SUPPLEMENTARY INFORMATION

Design and characterisation of mutant and wild-type huntingtin proteins produced from a toolkit of scalable eukaryotic expression systems

Rachel J. Harding[1]*, Peter Loppnau[1], Suzanne Ackloo[1], Alexander Lemak[2], Ashley Hutchinson[1], Brittany Hunt[1], Alex S. Holehouse[3], Jolene C. Ho[1], Lixin Fan[4], Leticia Toledo-Sherman[5], Alma Seitova[1], Cheryl H. Arrowsmith[1,2]*

[1]Structural Genomics Consortium, University of Toronto, MaRS South Tower, 101 College Street, Toronto, Ontario M5G 1L7, Canada.
[2]Princess Margaret Cancer Centre and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 1L7, Canada
[3]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University in Saint Louis, Saint Louis, Missouri 63130, USA
[4]Center for Cancer Research, National Cancer Institute, Advanced Technology Research Facility, Frederick, MD 21701, USA
[5]CHDI Foundation, 6080 Center Drive, Suite 700, Los Angeles, CA 90045, USA

***Corresponding authors:**

Rachel J. Harding
Tel,: 416-946-3795
Email: Rachel.harding@utoronto.ca

Cheryl Arrowsmith
Tel.: 416-946-0881
E-mail: Cheryl.Arrowsmith@uhnresearch.ca

Table S1. Database search results for HTT$^{1\text{-}3144}$ Q23 expressed in EXPI293F cells. Of the peptides which do not correspond to HTT protein, most either have very high scores in the CRAPome (Mellacheruvu et al., 2013), suggesting that these are non-specific contaminants of the co-immunoprecipitation step, not true HTT interactors, or they are very low abundance.

Protein Accession	Total no. spectra	No. unique peptides	Protein Accession	Total no. spectra	No. unique peptides	Protein Accession	Total no. spectra	No. unique peptides	Protein Accession	Total no. spectra	No. unique peptides
P42858|HD	7681	1150	P62805|H4	12	2	P09131|P3	5	1	Q7RTP6|MICA3	2	1
Q13748|TBA3C	175	29	Q14568|HS902	12	5	Q13576|IQGA2	5	2	Q86Y39|NDUAB	2	1
Q13885|TBB2A	111	26	Q5CZC0|FSIP2	12	5	Q15746|MYLK	5	1	Q8IWK6|AGRA3	2	1
Q9BVA1|TBB2B	111	26	Q5S007|LRRK2	12	2	Q86VV8|RTTN	5	3	Q8IYE0|CC146	2	1
P07900|HS90A	110	24	Q7Z5J8|ANKAR	12	3	Q8NBX0|SCPDL	5	2	Q8N2C7|UNC80	2	1
Q6PEY2|TBA3E	64	13	Q8TE73|DYH5	12	6	Q8TD26|CHD6	5	2	Q8N4X5|AF1L2	2	1
Q9BYE2|TMPSD	56	7	O96JB1|DYH8	12	4	Q9ULT8|HECD1	5	3	Q8N5H7|SH2D3	2	1
Q9NY65|TBA8	54	9	Q96PX9|PKH4B	12	2	O00189|AP4M1	4	1	Q8TBP6|S2540	2	1
Q71U36|TBA1A	51	12	Q9UDT6|CLIP2	12	3	O95613|PCNT	4	2	Q8WXX0|DYH7	2	1
P04350|TBB4A	48	12	O60890|OPHN1	11	3	P35475|IDUA	4	1	Q969F2|NKD2	2	1
P60709|ACTB	45	13	Q02224|CENPE	11	3	P55072|TERA	4	1	Q96JM4|LRIQ1	2	1
P11142|HSP7C	44	12	O6ZQQ6|WDR87	11	2	Q14997|PSME4	4	2	Q96KP1|EXOC2	2	2
P63261|ACTG	44	13	Q7Z333|SETX	11	4	Q15569|TESK1	4	1	Q96RT1|ERBIN	2	2
P23396|RS3	43	11	Q86UQ4|ABCAD	11	4	Q15643|TRIPB	4	1	Q9BXT5|TEX15	2	1
P54652|HSP72	43	12	Q9Y4F4|TGRM1	11	3	Q8IY17|PLPL6	4	1	Q9C099|LRCC1	2	2
P15880|RS2	38	11	P21817|RYR1	10	2	Q8N2E2|VWDE	4	2	Q9H649|NSUN3	2	1
Q13509|TBB3	36	8	Q12955|ANK3	10	5	Q8N3K9|CMYA5	4	1	Q9NT22|EMIL3	2	2
tr|A0A0B4J269	32	7	Q1XH10|SKDA1	10	2	Q8NHH9|ATLA2	4	1	Q9NYZ4|SIGL8	2	1
P0DMV8|HS71A	31	9	Q5T1B0|AXDN1	10	2	Q8TEX9|IPO4	4	2	Q9P219|DAPLE	2	2
P0DMV9|HS71B	31	9	Q86U86|PB1	10	3	Q93008|USP9X	4	2	Q9UI47|CTNA3	2	1
P07437|TBB5	30	7	Q86VI3|IQGA3	10	3	Q9BXS5|AP1M1	4	1	Q9Y4B5|MTCL1	2	1
Q9BWH6|RPAP1	27	4	Q8IZD9|DOCK3	10	3	Q9H9L3|I20L2	4	1	Q9Y4D8|HECD4	2	1
Q5VYK3|ECM29	26	6	Q13136|LIPA1	9	4	Q9HBW0|LPAR2	4	1	tr|A0A075B6H5	2	1
Q8WXG9|GPR98	26	4	O15751|HERC1	9	3	Q9HCF6|TRPM3	4	1	tr|K7EMT4|K7EMT4	2	1
Q562R1|ACTBL	24	7	Q6PI48|SYDM	9	3	Q9P281|BAHC1	4	2	A4UGR9|XIRP2	1	1
Q9UPN3|MACF1	23	4	Q9HBJ7|UBP29	9	3	Q9UM54|MYO6	4	2	O43918|AIRE	1	1
A6NKT7|RGPD3	22	6	Q9P0L2|MARK1	9	1	Q9Y6Q5|AP1M2	4	1	O60333|KIF1B	1	1
P17066|HSP76	21	7	Q9P217|ZSWM5	9	3	A6NKB5|PCX2	3	2	O60673|REV3L	1	1
Q03001|DYST	21	3	tr|Q5TEC6|Q5TEC6	9	3	O94991|SLIK5	3	2	O75151|PHF2	1	1
P08238|HS90B	20	7	O15078|CE290	8	3	P02786|TFR1	3	1	P01137|TGFB1	1	1
Q9BQG0|MBB1A	20	5	O60391|NMD3B	8	2	P14625|ENPL	3	1	P08779|K1C16	1	1
P11021|GRP78	19	6	O75534|CSDE1	8	2	P25054|APC	3	2	P19012|K1C15	1	1
Q58FG1|HS904	19	4	P49137|MAPK2	8	2	P49759|CLK1	3	1	P22492|H1T	1	1
Q8NET8|TRPV3	19	3	P62269|RS18	8	2	P98161|PKD1	3	1	P38935|SMBP2	1	1
O60814|H2B1K	18	3	Q16695|H31T	8	2	Q03169|TNAP2	3	2	P41218|MNDA	1	1
P06899|H2B1J	18	3	Q8NCG5|CHST4	8	2	Q14674|ESPL1	3	1	P42345|MTOR	1	1
P23527|H2B1O	18	3	Q8WUA7|TB22A	8	2	Q15326|ZMY11	3	2	P50548|ERF	1	1
P33778|H2B1B	18	3	Q9NRC6|SPTN5	8	2	Q6AI39|BICRL	3	1	P80192|M3K9	1	1
P57053|H2BFS	18	3	Q9NRD9|DUOX1	8	2	Q6PL18|ATAD2	3	1	Q02218|ODO1	1	1
P58876|H2B1D	18	3	O15381|NVL	7	2	Q6UWY2|PRS57	3	1	Q03164|KMT2A	1	1
P62807|H2B1C	18	3	Q15021|CND1	7	2	Q7Z6Z7|HUWE1	3	1	Q06141|REG3A	1	1
Q16778|H2B2E	18	3	Q5T9S5|CCD18	7	2	Q86XX4|FRAS1	3	2	Q13395|TARB1	1	1
Q5QNW6|H2B2F	18	3	Q5VZP5|DUS27	7	2	Q92824|PCSK5	3	2	Q14315|FLNC	1	1
Q8N257|H2B3B	18	3	Q6WRI0|IGS10	7	2	Q9BWT7|CAR10	3	2	Q14690|RRP5	1	1
Q93079|H2B1H	18	3	Q709C8|VP13C	7	2	Q9BWU0|NADAP	3	1	Q2Y0W8|S4A8	1	1
Q99877|H2B1N	18	3	Q96JC1|VPS39	7	2	Q9H6E4|CC134	3	2	Q3V6T2|GRDN	1	1
Q99879|H2B1M	18	3	Q9BSJ2|GCP2	7	2	Q9HB40|RISC	3	2	Q5TGY3|AHDC1	1	1
Q99880|H2B1L	18	3	Q9P1Y6|PHRF1	7	2	Q9HCE0|EPG5	3	1	Q6P1M3|L2GL2	1	1
Q12931|TRAP1	16	5	Q9P1Z3|HCN3	7	2	Q9NR09|BIRC6	3	2	Q6P1S2|CC033	1	1
Q8IVF2|AHNK2	16	1	Q9P2E5|CHPF2	7	2	Q9NU22|MDN1	3	2	Q6U841|S4A10	1	1
O14647|CHD2	15	2	Q9UL51|HCN2	7	2	Q9NUT2|ABCB8	3	2	Q6YHK3|CD109	1	1
P04908|H2A1B	14	4	Q9Y3Q4|HCN4	7	2	Q9UK17|KCND3	3	2	Q6ZN66|GBP6	1	1
P0C0S5|H2AZ	14	4	A6NE52|WDR97	6	1	Q9UNA1|RHG26	3	1	Q7Z2D5|PLPR4	1	1
P0C0S8|H2A1	14	4	O43314|VIP2	6	1	Q9Y2H2|SAC2	3	1	Q7Z417|NUFP2	1	1
P16104|H2AX	14	4	O43896|KIF1C	6	2	Q9Y462|ZN711	3	1	Q7Z4S6|KI21A	1	1
P20671|H2A1D	14	4	P05164|PERM	6	2	O15230|LAMA5	2	1	Q86XH1|IQCA1	1	1
P34931|HS71L	14	5	P08670|VIME	6	1	O75665|OFD1	2	2	Q8NDM7|CFA43	1	1
Q16777|H2A2C	14	4	P19823|ITIH2	6	2	O95153|RIMB1	2	1	Q8NEN0|ARMC2	1	1
Q6FI13|H2A2A	14	4	Q13608|PEX6	6	1	P14136|GFAP	2	1	Q8NF91|SYNE1	1	1
Q71UI9|H2AV	14	4	Q13683|ITA7	6	1	P18124|RL7	2	1	Q8NI77|KI18A	1	1
Q7L7L0|H2A3	14	4	Q2LD37|K1109	6	1	P23458|JAK1	2	1	Q8TF72|SHRM3	1	1
Q93077|H2A1C	14	4	Q5JV73|FRPD3	6	3	P42704|LPPRC	2	1	Q96M83|CCDC7	1	1
Q96KK5|H2A1H	14	4	Q5JWR5|DOP1	6	1	P46100|ATRX	2	1	Q9H2X0|CHRD	1	1
Q96QV6|H2A1A	14	4	Q6N069|NAA16	6	2	P61163|ACTZ	2	2	Q9H799|CE042	1	1
Q99878|H2A1J	14	4	Q86XA9|HTR5A	6	1	P68104|EF1A1	2	1	Q9HBB8|CDHR5	1	1
Q9BTM1|H2AJ	14	4	Q8NCM8|DYHC2	6	1	Q03112|EVI1	2	1	Q9HBG6|IF122	1	1
Q9H2M9|RBGPR	14	2	Q8TD57|DYH3	6	2	Q05639|EF1A2	2	1	Q9P1Y5|CAMP3	1	1
Q9NY93|DDX56	14	3	Q92997|DVL3	6	1	Q0VDD8|DYH14	2	1	Q9ULV0|MYO5B	1	1
Q9P225|DYH2	14	4	Q96A08|H2B1A	6	1	Q13315|ATM	2	2	Q9Y5B0|CTDP1	1	1
Q9UKN7|MYO15	14	3	Q96QT4|TRPM7	6	1	Q14999|CUL7	2	1	Q9Y5I0|PCDAD	1	1
Q58FF7|H90B3	13	5	Q99996|AKAP9	6	3	Q14D04|MELT	2	1	Q9Y5Y9|SCNAA	1	1
Q8IVL1|NAV2	13	3	Q9HCK8|CHD8	6	3	Q5VTE0|EF1A3	2	1	Q9Y618|NCOR2	1	1
Q9H853|TBA4B	13	3	Q9NYC9|DYH9	6	1	Q6V0I7|FAT4	2	2	Q9Y6F7|CDY2	1	1
P02549|SPTA1	12	2	Q9P2D1|CHD7	6	3	Q6ZR08|DYH12	2	1			

Figure S1 – FLAG-resin purified C-terminally FLAG-tagged HTT samples derived from EXPI293 expression, either by transient transfection or baculoviral transduction in EXPI293 cells or baculoviral transduction in Sf9 cells, were subject to Western Blot analysis with ab109115 which binds an epitope at amino acids 1-100.

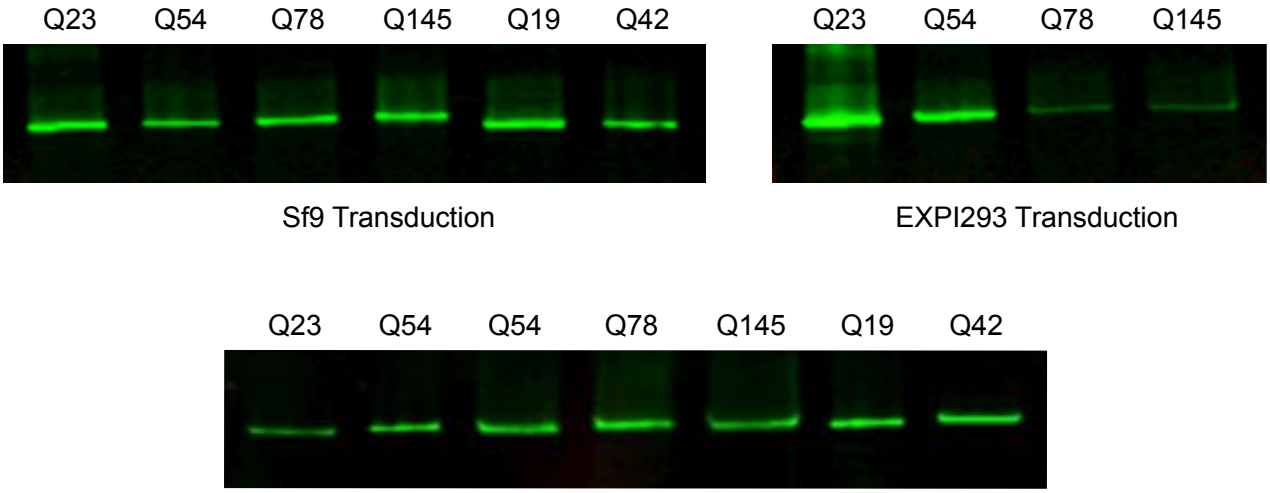


Figure S2. Sequence coverage maps of Sf9 HTT$^{1\text{-}3144}$ Q23 for A) five enzymes combined (97%) and B) trypsin and lysargiNase combined (66%) as obtained from bottom-up proteomics experiments.

A)

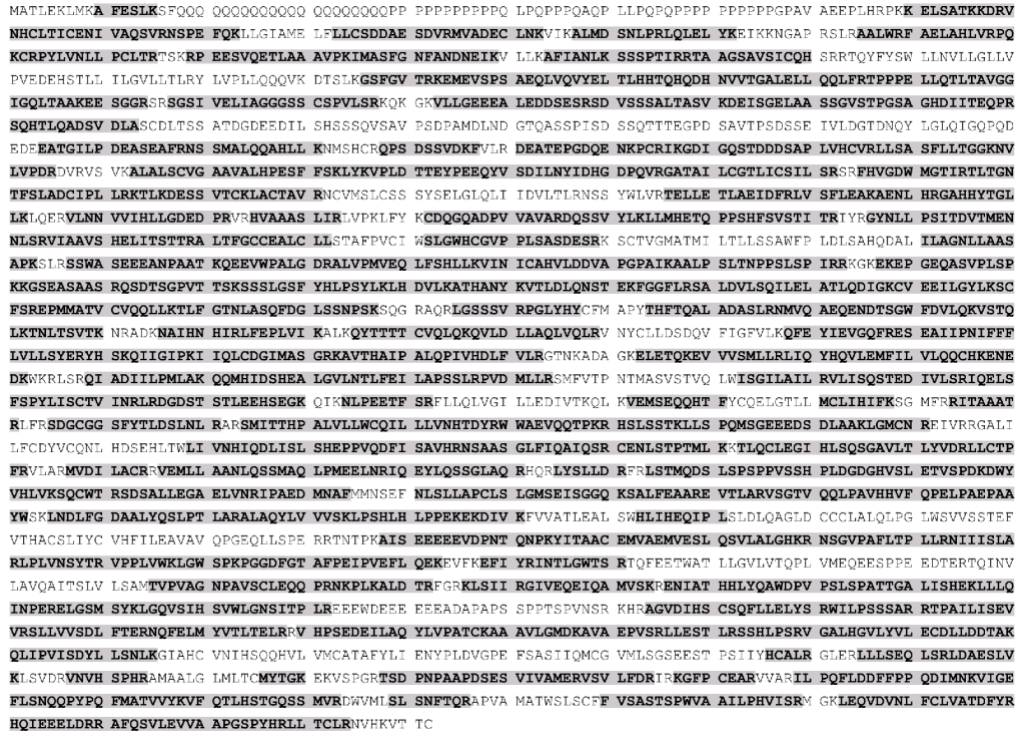
```
   1 MATLEKLMKA FESLKSFQQQ QQQQQQQQQQ QQQQQQQQPP PPPPPPPPPQ LPQPPPQAQP LLPQPQPPPP PPPPPPGPAV AEEPLHRPKK ELSATKKDRV
 101 NHCLTICENI VAQSVRNSPE FQKLLGIAME LFLLCSDDAE SDVRMVADEC LNKVIKALMD SNLPRLQLEL YKEIKKNGAP RSLRAALWRF AELAHLVRPQ
 201 KCRPYLVNLL PCLTRTSKRP EESVQETLAA AVPKIMASFG NFANDNEIKV LLKAFIANLK SSSPTIRRTA AGSAVSICQH SRRTQYFYSW LLNVLLGLLV
 301 PVEDEHSTLL ILGVLLTLRY LVPLLQQQVK DTSLKGSFGV TRKEMEVSPS AEQLVQVYEL TLHHTQHQDH NVVTGALELL QQLFRTPPPE LLQTLTAVGG
 401 IGQLTAAKEE SGGRSRSGSI VELIAGGGSS CSPVLSRKQK GKVLLGEEEA LEDDSESRSD VSSSALTASV KDEISGELAA SSGVSTPGSA GHDIITEQPR
 501 SQHTLQADSV DLASCDLTSS ATDGDEEDIL SHSSSQVSAV PSDPAMDLND GTQASSPISD SSQTTTEGPD SAVTPSDSSE IVLDGTDNQY LGLQIGQPQD
 601 EDEEATGILP DEASEAFRNS SMALQQAHLL KNMSHCRQPS DSSVDKFVLR DEATEPGDQE NKPCRIKGDI GQSTDDDSAP LVHCVRLLSA SFLLTGGKNV
 701 LVPDRDVRVS VKALALSCVG AAVALHPESF FSKLYKVPLD TTEYPEEQYV SDILNYIDHG DPQVRGATAI LCGTLICSIL SRSRFHVGDW MGTIRTLTGN
 801 TFSLADCIPL LRKTLKDESS VTCKLACTAV RNCVMSLCSS SYSELGLQLI IDVLTLRNSS YWLVRTELLE TLAEIDFRLV SFLEAKAENL HRGAHHYTGL
 901 LKLQERVLNN VVIHLLGDED PRVRHVAAAS LIRLVPKLFY KCDQGQADPV VAVARDQSSV YLKLLMHETQ PPSHFSVSTI TRIYRGYNLL PSITDVTMEN
1001 NLSRVIAAVS HELITSTTRA LTFGCCEALC LLSTAFPVCI WSLGWHCGVP PLSASDESRK SCTVGMATMI LTLLSSAWFP LDLSAHQDAL ILAGNLLAAS
1101 APKSLRSSWA SEEEANPAAT KQEEVWPALG DRALVPMVEQ LFSHLLKVIN ICAHVLDDVA PGPAIKAALP SLTNPPSLSP IRRKGKEKEP GEQASVPLSP
1201 KKGSEASAAS RQSDTSGPVT TSKSSSLGSF YHLPSYLKLH DVLKATHANY KVTLDLQNST EKFGGFLRSA LDVLSQILEL ATLQDIGKCV EEILGYLKSC
1301 FSREPMMATV CVQQLLKTLF GTNLASQFDG LSSNPSKSQG RAQRLGSSSV RPGLYHYCFM APYTHFTQAL ADASLRNMVQ AEQENDTSGW FDVLQKVSTQ
1401 LKTNLTSVTK NRADKNAIHN HIRLFEPLVI KALKQYTTTT CVQLQKQVLD LLAQLVQLRV NYCLLDSDQV FIGFVLKQFE YIEVGQFRES EAIIPNIFFF
1501 LVLLSYERYH SKQIIGIPKI IQLCDGIMAS GRKAVTHAIP ALQPIVHDLF VLRGTNKADA GKELETQKEV VVSMLLRLIQ YHQVLEMFIL VLQQCHKENE
1601 DKWKRLSRQI ADIILPMLAK QQMHIDSHEA LGVLNTLFEI LAPSSLRPVD MLLRSMFVTP NTMASVSTVQ LWISGILAIL RVLISQSTED IVLSRIQELS
1701 FSPYLISCTV INRLRDGDST STLEEHSEGK QIKNLPEETF SRFLLQLVGI LLEDIVTKQL KVEMSEQQHT FYCQELGTLL MCLIHIFKSG MFRRITAAAT
1801 RLFRSDGCGG SFYTLDSLNL RARSMITTHP ALVLLWCQIL LLVNHTDYRW WAEVQQTPKR HSLSSTKLLS PQMSGEEEDS DLAAKLGMCN REIVRRGALI
1901 LFCDYVCQNL HDSEHLTWLI VNHIQDLISL SHEPPVQDFI SAVHRNSAAS GLFIQAIQSR CENLSTPTML KKTLQCLEGI HLSQSGAVLT LYVDRLLCTP
2001 FRVLARMVDI LACRRVEMLL AANLQSSMAQ LPMEELNRIQ EYLQSSGLAQ RHQRLYSLLD RFRLSTMQDS LSPSPPVSSH PLDGDGHVSL ETVSPDKDWY
2101 VHLVKSQCWT RSDSALLEGA ELVNRIPAED MNAFMMNSEF NLSLLAPCLS LGMSEISGGQ KSALFEAARE VTLARVSGTV QQLPAVHHVF QPELPAEPAA
2201 YWSKLNDLFG DAALYQSLPT LARALAQYLV VVSKLPSHLH LPPEKEKDIV KFVVATLEAL SWHLIHEQIP LSLDLQAGLD CCCLALQLPG LWSVVSSTEF
2301 VTHACSLIYC VHFILEAVAV QPGEQLLSPE RRTNTPKAIS EEEEEVDPNT QNPKYITAAC EMVAEMVESL QSVLALGHKR NSGVPAFLTP LLRNIIISLA
2401 RLPLVNSYTR VPPLVWKLGW SPKPGGDFGT AFPEIPVEFL QEKEVFKEFI YRINTLGWTS RTQFEETWAT LLGVLVTQPL VMEQEESPPE EDTERTQINV
2501 LAVQAITSLV LSAMTVPVAG NPAVSCLEQQ PRNKPLKALD TRFGRKLSII RGIVEQEIQA MVSKRENIAT HHLYQAWDPV PSLSPATTGA LISHEKLLLQ
2601 INPERELGSM SYKLGQVSIH SVWLGNSITP LREEEWDEEE EEEADAPAPS SPPTSPVNSR KHRAGVDIHS CSQFLLELYS RWILPSSSAR RTPAILISEV
2701 VRSLLVVSDL FTERNQFELM YVTLTELRRV HPSEDEILAQ YLVPATCKAA AVLGMDKAVA EPVSRLLEST LRSSHLPSRV GALHGVLYVL ECDLLDDTAK
2801 QLIPVISDYL LSNLKGIAHC VNIHSQQHVL VMCATAFYLI ENYPLDVGPE FSASIIQMCG VMLSGSEEST PSIIYHCALR GLERLLLSEQ LSRLDAESLV
2901 KLSVDRVNVH SPHRAMAALG LMLTCMYTGK EKVSPGRTSD PNPAAPDSES VIVAMERVSV LFDRIRKGFP CEARVVARIL PQFLDDFFPP QDIMNKVIGE
3001 FLSNQQPYPQ FMATVVYKVF QTLHSTGQSS MVRDWVMLSL SNFTQRAPVA MATWSLSCFF VSASTSPWVA AILPHVISRM GKLEQVDVNL FCLVATDFYR
3101 HQIEEELDRR AFQSVLEVVA APGSPYHRLL TCLRNVHKVT TC
```

B)

```
   1 MATLEKLMKA FESLKSFQQQ QQQQQQQQQQ QQQQQQQQPP PPPPPPPPPQ LPQPPPQAQP LLPQPQPPPP PPPPPPGPAV AEEPLHRPKK ELSATKKDRV
 101 NHCLTICENI VAQSVRNSPE FQKLLGIAME LFLLCSDDAE SDVRMVADEC LNKVIKALMD SNLPRLQLEL YKEIKKNGAP RSLRAALWRF AELAHLVRPQ
 201 KCRPYLVNLL PCLTRTSKRP EESVQETLAA AVPKIMASFG NFANDNEIKV LLKAFIANLK SSSPTIRRTA AGSAVSICQH SRRTQYFYSW LLNVLLGLLV
 301 PVEDEHSTLL ILGVLLTLRY LVPLLQQQVK DTSLKGSFGV TRKEMEVSPS AEQLVQVYFL TLHHTQHQDH NVVTGALELL QQLFRTPPPE LLQTLTAVGG
 401 IGQLTAAKEE SGGRSRSGSI VELIAGGGSS CSPVLSRKQK GKVLLGEEEA LEDDSESRSD VSSSALTASV KDEISGELAA SSGVSTPGSA GHDIITEQPR
 501 SQHTLQADSV DLASCDLTSS ATDGDEEDIL SHSSSQVSAV PSDPAMDLND GTQASSPISD SSQTTTEGPD SAVTPSDSSE IVLDGTDNQY LGLQIGQPQD
 601 EDEEATGILP DEASEAFRNS SMALQQAHLL KNMSHCRQPS DSSVDKFVLR DEATEPGDQE NKPCRIKGDI GQSTDDDSAP LVHCVRLLSA SFLLTGGKNV
 701 LVPDRDVRVS VKALALSCVG AAVALHPESF FSKLYKVPLD TTEYPEEQYV SDILNYIDHG DPQVRGATAI LCGTLICSIL SRSRFHVGDW MGTIRTLTGN
 801 TFSLADCIPL LRKTLKDESS VTCKLACTAV RNCVMSLCSS SYSELGLQLI IDVLTLRNSS YWLVRTELLE TLAEIDFRLV SFLEAKAENL HRGAHHYTGL
 901 LKLQERVLNN VVIHLLGDED PRVRHVAAAS LIRLVPKLFY KCDQGQADPV VAVARDQSSV YLKLLMHETQ PPSHFSVSTI TRIYRGYNLL PSITDVTMEN
1001 NLSRVIAAVS HELITSTTRA LTFGCCEALC LLSTAFPVCI WSLGWHCGVP PLSASDESRK SCTVGMATMI LTLLSSAWFP LDLSAHQDAL ILAGNLLAAS
1101 APKSLRSSWA SEEEANPAAT KQEEVWPALG DRALVPMVEQ LFSHLLKVIN ICAHVLDDVA PGPAIKAALP SLTNPPSLSP IRRKGKEKEP GEQASVPLSP
1201 KKGSEASAAS RQSDTSGPVT TSKSSSLGSF YHLPSYLKLH DVLKATHANY KVTLDLQNST EKFGGFLRSA LDVLSQILEL ATLQDIGKCV EEILGYLKSC
1301 FSREPMMATV CVQQLLKTLF GTNLASQFDG LSSNPSKSQG RAQRLGSSSV RPGLYHYCFM APYTHFTQAL ADASLRNMVQ AEQENDTSGW FDVLQKVSTQ
1401 LKTNLTSVTK NRADKNAIHN HIRLFEPLVI KALKQYTTTT CVQLQKQVLD LLAQLVQLRV NYCLLDSDQV FIGFVLKQFE YIEVGQFRES EAIIPNIFFF
1501 LVLLSYERYH SKQIIGIPKI IQLCDGIMAS GRKAVTHAIP ALQPIVHDLF VLRGTNKADA GKELETQKEV VVSMLLRLIQ YHQVLEMFIL VLQQCHKENE
1601 DKWKRLSRQI ADIILPMLAK QQMHIDSHEA LGVLNTLFEI LAPSSLRPVD MLLRSMFVTP NTMASVSTVQ LWISGILAIL RVLISQSTED IVLSRIQELS
1701 FSPYLISCTV INRLRDGDST STLEEHSEGK QIKNLPEETF SRFLLQLVGI LLEDIVTKQL KVEMSEQQHT FYCQELGTLL MCLIHIFKSG MFRRITAAAT
1801 RLFRSDGCGG SFYTLDSLNL RARSMITTHP ALVLLWCQIL LLVNHTDYRW WAEVQQTPKR HSLSSTKLLS PQMSGEEEDS DLAAKLGMCN REIVRRGALI
1901 LFCDYVCQNL HDSEHLTWLI VNHIQDLISL SHEPPVQDFI SAVHRNSAAS GLFIQAIQSR CENLSTPTML KKTLQCLEGI HLSQSGAVLT LYVDRLLCTP
2001 FRVLARMVDI LACRRVEMLL AANLQSSMAQ LPMEELNRIQ EYLQSSGLAQ RHQRLYSLLD RFRLSTMQDS LSPSPPVSSH PLDGDGHVSL ETVSPDKDWY
2101 VHLVKSQCWT RSDSALLEGA ELVNRIPAED MNAFMMNSEF NLSLLAPCLS LGMSEISGGQ KSALFEAARE VTLARVSGTV QQLPAVHHVF QPELPAEPAA
2201 YWSKLNDLFG DAALYQSLPT LARALAQYLV VVSKLPSHLH LPPEKEKDIV KFVVATLEAL SWHLIHEQIP LSLDLQAGLD CCCLALQLPG LWSVVSSTEF
2301 VTHACSLIYC VHFILEAVAV QPGEQLLSPE RRTNTPKAIS EEEEEVDPNT QNPKYITAAC EMVAEMVESL QSVLALGHKR NSGVPAFLTP LLRNIIISLA
2401 RLPLVNSYTR VPPLVWKLGW SPKPGGDFGT AFPEIPVEFL QEKEVFKEFI YRINTLGWTS RTQFEETWAT LLGVLVTQPL VMEQEESPPE EDTERTQINV
2501 LAVQAITSLV LSAMTVPVAG NPAVSCLEQQ PRNKPLKALD TRFGRKLSII RGIVEQEIQA MVSKRENIAT HHLYQAWDPV PSLSPATTGA LISHEKLLLQ
2601 INPERELGSM SYKLGQVSIH SVWLGNSITP LREEEWDEEE EEEADAPAPS SPPTSPVNSR KHRAGVDIHS CSQFLLELYS RWILPSSSAR RTPAILISEV
2701 VRSLLVVSDL FTERNQFELM YVTLTELRRV HPSEDEILAQ YLVPATCKAA AVLGMDKAVA EPVSRLLEST LRSSHLPSRV GALHGVLYVL ECDLLDDTAK
2801 QLIPVISDYL LSNLKGIAHC VNIHSQQHVL VMCATAFYLI ENYPLDVGPE FSASIIQMCG VMLSGSEEST PSIIYHCALR GLERLLLSEQ LSRLDAESLV
2901 KLSVDRVNVH SPHRAMAALG LMLTCMYTGK EKVSPGRTSD PNPAAPDSES VIVAMERVSV LFDRIRKGFP CEARVVARIL PQFLDDFFPP QDIMNKVIGE
3001 FLSNQQPYPQ FMATVVYKVF QTLHSTGQSS MVRDWVMLSL SNFTQRAPVA MATWSLSCFF VSASTSPWVA AILPHVISRM GKLEQVDVNL FCLVATDFYR
3101 HQIEEELDRR AFQSVLEVVA APGSPYHRLL TCLRNVHKVT TC
```

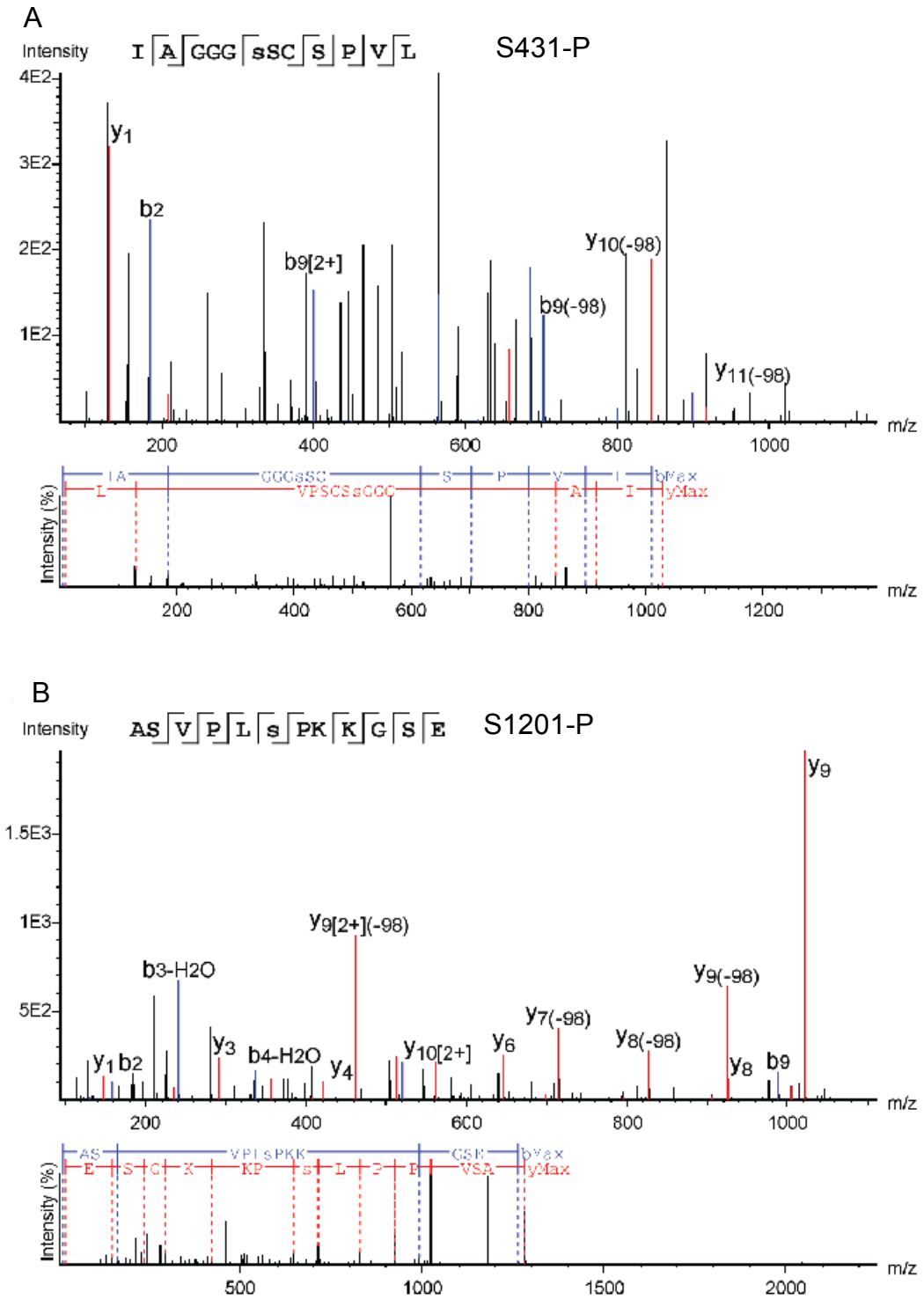
Figure S3. Sequence coverage maps of A) HEK293 Q23 HTT from solution (77%) and B) from a gel band (62%) as obtained from bottom-up proteomics experiments.

A)

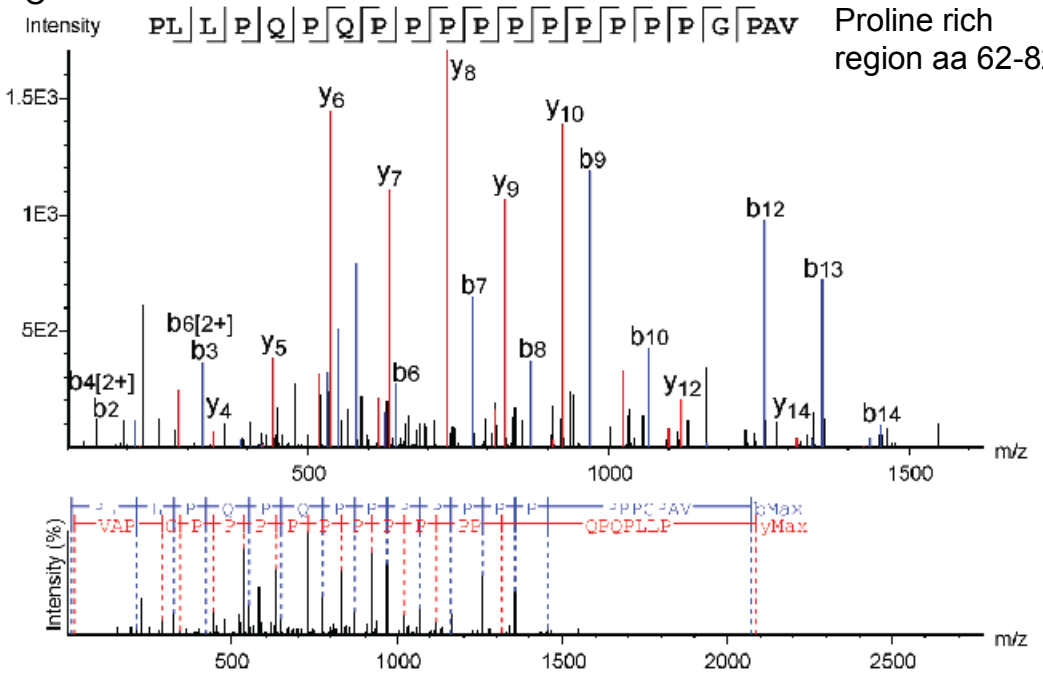


B)

Figure S4. Exemplary spectra of peptides identified after digesting HTT$^{1-3144}$ Q23 from Sf9 with pepsin (pH ~ 1). Full data can be found through PRIDE (Jones et al., 2006) with accession PXD010865.

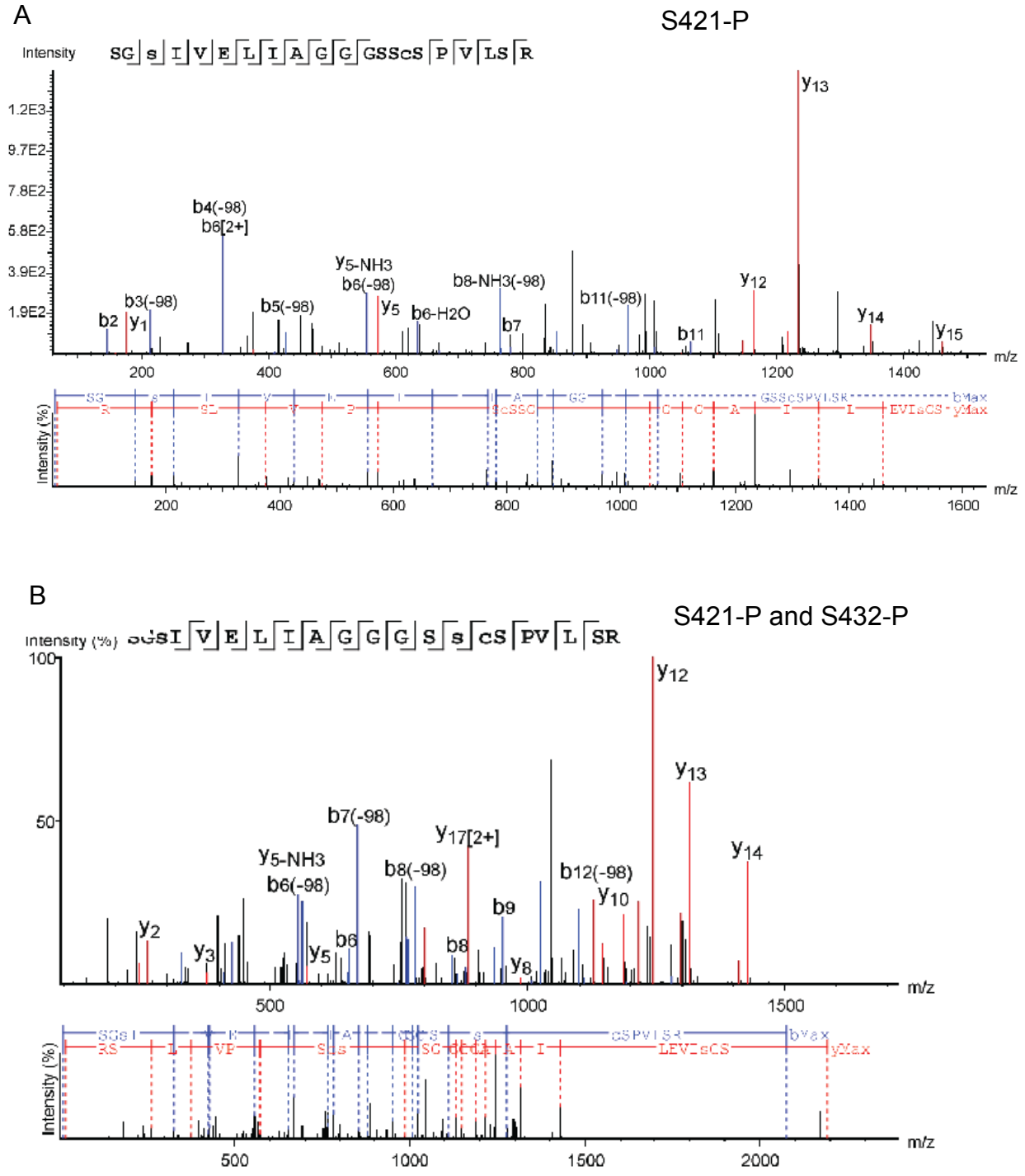


C

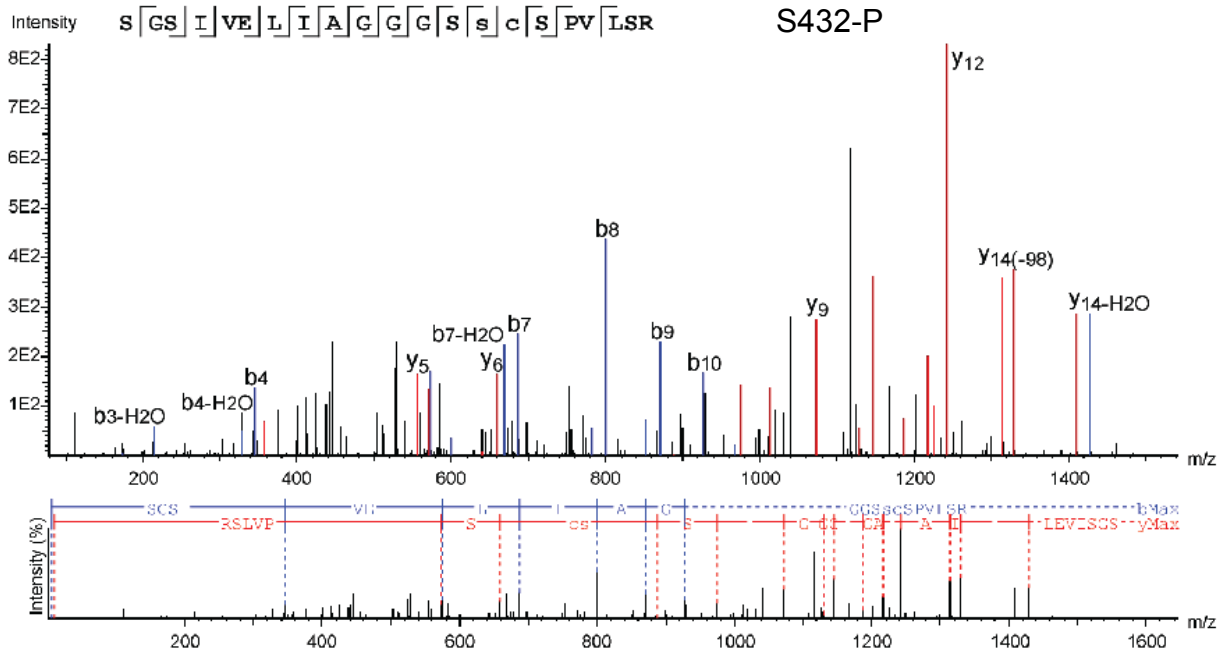


Proline rich region aa 62-82

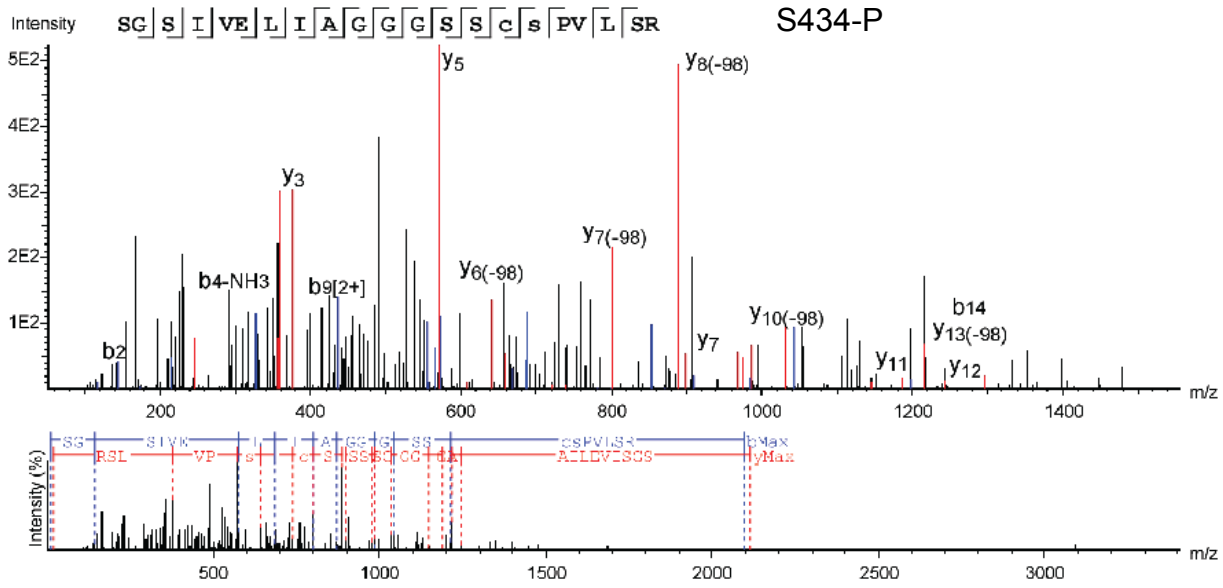
Figure S5. Exemplary spectra for peptides identified after digesting HTT$^{1-3144}$ Q23 from Sf9 with trypsin or lysarginase. Full data can be found through PRIDE (Jones et al., 2006) with accession PXD010865.



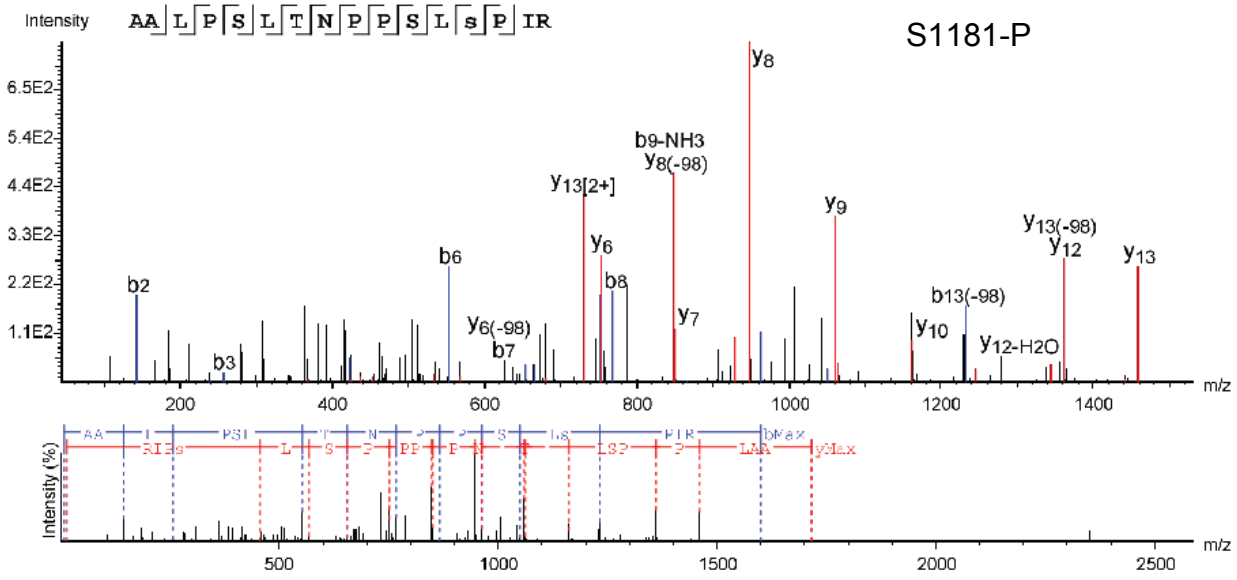
C


Intensity
S[GS]I[VE]L I[A]G[G]G[S]s[c]S[PV]LSR
S432-P
y12
b8
y14(-98)
y14-H2O
y9
b7-H2O
b7
b9
y5
y6
b10
b4
b4-H2O
b3-H2O
m/z
Intensity (%)
bMax
yMax

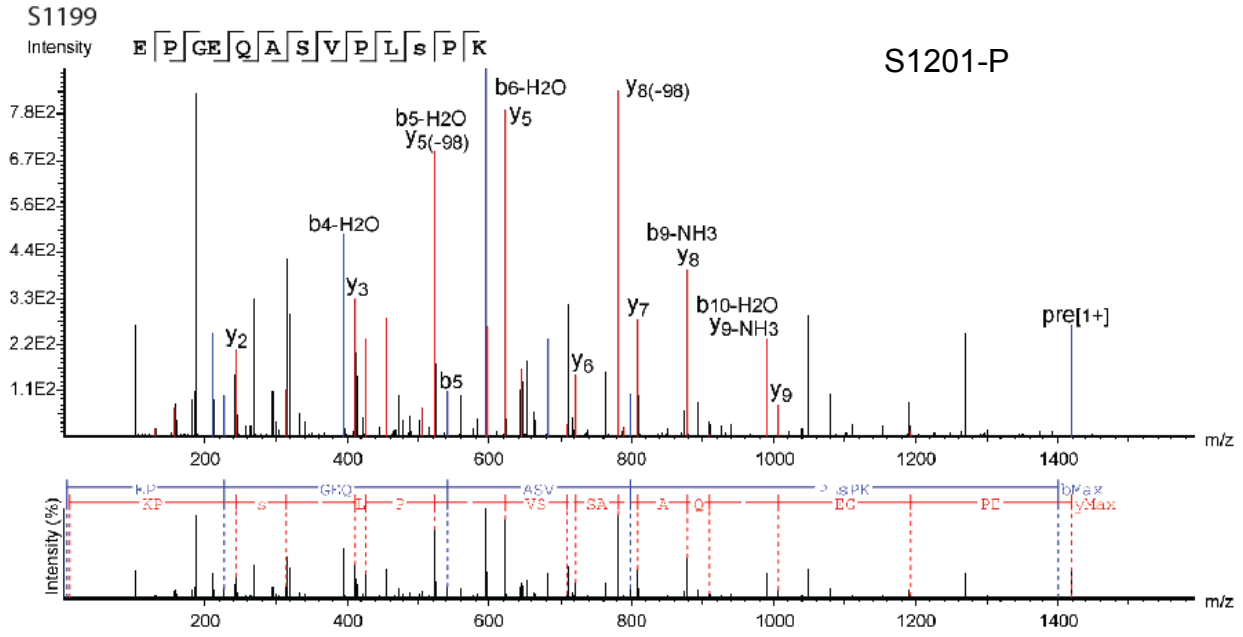
D


Intensity
SG[S]I[VE]L I[A]G[G]G[S]S[c]s[PV]L[SR]
S434-P
y5
y8(-98)
y3
y7(-98)
b4-NH3
b9[2+]
y6(-98)
y10(-98)
b14
y13(-98)
y7
y11
y12
b2
m/z
Intensity (%)
bMax
yMax

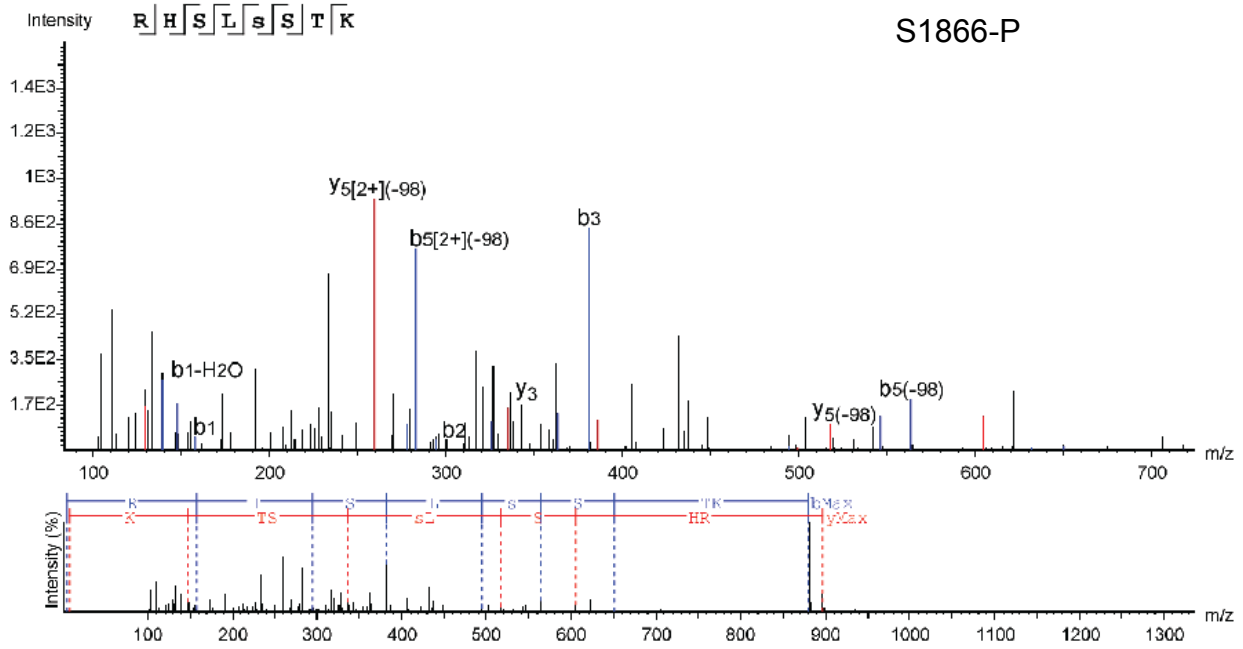
E


Intensity
AA L P S L T N P P S L s P IR
S1181-P
6.5E2
5.4E2
4.4E2
3.3E2
2.2E2
1.1E2
y8
b9-NH3
y8(-98)
y13[2+]
y9
y13(-98)
y12
y13
y6
b6
b8
b2
y6(-98)
b7
y7
b13(-98)
y10
y12-H2O
b3
m/z
200
400
600
800
1000
1200
1400
Intensity (%)
m/z
500
1000
1500
2000
2500

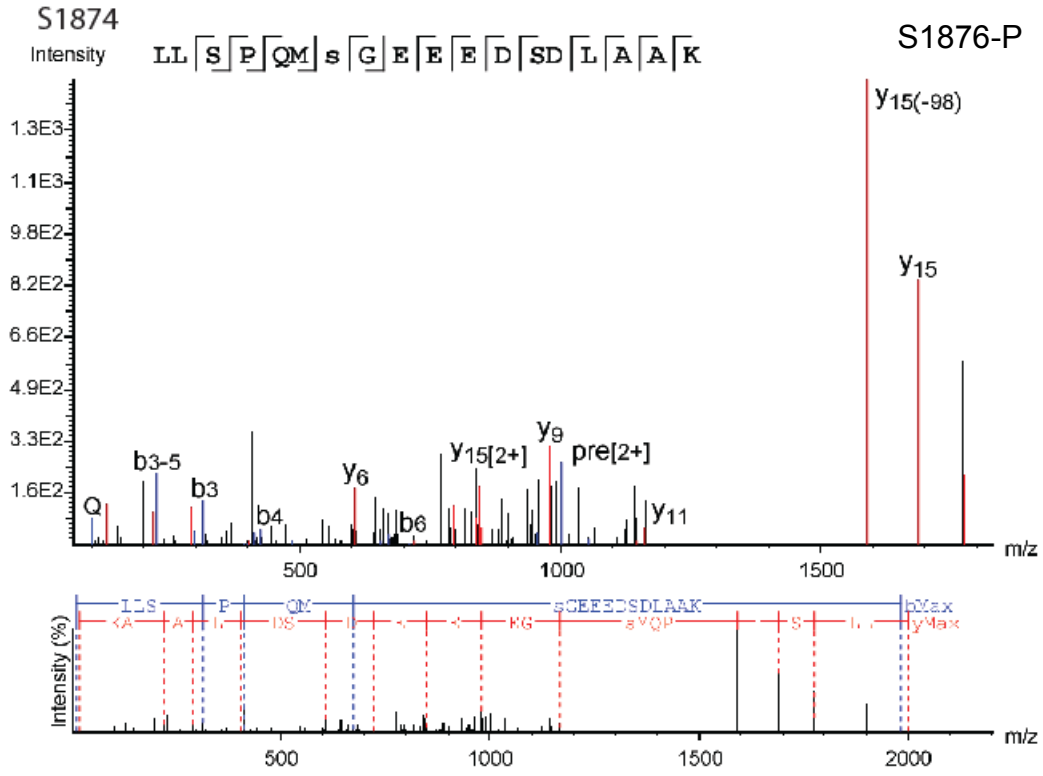
F


S1199
Intensity
E P GE Q A S V P L s P K
S1201-P
b6-H2O
y8(-98)
b5-H2O
y5
y5(-98)
b4-H2O
b9-NH3
y8
y3
y7
b10-H2O
y9-NH3
pre[1+]
y2
y6
b5
y9
7.8E2
6.7E2
5.6E2
4.4E2
3.3E2
2.2E2
1.1E2
m/z
200
400
600
800
1000
1200
1400
Intensity (%)

G


Intensity
R H S L s S T K
S1866-P
y5[2+](-98)
b3
b5[2+](-98)
b1-H2O
y3
b5(-98)
y5(-98)
b1
b2
1.4E3
1.2E3
1E3
8.6E2
6.9E2
5.2E2
3.5E2
1.7E2
m/z
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
Intensity (%)

H


S1874
Intensity
LL[S]P[QM]s[G]E[E]E[D][SD]L[A]A[K]
S1876-P
y15(-98)
y15
y9
y15[2+]
pre[2+]
y6
b3-5
b3
b4
b6
y11
Q
1.3E3
1.1E3
9.8E2
8.2E2
6.6E2
4.9E2
3.3E2
1.6E2
500
1000
1500
m/z
Intensity (%)
2000

I

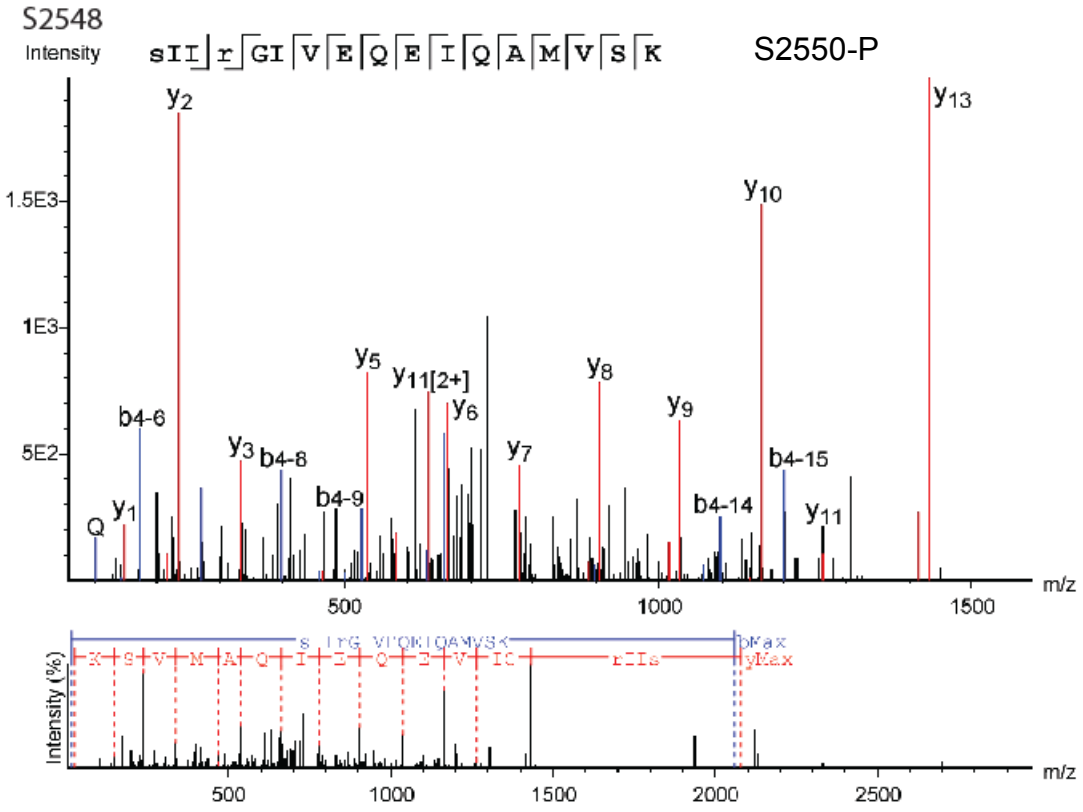
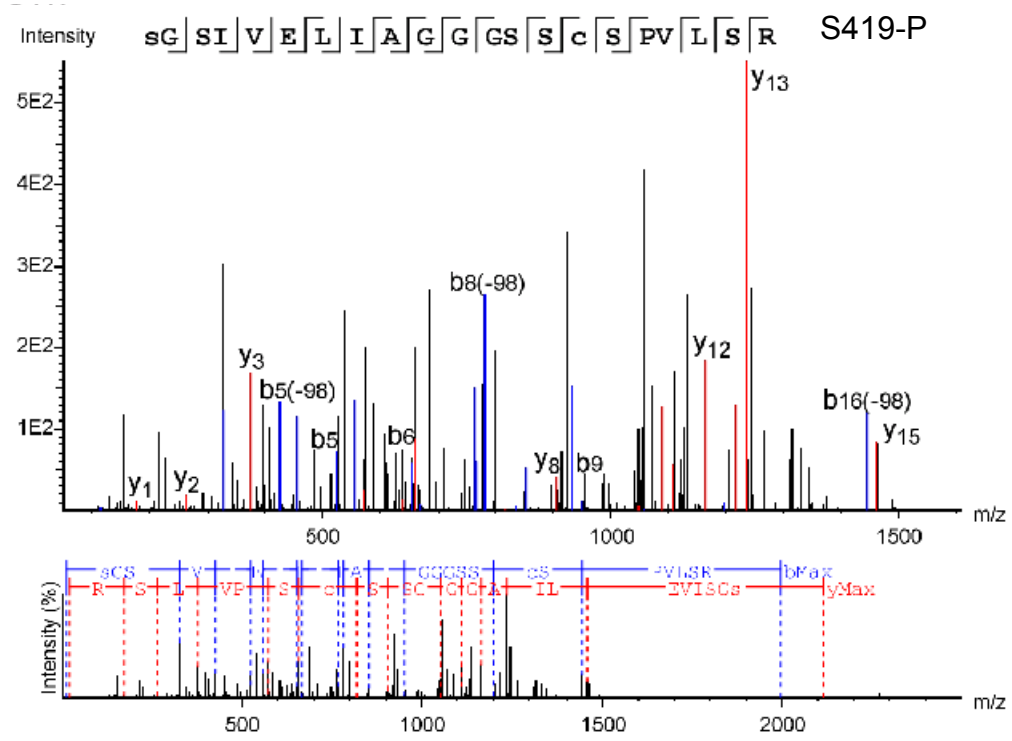
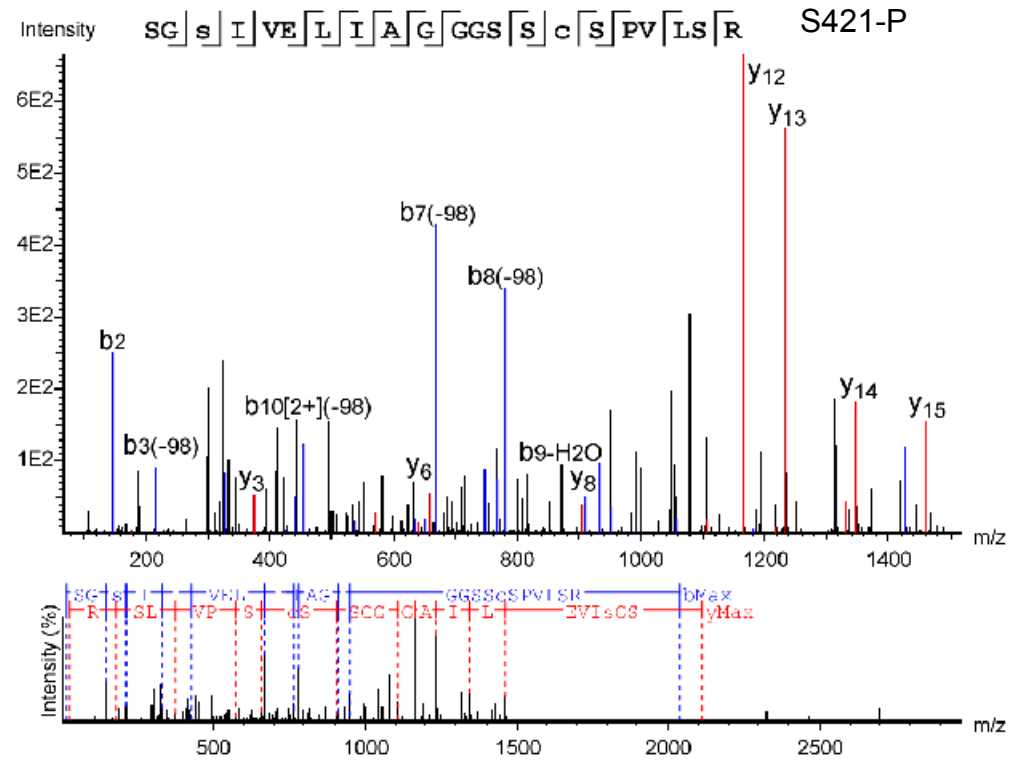

S2548
Intensity
sII[r[GI[V[E[Q[E[I[Q[A[M[V[S[K
S2550-P
y2
y13
y10
1.5E3
1E3
5E2
y5
y11[2+]
y8
y6
y9
b4-6
y3
b4-8
y7
b4-15
b4-9
b4-14
y11
y1
Q
m/z
500
1000
1500
Intensity (%)
bMax
yMax
m/z
500
1000
1500
2000
2500

Figure S6. Exemplary spectra of peptides identified after digesting HTT$^{1-3144}$ Q23 from EXPI293F with trypsin. Full data can be found through PRIDE (Jones et al., 2006) with accession PXD010865.

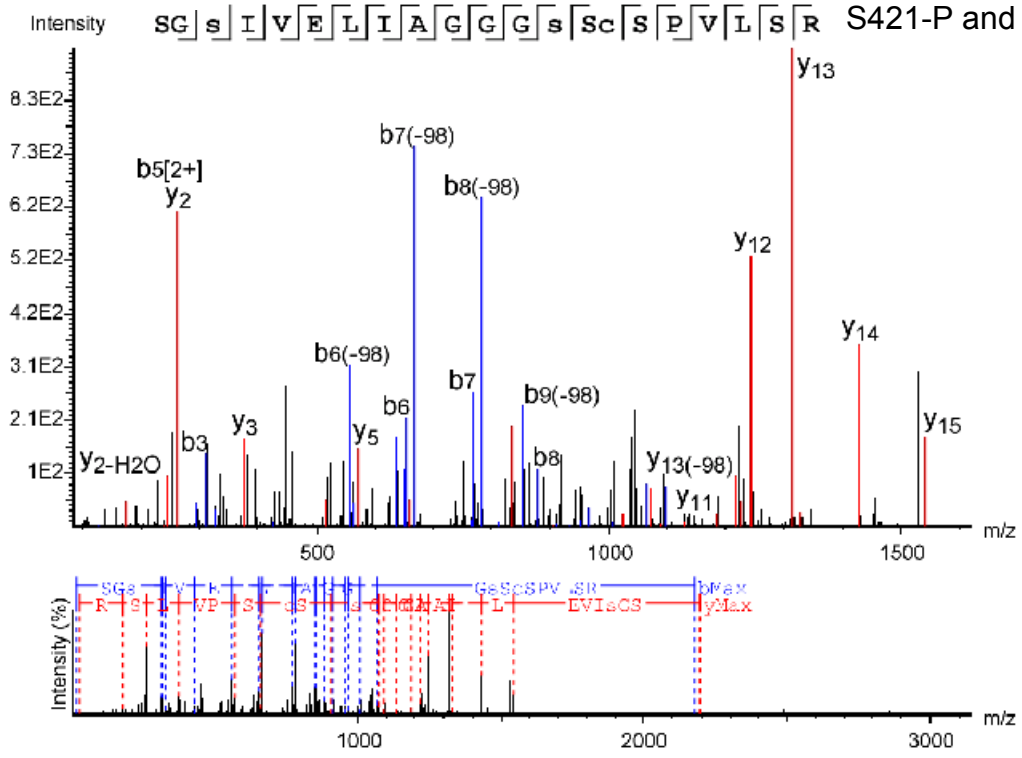
A



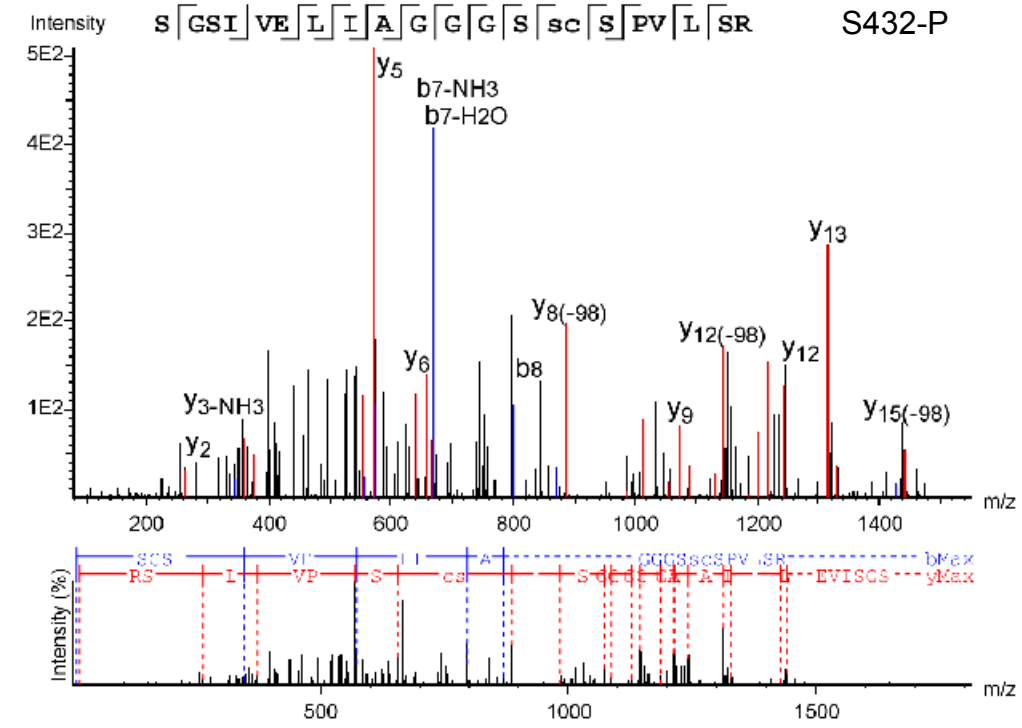
B



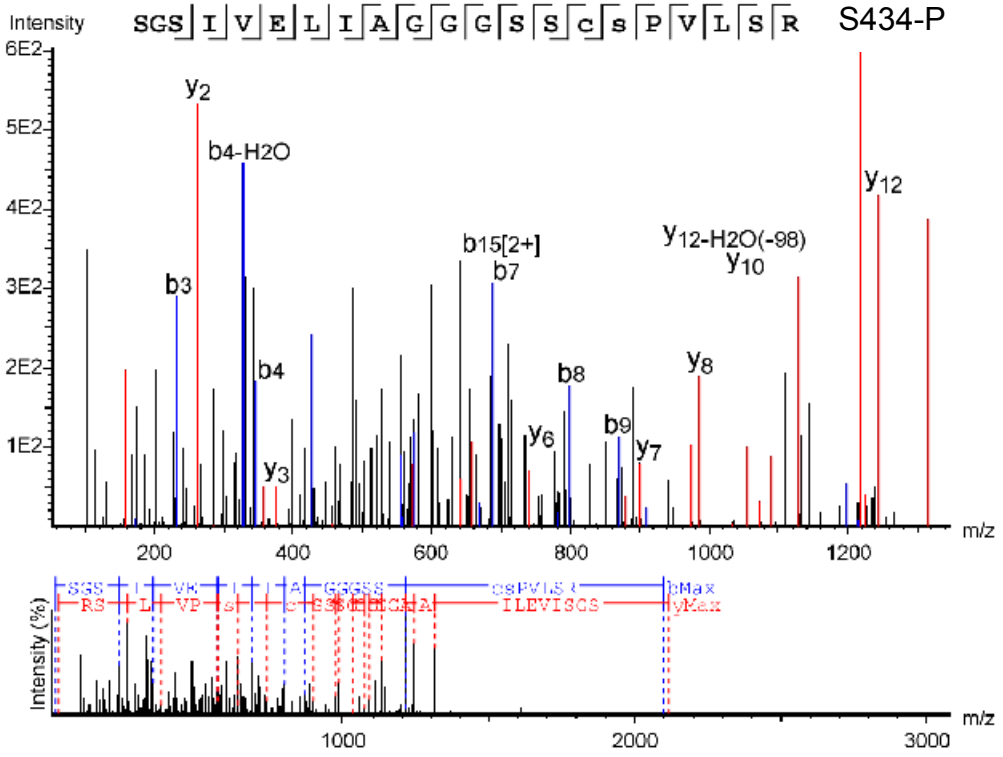
C


Intensity
SG s I V E L I A G G G s Sc S P V L S R
S421-P and S431-P
8.3E2
7.3E2
6.2E2
5.2E2
4.2E2
3.1E2
2.1E2
1E2
y13
b7(-98)
b5[2+]
y2
b8(-98)
y12
y14
b6(-98)
b7
b9(-98)
b6
y3
y5
y15
y2-H2O
b3
b8
y13(-98)
y11
m/z
500
1000
1500
Intensity (%)
1000
2000
3000

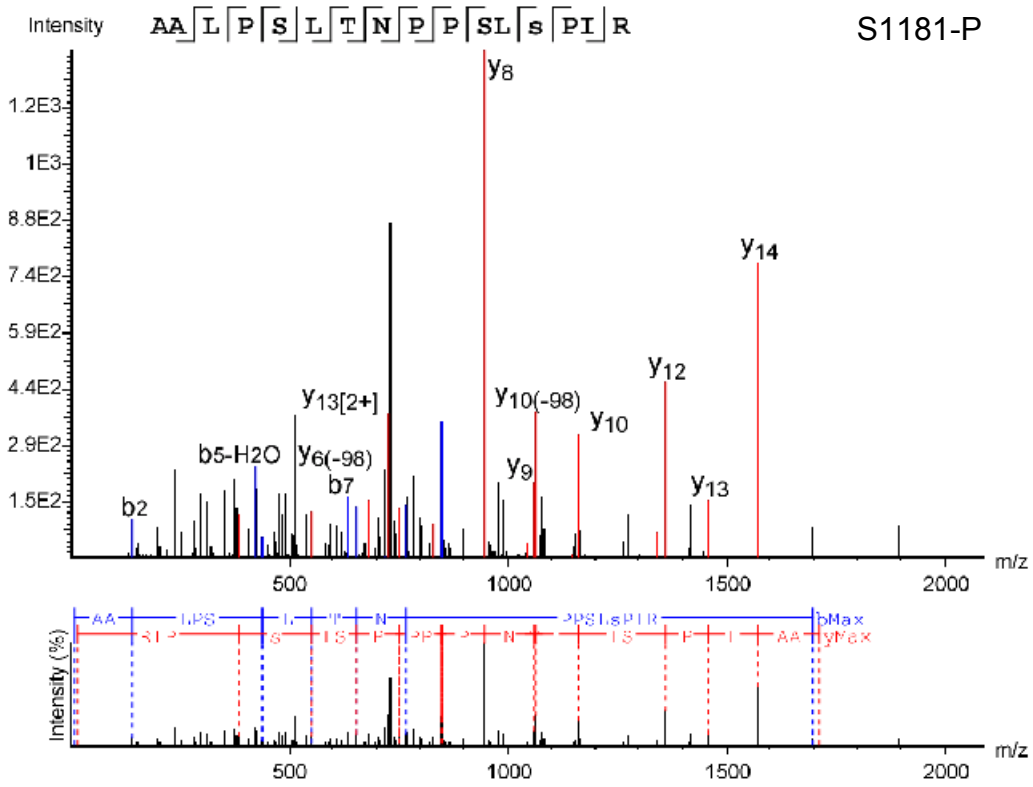
D


Intensity
S GSI VE L I A G G G S sc S PV L SR
S432-P
5E2
4E2
3E2
2E2
1E2
y5
b7-NH3
b7-H2O
y13
y8(-98)
y12(-98)
y12
y6
b8
y3-NH3
y9
y15(-98)
y2
m/z
200
400
600
800
1000
1200
1400
Intensity (%)
500
1000
1500

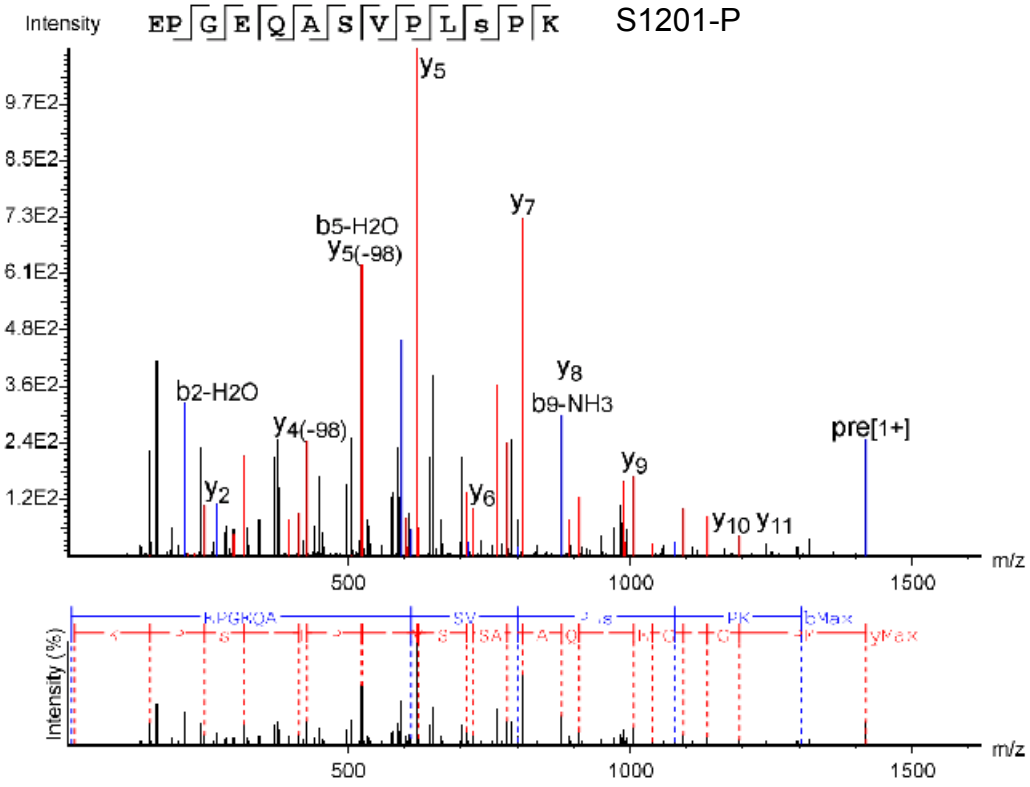
E


Intensity
SGS I V E L I A G G G S S c s P V L S R
S434-P
y2
b4-H2O
b3
b4
y3
b15[2+]
b7
b8
y6
b9
y7
y8
y12-H2O(-98)
y10
y12
m/z
Intensity (%)
pMax
yMax
m/z

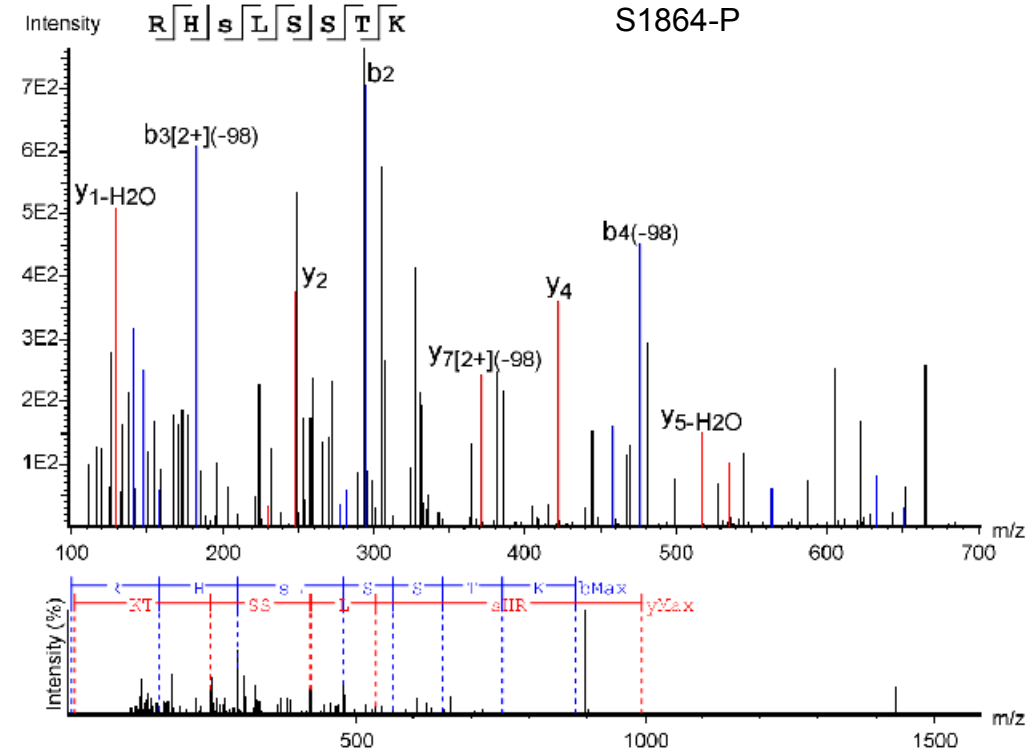
F


Intensity
AA L P S L T N P P SL s PI R
S1181-P
y8
y14
y12
y13[2+]
y10(-98)
y10
b5-H2O
y6(-98)
b7
y9
y13
b2
m/z
Intensity (%)
pMax
yMax
m/z

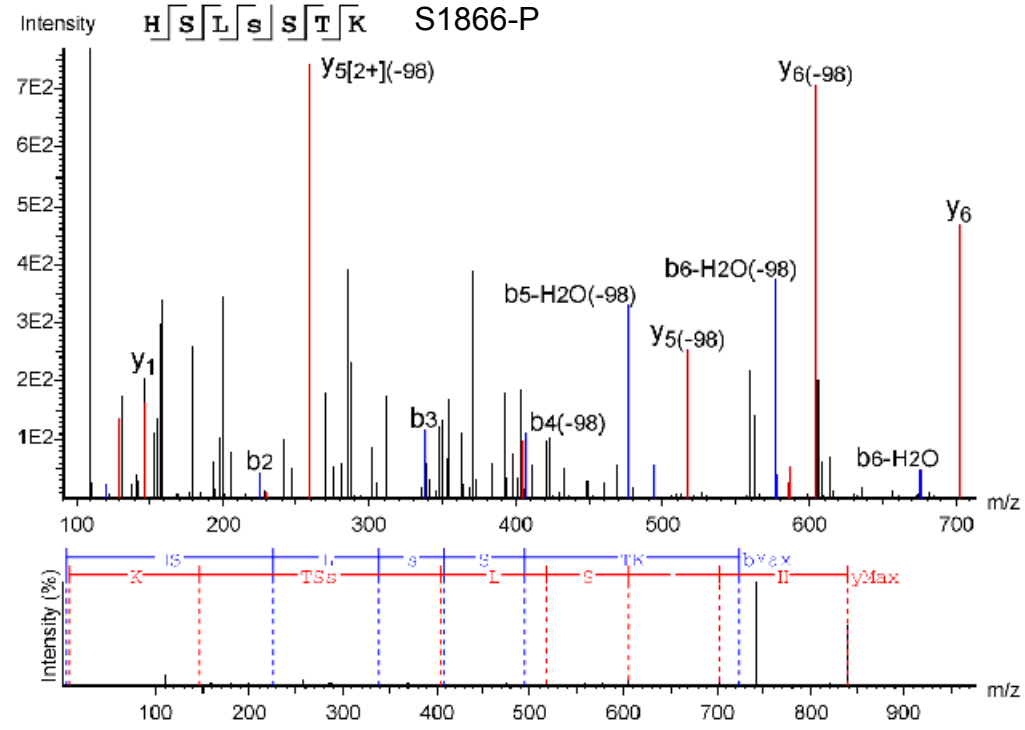
G


Intensity
EP[G]E[Q]A[S]V[P]L[s]P[K]
S1201-P
y5
y7
b5-H2O
y5(-98)
y8
b9-NH3
b2-H2O
y4(-98)
pre[1+]
y9
y2
y6
y10 y11
9.7E2
8.5E2
7.3E2
6.1E2
4.8E2
3.6E2
2.4E2
1.2E2
500
1000
1500
m/z
Intensity (%)

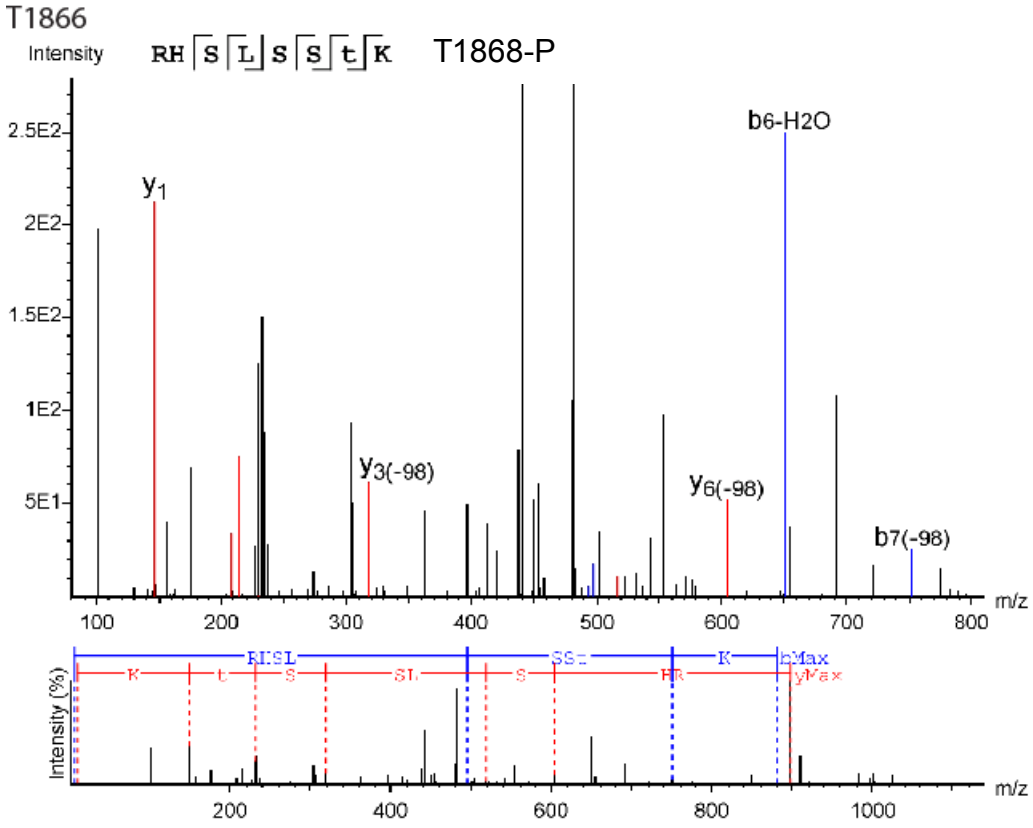
H


Intensity
R[H]s[L]S[S]T[K]
S1864-P
b2
b3[2+](-98)
y1-H2O
b4(-98)
y2
y4
y7[2+](-98)
y5-H2O
7E2
6E2
5E2
4E2
3E2
2E2
1E2
100
200
300
400
500
600
700
m/z
Intensity (%)

I


Intensity
H S L s S T K
S1866-P
y5[2+](-98)
y6(-98)
y6
b6-H2O(-98)
b5-H2O(-98)
y5(-98)
y1
b3
b4(-98)
b2
b6-H2O
m/z
Intensity (%)

J


T1866
Intensity
RH S L s S t K
T1868-P
b6-H2O
y1
y3(-98)
y6(-98)
b7(-98)
m/z
Intensity (%)

K

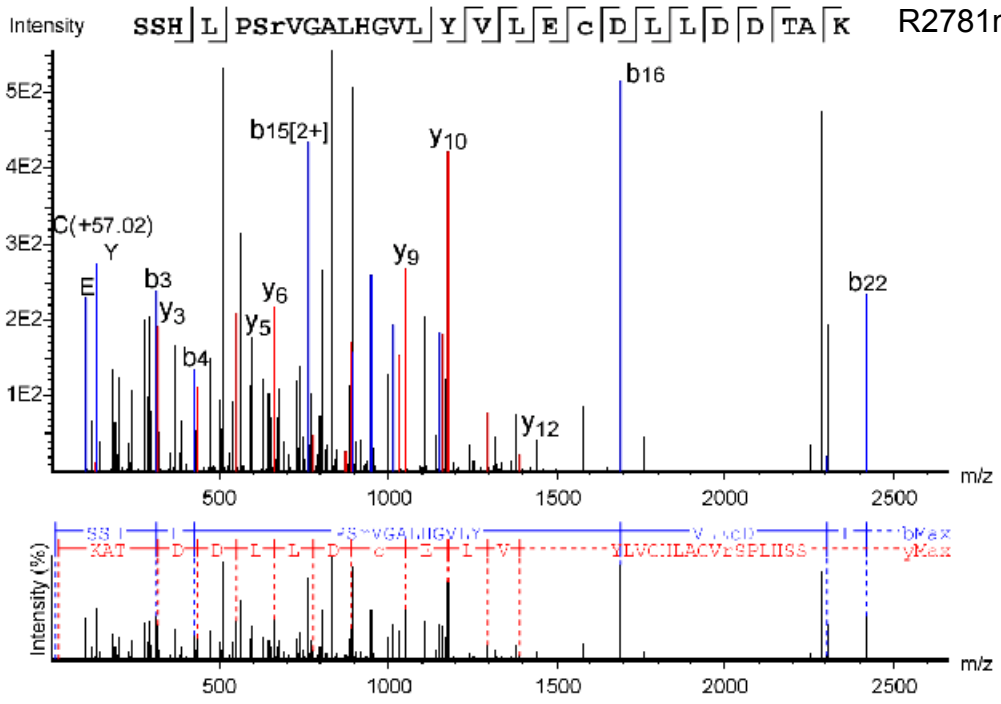

Intensity
SSH L PSrVGALHGVL Y V L E C D L L D D TA K
R2781me
5E2
4E2
3E2
2E2
1E2
C(+57.02)
b16
b15[2+]
y10
y9
b22
b3
y6
y3
y5
b4
y12
m/z
500
1000
1500
2000
2500
Intensity (%)

Figure S7 – Mapping HTT posttranslational modifications identified from HTT$^{1-3144}$ samples from Sf9 and EXPI293 cells onto the HTT structure shows solvent accessibility of nearly all sites. HTT is viewed top-down looking through the void in the N-terminal HEAT domain on the right hand side. Phosphorylation sites are shown in pink and all other modification sites are shown in green. As these samples were expressed in the absence of the stabilising HAP40 protein, it is likely that the more conformationally flexible apo HTT protein molecule would have greater exposure of different domain surfaces that would permit more sites to be modified than might be estimated from assessing the HTT-HAP40 molecule.