Supplementary Material (ESI) for Molecular BioSystems
This journal is (c) The Royal Society of Chemistry, 2010

Supplementary Material

Residue-specific global fluorination of *Candida antarctica* lipase B in *Pichia pastoris*

Nediljko Budisa[a], Waltraud Wenger[a], Birgit Wiltschi*[abc]

[a] Max-Planck-Institute of Biochemistry
Am Klopferspitz 18
82152 Martinsried
Germany

[b] bioss – Centre for Biological Signalling Studies
Albert-Ludwigs-University Freiburg
Engesserstrasse 4b
79108 Freiburg
Germany
E-mail: birgit.wiltschi@biologie.uni-freiburg.de
Phone: ++49 761 203 97656
Fax: ++49 761 203 97660

[c] Faculty of Biology
Albert-Ludwigs-University Freiburg
Schänzlestrasse 1
79104 Freiburg
Germany

*corresponding author (birgit.wiltschi@biologie.uni-freiburg.de)

Figure S1 Amino acid sequence of CalB N74D. The mature protein is highlighted in gray with the N74D mutation boxed. The EKR↓EA*EA* peptide in the alpha-mating factor pre–pro leader sequence that is cleaved by Kex2p (↓) and Ste13p (*) is in bold, and EF originating from the *Eco*RI site on plasmid pPICZαA is underlined.

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGLLFINTTIASIAAKEEGV
SL**EKREAEA**EFLPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSNWIPLSAQLGYTPCWISPP
PFMLDDTQVNTEYMVNAITTLYAGSGNNKLPVLTWSQGGLVAQWGLTFFPSIRSKVDRLMAFAPDYKGTVLAGPLDALAV
SAPSVWQQTTGSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSNSPLDSSYLFNGKNVQAQAVCGPLFVIDHAGSLT
SQFSYVVGRSALRSTTGQARSADYGITDCNPLPANDLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARPFAVGKR
TCSGIVTP

Figure S2 Amino acid composition of correctly Kex2p/Ste13p processed CalB N74D (319 residues). Tryptophan (W), tyrosine (Y), or phenylalanine (F) residues are highlighted. Calculation by ExPaSy ProtParam tool.

Ala	(A)	36	11.3%
Arg	(R)	8	2.5%
Asn	(N)	13	4.1%
Asp	(D)	15	4.7%
Cys	(C)	6	1.9%
Gln	(Q)	18	5.6%
Glu	(E)	5	1.6%
Gly	(G)	26	8.2%
His	(H)	1	0.3%
Ile	(I)	11	3.4%
Leu	(L)	31	9.7%
Lys	(K)	9	2.8%
Met	(M)	4	1.3%
Phe	**(F)**	**11**	**3.4%**
Pro	(P)	30	9.4%
Ser	(S)	31	9.7%
Thr	(T)	27	8.5%
Trp	**(W)**	**5**	**1.6%**
Tyr	**(Y)**	**9**	**2.8%**
Val	(V)	23	7.2%

Sequence distribution of the tryptophan residues:

EFLPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSN**W**IPLSAQLGYTPC**W**ISPPPFMLDDTQV
NTEYMVNAITTLYAGSGNNKLPVLT**W**SQGGLVAQ**W**GLTFFPSIRSKVDRLMAFAPDYKGTVLAGPLDALAVSAPSV**W**QQT
TGSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSNSPLDSSYLFNGKNVQAQAVCGPLFVIDHAGSLTSQFSYVVGR
SALRSTTGQARSADYGITDCNPLPANDLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARPFAVGKRTCSGIVTP

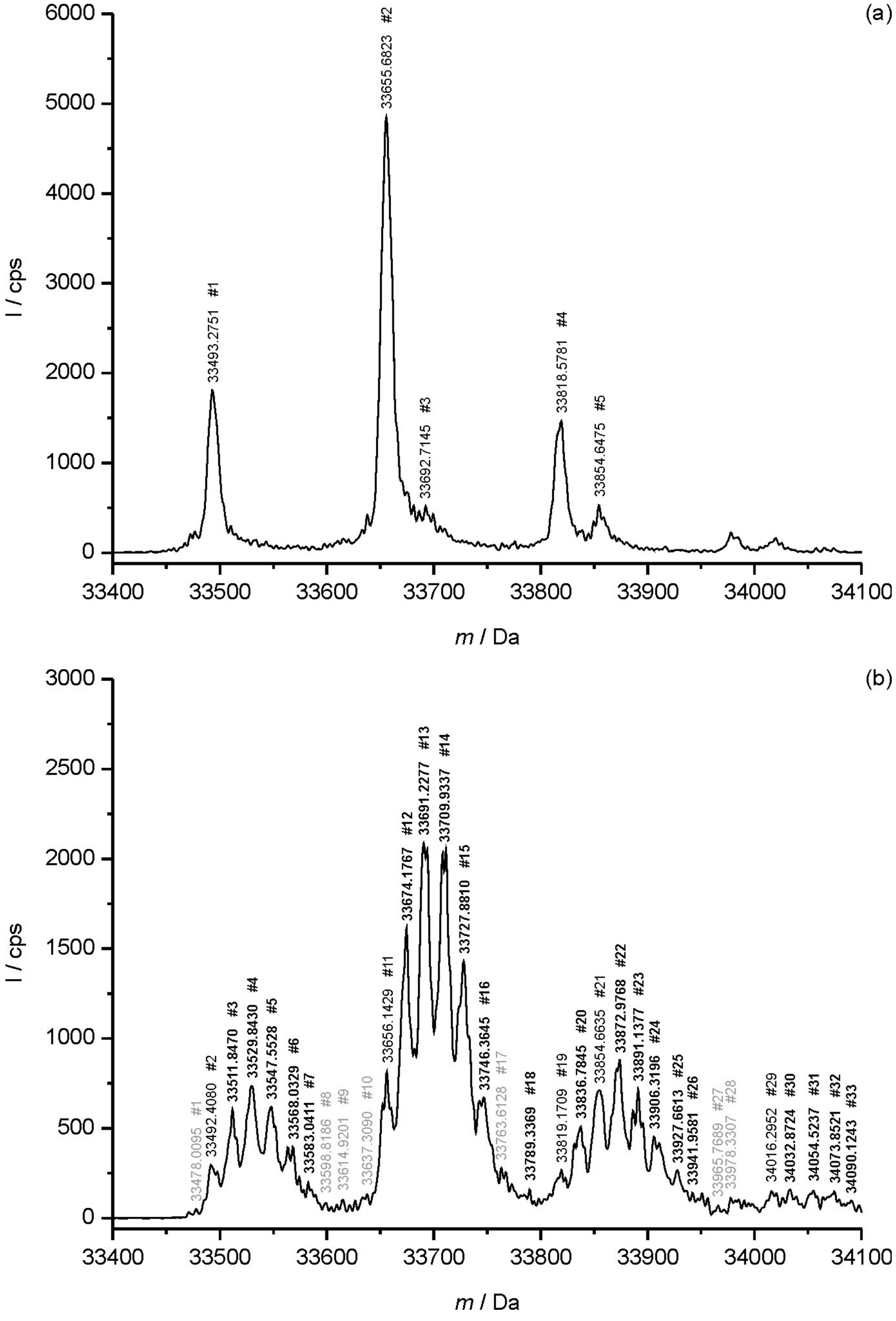
Sequence distribution of the tyrosine residues:

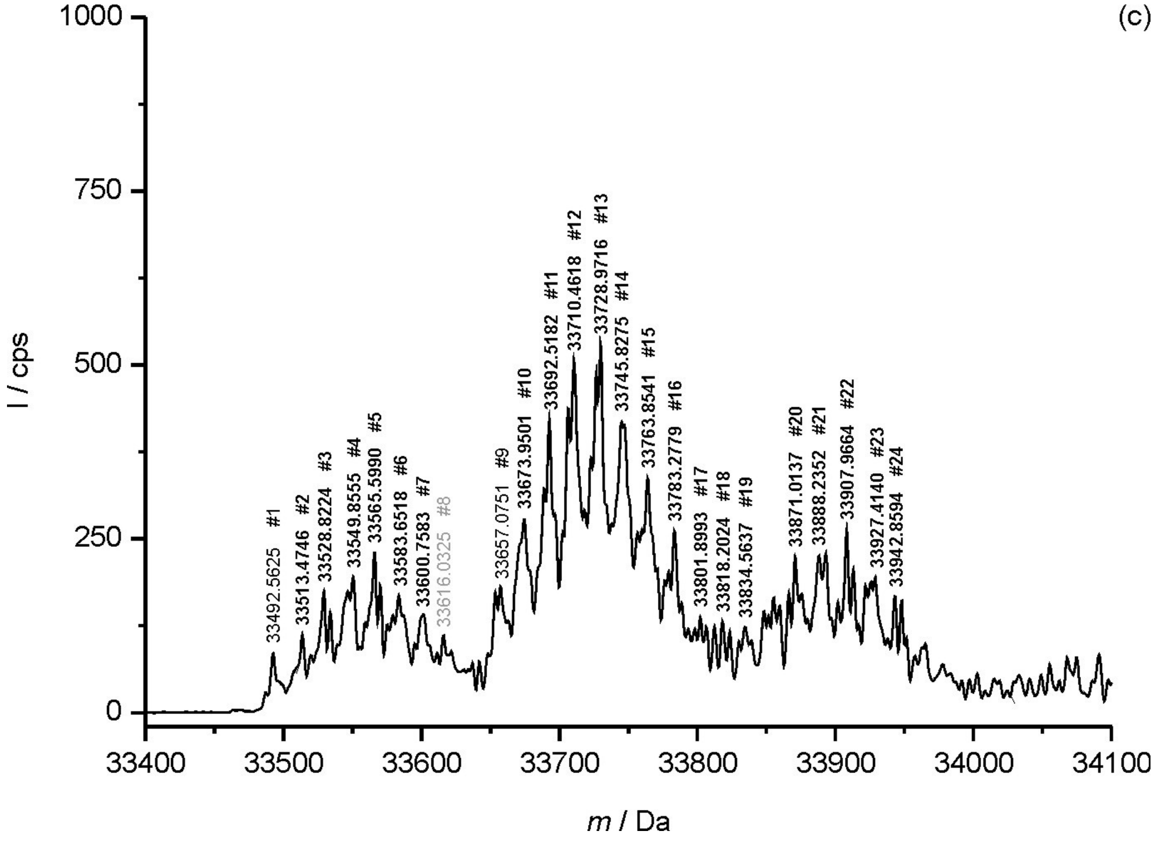
EFLPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSNWIPLSAQLG**Y**TPCWISPPPFMLDDTQV
NTE**Y**MVNAITTL**Y**AGSGNNKLPVLTWSQGGLVAQWGLTFFPSIRSKVDRLMAFAPD**Y**KGTVLAGPLDALAVSAPSVWQQT
TGSALTTALRNAGGLTQIVPTTNL**Y**SATDEIVQPQVSNSPLDSS**Y**LFNGKNVQAQAVCGPLFVIDHAGSLTSQFS**Y**VVGR
SALRSTTGQARSAD**Y**GITDCNPLPANDLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMP**Y**ARPFAVGKRTCSGIVTP

Sequence distribution of the phenylalanine residues:

E**F**LPSGSDPA**F**SQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQS**F**DSNWIPLSAQLGYTPCWISPPP**F**MLDDTQV
NTEYMVNAITTLYAGSGNNKLPVLTWSQGGLVAQWGLT**FF**PSIRSKVDRLMA**F**APDYKGTVLAGPLDALAVSAPSVWQQT
TGSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSNSPLDSSYL**F**NGKNVQAQAVCGPL**F**VIDHAGSLTSQ**F**SYVVGR
SALRSTTGQARSADYGITDCNPLPANDLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARP**F**AVGKRTCSGIVTP

Figure S3 Full range mass spectra of the different variant preparations. (a) WYF-CalB N74D; (b) 5FW-CalB N74D; (c) mFY-CalB N74D; (d) pFF-CalB N74D. Mass peaks assigned to fluorinated protein species are shown in bold black. Peaks corresponding to parent protein are in black and undefined peaks are indicated in gray. For identification of the mass peaks refer to Table S1. I, intensity; cps, counts per second; m, mass.




(c)
I / cps
1000
750
500
250
0
33400
33500
33600
33700
33800
33900
34000
34100
m / Da
33492.5625 #1
33513.4746 #2
33528.8224 #3
33549.8555 #4
33565.5990 #5
33583.6518 #6
33600.7583 #7
33616.0325 #8
33657.0751 #9
33673.9501 #10
33692.5182 #11
33710.4618 #12
33728.9716 #13
33745.8275 #14
33763.8541 #15
33783.2779 #16
33801.8993 #17
33818.2024 #18
33834.5637 #19
33871.0137 #20
33888.2352 #21
33907.9664 #22
33927.4140 #23
33942.8594 #24

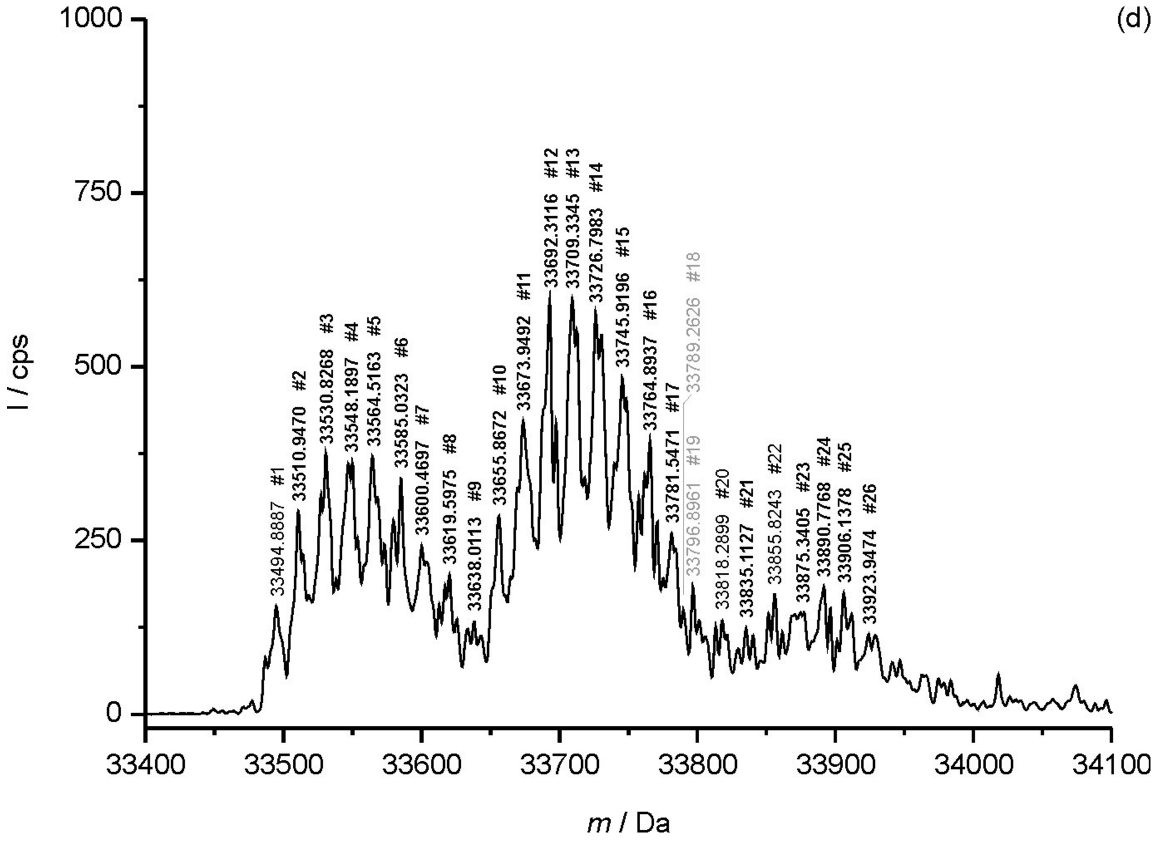

(d)
I / cps
1000
750
500
250
0
33400
33500
33600
33700
33800
33900
34000
34100
m / Da
33494.8887 #1
33510.9470 #2
33530.8268 #3
33548.1897 #4
33564.5163 #5
33585.0323 #6
33600.4697 #7
33619.5975 #8
33638.0113 #9
33655.8672 #10
33673.9492 #11
33692.3116 #12
33709.3345 #13
33726.7983 #14
33745.9196 #15
33764.8937 #16
33781.5471 #17
33789.2626 #18
33796.8961 #19
33818.2899 #20
33835.1127 #21
33855.8243 #22
33875.3405 #23
33890.7768 #24
33906.1378 #25
33923.9474 #26

Figure S4 Proteinase K digestion of (a) WYF-CalB N74D, (b) 5FW-CalB N74D, (c) mFY-CalB N74D, and (d) pFF-CalB N74D. Equal amounts of the individual CalB N74D variants were loaded undigested (lane 1) or digested with 4 µg (lane 2), 0.4 µg (lane 3), 0.04 µg (lane 4), 0.004 µg (lane 5), or 0.0004 µg proteinase K (lane 6). 4 µg proteinase K were loaded onto lane 7. The bands of the molecular weight marker (M) are indicated on the left margin each. The calculated molecular weight of CalB N74D is 33 kDa and that of proteinase K is 29 kDa. Experimental details are given in the Materials and methods section.

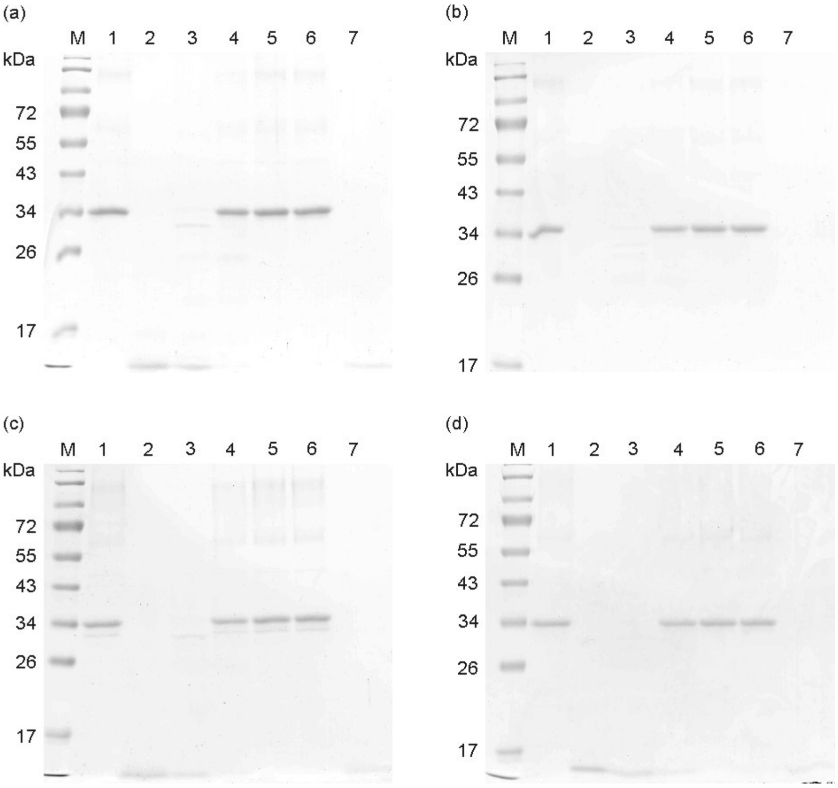


Table S1 Interpretation of the mass peaks in Fig. S3. The EKR↓EA*EA*EF peptide in the alpha-mating factor pre–pro leader sequence is cleaved by Kex2p (↓) and Ste13p (*), EF originates from cloning of the CalB N74D coding sequence into the *Eco*RI site on plasmid pPICZαA. Masses were calculated assuming reduced cysteine residues; Δm denotes the mass difference between found mass and calculated mass; ?, uncertain interpretation as Δm is lower than -7 Da or is positive; peak numbers are as indicated in the corresponding mass spectra in Fig. S3a-d. Hex, attached hexose sugar residue of about 162 Da[1]; Kex2p(i), incorrect Kex2p digestion; S-S, disulfide bridge. Mass peaks allocated to CalB N74D variants containing an analog are indicated in bold, peaks corresponding to parent proteins are in standard print and unidentified peaks are shown in gray.

CalB N74D variant	found mass (Da)	peak	calculated mass (Da)	Δm (Da)	interpretation	N-terminus
WYF (Fig. S3a)	33493.2751	#1	33499.6024	-6	WYF, 3 S-S	EAEF
	33655.6823	#2	33661.7432	-6 (+162 to #1)	WYF, 3 S-S, 1 Hex	EAEF
	33692.7145	#3	33699.7947	-7	WYF, 3 S-S	EAEAEF
	33818.5781	#4	33823.8841	-5 (+163 to #2)	WYF, 3 S-S, 2 Hex	EAEF
	33854.6475	#5	33855.9807	-1	WYF, Kex2p(i)	REAEAEF
5FW (Fig. S3b)	33478.0095	#1	-			
	33492.4080	#2	33499.6024	-7	WYF, 3 S-S	EAEF
	33511.8470	**#3**	**33517.5928**	**-6**	**1 5FW, 3 S-S**	**EAEF**
	33529.8430	**#4**	**33535.5833**	**-6**	**2 5FW, 3 S-S**	**EAEF**
	33547.5528	**#5**	**33553.5738**	**-6**	**3 5FW, 3 S-S**	**EAEF**
	33568.0329	**#6**	**33571.5642**	**-4**	**4 5FW, 2 S-S**	**EAEF**
	33583.0411	**#7**	**33589.5547**	**-7**	**5 5FW, 3 S-S**	**EAEF**
	33598.8186	#8	-			
	33614.9201	#9	-			
	33637.3090	#10	-			
	33656.1429	#11	33661.7432	-6 (+164 to #2)	WYF, 3 S-S, 1 Hex	EAEF
	33674.1767	**#12**	**33679.7337**	**-6 (+163 to #3)**	**1 5FW, 3 S-S, 1 Hex**	**EAEF**
	33691.2277	**#13**	**33697.7242**	**-6 (+162 to #4)**	**2 5FW, 3 S-S, 1 Hex**	**EAEF**
	33709.9337	**#14**	**33715.7146**	**-6 (+162 to #5)**	**3 5FW, 3 S-S, 1 Hex**	**EAEF**
	33727.8810	**#15**	**33733.7051**	**-6 (+159 to #6)**	**4 5FW, 2 S-S, 1 Hex**	**EAEF**
	33746.3645	**#16**	**33751.6955**	**-5 (+163 to #7)**	**5 5FW, 3 S-S, 1 Hex**	**EAEF**
	33763.6128	#17	-			
	33789.3369	**#18**	**33789.7470**	**0**	**5 5FW**	**EAEAEF**
	33819.1709	#19	33823.8841	-5 (+163 to #11)	WYF, 3 S-S, 2 Hex	EAEF
	33836.7845	**#20**	**33841.8745**	**-5 (+162 to #12)**	**1 5FW, 3 S-S, 2 Hex**	**EAEF**
	33854.6635	#21	33855.9807	-1	WYF, Kex2p(i)	REAEAEF
	33854.6635	**#21**	**33859.8650**	**-5 (+163 to #13)**	**2 5FW, 3 S-S, 2 Hex**	**EAEF**
	33872.9768	**#22**	**33873.9712**	**-1**	**1 5FW, Kex2p(i)**	**REAEAEF**
	33891.1377	**#23**	**33891.9617**	**-1**	**2 5FW, Kex2p(i)**	**REAEAEF**

	33906.3196	**#24**	**33909.9521**	**-4**	**3 5FW, 2 S-S, Kex2p(i)**	**REAEAEF**	
	33927.6613	**#25**	**33927.9426**	**0**	**4 5FW, Kex2p(i)**	**REAEAEF**	
	33941.9581	**#26**	**33945.9331**	**-4**	**5 5FW, 2 S-S, Kex2p(i)**	**REAEAEF**	
	33965.7689	#27	-				
	33978.3307	#28	-				
	34016.2952	#29	34018.1216	-2 (+162 to #21)	WYF, 1 S-S, Kex2p(i), 1 Hex	REAEAEF	
	34032.8724	**#30**	**34036.1121**	**-3 (+160 to #22)**	**1 5FW, 2 S-S, Kex2p(i), 1 Hex**	**REAEAEF**	
	34054.5237	**#31**	**34054.1025**	**0 (+163 to #23)**	**2 5FW, Kex2p(i), 1 Hex**	**REAEAEF**	
	34073.8521	**#32**	**34072.0930**	**+2 (+168 to #24)**	**3 5FW, Kex2p(i), 1 Hex**	**REAEAEF**	**?**
	34090.1243	**#33**	**34090.0834**	**0 (+162 to #25)**	**4 5FW, Kex2p(i), 1 Hex**	**REAEAEF**	
mFY (Fig. S3c)	33492.5625	#1	33499.6024	-7	WFY, 3 S-S	EAEF	
	33513.4746	**#2**	**33517.5928**	**-4**	**1 mFY, 2 S-S**	**EAEF**	
	33528.8224	**#3**	**33535.5833**	**-7**	**2 mFY, 3 S-S**	**EAEF**	
	33549.8555	**#4**	**33553.5738**	**-4**	**3 mFY, 2 S-S**	**EAEF**	
	33565.5990	**#5**	**33571.5642**	**-6**	**4 mFY, 3 S-S**	**EAEF**	
	33583.6518	**#6**	**33589.5547**	**-6**	**5 mFY, 3 S-S**	**EAEF**	
	33600.7583	**#7**	**33607.5452**	**-7**	**6 mFY, 3 S-S**	**EAEF**	
	33616.0325	#8	-				
	33657.0751	#9	33661.7432	-5 (+165 to #1)	WFY, 3 S-S, 1 Hex	EAEF	
	33673.9501	**#10**	**33679.7337**	**-6 (+160 to #2)**	**1 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33692.5182	**#11**	**33697.7242**	**-5 (+164 to #3)**	**2 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33710.4618	**#12**	**33715.7146**	**-5 (+161 to #4)**	**3 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33728.9716	**#13**	**33733.7051**	**-5 (+163 to #5)**	**4 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33745.8275	**#14**	**33751.6955**	**-6 (+162 to #6)**	**5 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33763.8541	**#15**	**33769.6860**	**-6 (+163 to #7)**	**6 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33783.2779	**#16**	**33787.6765**	**-4**	**7 mFY, 2 S-S, 1 Hex**	**EAEF**	
	33801.8993	**#17**	**33805.6669**	**-4**	**8 mFY, 2 S-S, 1 Hex**	**EAEF**	
	33818.2024	**#18**	**33823.6574**	**-5**	**9 mFY, 3 S-S, 1 Hex**	**EAEF**	
	33818.2024	#18	33823.8841	-6 (+161 to #9)	WYF, 3 S-S, 2 Hex	EAEF	
	33834.5637	**#19**	**33841.8745**	**-7 (+161 to #10)**	**1 mFY, 3 S-S, 2 Hex**	**EAEF**	
	33871.0137	**#20**	**33873.9712**	**-3**	**1 mFY, 2 S-S**	**REAEAEF**	
	33888.2352	**#21**	**33891.9617**	**-4**	**2 mFY, 2 S-S**	**REAEAEF**	
	33907.9664	**#22**	**33909.9521**	**-2**	**3 mFY, 1 S-S**	**REAEAEF**	
	33907.9664	**#22**	**33913.8364**	**-6 (+162 to #14)**	**5 mFY, 3 S-S, 2 Hex**	**EAEF**	
	33927.4140	**#23**	**33927.9426**	**-1**	**4 mFY**	**REAEAEF**	
	33942.8594	**#24**	**33945.9331**	**-3**	**5 mFY, 2 S-S**	**REAEAEF**	

pFF (Fig. S3d)	33494.8887	#1	33499.6024	-5	WYF, 3 S-S	EAEF	
	33510.9470	**#2**	**33517.5928**	**-7**	**1 pFF, 3 S-S**	**EAEF**	
	33530.8268	**#3**	**33535.5833**	**-5**	**2 pFF, 3 S-S**	**EAEF**	
	33548.1897	**#4**	**33553.5738**	**-5**	**3 pFF, 3 S-S**	**EAEF**	
	33564.5163	**#5**	**33571.5642**	**-7**	**4 pFF, 3 S-S**	**EAEF**	
	33585.0323	**#6**	**33589.5547**	**-5**	**5 pFF, 3 S-S**	**EAEF**	
	33600.4697	**#7**	**33607.5452**	**-7**	**6 pFF, 3 S-S**	**EAEF**	
	33619.5975	**#8**	**33625.5356**	**-6**	**7 pFF, 3 S-S**	**EAEF**	
	33638.0113	**#9**	**33643.5261**	**-6**	**8 pFF, 3 S-S**	**EAEF**	
	33655.8672	**#10**	**33661.5165**	**-6**	**9 pFF, 3 S-S**	**EAEF**	
	33655.8672	#10	33661.7432	-6 (+161 to #1)	WYF, 3 S-S, 1 Hex	EAEF	
	33673.9492	**#11**	**33679.5070**	**-6**	**10 pFF, 3 S-S**	**EAEF**	
	33673.9492	**#11**	**33679.7337**	**-6 (+162 to #2)**	**1 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33692.3116	**#12**	**33697.4975**	**-5**	**11 pFF, 3 S-S**	**EAEF**	
	33692.3116	**#12**	**33697.7242**	**-5 (+161 to #3)**	**2 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33709.3345	**#13**	**33715.7146**	**-6**	**3 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33726.7983	**#14**	**33733.7051**	**-7**	**4 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33745.9196	**#15**	**33751.6955**	**-6**	**5 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33764.8937	**#16**	**33769.6860**	**-5**	**6 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33781.5471	**#17**	**33787.6765**	**-6**	**7 pFF, 3 S-S, 1 Hex**	**EAEF**	
	33789.2626	#18					
	33796.8961	#19					
	33818.2899	#20	33823.8841	-6 (+162 to #10)	WYF, 3 S-S, 2 Hex	EAEF	
	33835.1127	**#21**	**33841.8745**	**-7 (+161 to #11)**	**1 pFF, 3 S-S, 2 Hex**	**EAEF**	
	33855.8243	#22	33855.9807	0	WYF, Kex2p(i)	REAEAEF	
	33855.8243	**#22**	**33859.8650**	**-4 (+164 to #12)**	**2 pFF, 2 S-S, 2 Hex**	**EAEF**	
	33875.3405	**#23**	**33877.8555**	**-3 (+166 to #13)**	**3 pFF, 2 S-S, 2 Hex**	**EAEF**	
	33890.7768	**#24**	**33895.8459**	**-5 (+164 to #14)**	**4 pFF, 3 S-S, 2 Hex**	**EAEF**	
	33906.1378	**#25**	**33913.8364**	**-8 (+160 to #15)**	**5 pFF, 3 S-S, 2 Hex**	**EAEF**	**?**
	33906.1378	**#25**	**33909.9521**	**-4**	**3 pFF, 2 S-S, Kex2p(i)**	**REAEAEF**	
	33923.9474	**#26**	**33931.8269**	**-8 (+159 to #16)**	**6 pFF, 3 S-S, 2 Hex**	**EAEF**	**?**
	33923.9474	**#26**	**33927.9426**	**-4**	**4 pFF, 2 S-S, Kex2p(i)**	**REAEAEF**	

1 D. W. Speicher, *Curr. Protoc. Protein Sci.*, 2002, **Chapter 16**, Unit 16.11.