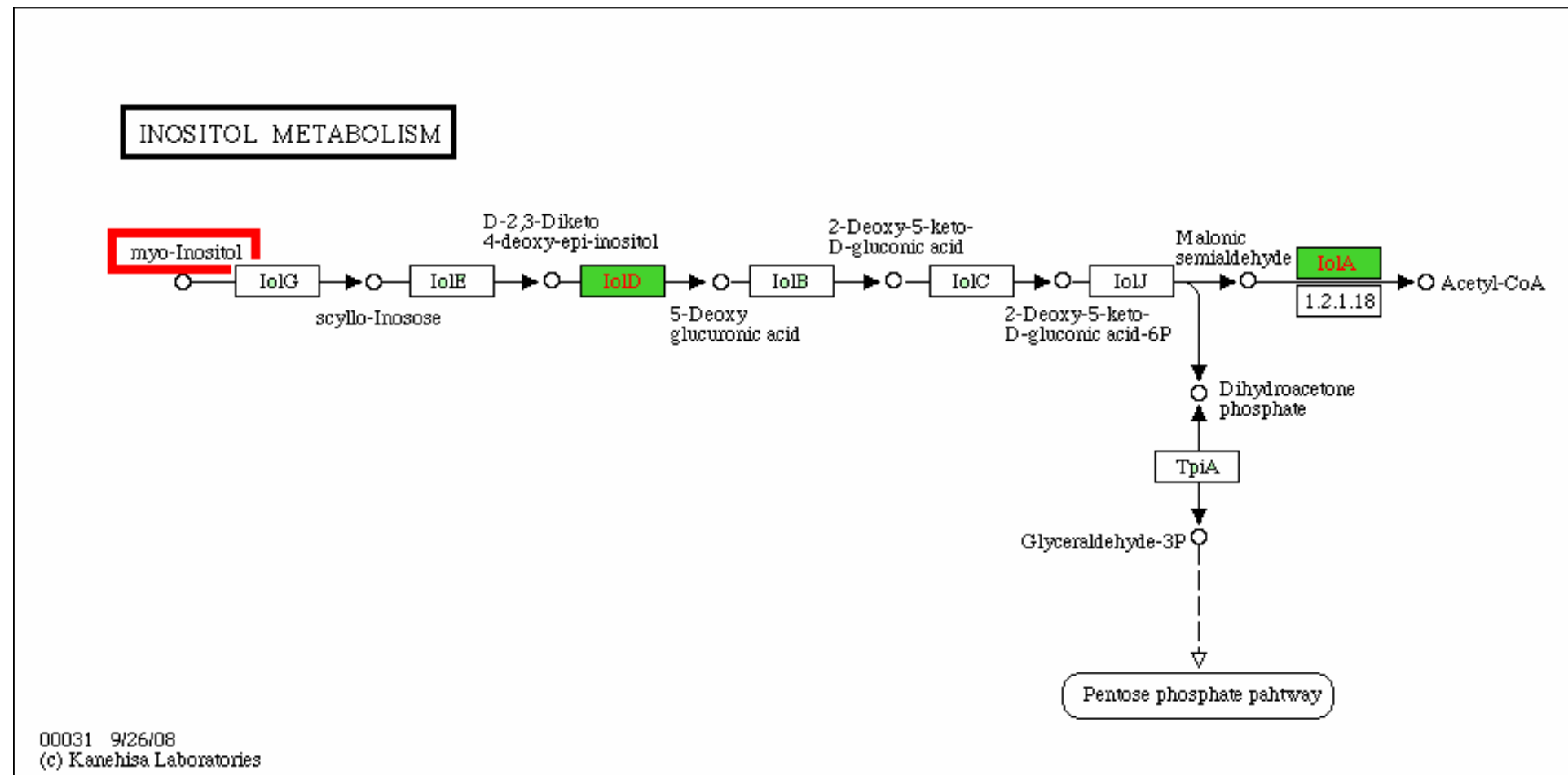
SF128: Presence of PUPylated proteins in the TCA cycle



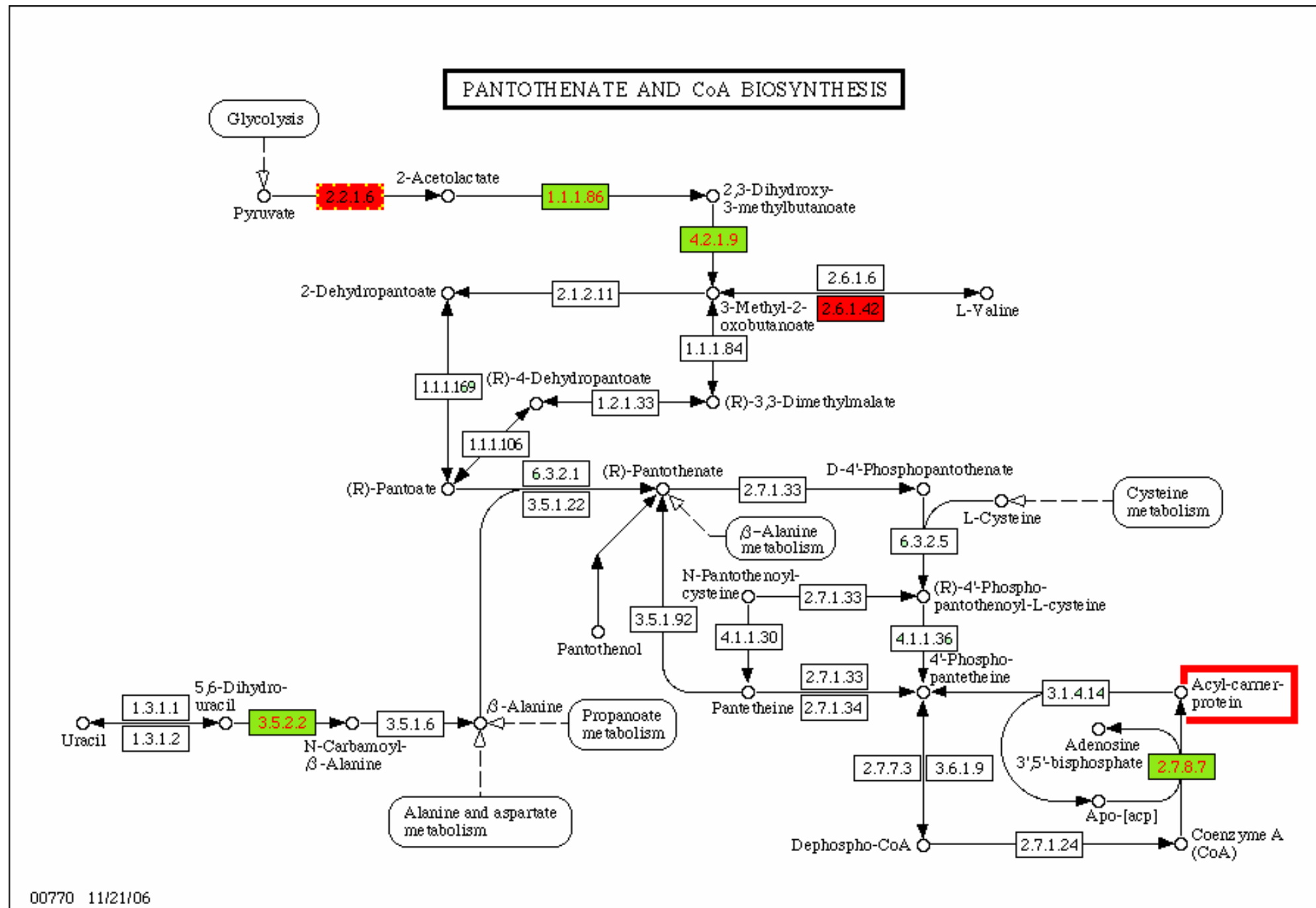
SF129: Presence of PUPylated proteins in glycolysis and gluconeogenesis



SF130: Presence of PUPylated proteins in inositol metabolism

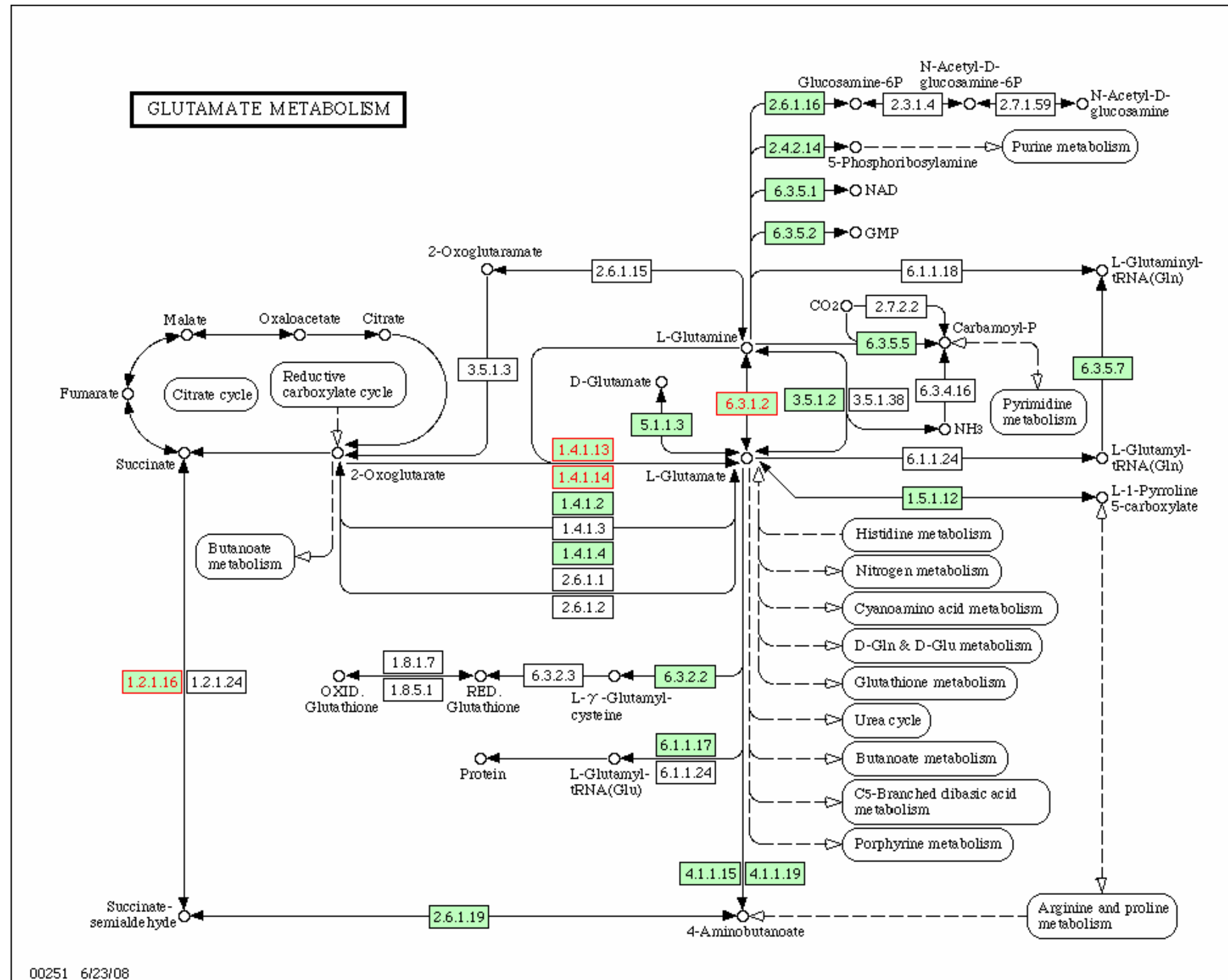


SF131: Presence of PUPylated proteins in pantothenate and coenzyme A biosynthesis

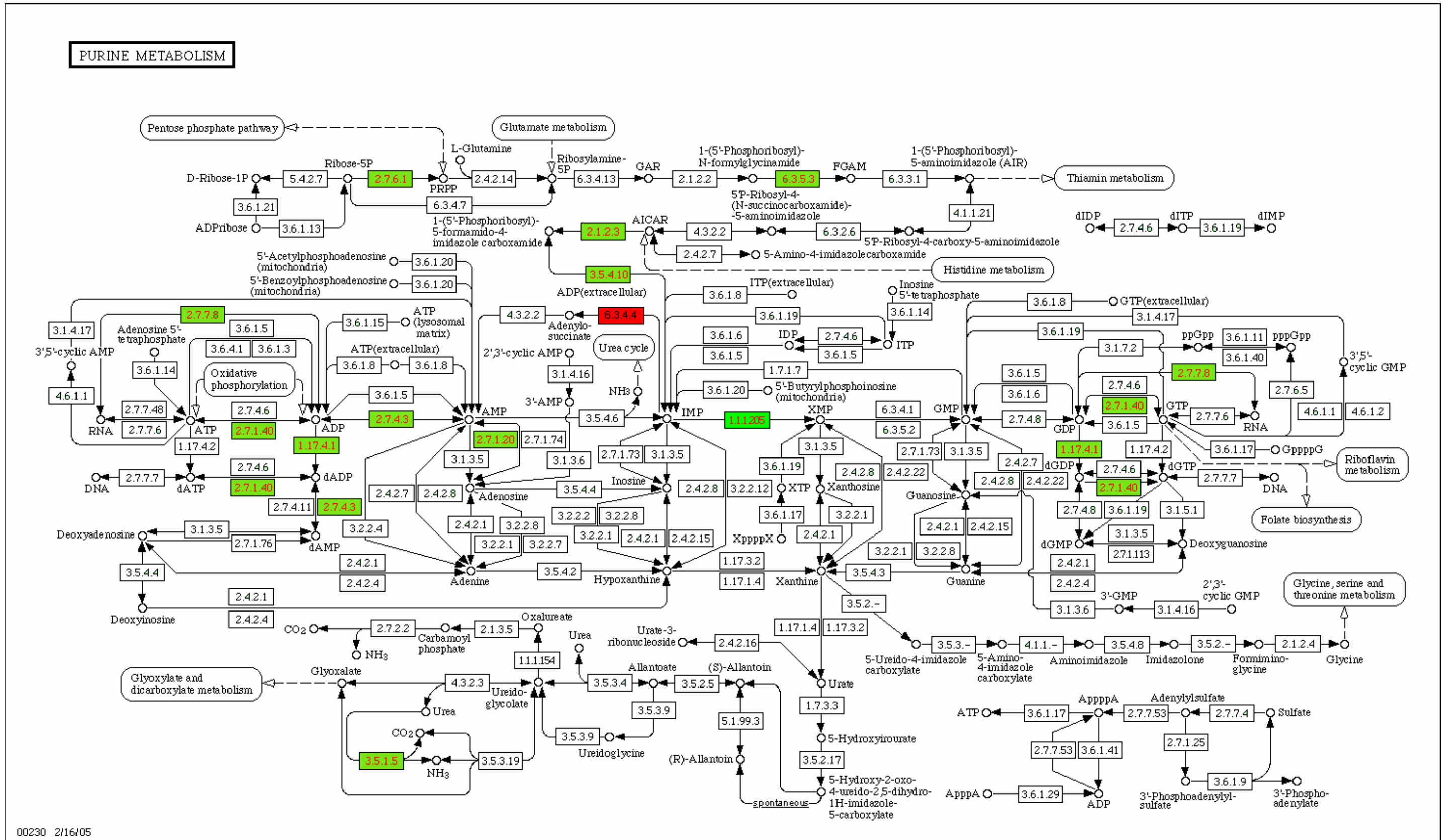


SF132: Presence of PUPylated proteins in glutamate metabolism

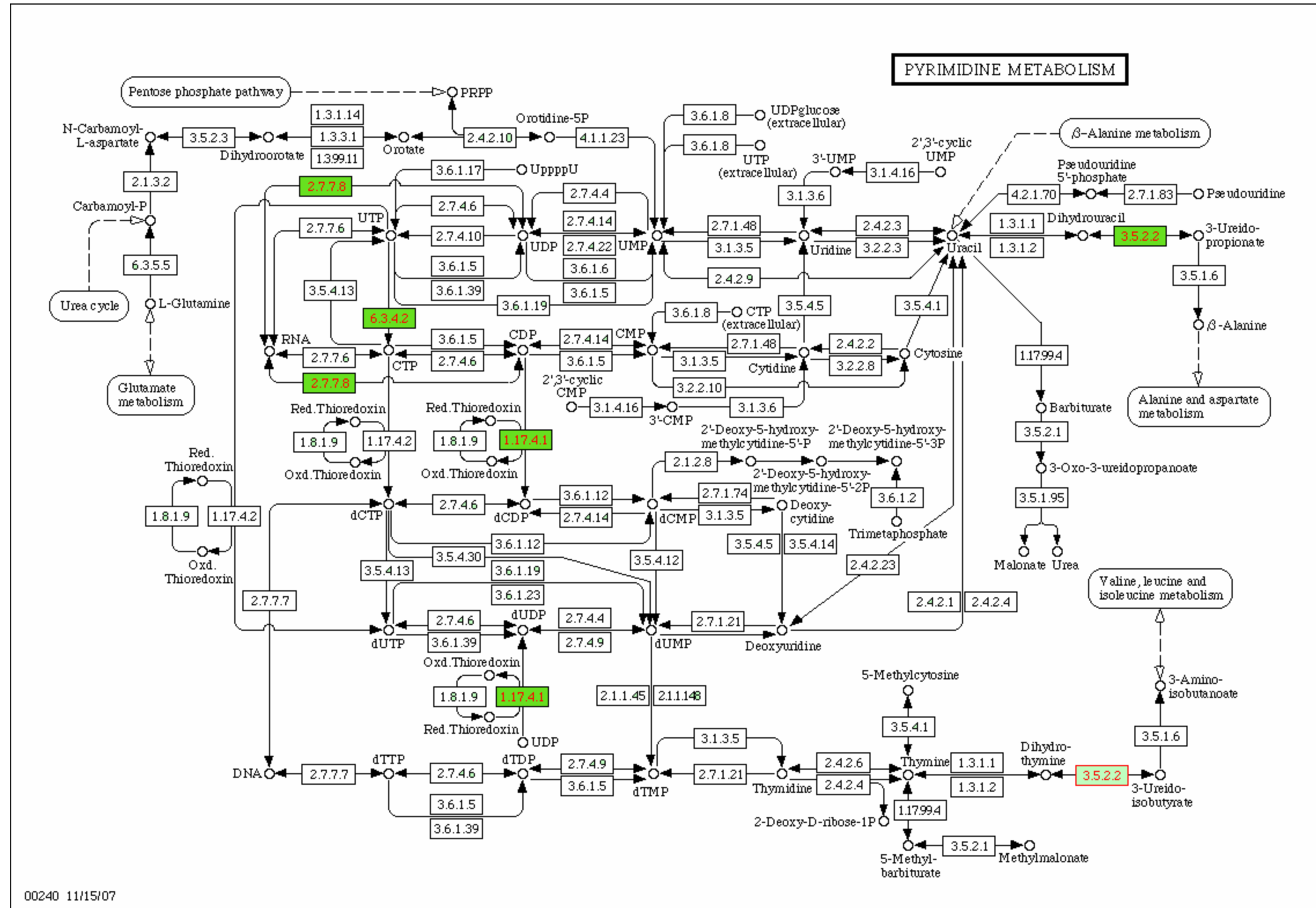
(light green = candidate PUPylated proteins, red border = confirmed PUPylated proteins)



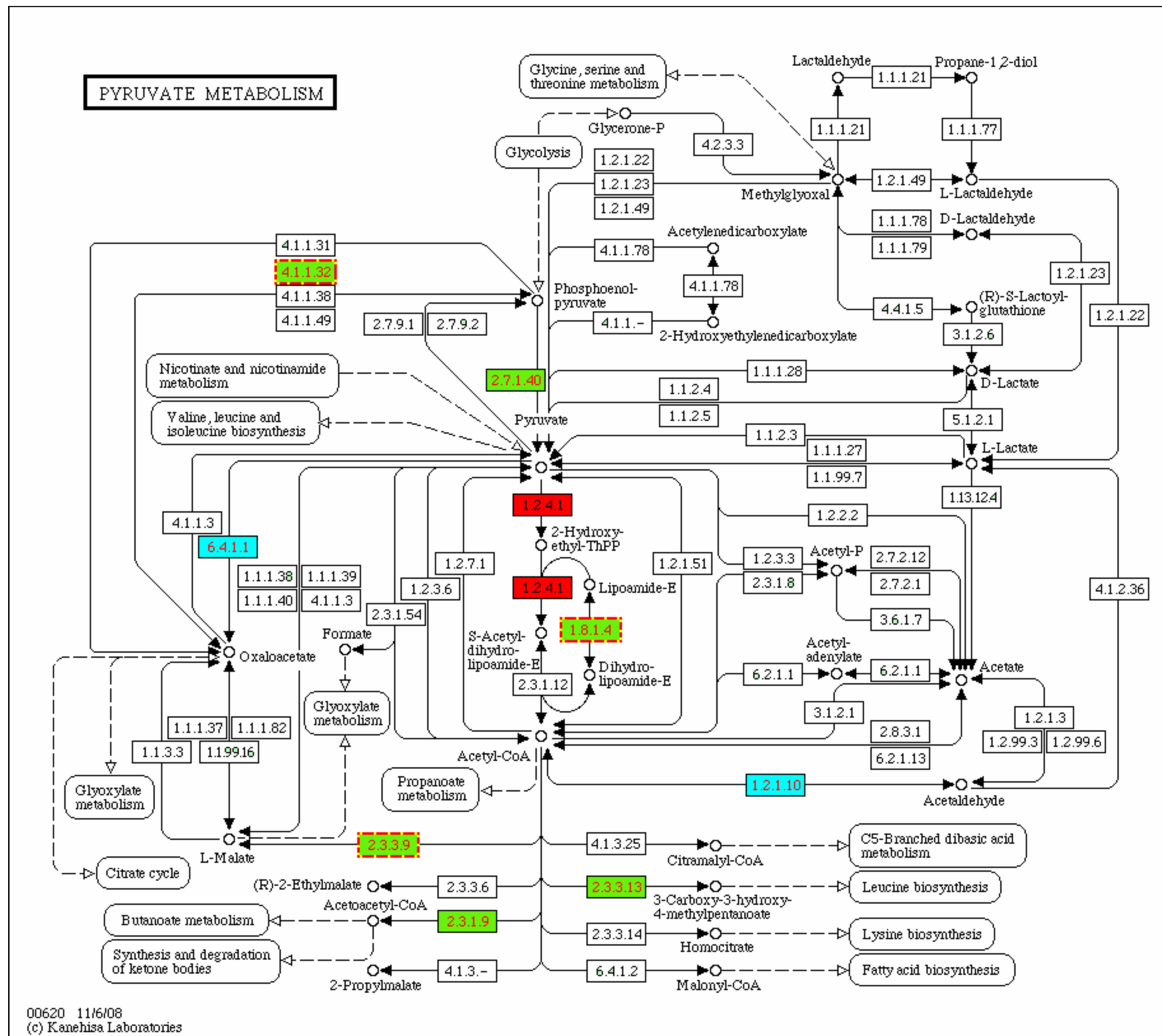
SF133: Presence of PUPylated proteins in purine metabolism



SF134: Presence of PUPylated proteins in pyrimidine metabolism



SF135: Presence of PUPylated proteins in pyruvate metabolism



SF136: Presence of PUPylated proteins in alanine and aspartate metabolism

