SUPPLEMENTARY INFORMATION

Table S1. Primers used for introduction of 6-Histidine affinity purification tag at the 3'-end of *petA* in *F. diplosiphon* SF33 (restriction sites are underlined, 6-Histidine tag is marked in bold, italicized letters).

Primer	**Sequence (5' to 3')**
petA-F-NcoI	GCTCCATGGGCTTGCTCAACATACTTCCACATG
petA-His-R3	***GTGATGGTGATGGTGATG***GAAATTCATTTCTGCGGCTTGGACTTTCTCAACCTGCTTCTT
petA-R-EcoRI	GCGGAATTCCCTACAAGGAAGAGACAAGCGTAT
petA-His-F3	***CATCACCATCACCATCAC***TAAATTCTTTGCTAACTTACTTATTTAAAGACAGGCAATTTG
petAseq2	ACTTCCACTTCTGTCGGCTTTGCA
pJCF276seq2NcoI	ACAGACGGCATGATGAACCTGAAT
pJCF276seqEcoRI	TGTACCTATAACCAGACCGTTCAGC

Table S2. Inter-monomer Interactions in the Cyt b_6f Complex (PDB ID 4OGQ). Abbreviations, res1, res2, residues 1 and 2; Ch.1, Ch.2, Chains 1 and 2; aaProperty, amino acid property (phil, hydrophilic; phob, hydrophobic); #, amino acid number.

Monomer1				***Monomer2***					***aaProperty***	
Res1	***#***	***Atom1***	***Ch.1***	***Res2***	***#***	***Atom2***	***Ch.2***	***Dist(Å)***	***Prop1***	***Prop2***
TRP	7	CZ3	A	LEU	116	CG	A	3.86	phil	phob
TRP	7	CZ2	A	GLU	115	CD	A	3.92	phil	phil
PHE	8	CE1	A	LEU	116	CD1	A	3.96	phob	phob
GLU	10	OE2	A	LYS	112	NZ	A	3.58	phil	phil
ARG	11	NE	A	GLU	115	OE2	A	2.72	phil	phil
ARG	11	CG	A	PRO	113	O	A	3.70	phil	phob
ARG	11	NH2	A	LYS	112	CG	A	3.56	phil	phil
ARG	11	O	A	LYS	208	CE	A	3.54	phil	phil
LEU	12	CD2	A	MET	205	CE	A	3.74	phob	phob
LEU	12	CD2	A	LEU	116	CD1	A	3.72	phob	phob
LEU	12	CD2	A	PHE	201	CZ	A	3.99	phob	phob
LEU	12	O	A	LYS	208	CE	A	3.82	phob	phil
GLU	13	OE2	A	LYS	208	NZ	A	2.95	phil	phil
PHE	48	CZ	A	PHE	189	CZ	A	3.66	phob	phob
PHE	48	CE2	A	TRP	193	CE2	A	3.86	phob	phil
PHE	52	CB	A	PHE	189	CE2	A	3.58	phob	phob
PHE	52	CE1	A	VAL	190	CG2	A	3.36	phob	phob
THR	55	OG1	A	SER	185	OG	A	2.61	phil	phil
THR	55	O	A	THR	181	CG2	A	3.42	phil	phil
PHE	56	O	A	THR	181	CG2	A	3.41	phob	phil
PHE	56	CE1	A	SER	185	OG	A	3.34	phob	phil
PHE	56	CD2	A	ARG	182	NE	A	3.55	phob	phil
TYR	57	N	A	THR	181	CG2	A	3.88	phil	phil
TYR	57	OH	A	ARG	182	NH2	A	2.93	phil	phil
LYS	59	CE	A	GLN	177	CG	A	3.50	phil	phil
LYS	59	CE	A	THR	181	OG1	A	3.92	phil	phil
PRO	60	CG	A	PRO	60	CG	A	3.77	phob	phob
THR	61	OG1	A	THR	61	OG1	A	3.16	phil	phil
LYS	112	CG	A	ARG	11	NH2	A	3.56	phil	phil
LYS	112	NZ	A	GLU	10	OE2	A	3.58	phil	phil
PRO	113	O	A	ARG	11	CG	A	3.70	phob	phil
GLU	115	OE2	A	ARG	11	NE	A	2.72	phil	phil
GLU	115	CD	A	TRP	7	CZ2	A	3.92	phil	phil
LEU	116	CD1	A	LEU	12	CD2	A	3.72	phob	phob
LEU	116	CG	A	TRP	7	CZ3	A	3.86	phob	phil
LEU	116	CD1	A	PHE	8	CE1	A	3.96	phob	phob
GLN	177	CG	A	LYS	59	CE	A	3.50	phil	phil
THR	181	OG1	A	LYS	59	CE	A	3.91	phil	phil
THR	181	CG2	A	THR	55	O	A	3.42	phil	phil
THR	181	CG2	A	PHE	56	O	A	3.41	phil	phob
THR	181	CG2	A	TYR	57	N	A	3.88	phil	phil
ARG	182	NH2	A	TYR	57	OH	A	2.93	phil	phil
ARG	182	NE	A	PHE	56	CD2	A	3.55	phil	phob

SER	185	OG	A	PHE	56	CE1	A	3.34	phil	phob
SER	185	OG	A	THR	55	OG1	A	2.61	phil	phil
THR	188	O	A	PHE	189	CD1	A	3.36	phil	phob
PHE	189	CE2	A	PHE	52	CB	A	3.58	phob	phob
PHE	189	CZ	A	PHE	48	CZ	A	3.66	phob	phob
PHE	189	CD1	A	THR	188	O	A	3.36	phob	phil
VAL	190	CG2	A	PHE	52	CE1	A	3.36	phob	phob
PRO	192	O	A	TRP	193	CH2	A	3.91	phob	phil
TRP	193	CZ3	A	TRP	193	CZ3	A	3.48	phil	phil
TRP	193	CH2	A	PRO	192	O	A	3.91	phil	phob
TRP	193	CE2	A	PHE	48	CE2	A	3.86	phil	phob
TRP	193	CH2	A	ALA	196	CB	A	3.62	phil	phob
ALA	196	CB	A	TRP	193	CH2	A	3.62	phob	phil
PHE	201	CZ	A	LEU	12	CD2	A	3.99	phob	phob
MET	205	CE	A	LEU	12	CD2	A	3.74	phob	phob
LYS	208	NZ	A	GLU	13	OE2	A	2.95	phil	phil
LYS	208	CE	A	LEU	12	O	A	3.82	phil	phob
LYS	208	CE	A	ARG	11	O	A	3.54	phil	phil
TRP	7	CZ2	A	PHE	113	CE1	B	3.52	phil	phob

Table S3. Inter-monomer Interactions in the Cyt *bc*$_1$ Complex (PDB ID 3CX5). Abbreviations, res1, res2, residues 1 and 2; Ch.1, Ch.2, Chains 1 and 2; aaProperty, amino acid property (phil, hydrophilic; phob, hydrophobic); #, amino acid number.

Monomer1				***Monomer2***					***aaProperty***	
Res1	***#***	***Atom1***	***Ch.1***	***Res2***	***#***	***Atom2***	***Ch.2***	***Dist(Å)***	***Prop1***	***Prop2***
VAL	8	O	C	MET	199	CE	N	3.80	phob	phob
VAL	8	CB	C	ILE	203	CD1	N	3.99	phob	phob
TYR	9	OH	C	ALA	200	CB	N	3.47	phil	phob
TYR	9	OH	C	MET	196	O	N	2.62	phil	phob
TYR	9	CD2	C	MET	199	CE	N	3.91	phil	phob
TYR	9	OH	C	VAL	116	CG2	N	3.50	phil	phob
TYR	9	CE1	C	THR	112	CG2	N	3.77	phil	phil
LEU	12	CD2	C	MET	199	CE	N	3.70	phob	phob
ILE	48	CD1	C	LEU	185	CG	N	3.67	phob	phob
MET	52	O	C	GLN	177	CB	N	3.36	phob	phil
MET	52	CE	C	ARG	178	CB	N	3.40	phob	phil
TYR	54	O	C	GLN	177	NE2	N	3.27	phil	phil
SER	55	OG	C	ASN	57	ND2	N	2.83	phil	phil
SER	55	OG	C	GLN	177	NE2	N	3.04	phil	phil
ASN	57	ND2	C	LEU	60	CD1	N	3.88	phil	phob
ASN	57	ND2	C	SER	55	OG	N	2.82	phil	phil
ASN	57	OD1	C	ASN	57	OD1	N	3.70	phil	phil
LEU	60	CD2	C	ASN	57	OD1	N	3.88	phob	phil
LEU	60	CD2	C	LEU	60	CD1	N	3.69	phob	phob
THR	112	CG2	C	TYR	9	CE1	N	3.68	phil	phil
VAL	116	CG2	C	TYR	9	OH	N	3.61	phob	phil
GLN	177	NE2	C	TYR	54	O	N	3.33	phil	phil
GLN	177	CB	C	MET	52	O	N	3.26	phil	phob
GLN	177	NE2	C	SER	55	OG	N	2.91	phil	phil
ARG	178	N	C	MET	52	O	N	3.70	phil	phob
ALA	181	O	C	TYR	184	OH	N	2.76	phob	phil
LEU	182	CD1	C	MET	52	CE	N	3.91	phob	phob
TYR	184	OH	C	ALA	181	O	N	2.70	phil	phob
TYR	184	CD1	C	TYR	184	CD1	N	3.70	phil	phil
TYR	184	OH	C	LEU	185	CB	N	3.50	phil	phob
LEU	185	CB	C	TYR	184	OH	N	3.54	phob	phil
LEU	185	CD2	C	PHE	188	CZ	N	3.89	phob	phob
PHE	188	CZ	C	LEU	185	CD2	N	3.83	phob	phob
MET	196	O	C	TYR	9	OH	N	2.67	phob	phil
MET	199	CE	C	VAL	8	O	N	3.81	phob	phob
MET	199	CE	C	LEU	12	CD1	N	3.59	phob	phob
MET	199	CE	C	TYR	9	CD2	N	3.99	phob	phil
ALA	200	CB	C	TYR	9	OH	N	3.46	phob	phil
ILE	203	CD1	C	VAL	8	CB	N	3.90	phob	phob

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Bent trans-membrane helix of the ISP subunit (pink) in the cyt b_6f complex (PDB ID 4OGQ). The PetL subunit (light brown) interacts with the ISP TMH on the n-side, through an ordered UDM detergent molecule (black/red sticks).

Figure S2. Low resolution mass spectroscopy (ESI-MS) analysis of the truncated ISP subunit from *F. diplosiphon*.

Figure S3. High resolution mass spectroscopy (ESI-MS) analysis of the truncated ISP subunit from *F. diplosiphon.* Ion-isolation of 14-charge precursor (inset) and electron-transfer dissociation tandem mass spectrum are shown.

Figure S4. Low resolution mass spectroscopy (ESI-MS) analysis of the full-length ISP subunit from *F. diplosiphon*.

Figure S5. High resolution mass spectroscopy (ESI-MS) analysis of the PetG subunit from *F. diplosiphon.* Ion-isolation of 3-charge precursor ion (inset) and collisionally activated dissociation tandem mass spectrum of the ion are shown.

Figure S6. High resolution mass spectroscopy (ESI-MS) analysis of the PetM subunit from *F. diplosiphon.* Ion-isolation of 3-charge precursor ion (inset) and collisionally activated dissociation tandem mass spectrum of the ion are shown.

Figure S7. High resolution mass spectroscopy (ESI-MS) analysis of the PetL subunit from *F. diplosiphon.* Ion-isolation of 3-charge precursor ion (inset) and collisionally activated dissociation tandem mass spectrum of the ion are shown.

Figure S8. High resolution mass spectroscopy (ESI-MS) analysis of the PetN subunit from *F. diplosiphon.* Ion-isolation of 3-charge precursor ion (inset) and collisionally activated dissociation tandem mass spectrum of the ion are shown.

Figure S9. High-resolution top-down mass spectrometry of small sub-units of *F. diplosiphon* SF33 b_6f complex. Selected fractions from LC-MS+ analysis were analyzed by static nanospray according to *Materials and Methods*. Product ion assignments for known proteins were made using ProSightPC operated in single protein mode with a 15-ppm mass accuracy threshold and with the delta mass feature deactivated. Product ion assignments from collisional activated dissociation (CAD) experiments are labeled with 135˚ slashes to the top left (*b*-ions) or bottom right (*y*-ions) and from electron transfer dissociation (ETD) experiments with 90˚ slashes to the top left (*c*-ions) or bottom right (*z*-ions), as described in the main text [64] . (**A**) The 14579 Da component (**Table 1**) matched to 46 product-ions of electron transfer dissociation of a precursor-ion of monoisotopic mass 14567.0896 Da for a P-Score of $2.2\mathrm{x}10^{-43}$ identifying this species as an N-terminal truncation product of the Rieske subunit. (**B**) Cleavage site of ISP subunit in cyt b_6f complex (PDB ID 2E74). Tyr41 and Phe42 are located proximal to the membrane-water interface. (**C**) The 4026 Da component matched to 41 product-ions of collisionally activated dissociation (CAD) of a precursor-ion of monoisotopic mass of 4023.1578 Da for a P-Score of 3.32×10^{-64} identifying this species as *F. diplosiphon* SF33 PetG. Shading of the initiating Met residue indicates formylation. (**D**) The 3578 Da component matched to 16 product-ions of collisionally activated dissociation (CAD) of a precursor-ion of monoisotopic mass of 3575.8351 Da for a P-Score of 7.8×10^{-23} when matched to the *Nostoc azollae* PetM sequence, identifying this species as *F. diplosiphon* SF33 PetM. (**E**) The 3255 Da component matched to 27 product-ions of collisionally activated dissociation (CAD) of a precursor-ion of monoisotopic mass of 3256.0387 Da for a P-Score of 2.6×10^{-53} when matched to a custom PetL sequence, identifying this species as *F. diplosiphon* SF33 PetL. (**F**) The 3261 Da component matched to 33 product-ions of collisionally activated dissociation (CAD) of a precursor-ion of monoisotopic mass of 3259.7588 Da for a P-Score of 5.6×10^{-46} identifying this species as PetN of *F. diplosiphon* SF33.

Figure S10. Alignment of cyt b_6 polypeptide sequences from unicellular prokaryotes (*Synechocystis* PCC 6803, and *Synechococcus* PCC 7002), filamentous prokaryotes (*Nostoc* PCC 7120, *Fremyella diplosiphon* SF33, and *Mastigocladus laminosus*), a unicellular eukaryotic alga (*Chlamydomonas reinhardtii*), and higher eukaryotes (*Arabidopsis thaliana*, and *Spinacia oleracea*) sources. Secondary structure assignment and residue numbering are based on the cyt b_6f complex crystal structure from *Nostoc* PCC 7120 (PDB ID 4OGQ, chain A). Residues involved in inter-monomer contacts are highlighted with arrows. Color code: red- complete conservation, yellow- partial conservation, and white- no conservation.

Figure S11. Alignment of subIV polypeptide sequences from unicellular prokaryotes (*Synechocystis* PCC 6803, and *Synechococcus* PCC 7002), filamentous prokaryotes (*Nostoc* PCC 7120, *Fremyella diplosiphon* SF33, and *Mastigocladus laminosus*), a unicellular eukaryotic alga (*Chlamydomonas reinhardtii*), and higher eukaryotes (*Arabidopsis thaliana*, and *Spinacia oleracea*) sources. Secondary structure assignment and residue numbering are based on the cyt b_6f complex crystal structure from *Nostoc* PCC 7120 (PDB ID 4OGQ, chain B). Residues involved in inter-monomer contacts are highlighted with arrows. Color code: red- complete conservation, yellow- partial conservation, and white- no conservation.

Figure S12. Alignment of cyt b_6 polypeptide sequence (from the cyt b_6f complex) with the homologous N-terminal domain of the cyt *b* subunit (of the cyt bc_1 complex). Sequences of the cyt *b* (bc_1) polypeptide represent prokaryotic anoxygenic photosynthetic bacteria (*Rhodobacter capsulatus*, and *Rhodobacter sphaeroides*), a unicellular eukaryote (*Saccharomyces cerevisiae*), and higher eukaryotes (*Gallus gallus*, and *Bos taurus*). Residue numbering is follows the sequence of the cyt b_6 polypeptide from *Nostoc* PCC 7120. Residues involved in inter-monomer interactions in the cyt b_6f complex of *Nostoc* PCC 7120 (PDB ID 4OGQ) are highlighted with arrows. Color code: red- complete conservation, yellow- partial conservation, and white- no conservation. For clarity, partial sequences of the cyt *b* polypeptide, which shares homology with the cyt b_6 subunit, are shown in the figure.

Figure S13. Alignment of subIV polypeptide sequence (from the cyt b_6f complex) with the homologous C-terminal domain of the cyt *b* subunit (of the cyt bc_1 complex). Sequences of the cyt *b* (bc_1) polypeptide represent prokaryotic anoxygenic photosynthetic bacteria (*Rhodobacter capsulatus*, and *Rhodobacter sphaeroides*), a unicellular eukaryote (*Saccharomyces cerevisiae*), and higher eukaryotes (*Gallus gallus*, and *Bos taurus*). Residue numbering is follows the sequence of the subIV polypeptide from *Nostoc* PCC 7120. The residue involved in inter-monomer interactions in the cyt b_6f complex of *Nostoc* PCC 7120 (PDB ID 4OGQ) is highlighted with an arrow. Color code: red- complete conservation, yellow- partial conservation, and white- no conservation. For clarity, partial sequences of the cyt *b* polypeptide, which shares homology with the subIV polypeptide, are shown in the figure.

Figure S1

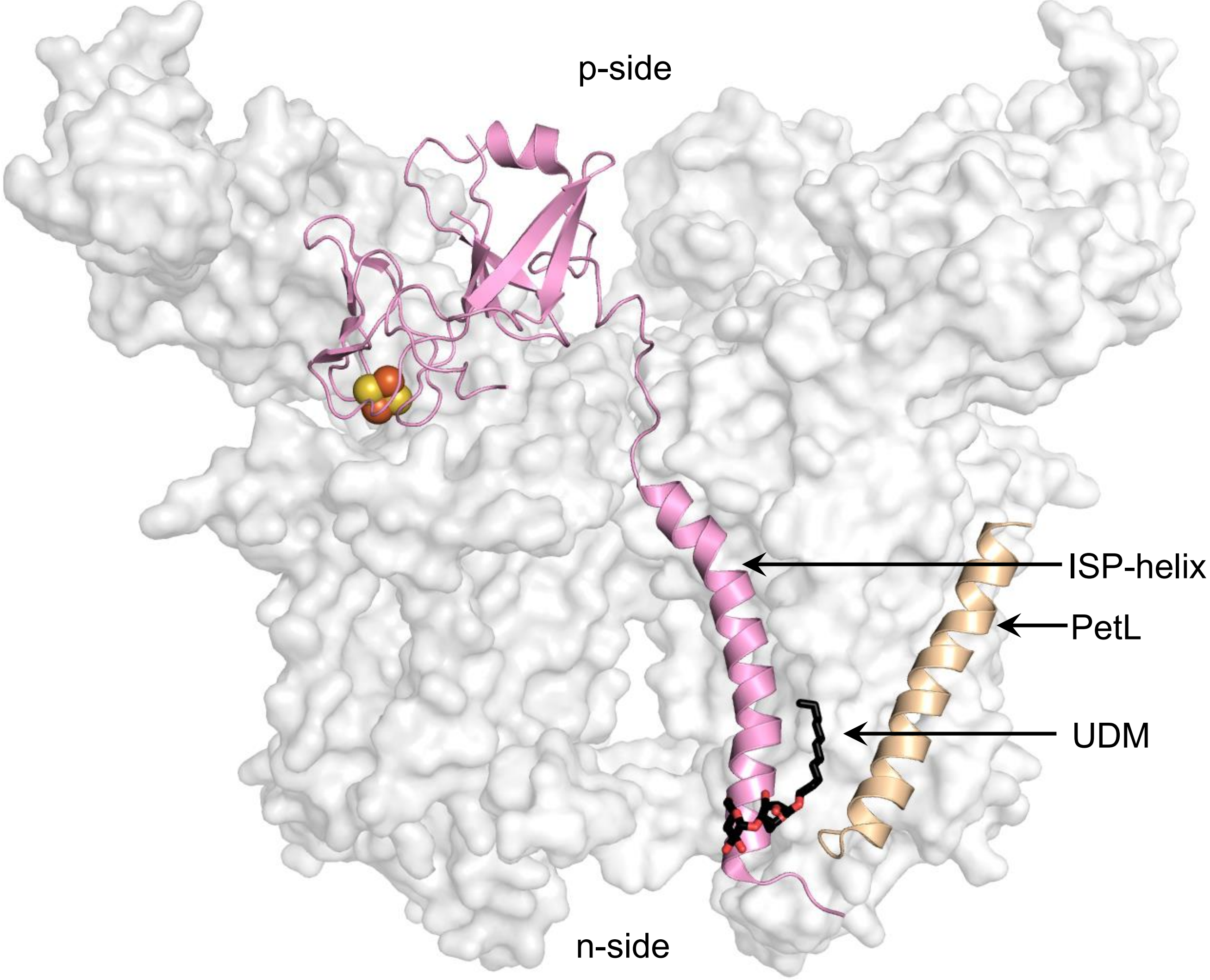
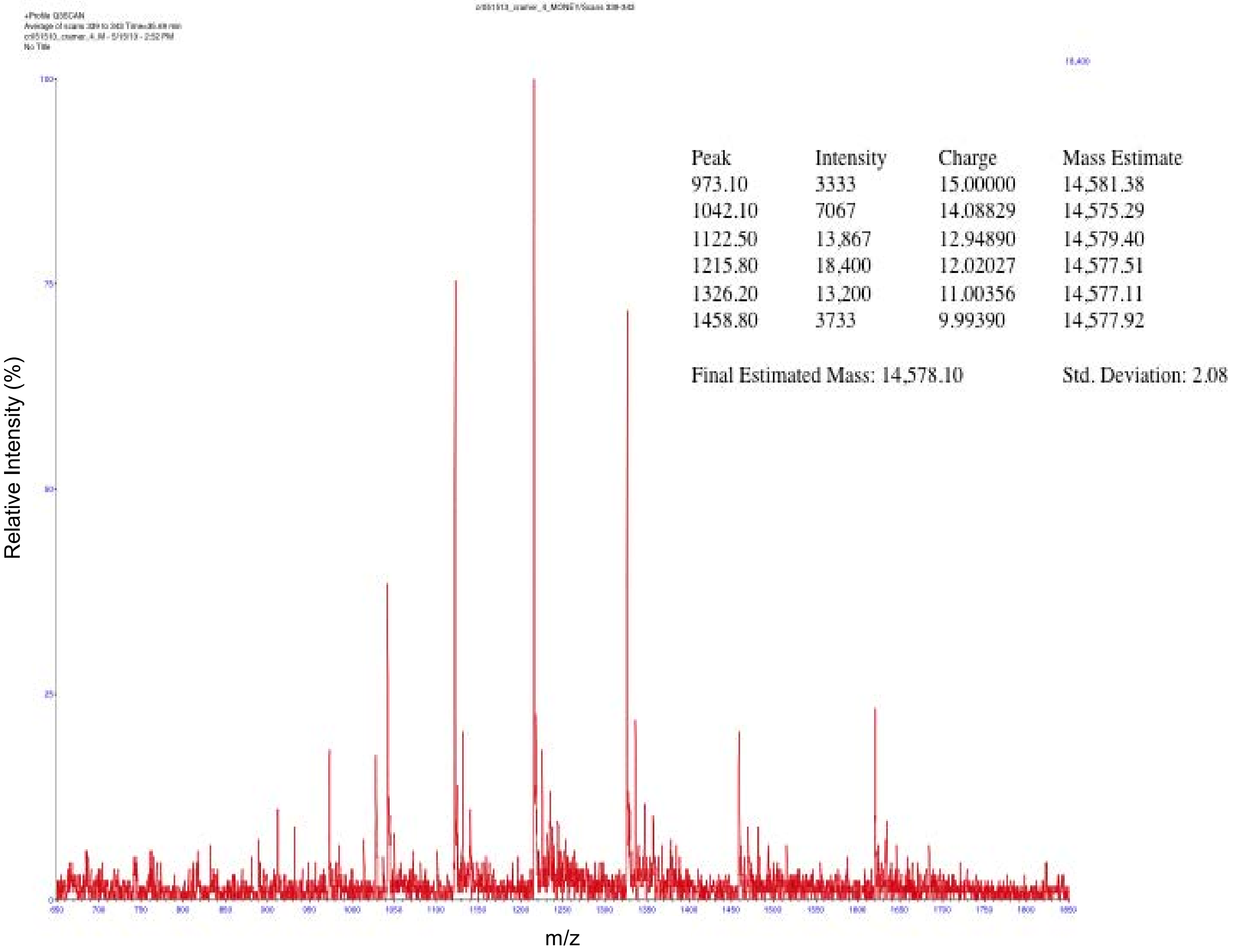


Figure S2



Peak	Intensity	Charge	Mass Estimate
973.10	3333	15.00000	14,581.38
1042.10	7067	14.08829	14,575.29
1122.50	13,867	12.94890	14,579.40
1215.80	18,400	12.02027	14,577.51
1326.20	13,200	11.00356	14,577.11
1458.80	3733	9.99390	14,577.92

Final Estimated Mass: 14,578.10 Std. Deviation: 2.08

Figure S3

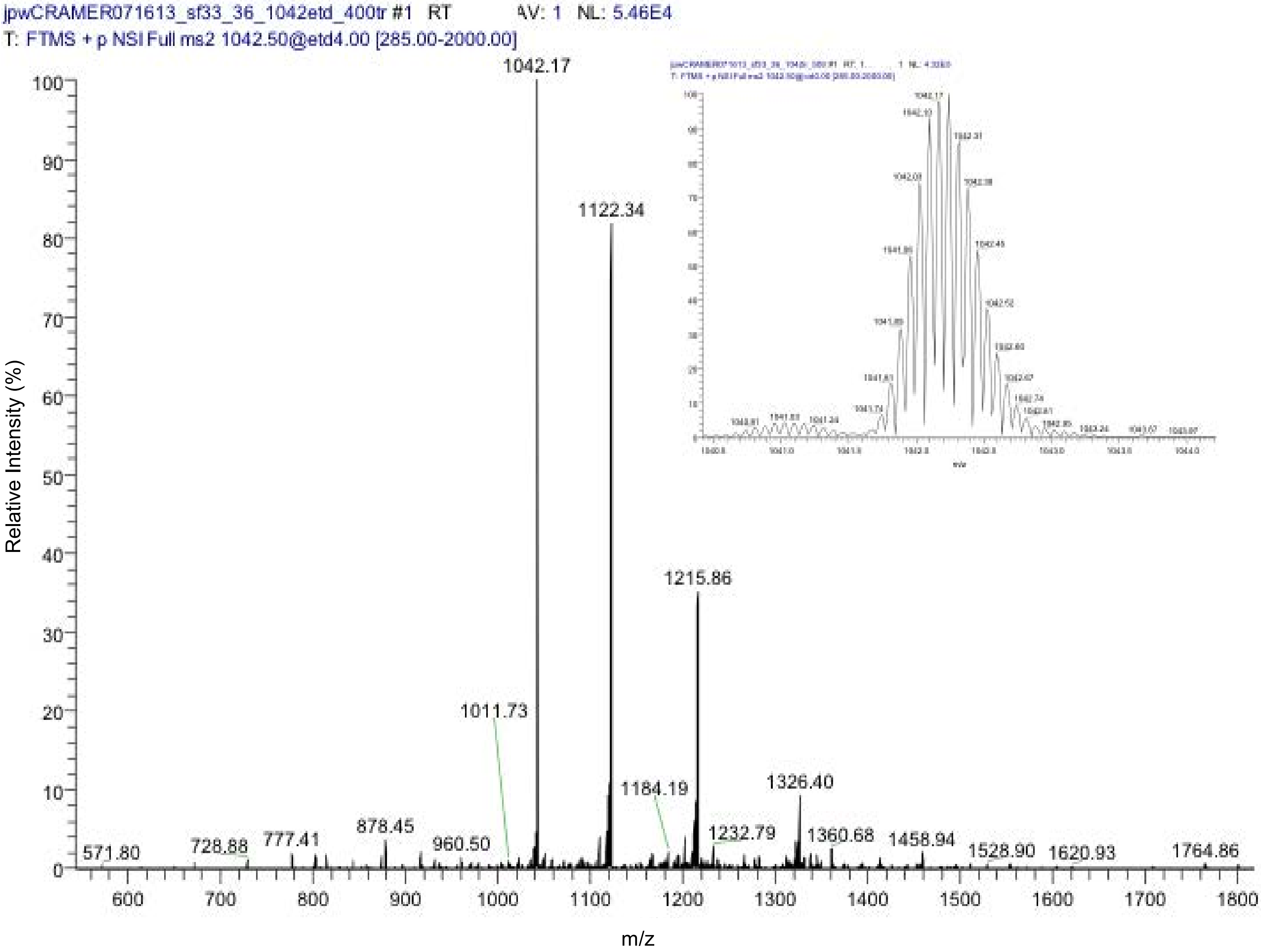
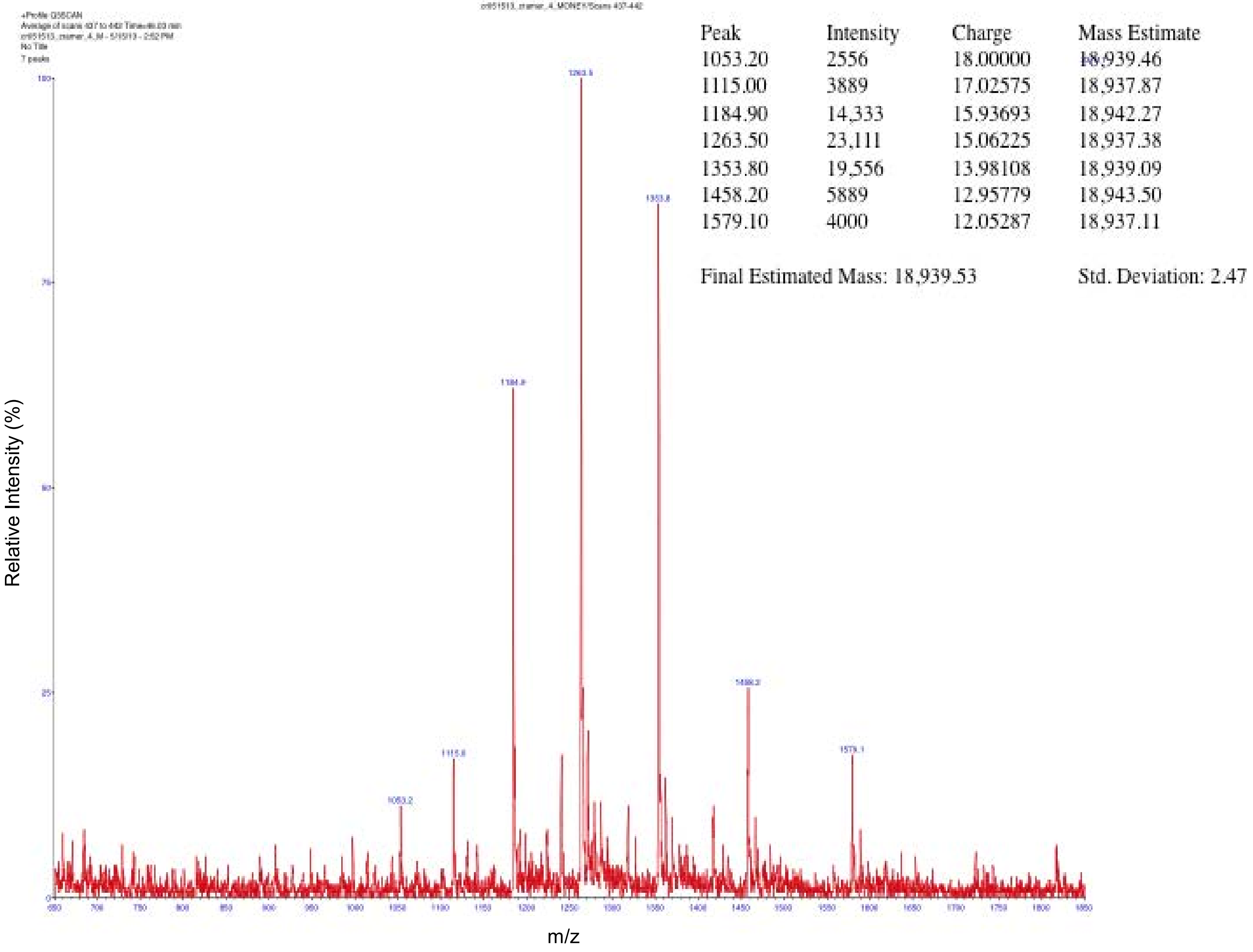

jpwCRAMER071613_sf33_36_1042etd_400tr #1 RT AV: 1 NL: 5.46E4
T: FTMS + p NSI Full ms2 1042.50@etd4.00 [285.00-2000.00]
Relative Intensity (%)
m/z
0
10
20
30
40
50
60
70
80
90
100
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1042.17
1122.34
1215.86
1011.73
1184.19
1326.40
1232.79
1360.68
1458.94
1528.90
1620.93
1764.86
571.80
728.88
777.41
878.45
960.50
1042.17
1042.10
1042.31
1042.03
1042.38
1041.96
1042.45
1041.89
1042.52
1041.81
1042.60
1042.67
1041.74
1042.74
1042.81
m/z

Figure S4



Peak	Intensity	Charge	Mass Estimate
1053.20	2556	18.00000	18,939.46
1115.00	3889	17.02575	18,937.87
1184.90	14,333	15.93693	18,942.27
1263.50	23,111	15.06225	18,937.38
1353.80	19,556	13.98108	18,939.09
1458.20	5889	12.95779	18,943.50
1579.10	4000	12.05287	18,937.11

Final Estimated Mass: 18,939.53 Std. Deviation: 2.47

Figure S5

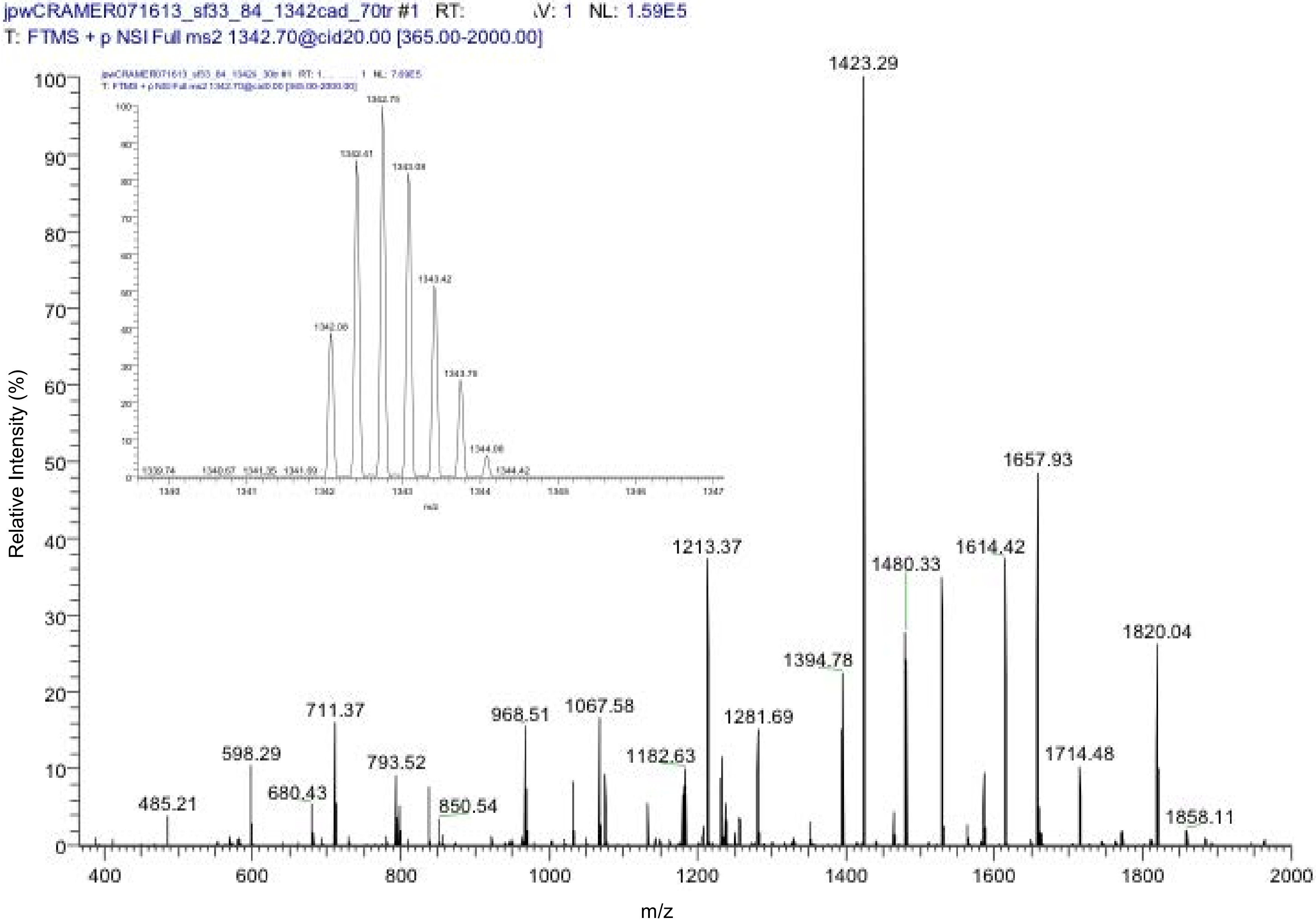

jpwCRAMER071613_sf33_84_1342cad_70tr #1 RT: .V: 1 NL: 1.59E5
T: FTMS + p NSI Full ms2 1342.70@cid20.00 [365.00-2000.00]
Relative Intensity (%)
m/z
0
10
20
30
40
50
60
70
80
90
100
400
600
800
1000
1200
1400
1600
1800
2000
485.21
598.29
680.43
711.37
793.52
850.54
968.51
1067.58
1182.63
1213.37
1281.69
1394.78
1423.29
1480.33
1614.42
1657.93
1714.48
1820.04
1858.11
1342.70
1342.41
1342.06
1343.42
1343.76
1344.08
1344.42
1339.74
1340.67
1341.35
1341.69

Figure S6

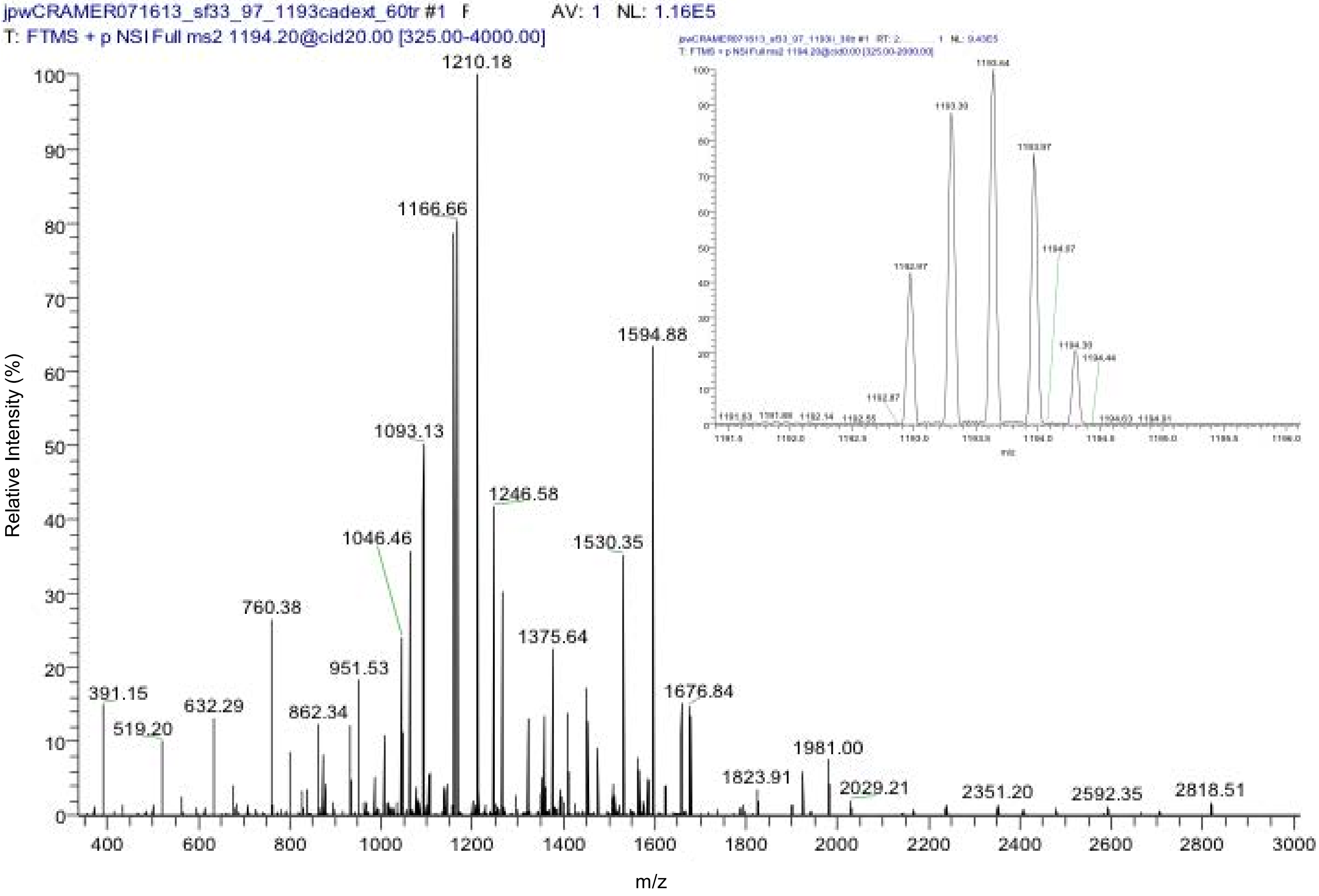

jpwCRAMER071613_sf33_97_1193cadext_60tr #1 F AV: 1 NL: 1.16E5
T: FTMS + p NSI Full ms2 1194.20@cid20.00 [325.00-4000.00]
Relative Intensity (%)
m/z
1210.18
1166.66
1594.88
1093.13
1246.58
1046.46
1530.35
760.38
1375.64
951.53
391.15
632.29
519.20
862.34
1676.84
1981.00
1823.91
2029.21
2351.20
2592.35
2818.51
1193.64
1193.30
1193.97
1192.97
1194.30
1194.44

Figure S7

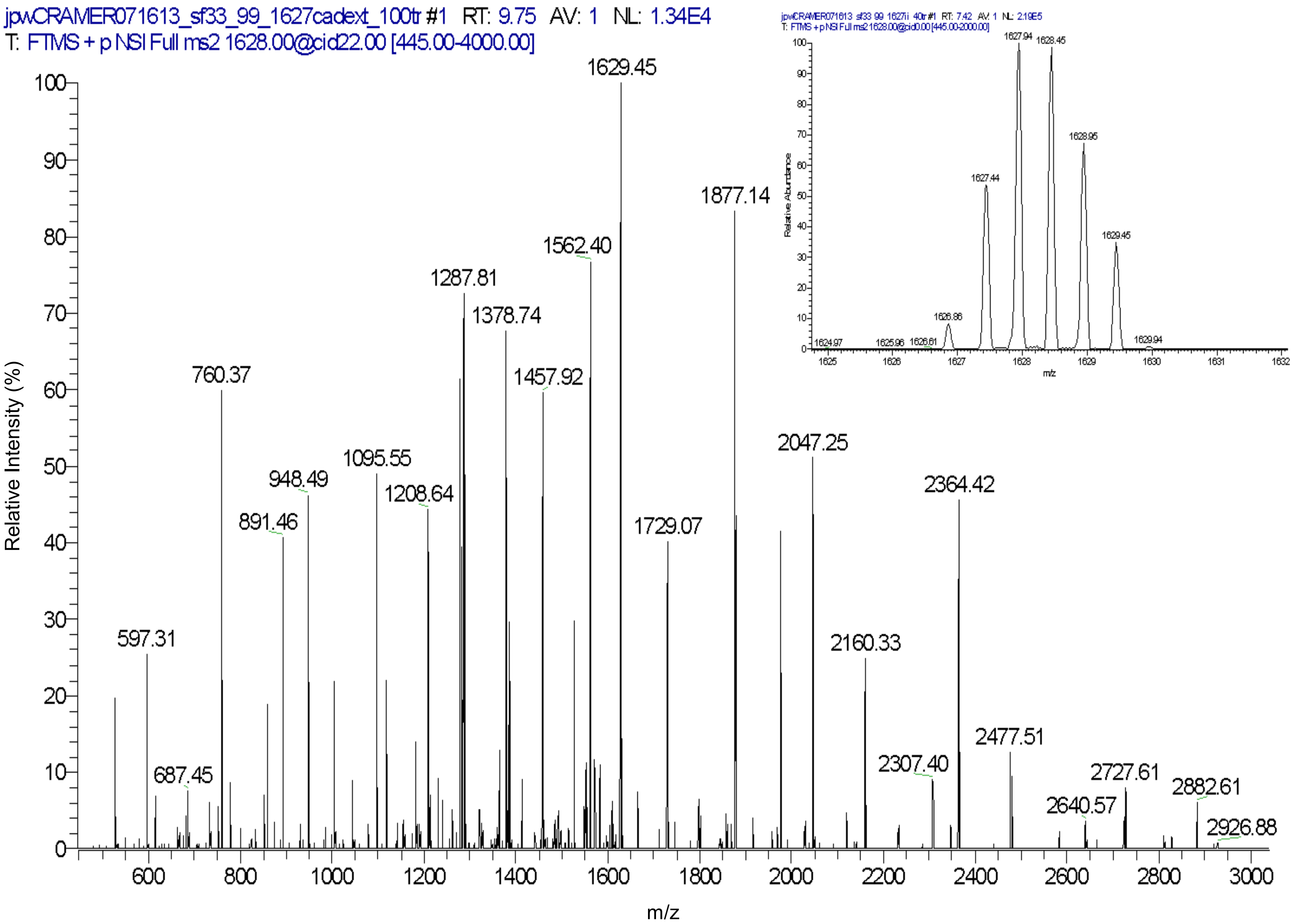
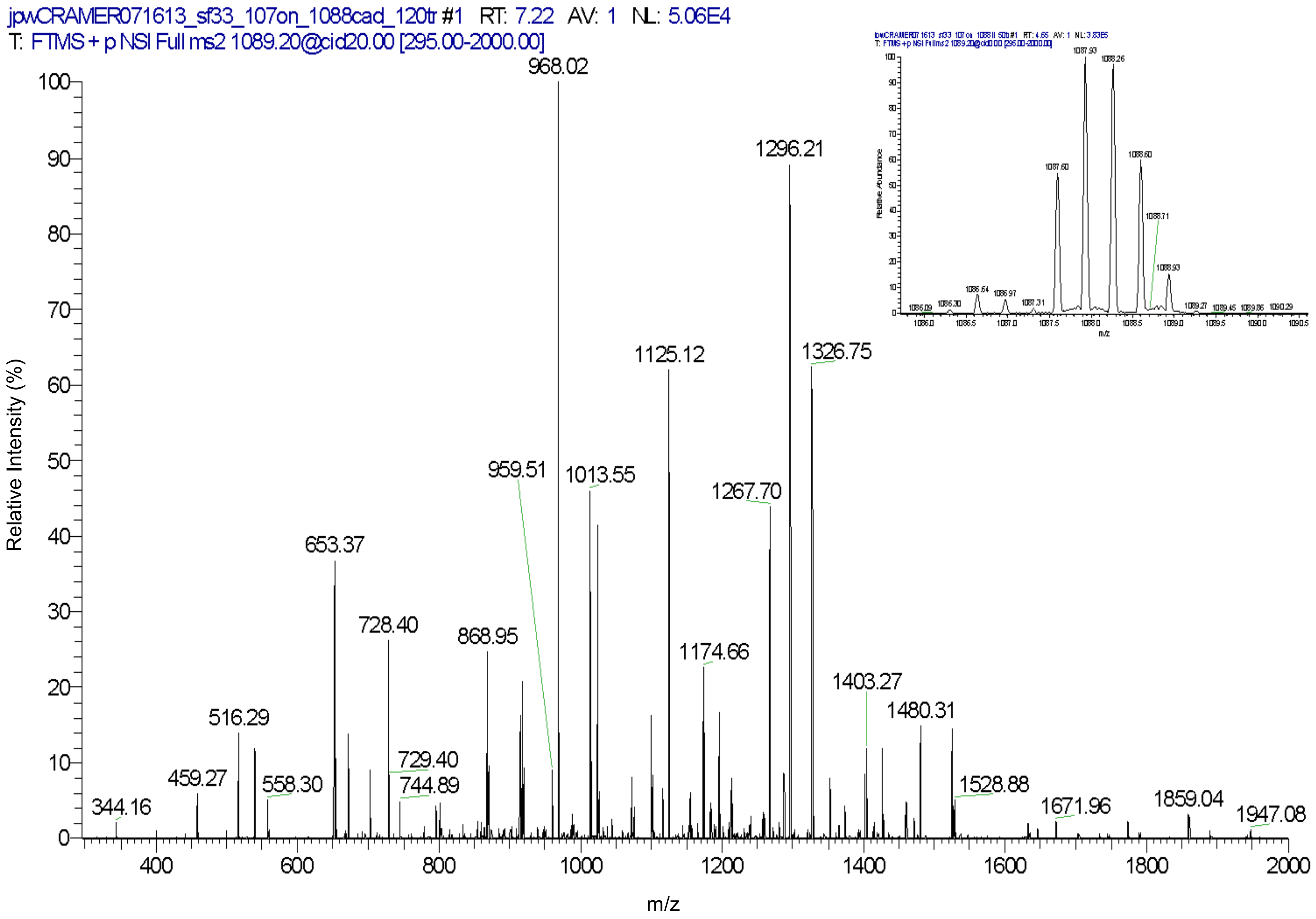

jpwCRAMER071613_sf33_99_1627cadext_100tr #1 RT: 9.75 AV: 1 NL: 1.34E4
T: FTMS + p NSI Full ms2 1628.00@cid22.00 [445.00-4000.00]
Relative Intensity (%)
m/z
100
90
80
70
60
50
40
30
20
10
0
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
1629.45
1877.14
1562.40
1287.81
1378.74
760.37
1457.92
2047.25
1095.55
948.49
2364.42
1208.64
891.46
1729.07
597.31
2160.33
2477.51
2307.40
687.45
2727.61
2882.61
2640.57
2926.88
jpwCRAMER071613 sf33 99 1627ii 40tr #1 RT: 7.42 AV: 1 NL: 2.19E5
T: FTMS + p NSI Full ms2 1628.00@cid0.00 [445.00-2000.00]
Relative Abundance
m/z
1627.94
1628.45
1628.95
1627.44
1629.45
1626.86
1624.97
1625.96
1626.61
1629.94
1625
1626
1627
1628
1629
1630
1631
1632

Figure S8



jpwCRAMER071613_sf33_107on_1088cad_120tr #1 RT: 7.22 AV: 1 NL: 5.06E4
T: FTMS + p NSI Full ms2 1089.20@cid20.00 [295.00-2000.00]
Relative Intensity (%)
100
90
80
70
60
50
40
30
20
10
0
968.02
1296.21
1125.12
1326.75
959.51
1013.55
1267.70
653.37
728.40
868.95
1174.66
1403.27
1480.31
516.29
729.40
459.27
558.30
744.89
1528.88
344.16
1671.96
1859.04
1947.08
400
600
800
1000
1200
1400
1600
1800
2000
m/z
jpwCRAMER071613_sf33_107on_1088ll 50tr #1 RT: 4.65 AV: 1 NL: 3.83E5
T: FTMS + p NSI Full ms2 1089.20@cid0.00 [295.00-2000.00]
Relative Abundance
1087.93
1088.26
1088.60
1087.60
1088.71
1088.93
1086.64
1086.97
1087.31
1086.09
1086.30
1089.27
1089.45
1089.86
1090.29
1086.0
1086.5
1087.0
1087.5
1088.0
1088.5
1089.0
1089.5
1090.0
1090.5
m/z

Figure S9

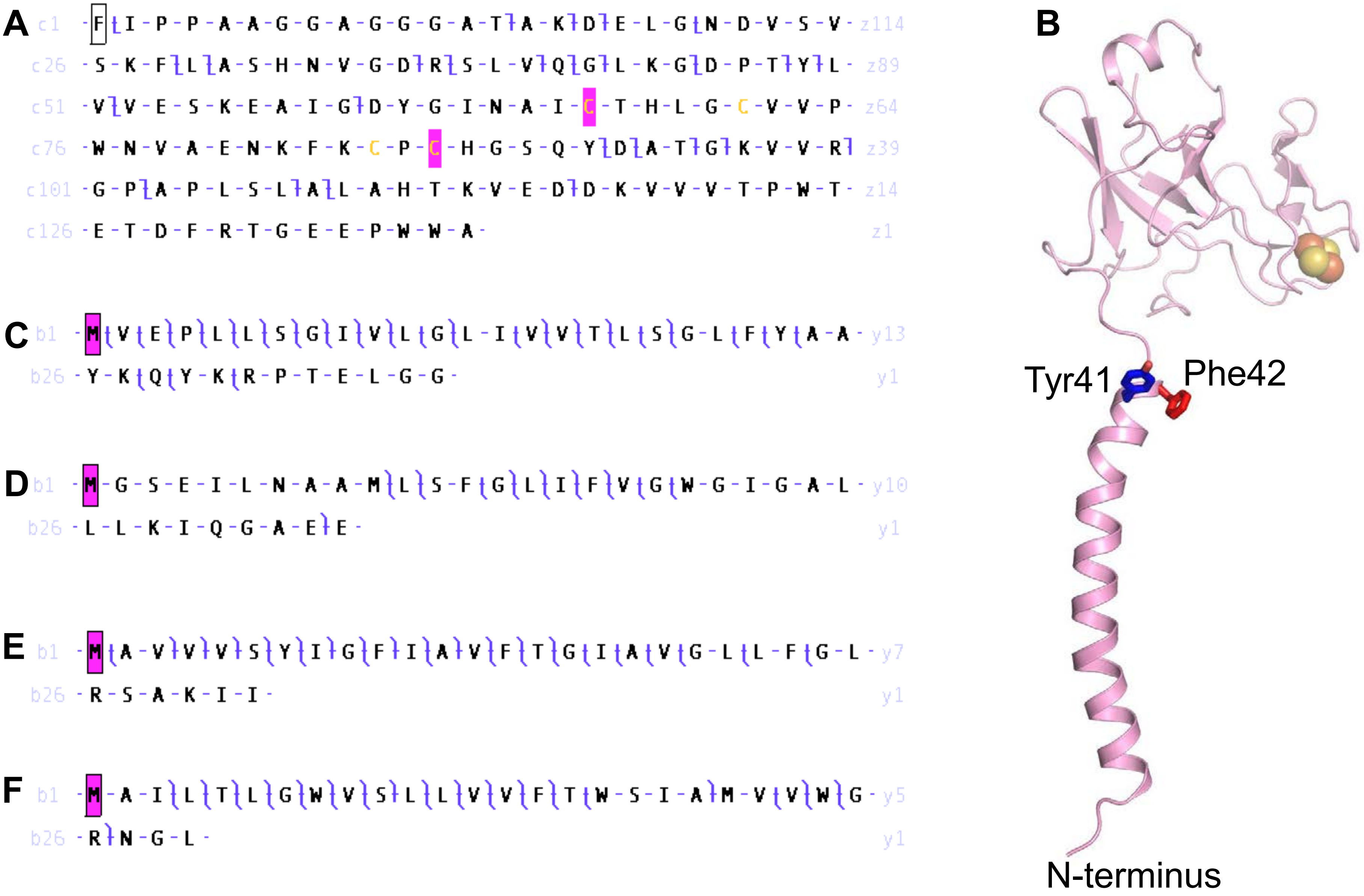

A
c1 F I P P A A G G A G G G A T A K D E L G N D V S V z114
c26 S K F L A S H N V G D R S L V Q G L K G D P T Y L z89
c51 V V E S K E A I G D Y G I N A I C T H L G C V V P z64
c76 W N V A E N K F K C P C H G S Q Y D A T G K V V R z39
c101 G P A P L S L A L A H T K V E D D K V V V T P W T z14
c126 E T D F R T G E E P W W A z1
B
Tyr41
Phe42
N-terminus
C
b1 M V E P L L S G I V L G L I V V T L S G L F Y A A y13
b26 Y K Q Y K R P T E L G G y1
D
b1 M G S E I L N A A M L S F G L I F V G W G I G A L y10
b26 L L K I Q G A E E y1
E
b1 M A V V V S Y I G F I A V F T G I A V G L L F G L y7
b26 R S A K I I y1
F
b1 M A I L T L G W V S L L V V F T W S I A M V V W G y5
b26 R N G L y1

Figure S10

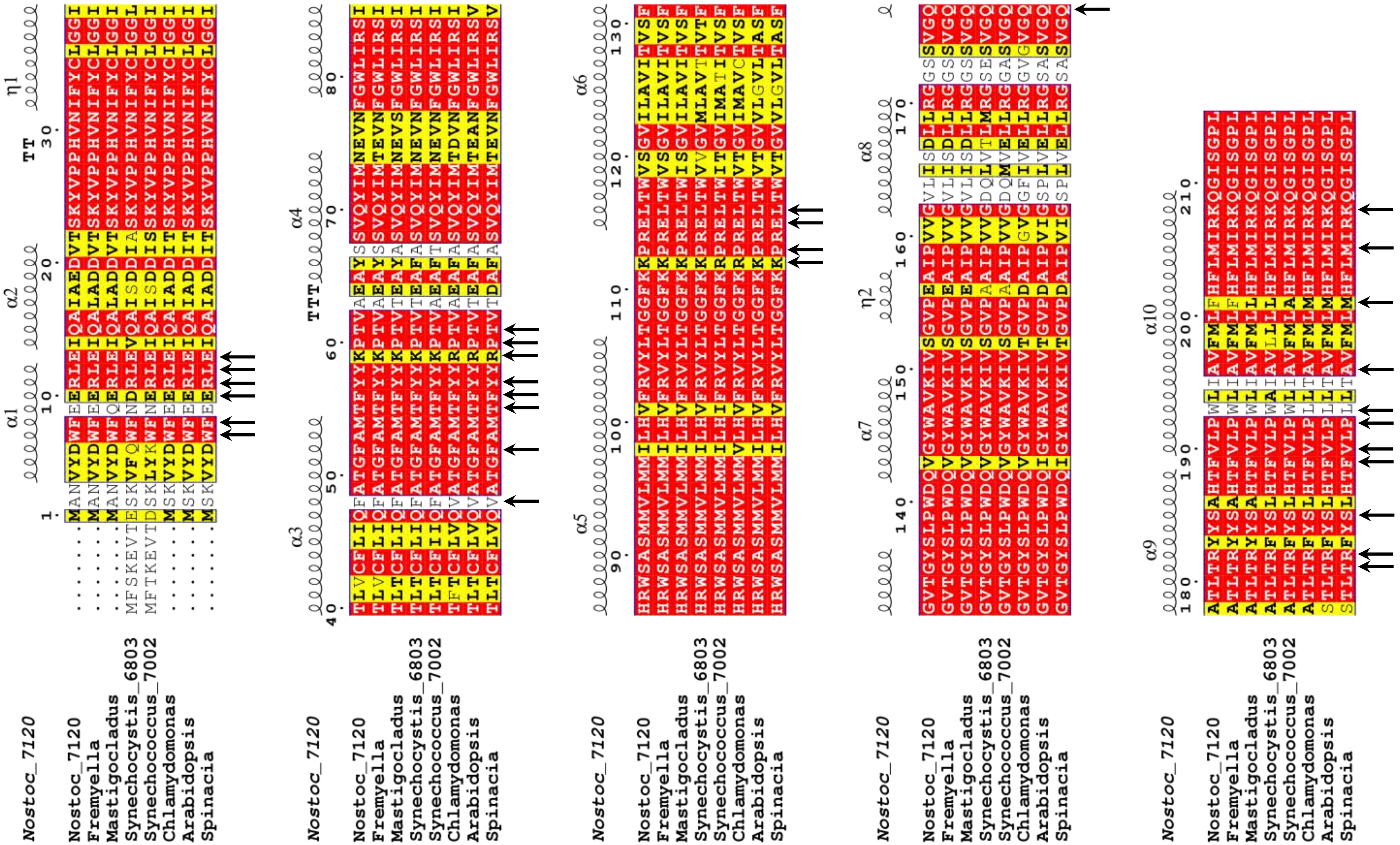

Nostoc_7120
Nostoc_7120
Fremyella
Mastigocladus
Synechocystis_6803
Synechococcus_7002
Chlamydomonas
Arabidopsis
Spinacia
α1
α2
η1
TT
α3
TTT
α4
α5
α6
α7
η2
α8
α9
α10
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210

Figure S11

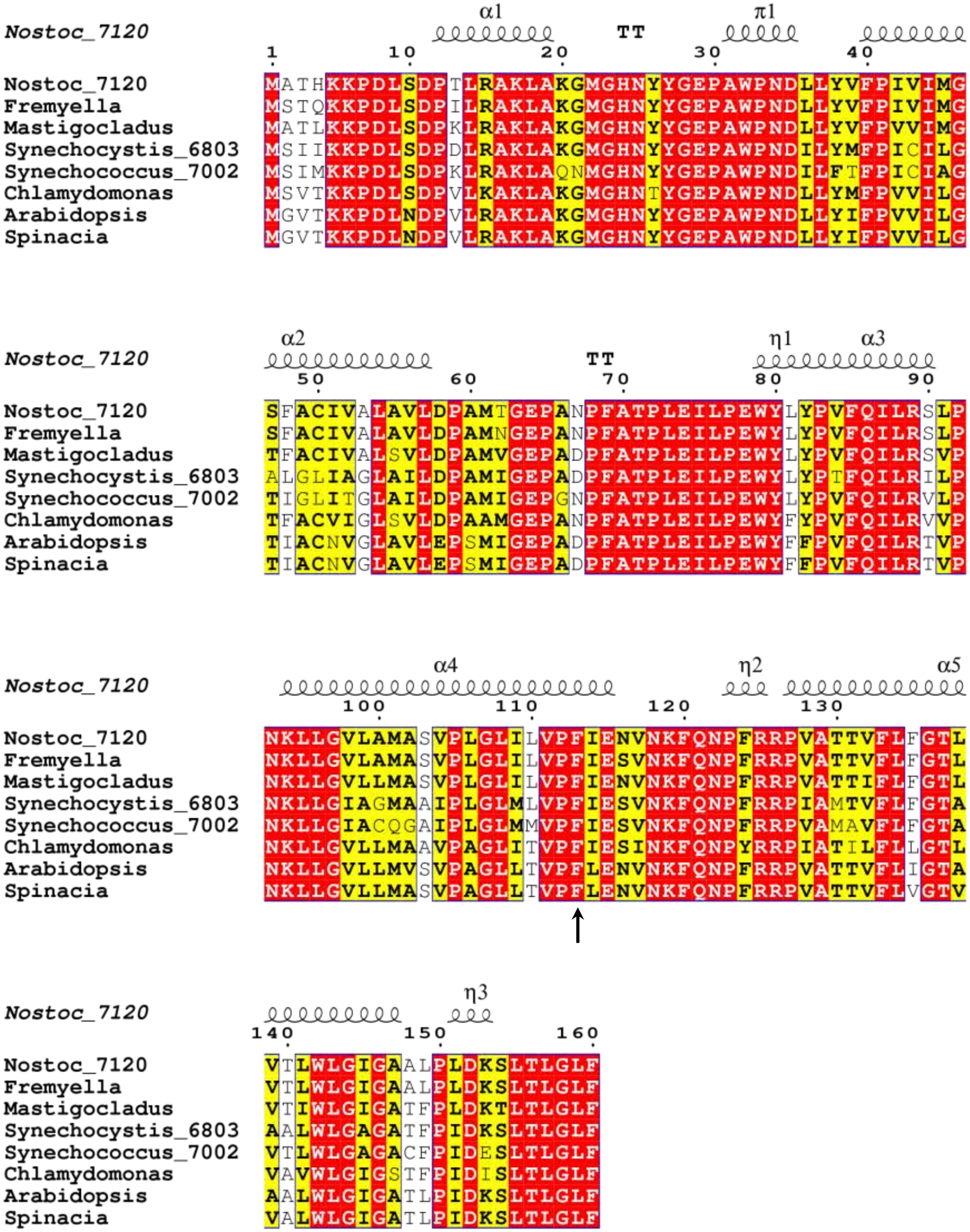
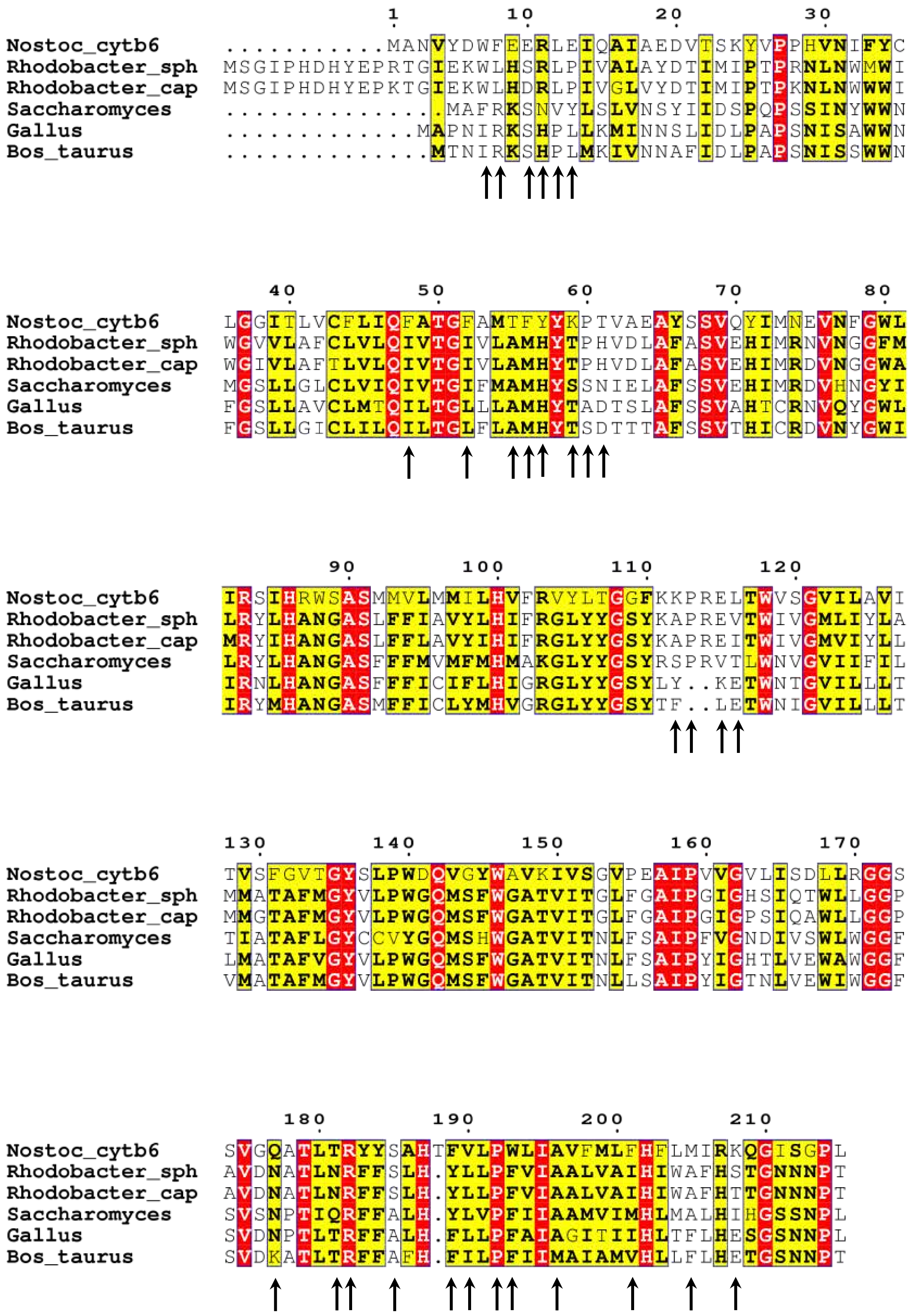
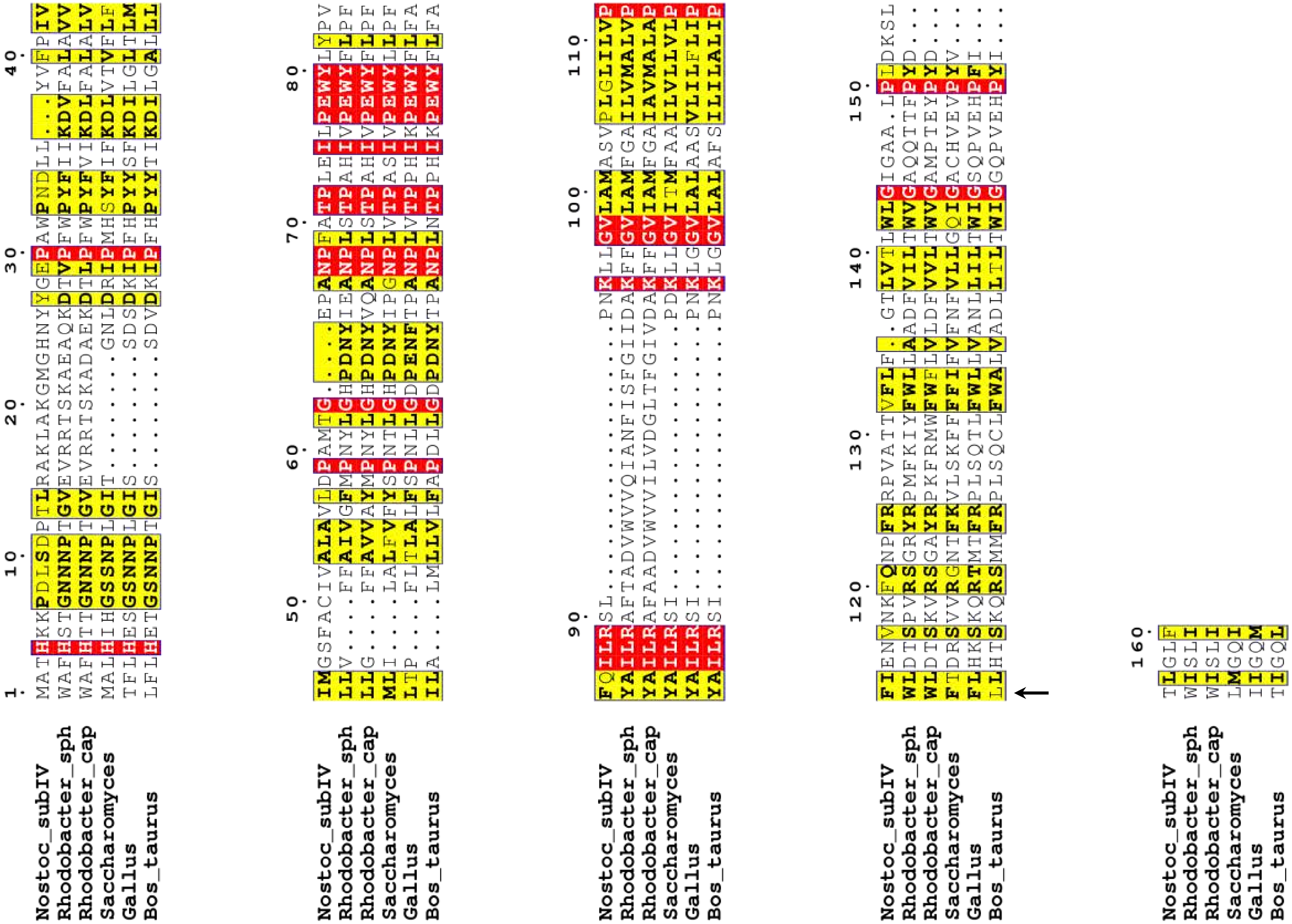

Nostoc_7120
α1
TT
π1
1
10
20
30
40
Nostoc_7120 MATHKKPDLSDPTLRAKLAKGMGHNYYGEPAWPNDLLYVFPIVIMG
Fremyella MSTQKKPDLSDPILRAKLAKGMGHNYYGEPAWPNDLLYVFPIVIMG
Mastigocladus MATLKKPDLSDPKLRAKLAKGMGHNYYGEPAWPNDLLYVFPVVIMG
Synechocystis_6803 MSIIKKPDLSDPDLRAKLAKGMGHNYYGEPAWPNDILYMFPICILG
Synechococcus_7002 MSIMKKPDLSDPKLRAKLAQNMGHNYYGEPAWPNDILFTFPICIAG
Chlamydomonas MSVTKKPDLSDPVLKAKLAKGMGHNTYGEPAWPNDLLYMFPVVILG
Arabidopsis MGVTKKPDLNDPVLRAKLAKGMGHNYYGEPAWPNDLLYIFPVVILG
Spinacia MGVTKKPDLNDPVLRAKLAKGMGHNYYGEPAWPNDLLYIFPVVILG
Nostoc_7120
α2
TT
η1
α3
50
60
70
80
90
Nostoc_7120 SFACIVALAVLDPAMTGEPANPFATPLEILPEWYLYPVFQILRSLP
Fremyella SFACIVALAVLDPAMNGEPANPFATPLEILPEWYLYPVFQILRSLP
Mastigocladus TFACIVALSVLDPAMVGEPADPFATPLEILPEWYLYPVFQILRSVP
Synechocystis_6803 ALGLIAGLAILDPAMIGEPADPFATPLEILPEWYLYPTFQILRILP
Synechococcus_7002 TIGLITGLAILDPAMIGEPGNPFATPLEILPEWYLYPVFQILRVLP
Chlamydomonas TFACVIGLSVLDPAAMGEPANPFATPLEILPEWYFYPVFQILRVVP
Arabidopsis TIACNVGLAVLEPSMIGEPADPFATPLEILPEWYFFPVFQILRTVP
Spinacia TIACNVGLAVLEPSMIGEPADPFATPLEILPEWYFFPVFQILRTVP
Nostoc_7120
α4
η2
α5
100
110
120
130
Nostoc_7120 NKLLGVLAMASVPLGLILVPFIENVNKFQNPFRRPVATTVFLFGTL
Fremyella NKLLGVLAMASVPLGLILVPFIESVNKFQNPFRRPVATTVFLFGTL
Mastigocladus NKLLGVLLMASVPLGLILVPFIENVNKFQNPFRRPVATTIFLFGTL
Synechocystis_6803 NKLLGIAGMAAIPLGLMLVPFIESVNKFQNPFRRPIAMTVFLFGTA
Synechococcus_7002 NKLLGIACQGAIPLGLMMVPFIESVNKFQNPFRRPVAMAVFLFGTA
Chlamydomonas NKLLGVLLMAAVPAGLITVPFIESINKFQNPYRRPIATILFLLGTL
Arabidopsis NKLLGVLLMVSVPAGLLTVPFLENVNKFQNPFRRPVATTVFLIGTA
Spinacia NKLLGVLLMASVPAGLLTVPFLENVNKFQNPFRRPVATTVFLVGTV
↑
Nostoc_7120
η3
140
150
160
Nostoc_7120 VTLWLGIGAALPLDKSLTLGLF
Fremyella VTLWLGIGAALPLDKSLTLGLF
Mastigocladus VTIWLGIGATFPLDKTLTLGLF
Synechocystis_6803 AALWLGAGATFPIDKSLTLGLF
Synechococcus_7002 VTLWLGAGACFPIDESLTLGLF
Chlamydomonas VAVWLGIGSTFPIDISLTLGLF
Arabidopsis AALWLGIGATLPIDKSLTLGLF
Spinacia VALWLGIGATLPIDKSLTLGLF

Figure S12



```
Nostoc_cytb6        ...........MANVYDWFEERLEIQAIAEDVTSKYVPPHVNIFYC
Rhodobacter_sph     MSGIPHDHYEPRTGIEKWLHSRLPIVALAYDTIMIPTPRNLNWMWI
Rhodobacter_cap     MSGIPHDHYEPKTGIEKWLHDRLPIVGLVYDTIMIPTPKNLNWWWI
Saccharomyces       ...............MAFRKSNVYLSLVNSYIIDSPQPSSINYWWN
Gallus              .............MAPNIRKSHPLLKMINNSLIDLPAPSNISAWWN
Bos_taurus          ..............MTNIRKSHPLMKIVNNAFIDLPAPSNISSWWN

Nostoc_cytb6        LGGITLVCFLIQFATGFAMTFYYKPTVAEAYSSVQYIMNEVNFGWL
Rhodobacter_sph     WGVVLAFCLVLQIVTGIVLAMHYTPHVDLAFASVEHIMRNVNGGFM
Rhodobacter_cap     WGIVLAFTLVLQIVTGIVLAMHYTPHVDLAFASVEHIMRDVNGGWA
Saccharomyces       MGSLLGLCLVIQIVTGIFMAMHYSSNIELAFSSVEHIMRDVHNGYI
Gallus              FGSLLAVCLMTQILTGLLLAMHYTADTSLAFSSVAHTCRNVQYGWL
Bos_taurus          FGSLLGICLILQILTGLFLAMHYTSDTTTAFSSVTHICRDVNYGWI

Nostoc_cytb6        IRSIHRWSASMMVLMMILHVFRVYLTGGFKKPRELTWVSGVILAVI
Rhodobacter_sph     LRYLHANGASLFFIAVYLHIFRGLYYGSYKAPREVTWIVGMLIYLA
Rhodobacter_cap     MRYIHANGASLFFLAVYIHIFRGLYYGSYKAPREITWIVGMVIYLL
Saccharomyces       LRYLHANGASFFFMVMFMHMAKGLYYGSYRSPRVTLWNVGVIIFIL
Gallus              IRNLHANGASFFFICIFLHIGRGLYYGSYLY..KETWNTGVILLLT
Bos_taurus          IRYMHANGASMFFICLYMHVGRGLYYGSYTF..LETWNIGVILLLT

Nostoc_cytb6        TVSFGVTGYSLPWDQVGYWAVKIVSGVPEAIPVVGVLISDLLRGGS
Rhodobacter_sph     MMATAFMGYVLPWGQMSFWGATVITGLFGAIPGIGHSIQTWLLGGP
Rhodobacter_cap     MMGTAFMGYVLPWGQMSFWGATVITGLFGAIPGIGPSIQAWLLGGP
Saccharomyces       TIATAFLGYCCVYGQMSHWGATVITNLFSAIPFVGNDIVSWLWGGF
Gallus              LMATAFVGYVLPWGQMSFWGATVITNLFSAIPYIGHTLVEWAWGGF
Bos_taurus          VMATAFMGYVLPWGQMSFWGATVITNLLSAIPYIGTNLVEWIWGGF

Nostoc_cytb6        SVGQATLTRYYSAHTFVLPWLIAVFMLFHFLMIRKQGISGPL
Rhodobacter_sph     AVDNATLNRFFSLH.YLLPFVIAALVAIHIWAFHSTGNNNPT
Rhodobacter_cap     AVDNATLNRFFSLH.YLLPFVIAALVAIHIWAFHTTGNNNPT
Saccharomyces       SVSNPTIQRFFALH.YLVPFIIAAMVIMHLMALHIHGSSNPL
Gallus              SVDNPTLTRFFALH.FLLPFAIAGITIIHLTFLHESGSNNPL
Bos_taurus          SVDKATLTRFFAFH.FILPFIIMAIAMVHLLFLHETGSNNPT
```

Figure S13


Nostoc_subIV
Rhodobacter_sph
Rhodobacter_cap
Saccharomyces
Gallus
Bos_taurus