Supplemental Table S2

Notes:

P-OH = hydroxyproline
m = oxidation of M
K-OH.Gal = galactose on hydroxylysine
K-OH.Gal.Glc = galactose.glucose on hydroxylysine
All product ions: ppm<10, S/N>=3
Systematic mass error is -8.9ppm for the spectra whose numbers are labeled in red.

	# of sequence in human recombinant						
Spectrum #	**start**	**end**	**sequence**	**total PTMs**	**localized sites**	**unlocalized sites**	**pseudolocalized**
1	422	438	DPTSSPSEIGPGMPANQ	OH	P435-OH		
2	443	466	EGIGGPRGEKGQKGEPAIIEPGML	2OH+2Gal.Glc	K452-OH.Gal.Glc, K455-OH.Gal.Glc		
3	499	525	GPPGRPGLPGADGLPGPPGTMLMLPFR	6OH	P501-OH, P504-OH, P507-OH, P513-OH, P515-OH, P516-OH		
4	526	554	FGGGGDAGSKGPMVSAQESQAAQAILQQAR	OH	K535-OH		
5	567	581	TGRPGPVGPPGSGGL	3OH	P570-OH, P575-OH, P576-OH		
6	555	594	LALRGPAGPMGLTGRPGPVGPPGSGGLKGEPGDVGPQGPR	4OH+Gal.Glc	P576-OH, K582-OH.Gal.Glc, P585-OH	P570-OH?, P572-OH?	P570-OH (Gly-X-Hyp motif)
7	595	608	GVQGPPGPAGKPGR	3OH	P599-OH, P600-OH, P606-OH		
8	619	630	GMPGQTGPKGDR	2OH+Gal.Glc	P621-OH, K627-OH.Gal.Glc		
9	631	645	GFDGLAGLPGEKGHR	2OH+Gal.Glc	P639-OH, K642-OH.Gal.Glc		
10	633	658	DGLAGLPGEKGHRGDPGPSGPPGPPG	5OH+Gal	P639-OH, K642-OH.Gal, P657-OH	P648-OH?, P650-OH?, P654-OH?, P656-OH?	P648-OH, P654-OH (Gly-X-Hyp motif)
11	646	672	GDPGPSGPPGPPGDDGERGDDGEVGPR	4OH	P648-OH,P653-OH,P654-OH,P656-OH,P657-OH		
12	673	681	GLPGEPGPR	2OH	P675-OH, P678-OH		
13	688	708	GPPGPPGPPGVTGMDGQPGPK	5OH	P690-OH, P693-OH, P696-OH, P705-OH	P707-OH?, K708-OH?	K708-OH (Hyl)
14	702	716	DGQPGPKGNVGPQGE	2OH+Gal.Glc	K708-OH.Gal.Glc	P705-OH?, P707-OH?	P705-OH (Gly-X-Hyp motif)
15	709	744	GNVGPQGEPGPPGQQGNPGAQGLPGPQGAIGPPGEK	6OH	P717-OH, P726-OH, P732-OH, P741-OH, K744-OH	P719-OH?, P720-OH?	P720-OH (Gly-X-Hyp motif)
16	745	767	GPLGKPGLPGMPGADGPPGHPGK	5OH	P750-OH, P753-OH, P756-OH, P762-OH, P765-OH		
17	768	792	EGPPGEKGGQGPPGPQGPIGYPGPR	4OH+Gal.Glc	K774-OH.Gal.Glc, P780-OH, P789-OH	P770-OH?, P771-OH?	P771-OH (Gly-X-Hyp motif)
18	793	801	GVKGADGIR	OH+Gal.Glc	K795-OH.Gal.Glc		
19	808	828	GEKGEDGFPGFKGDMGIKGDR	4OH+3Gal.Glc	K810-OH.Gal.Glc, P816-OH, K819-OH.Gal.Glc, K825-OH.Gal.Glc		
20	829	848	GEIGPPGPRGEDGPEGPKGR	3OH+Gal.Glc	P833-OH,P834-OH, K846-OH.Gal.Glc		
21	849	866	GGPNGDPGPLGPPGEKGK	3OH+Gal.Glc	P855-OH	P861-OH?, K864-OH.Gal.Glc?, K866-OH.Gal.Glc	P861-OH, K864-OH.Gal.Glc (Gly-X-Hyp and Gly-X-Hyl.Gal.Glc motif)
22	867	878	LGVPGLPGYPGR	3OH	P870-OH, P873-OH, P876-OH		
23	879	900	QGPKGSIGFPGFPGANGEKGGR	4OH+2Gal.Glc	K882-OH.Gal.Glc, P888-OH, P891-OH, K897-OH.Gal.Glc		
24	901	909	GTPGKPGPR	2OH	P903-OH	P906-OH?, P908-OH?	P906-OH (Gly-X-Hyp motif)
25	925	948	GITGKPGPKGNSGGDGPAGPPGER	3OH	P930-OH, P944-OH, P945-OH		
26	925	948	GITGKPGPKGNSGGDGPAGPPGER	2OH+Gal.Glc	K933-OH.Gal.Glc, P945-OH		
27	949	963	GPNGPQGPTGFPGPK	2OH	P960-OH	P962-OH?, K963-OH?	K963-OH (Hyl)
28	964	981	GPPGPPGKDGLPGHPGQR	4OH	P966-OH, P969-OH, P975-OH, P978-OH		
29	990	1014	TGPPGPPGVVGPQGPTGETGPMGER	4OH	P992-OH, P993-OH, P995-OH, P996-OH		
30	1015	1034	GHPGPPGPPGEQGLPGLAGK	5OH	P1017-OH, P1019-OH, P1020-OH, P1023-OH, P1029-OH		
31	1035	1049	EGTKGDPGPAGLPGK	3OH	K1038-OH,P1041-OH,P1047-OH		
32	1035	1049	EGTKGDPGPAGLPGK	2OH+Gal.Glc	K1038-OH.Gal.Glc, P1047-OH		
33	1050	1060	DGPPGLRGFPG	2OH	P1053-OH, P1059-OH		
34	1063	1074	GLPGPVGALGLK	OH	P1065-OH		
35	1063	1092	GLPGPVGALGLKGNEGPPGPPGPAGSPGER	4OH+Gal.Glc	K1074-OH.Gal.Glc, P1083-OH, P1089-OH	P1065-OH?, P1067-OH?	P1065-OH (Gly-X-Hyp motif)
36	1075	1092	GNEGPPGPPGPAGSPGER	3OH	P1080-OH, P1083-OH, P1089-OH		
37	1092	1118	RGPAGAAGPIGIPGRPGPQGPPGPAGE	3OH	P1104-OH, P1107-OH, P1113-OH		
38	1093	1125	GPAGAAGPIGIPGRPGPQGPPGPAGEKGAPGEK	2OH+Gal.Glc	K1119-OH.Gal.Glc	P1104-OH?, P1107-OH?, P1109-OH?, P1112-OH?, P1113-OH?, P1115-OH?	P1104-OH or P1107-OH or P1113-OH (Gly-X-Hyp motif)
39	1134	1160	DGLQGPVGLPGPAGPVGPPGEDGDKGE	3OH	P1143-OH, P1152-OH, K1158-OH		
40	1134	1167	DGLQGPVGLPGPAGPVGPPGEDGDKGEIGEPGQK	4OH	P1143-OH, P1164-OH, K1167-OH	P1151-OH?, P1152-OH?	P1152-OH (Gly-X-Hyp motif)
41	1171	1203	GDKGEQGPPGPTGPQGPIGQPGPSGADGEPGPR	4OH+Gal.Glc	K1173-OH.Gal.Glc, P1179-OH, P1191-OH	P1200-OH?, P1202-OH?	P1200-OH (Gly-X-Hyp motif)
42	1204	1218	GQQGLFGQKGDEGPR	OH+Gal.Glc	K1212-OH.Gal.Glc		
43	1219	1239	GFPGPPGPVGLQGLPGPPGEK	5OH	P1221-OH, P1224-OH, P1233-OH, P1236-OH, K1239-OH		
44	1244	1265	DVGQMGPPGPPGPRGPSGAPGA	3OH	P1251-OH, P1254-OH, P1263-OH		
45	1258	1284	GPSGAPGADGPQGPPGGIGNPGAVGEK	OH	P1278-OH		
46	1285	1308	GEPGEAGEPGLPGEGGPPGPKGER	4OH+Gal.Glc	P1287-OH, P1293-OH, P1302-OH, K1305-OH.Gal.Glc		
47	1309	1326	GEKGESGPSGAAGPPGPK	2OH+Gal.Glc	K1311-OH.Gal.Glc, P1323-OH		
48	1331	1358	DDGPKGSPGPVGFPGDPGPPGEPGPAGQ	5OH	P1338-OH, P1344-OH, P1347-OH, P1353-OH	P1349-OH?, P1350-OH?	P1350-OH (Gly-X-Hyp motif)
49	1346	1358	DPGPPGEPGPAGQ	3OH	P1350-OH, P1353-OH	P1347-OH?, P1349-OH?	P1347-OH (Gly-X-Hyp motif)
50	1366	1391	GDDGEPGQTGSPGPTGEPGPSGPPGK	2OH	P1377-OH, P1383-OH, P1389-OH (two isoforms eluted together)		
51	1393	1403	GPPGPAGPEGR	OH	P1395-OH		
52	1404	1421	QGEKGAKGEAGLEGPPGK	2OH+2Gal.Glc	K1407-OH.Gal.Glc, K1410-OH.Gal.Glc		
53	1411	1421	GEAGLEGPPGK	OH	P1419-OH		
54	1422	1440	TGPIGPQGAPGKPGPDGLR	2OH	P1431-OH, P1434-OH		
55	1441	1473	GIPGPVGEQGLPGSPGPDGPPGPMGPPGLPGLK	4OH	P1452-OH, P1470-OH, K1473-OH	P1443-OH?, P1445-OH?	P1443-OH (Gly-X-Hyp motif)
56	1482	1499	KGHPGLIGLIGPPGEQGE	3OH	K1482-OH, P1485-OH, P1494-OH		
57	1483	1503	GHPGLIGLIGPPGEQGEKGDR	3OH+Gal.Glc	P1485-OH, P1494-OH, K1500-OH.Gal.Glc		
58	1504	1515	GLPGPQGSSGPK	2OH	P1506-OH	P1514-OH?, K1515-OH?	K1515-OH (Hyl)
59	1546	1582	GSSGPTGPKGEAGHPGPPGPPGPPGEVIQPLPIQASR	4OH+Gal	K1554-OH.Gal	P1550-OH?, P1553-OH?, P1560-OH?, P1562-OH?, P1563-OH?, P1565-OH?, P1566-OH?	P1560-OH, P1562-OH, P1566-OH (Gly-X-Hyp motif)
60	1557	1571	AGHPGPPGPPGPPGE	4OH	P1560-OH, P1563-OH, P1566-OH, P1569-OH		

```
        10         20         30         40         50         60
MDVHTRWKAR SALRPGAPLL PPLLLLLLWA PPPSRAAQPA DLLKVLDFHN LPDGITKTTG

        70         80         90        100        110        120
FCATRRSSKG PDVAYRVTKD AQLSAPTKQL YPASAFPEDF SILTTVKAKK GSQAFLVSIY

       130        140        150        160        170        180
NEQGIQQIGL ELGRSPVFLY EDHTGKPGPE DYPLFRGINL SDGKWHRIAL SVHKKNVTLI

       190        200        210        220        230        240
LDCKKKTTKF LDRSDHPMID INGIIVFGTR ILDEEVFEGD IQQLLFVSDH RAAYDYCEHY

       250        260        270        280        290        300
SPDCDTAVPD TPQSQDPNPD EYYTEGDGEG ETYYYEYPYY EDPEDLGKEP TPSKKPVEAA

       310        320        330        340        350        360
KETTEVPEEL TPTPTEAAPM PETSEGAGKE EDVGIGDYDY VPSEDYYTPS PYDDLTYGEG

       370        380        390        400        410        420
EENPDQPTDP GAGAEIPTST ADTSNSSNPA PPPGEGADDL EGEFTEETIR NLDENYYDPY

       430        440        450        460        470        480
YDPTSSPSEI GPGMPANQDT IYEGIGGPRG EKGQKGEPAI IEPGMLIEGP PGPEGPAGLP

       490        500        510        520        530        540
GPPGTMGPTG QVGDPGERGP PGRPGLPGAD GLPGPPGTML MLPFRFGGGG DAGSKGPMVS

       550        560        570        580        590        600
AQESQAQAIL QQARLALRGP AGPMGLTGRP GPVGPPGSGG LKGEPGDVGP QGPRGVQGPP

       610        620        630        640        650        660
GPAGKPGRRG RAGSDGARGM PGQTGPKGDR GFDGLAGLPG EKGHRGDPGP SGPPGPPGDD

       670        680        690        700        710        720
GERGDDGEVG PRGLPGEPGP RGLLGPKGPP GPPGPPGVTG MDGQPGPKGN VGPQGEPGPP

       730        740        750        760        770        780
GQQGNPGAQG LPGPQGAIGP PGEKGPLGKP GLPGMPGADG PPGHPGKEGP PGEKGGQGPP

       790        800        810        820        830        840
GPQGPIGYPG PRGVKGADGI RGLKGTKGEK GEDGFPGFKG DMGIKGDRGE IGPPGPRGED

       850        860        870        880        890        900
GPEGPKGRGG PNGDPGPLGP PGEKGKLGVP GLPGYPGRQG PKGSIGFPGF PGANGEKGGR

       910        920        930        940        950        960
GTPGKPGPRG QRGPTGPRGE RGPRGITGKP GPKGNSGGDG PAGPPGERGP NGPQGPTGFP

       970        980        990       1000       1010       1020
GPKGPPGPPG KDGLPGHPGQ RGETGFQGKT GPPGPPGVVG PQGPTGETGP MGERGHPGPP

      1030       1040       1050       1060       1070       1080
GPPGEQGLPG LAGKEGTKGD PGPAGLPGKD GPPGLRGFPG DRGLPGPVGA LGLKGNEGPP

      1090       1100       1110       1120       1130       1140
GPPGPAGSPG ERGPAGAAGP IGIPGRPGPQ GPPGPAGEKG APGEKGPQGP AGRDGLQGPV
```

```
      1150        1160        1170        1180        1190        1200
GLPGPAGPVG PPGEDGDKGE IGEPGQKGSK GDKGEQGPPG PTGPQGPIGQ PGPSGADGEP

      1210        1220        1230        1240        1250        1260
GPRGQQGLFG QKGDEGPRGF PGPPGPVGLQ GLPGPPGEKG ETGDVGQMGP PGPPGPRGPS

      1270        1280        1290        1300        1310        1320
GAPGADGPQG PPGGIGNPGA VGEKGEPGEA GEPGLPGEGG PPGPKGERGE KGESGPSGAA

      1330        1340        1350        1360        1370        1380
GPPGPKGPPG DDGPKGSPGP VGFPGDPGPP GEPGPAGQDG PPGDKGDDGE PGQTGSPGPT

      1390        1400        1410        1420        1430        1440
GEPGPSGPPG KRGPPGPAGP EGRQGEKGAK GEAGLEGPPG KTGPIGPQGA PGKPGPDGLR

      1450        1460        1470        1480        1490        1500
GIPGPVGEQG LPGSPGPDGP PGPMGPPGLP GLKGDSGPKG EKGHPGLIGL IGPPGEQGEK

      1510        1520        1530        1540        1550        1560
GDRGLPGPQG SSGPKGEQGI TGPSGPIGPP GPPGLPGPPG PKGAKGSSGP TGPKGEAGHP

      1570        1580        1590        1600        1610        1620
GPPGPPGPPG EVIQPLPIQA SRTRRNIDAS QLLDDGNGEN YVDYADGMEE IFGSLNSLKL

      1630        1640        1650        1660        1670        1680
EIEQMKRPLG TQQNPARTCK DLQLCHPDFP DGEYWVDPNQ GCSRDSFKVY CNFTAGGSTC

      1690        1700        1710        1720        1730        1740
VFPDKKSEGA RITSWPKENP GSWFSEFKRG KLLSYVDAEG NPVGVVQMTF LRLLSASAHQ

      1750        1760        1770        1780        1790        1800
NVTYHCYQSV AWQDAATGSY DKALRFLGSN DEEMSYDNNP YIRALVDGCA TKKGYQKTVL

      1810        1820        1830
EIDTPKVEQV PIVDIMFNDF GEASQKFGFE VGPACFMG
```

K P: hydroxylation

K: gal.glc hydroxylysine

K: gal

K: Gal.Glc or Gal hydroxylysine

Red sequence: observed in MS

Black sequence: not observed in MS