

Post-translational modification of type IV collagen with 3-hydroxyproline affects its interactions with glycoprotein VI and nidogens 1 and 2

Nathan T. Montgomery<sup>1,2</sup>, Keith D. Zientek<sup>1</sup>, Elena N. Pokidysheva<sup>3</sup>, Hans Peter Bächinger<sup>1,2</sup>

**Supplemental File**

**Supplemental Methods**

*Karyotyping* - All karyotyping procedures and interpretation were performed by the Oregon Health and Science University Research Cytogenetics Laboratory. Confluent cells were exposed to colcemid (0.05 ug / mL) for 2 hours to arrest cells at metaphase and then were harvested. Cells were incubated in 75 mM KCl, 5% Fetal Bovine Serum for 10 minutes, then fixed in 3:1 methanol acetic acid. The cell suspension was dropped onto alcohol cleaned microscope slides and baked at 90 °C for 20 minutes. After cooling, the slides were trypsinized for 45 seconds, incubated with Wright stain for 80 seconds, rinsed with diH<sub>2</sub>O and dried. Chromosomes were imaged using bright field microscopy, and analyzed using Cytovision software (Applied Imaging, San Jose, CA). Karyotype designations were assigned according to MGI rules mouse nomenclature of chromosome abnormalities.

*Circular Dichroism* - Measurements were all made on an AVIV CD Spectrometer Model 202, using AVIV CDS.EXE Version 2.88 instrument control and data acquisition software running on a Windows XP computer. Samples were placed into 1 mm path length quartz cuvettes (Starna Cells, 21-Q-1). Samples were dialyzed into 10 mM acetic acid. Wavelength scans were conducted over a range from 200 to 300 nm with 0.5 nm steps, 60 second averaging per step and 1.0 nm bandwidth at 25 °C. Temperature scans observed CD signal at 220 nm (1.0 nm wavelength) and were conducted over a 25 to 55 °C range (High concentration, Figure 2a.) or a 4 to 65 °C range (Figure 2b). The temperature was increased at a rate of 10 °C / hour (36 second averaging time). Sample concentrations for WT and 6A6 (P3H2 KO) PFHR-9 type IV collagen were obtained by amino acid analysis (see Methods) and normalized for wavelength and temperature scans, used to obtain molar ellipticities. Melting curves were obtained from transformation of temperature scan data using a method found in Persikov et al. 2004 (1 ).

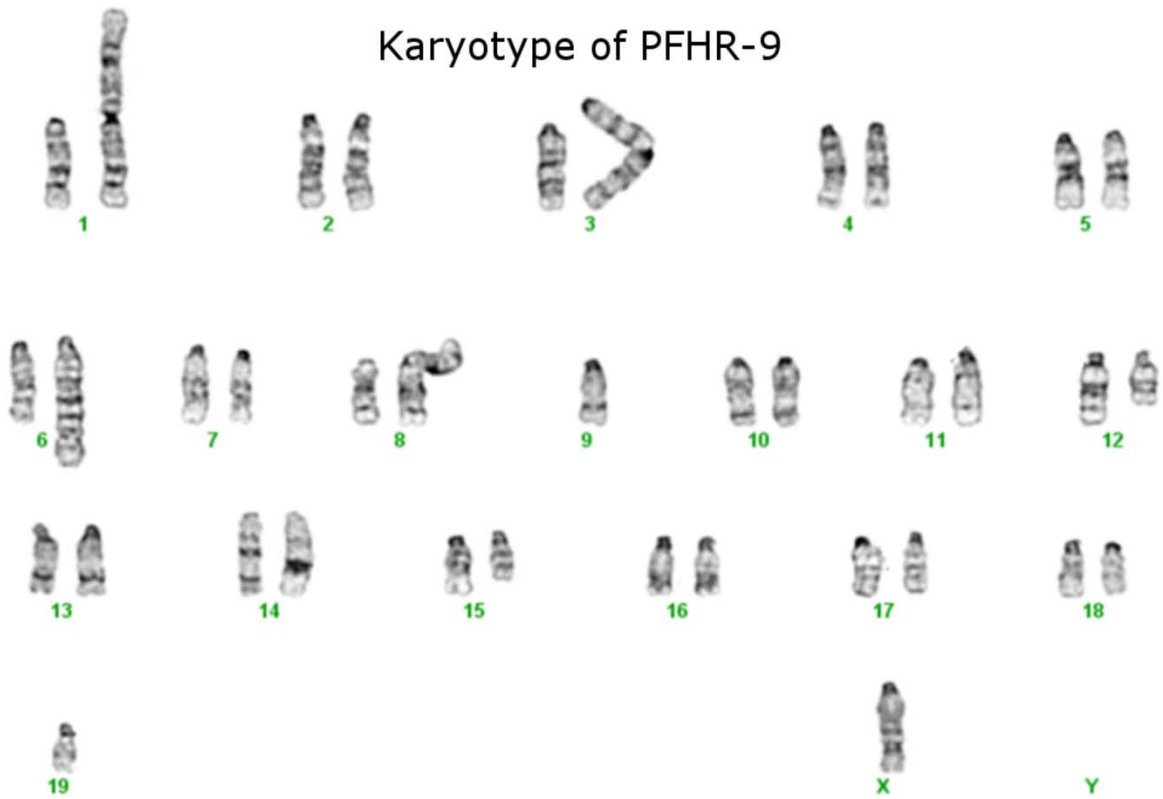
Supplemental Table 1. Amino acid composition of type IV collagen samples from PFHR-9 cells.

Amino Acid	Experimental Composition WT	Experimental Composition P3H2 KO	Amino Acids Per Triple Helix
D + N	268.9	259.9	244
T	154.8	156.0	148
S	216.1	212.8	225
E + Q	491.8	448	432
A	170.8	179.4	164
V	175.7	176.9	170
M	94.1	95.7	93
I	142.4	143.5	143
L	305.5	317	253
Y	53.2	59.3	52
F	161.0	172	156
H	57.5	56.7	56
R	166.0	160.3	157

**Supplemental Table 1.** *Amino acid composition of type IV collagen samples from PFHR-9 cells.*

Experimental concentration of amino acids was derived from peak integration and the average of three runs. Amino acid compositions are shown for D + N and E + Q, because the amide side chains residues (N or Q) are converted to the corresponding carboxylic acids (D or E) during acid hydrolysis. The predicted amino acid composition is derived from the amino acid sequence of a hetero trimeric mouse type IV collagen triple helix (using UniprotKB entries P02463-1 and P08122-1), lacking the N-terminal propeptide.

## Supplemental Figures



**Supplemental Figure 1.** *Karyotype of PFHR-9, an Aneuploid Mouse Teratocarcinoma Derived Cell Line.* PFHR-9 is diploid for chromosome 16 (containing *LEPREL1*, the gene of interest for KO), effectively triploid for chromosome 8 (containing *COL4A1* and *COL4A2*), and is missing a Y chromosome though the original tumor cells were from a male mouse. Detailed interpretation of karyotypes is found in this supplement and karyotyping was performed by the Oregon Health & Science University Cytogenetics Research Core.

## OHSU Research Cytogenetics Laboratory Karotyping Report for PFHR-9

Chromosome Results: **Abnormal**

Karyotype:

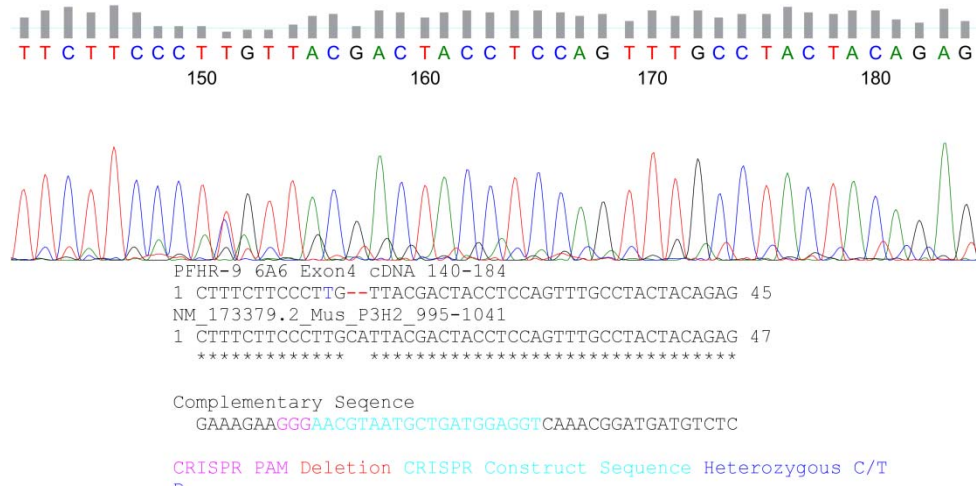
**37,X,Rb(1.der(1)T(1H5;3H2)),Add(3G),Dic(3;9)(9A1→9B::9D→9F2::9F2→9B::3A1→3H4),Der(6)T(6G3;19B)Is(?;6D),+8,Rb(8.8),Del(12E-12F2),InDp(14D2;14E2),-19[cp10]/37,X,sl,T(14E4;15D3)[cp7]/37,X,sl,-Rb(1.der(1)T(1H5;3H2)),+Rb(der(1)T(1H5;3H2).der(1)T(1H5;3H2)),+14,-InDp(14D2;14E2)[cp3]**

Interpretation and Comments:

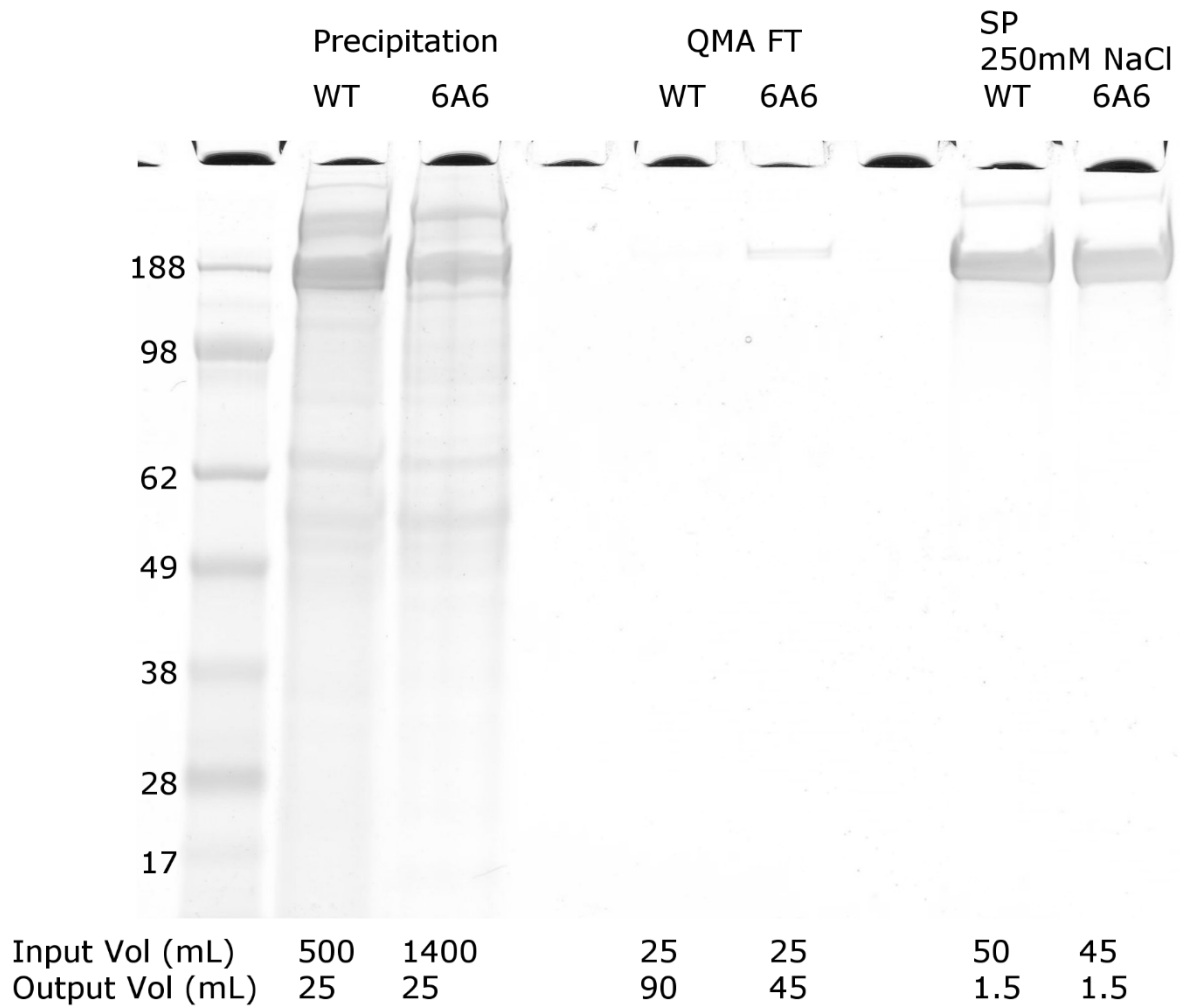
All twenty metaphase cells examined comprised three related, complex abnormal clones, all with 37 chromosomes and a single sex chromosome. The stemline contains seven structural abnormalities:

- 1) a robertsonian translocation (fusion at centromeres) involving an intact copy of chromosome 1 and a derivative chromosome 1
- 2) a chromosome 3 with additional material of unknown origin attached at band G
- 3) a dicentric chromosome involving chromosome 3 and chromosome 9 (with duplicated and deleted material on chromosome 9)
- 4) a derivative chromosome 6 from a translocation with chromosome 19 and an insertion of material of unknown origin at band D
- 5) a robertsonian translocation involving two copies of chromosome 8 (in addition to a free copy of chromosome 8, thus, essentially trisomy 8)
- 6) a deletion on chromosome 12
- 7) an inverted duplication on chromosome 14

Two sidelines were present in this cell line as well. Seven cells had a translocation between chromosome 14 and chromosome 15. Three other cells lost the inverted and duplicated chromosome 14 while gaining a normal chromosome 14, and replaced the robertsonian translocation between chromosome 1 and derivative chromosome 1 with one between two copies of the derivative chromosome 1.



**Supplemental Figure 2.** Sequencing of exon 4 of *PFHR-9 6A6 LEPREL1* gene. An electropherogram of the region surrounding the disrupted portion of *LEPREL1* (coding for P3H2) exon 4 is shown. The location of the deletion (red), and location of a heterozygous base (C/T) (blue) are color coded. The complementary sequence used to generate the CRISPR construct (teal) and the corresponding CRISPR PAM sequence (purple) are also color coded.



**Supplemental Figure 3.** Summary Gel of Purification of type IV collagen from PFHR-9 Cell Culture Media from WT and 6A6 (P3H2 KO) Cells. Input volume in precipitation step is volume of thawed PFHR-9 culture media and output volume is the resuspension volume of the ammonium sulfate precipitate. QMA flow through input volume is the input volume of clarified resuspended ammonium sulfate precipitate and output volume is the volume of pooled flow-thru from the QMA column. SP 250 mM NaCl input volume is the volume of pooled QMA flow-thru loaded onto the SP Sepharose column, output volume is the pooled volume from the most concentrated fractions eluted at 250 mM NaCl. Equal volumes of sample were loaded onto each lane of a 4-12 % BOLT MES gel.

1			50
Human_Col4A1	(1)	MGPRLSVWLLLLPAALLLHEEHSRAAAKGGCAGSGCGKCDCHGVKQKGE	
mouse_Col4A1	(1)	MGPRLSVWLLLLFAALLLHEERSRAAAKGDCCGSGCGKCDCHGVKQKGE	
		51	100
Human_Col4A1	(51)	RGLPGLQGVIGFPGMQGPEGPQGGPPGQKGDTEGEPGLPGTKGTRGPPGASG	
mouse_Col4A1	(51)	RGLPGLQGVIGFPGMQGPEGPHGPPGQKGDAGEPGLPGTKGTRGPPGAAG	
		101	150
Human_Col4A1	(101)	YPGNPGLPGIPGQDGGPPGPPGIPGCNGTKGERGRLGPPGLPGFAGNPGPP	
mouse_Col4A1	(101)	YPGNPGLPGIPGQDGGPPGPPGIPGCNGTKGERGRLGPPGLPGFSGNPGPP	
		151	200
Human_Col4A1	(151)	GLPGMKGDPGEILGHVPGMLLKGERGFPGIPGTPGPPGLPGLQGPVGGPPG	
mouse_Col4A1	(151)	GLPGMKGDPGEILGHVPGTLLKGERGFPGIPGMPGSPGLPGLQGPVGGPPG	
		201	250
Human_Col4A1	(201)	FTGPPGPPGPPGPPGPEKQMGLSFQGPKGDKDQGVSGPPGVPGQAQVQE	
mouse_Col4A1	(201)	FTGPPGPPGPPGPPGPEKQMGSSFQGPKGDKGEQGVSGPPGVPGQAQVKE	
		251	300
Human_Col4A1	(251)	KGDFATKGEKGQKGEPPGFQGMVGVGEKGEPPGKPGPRGKPGKDGDKGEKGS	
mouse_Col4A1	(251)	KGDFAPTGEKGQKGEPPGFVPGVGEKGEPPGKQGPGRGKPGKDGDKGERGS	
		301	350
Human_Col4A1	(301)	PGFPGEPPGYPLIGRQGPQGEKGEAGPPGPPGIVIGTGPLGEKGERGYPG	
mouse_Col4A1	(301)	PGIPGDSGYPLPGRQGPQGEKGEAGLPGPPGTIVIGTMPLGEKGDRTGYPG	
		351	400
Human_Col4A1	(351)	TPGPRGEPGPKGFPLPGQPGPPGLVPPGQAGAPGFPPGERGEKGDRTGFPG	
mouse_Col4A1	(351)	APGLRGEPPGPKGFPGTTPGQPGPPGFPTPGQAGAPGFPPGERGEKGDQGFPG	
		401	450
Human_Col4A1	(401)	TSLPGPSGRDGLPGPPGSPGPPGQPGYTNGIVECQPGPPGDQGGPPGIPGQ	
mouse_Col4A1	(401)	VSLPGPSGRDGAAPGPPGPPGPPGQPGHTNGIVECQPGPPGDQGGPPGTTPGQ	
		451	500
Human_Col4A1	(451)	PGFIGEIGEKQKGESCLICDIDGYRGPPGPPGPPGEGIFGPPGQPGAKGDR	
mouse_Col4A1	(451)	PGLTGEVQKQKGESCLACDTEGLRGPVPPGPPGPPGEGIFGPPGQPGAKGDR	
		501	550
Human_Col4A1	(501)	GLPGRDGVAGVPPGQGTPLIGQPGAKGEPGEFYFDLRLKGDKDPGFPPG	
mouse_Col4A1	(501)	GLPGRDGLEGLPVPQGSPLIGQPGAKGEPGEIFFDMRLKGDKDPGFPPG	
		551	600
Human_Col4A1	(551)	QPGMTGRAGSPGRDGHPLPGPKGSPGVSGLKGERGPPGGVGFPPGSRGDT	
mouse_Col4A1	(551)	QPGMPGRAGTTPGRDGHPLPGPKGSPGSIGLKGERGPPGGVGFPPGSRGDI	
		601	650
Human_Col4A1	(601)	GPPGPPGYGAPGIGDKQAGFPGGPGSPGLPGPKGEPGKIVPLPGPPGA	
mouse_Col4A1	(601)	GPPGPPGVGPIGPVGEKQAGFPGGPGSPGLPGPKGEPGKVVPLPGPPGA	
		651	700
Human_Col4A1	(651)	EGLPGSPGFPPGQDGRGFPGTTPRPLPGEKGAVGQPGIGFPPGPPGPKGV	
mouse_Col4A1	(651)	AGLPGSPGFPPGQDGRGFPGTTPRPLPGEKGAVGQPGIGFPPGLPGPKGV	
		701	750
Human_Col4A1	(701)	DGLPGDMGPPGTPGRPGFNGLPGNPGVQGGKGEPPVGLPGLKGLPGLPGI	
mouse_Col4A1	(701)	DGLPGEIIGRPGSPGRPGFNGLPGNPGVQGGKGEPPVGLPGLKGLPGLPGI	
		751	800
Human_Col4A1	(751)	PGTPGEKGSIGVPGVPGEHGAIGPPGLQGIRGEPGPPGLPGSVGSPGVPG	
mouse_Col4A1	(751)	PGTPGEKGSIGGPGVPGEQGLTGPPGLQGIRGDPGPPGVQGPAGPPGVPG	
		801	850
Human_Col4A1	(801)	IGPPGARPPGGQGGPPGLSGPPGKKEKGFPGFPGLDMPGPKGDKGAQGL	
mouse_Col4A1	(801)	IGPPGAMPPGGQGGPPGSSGPPGKKEKGFPGFPGLDMPGPKGDKGSQGL	
		851	900
Human_Col4A1	(851)	PGITGQSGPLPGLPGQAGPIPGFPGSKGEMGVMGTPGQPGSPGPPVAGP	
mouse_Col4A1	(851)	PGLTGQSGPLPGLPGQGTTPGVPGFPGSKGEMGVMGTPGQPGSPGPPAGTPG	
		901	950
Human_Col4A1	(901)	LPGEKGDHGFPGSSGPRGDPGLKGDKGDVGLPGKPGSMDKVDMSGSMKQK	
mouse_Col4A1	(901)	LPGEKGDHGLPGSSGPRGDPGFKGDKGDVGLPGMPGSMHVDMSGSMKQK	

		951		1000
Human_Col4A1	(951)	GDQGEKGGQIGP I G E K G S R G D P G T P G V P G K D G Q A G Q P G Q P G P K G D P G I S G T		
mouse_Col4A1	(951)	GDQGEKGGQIGP T G D K G S R G D P G T P G V P G K D G Q A G H P G Q P G P K G D P G L S G T		
		1001		1050
Human_Col4A1	(1001)	P G A P G L P G P K G S V G G M G L P G T P G E K G V P G I P G P Q G S P G L P G D K G A K G E K G		
mouse_Col4A1	(1001)	P G S P G L P G P K G S V G G M G L P G S P G E K G V P G I P G S Q G V P G S P G E K G A K G E K G		
		1051		1100
Human_Col4A1	(1051)	Q A G P P G I G I P G L R G E K G D Q G I A G F P G S P G E K G E K G S I G I P G M P G S P G L K G		
mouse_Col4A1	(1051)	Q S G L P G I G I P G R P G D K G D Q G L A G F P G S P G E K G E K G S A G T P G M P G S P G P R G		
		1101		1150
Human_Col4A1	(1101)	S P G S V G Y P G S P G L P G E K G D K G L P G L D G I P G V K G E A G L P G T P G T P G P A G Q K		
mouse_Col4A1	(1101)	S P G N I G H P G S P G L P G E K G D K G L P G L D G V P G V K G E A G L P G T P G T P G P A G Q K		
		1151		1200
Human_Col4A1	(1151)	G E P G S D G I P G S A G E K G E P G L P G R G F P G F P G A K G D K G S K G E V G F P G L A G S P		
mouse_Col4A1	(1151)	G E P G S D G I P G S A G E K G E Q G V P G R G F P G F P G S K G D K G S K G E V G F P G L A G S P		
		1201		1250
Human_Col4A1	(1201)	G I P G S K G E Q G F M G P P G P Q G Q P G L P G S P G H A T E G P K G D R G P Q G Q P G L P G L P		
mouse_Col4A1	(1201)	G I P G V K G E Q G F M G P P G P Q G Q P G L P G T P G H P V E G P K G D R G P Q G Q P G L P G H P		
		1251		1300
Human_Col4A1	(1251)	G P M G P P G L P G I D G V K G D K G N P G W P G A P G V P G P K G D P G F Q G M P G I G G S P G I		
mouse_Col4A1	(1251)	G P M G P P G F P G I N G P K G D K G N Q G W P G A P G V P G P K G D P G F Q G M P G I G G S P G I		
		1301		1350
Human_Col4A1	(1301)	T G S K G D M G P P G V P G F Q G P K G L P G L Q G I K G D Q G D Q G V P G A K G L P G P P G P P G		
mouse_Col4A1	(1301)	T G S K G D M G L P G V P G F Q G Q K G L P G L Q G V K G D Q G D Q G V P G P K G L Q G P P G P P G		
		1351		1400
Human_Col4A1	(1351)	P Y D I I K G E P G L P G P E G P P G L K G L Q G L P G P K G Q Q G V T G L V G I P G P P G I P G F		
mouse_Col4A1	(1351)	P Y D V I K G E P G L P G P E G P P G L K G L Q G P P G P K G Q Q G V T G S V G L P G P P G V P G F		
		1401		1450
Human_Col4A1	(1401)	D G A P G Q K G E M G P A G T P G P R G F P G P P G P D G L P G S M G P P G T P S V D H G F L V T R		
mouse_Col4A1	(1401)	D G A P G Q K G E T G P F G P P G P R G F P G P P G P D G L P G S M G P P G T P S V D H G F L V T R		
		1451		1500
Human_Col4A1	(1451)	H S Q T I D D P Q C P S G T K I L Y H G Y S L L Y V Q G N E R A H G Q D L G T A G S C L R K F S T M		
mouse_Col4A1	(1451)	H S Q T T D D P L C P P G T K I L Y H G Y S L L Y V Q G N E R A H G Q D L G T A G S C L R K F S T M		
		1501		1550
Human_Col4A1	(1501)	P F L F C N I N N V C N F A S R N D Y S Y W L S T P E P M P M S M A P I T G E N I R P F I S R C A V		
mouse_Col4A1	(1501)	P F L F C N I N N V C N F A S R N D Y S Y W L S T P E P M P M S M A P I S G D N I R P F I S R C A V		
		1551		1600
Human_Col4A1	(1551)	C E A P A M V M A V H S Q T I Q I P P C P S G W S S L W I G Y S F V M H T S A G A E G S G Q A L A S		
mouse_Col4A1	(1551)	C E A P A M V M A V H S Q T I Q I P Q C P N G W S S L W I G Y S F V M H T S A G A E G S G Q A L A S		
		1601		1650
Human_Col4A1	(1601)	P G S C L E E F R S A P F I E C H G R G T C N Y Y A N A Y S F W L A T I E R S E M F K K P T P S T L		
mouse_Col4A1	(1601)	P G S C L E E F R S A P F I E C H G R G T C N Y Y A N A Y S F W L A T I E R S E M F K K P T P S T L		
		1651		1669
Human_Col4A1	(1651)	K A G E L R T H V S R C Q V C M R R T		
mouse_Col4A1	(1651)	K A G E L R T H V S R C Q V C M R R T		

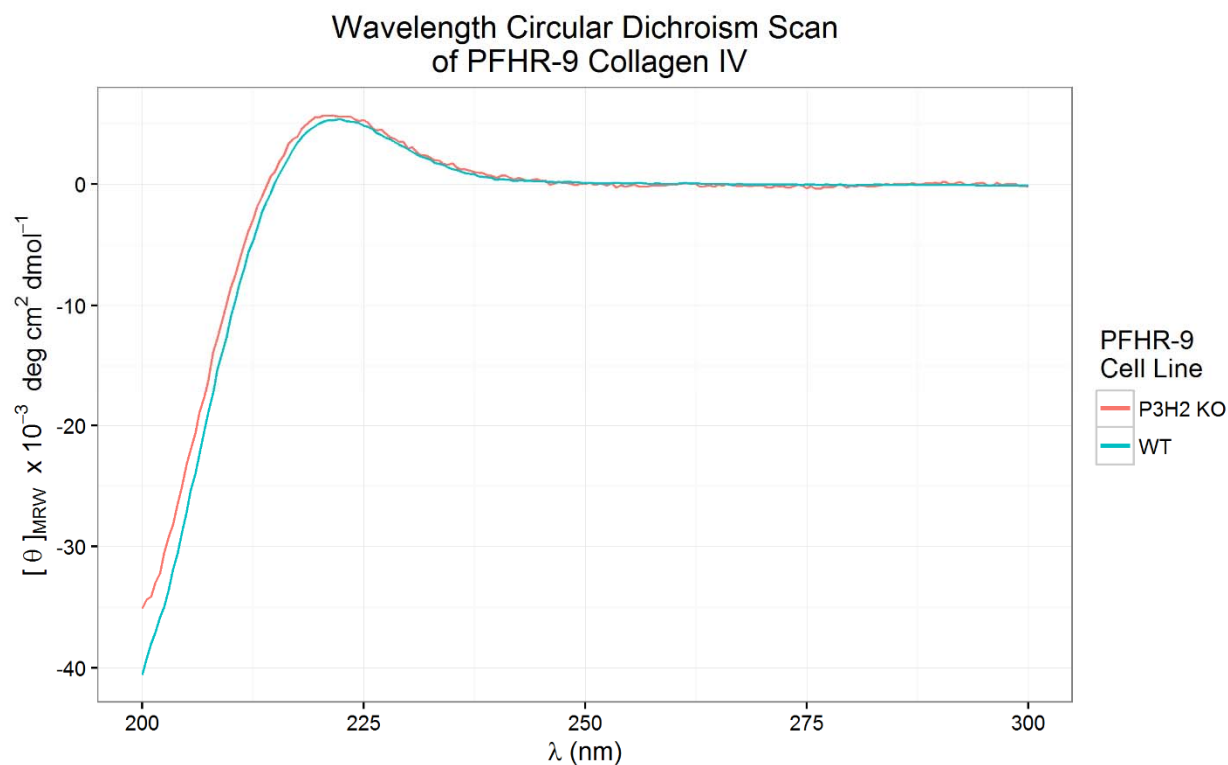
**Supplemental Figure 4.** Sequence alignment of mouse and human Type IV collagen  $\alpha 1$  chain. X-position prolines predicted to be 3-hydroxylated in mouse, based on the mass spectrometry data reported here, are highlighted in blue. Sequences with identity between mouse and human are highlighted in yellow, similar amino acids (i.e. I/L, T/S) are highlighted in green.



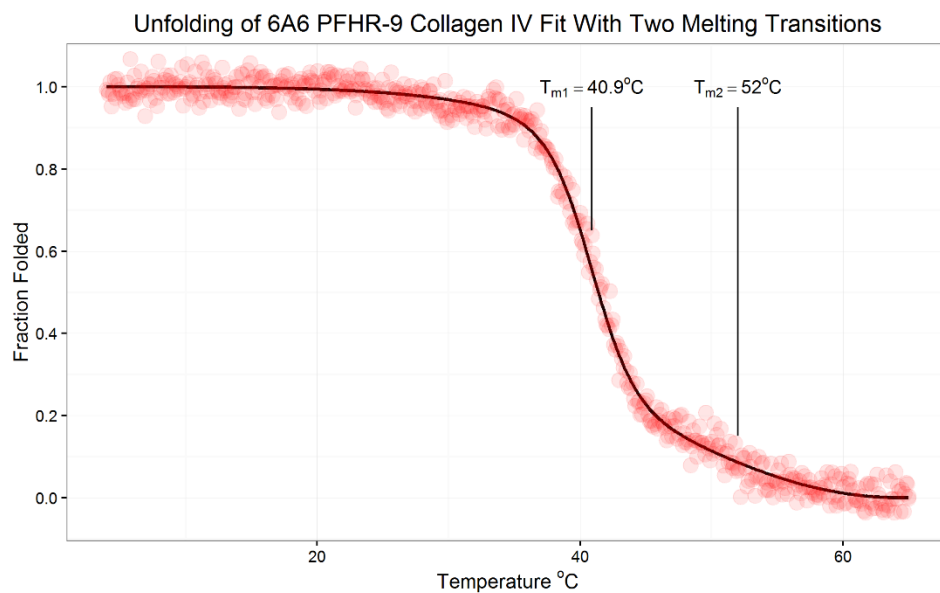
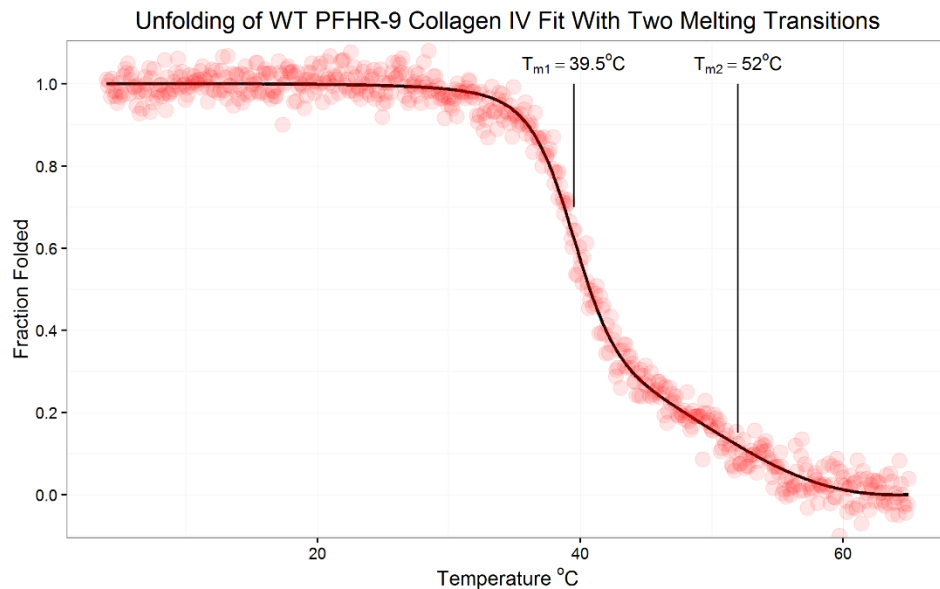
1	50	
human_Col4A2	(1)	MGRDQRAVAGPALRRWLLLGTVTVGF <del>LAQSVLAGVKKF</del> DVPCGGRDCSSG
mouse_Col4A2	(1)	MDRVRFKASGPPLRGWLLLATVTVGL <del>LAQSVLGGVKKL</del> DVPCGGRDCSSG
		51 100
human_Col4A2	(51)	CQCYPEKGGRGQPGPVGPGQYNGPPGLQGF <del>PGLQGRKGDKGERG</del> APGVTG
mouse_Col4A2	(51)	CQCYPEKGGARGQPGAVGPGQYNGP <del>PGLQGF</del> PGLQGRKGDKGERGVPGPTG
		101 150
human_Col4A2	(101)	PKGDVGARGVSGFFGADGIPGHPGQGGPRGRPGYDGCNGTQGD <del>SGPQGP</del> P
mouse_Col4A2	(101)	PKGDVGARGVSGFFGADGIPGHPGQGGPRGRPGYDGCNGTRGD <del>AGPQGP</del> S
		151 200
human_Col4A2	(151)	GSEGF <del>TGPPGPGQPKGQKGE</del> PYALPKEERDRYRGEPEPEGLVGFQGP <del>PPGR</del>
mouse_Col4A2	(151)	GSGGF <del>PGLPGPQPKGQKGE</del> PYALSKEERDKYRGEPEPEGLVGYQGP <del>PPGR</del>
		201 250
human_Col4A2	(201)	PGHVGQMG <del>PVGAPGRPGPPGPKGQ</del> QGNRGLGFYGVKGEKGDVQ <del>PGP</del>
mouse_Col4A2	(201)	PGPIGQMG <del>PMGAPGRPGPPGPKGQ</del> PNRGLGFYGVKGEKGDIG <del>QPGP</del>
		251 300
human_Col4A2	(251)	NGIPSDTLHPIT <del>IAPT</del> GVTFHPDQYKGEKGS <del>E</del> EGEPGIRGI <del>ISL</del> KGEEGIMGF
mouse_Col4A2	(251)	NGIPSDIT--LVGPTT <del>STIHPDL</del> YKGEKGD <del>E</del> EGEQGIPGV <del>IS</del> KGEEGIMGF
		301 350
human_Col4A2	(301)	PGLRGYPGLS <del>GEKGS</del> SPGQKGS <del>RGLDGY</del> QGPDPGRGPKGEAGDP <del>GGP</del> LPA
mouse_Col4A2	(299)	PGIRGF <del>PGLDGEKGV</del> VGQKGS <del>RGLDGF</del> QGPSGPRGPKGERGE <del>QGP</del> PPSV
		351 400
human_Col4A2	(351)	YSPHPSLAKGARGDPGFPGA <del>QGE</del> PGSQ <del>GE</del> PGDPGLP <del>GGP</del> GLSIGDGD <del>QRR</del>
mouse_Col4A2	(349)	YSPHPSLAKGARGDPGFQGAH <del>GE</del> PGSR <del>GE</del> PEPGTAG <del>PP</del> PSV <del>GE</del> DEDSMR
		401 450
human_Col4A2	(401)	GLPGEMGPKGFI <del>DPG</del> IPALYGGPPGPDGKRGPPG <del>PGLP</del> GGP <del>GG</del> PDGFLF
mouse_Col4A2	(399)	GLPGEMGPKGFS <del>EPG</del> SPARYL <del>GPP</del> GADGRPGPQGV <del>PG</del> PAGPPG <del>PDG</del> FLF
		451 500
human_Col4A2	(451)	GLKGAKGRAGFP <del>GLP</del> GS <del>PG</del> ARGPKG <del>WKGD</del> AGECRCTEGDEA <del>IKGL</del> PGLPG
mouse_Col4A2	(449)	GLKGS <del>EGR</del> VGY <del>PG</del> PSGF <del>PG</del> TRGQK <del>GW</del> GEAGDC <del>CGQ</del> ---VIGGLPGLPG
		501 550
human_Col4A2	(501)	PKGFAGINGEP <del>GRK</del> GD <del>RD</del> PGQH <del>GLP</del> GF <del>PGL</del> KGV <del>PGNI</del> GAPGPKGAKGDS
mouse_Col4A2	(496)	PKGF <del>PGV</del> NGEL <del>GKK</del> GD <del>QD</del> PGLH <del>GI</del> PGF <del>PG</del> FKGAP <del>GVA</del> GAPGPKGIK <del>GD</del> S
		551 600
human_Col4A2	(551)	RTITTKGERGQ <del>PGV</del> PGV <del>PG</del> MK <del>GDD</del> GS <del>PR</del> RD <del>GLD</del> GF <del>PGL</del> PG <del>PP</del> GD <del>G</del> IK <del>GP</del> P
mouse_Col4A2	(546)	RTITTKGERGQ <del>PGI</del> PGV <del>HGM</del> K <del>GDD</del> GV <del>PR</del> RD <del>GLD</del> GF <del>PGL</del> PG <del>PP</del> GD <del>G</del> IK <del>GP</del> P
		601 650
human_Col4A2	(601)	GD <del>PGY</del> PGIP <del>GT</del> KGT <del>PG</del> EM <del>GP</del> PG <del>LGL</del> P <del>LK</del> Q <del>R</del> GF <del>PG</del> DAG <del>LPG</del> PP <del>GF</del> L <del>GP</del> P
mouse_Col4A2	(596)	GDAGL <del>PGV</del> PGT <del>KGF</del> PGDI <del>PPP</del> QGL <del>PG</del> PK <del>GER</del> GF <del>PG</del> DAG <del>LPG</del> PP <del>GF</del> PG <del>PP</del>
		651 700
human_Col4A2	(651)	GPAGTPGQID <del>CDT</del> DV <del>KRA</del> VGG <del>DR</del> QEAIQ <del>PG</del> CI <del>GG</del> PK <del>GLP</del> GL <del>PG</del> PP <del>GP</del> TGA
mouse_Col4A2	(646)	GP <del>PG</del> TPGQR <del>DC</del> DTG <del>VKR</del> PI <del>GG</del> GQ <del>VVV</del> Q <del>PG</del> CI <del>EG</del> PT <del>GS</del> PG <del>Q</del> PP <del>GP</del> TGA
		701 750
human_Col4A2	(701)	KGLRGI <del>PG</del> FAGADGG <del>PG</del> PR <del>GL</del> PGDAG <del>REG</del> FP <del>GG</del> PGFI <del>GPR</del> GSKGAV <del>GL</del> PG
mouse_Col4A2	(696)	KGV <del>RG</del> MPGF <del>PG</del> AS <del>GE</del> QL <del>GK</del> GF <del>PG</del> D <del>PG</del> REG <del>FP</del> GG <del>PM</del> GPRGSKGTT <del>GL</del> PG
		751 800
human_Col4A2	(751)	PDGSPGI <del>GL</del> PG <del>D</del> GG <del>PE</del> RG <del>LPE</del> VLGA <del>Q</del> PG <del>PR</del> DAG <del>VPG</del> Q <del>PL</del> K <del>GL</del> PG
mouse_Col4A2	(746)	PDGPPGI <del>GL</del> PG <del>A</del> GG <del>PD</del> RG <del>IP</del> GE <del>VL</del> GA <del>Q</del> PG <del>TR</del> DAG <del>LPG</del> Q <del>PL</del> K <del>GL</del> PG
		801 850
human_Col4A2	(801)	DRGPPGFRGSQ <del>GM</del> PG <del>MP</del> GL <del>KG</del> Q <del>PGL</del> PG <del>PS</del> Q <del>PG</del> LY <del>GG</del> PP <del>LHG</del> FP <del>GA</del> PG <del>QE</del>
mouse_Col4A2	(796)	ETGAPGFRGSQ <del>GM</del> PG <del>MP</del> GL <del>KG</del> Q <del>PGF</del> PG <del>PS</del> Q <del>PG</del> QS <del>GG</del> PP <del>QH</del> GF <del>PG</del> T <del>PGRE</del>
		851 900
human_Col4A2	(851)	G <del>PL</del> GL <del>PGI</del> PG <del>RE</del> GL <del>PG</del> DR <del>DP</del> GD <del>TG</del> AP <del>GP</del> V <del>GM</del> K <del>GLS</del> GD <del>RG</del> DAG <del>FT</del> GE <del>QH</del>
mouse_Col4A2	(846)	G <del>PL</del> Q <del>PGS</del> P <del>GL</del> GL <del>PG</del> DR <del>EP</del> GD <del>PG</del> V <del>GP</del> V <del>GM</del> K <del>GLS</del> GD <del>RG</del> DAG <del>MS</del> GER <del>GH</del>
		901 950
human_Col4A2	(901)	PGSPGF <del>FKG</del> ID <del>GM</del> PGT <del>PL</del> K <del>GD</del> RG <del>SP</del> GM <del>DG</del> F <del>QG</del> MP <del>GL</del> K <del>GR</del> PF <del>FG</del> SK <del>GE</del> AG
mouse_Col4A2	(896)	PGSPGF <del>FKG</del> MAG <del>MP</del> GI <del>PG</del> Q <del>K</del> GD <del>RG</del> SP <del>GM</del> D <del>G</del> F <del>QG</del> ML <del>GL</del> K <del>GR</del> Q <del>GF</del> PG <del>T</del> K <del>GE</del> AG

		951		1000
human_Col4A2	(951)	FFG	I	PGLKGLAGEPGFKGSRGDPGPPGPPPVILPGMKDIKGEKGDDEGPMG
mouse_Col4A2	(946)	FFG	V	PGLKGLPGEPEVKGNRGRGPPGPPPLILPGMKDIKGEKGDDEGPMG
		1001		1050
human_Col4A2	(1001)	LKGYL	GAKGIQGMPI	PGLSGI
mouse_Col4A2	(996)	LKGYL	LGLKGIQGMPI	PGLSGI
		1051		1100
human_Col4A2	(1051)	PGV	A	GPPGITGFP
mouse_Col4A2	(1046)	PGV	S	GPPGITGFP
		1101		1150
human_Col4A2	(1101)	GRP	L	KGERGT
mouse_Col4A2	(1096)	GSP	L	KGERGT
		1151		1200
human_Col4A2	(1151)	TGF	P	LTPG
mouse_Col4A2	(1146)	TGF	P	LTPG
		1201		1250
human_Col4A2	(1201)	LP	G	TG
mouse_Col4A2	(1196)	LP	G	TG
		1251		1300
human_Col4A2	(1251)	GAP	G	ERG
mouse_Col4A2	(1246)	GTP	G	ERG
		1301		1350
human_Col4A2	(1301)	PG	S	AAL
mouse_Col4A2	(1296)	PG	P	NAL
		1351		1400
human_Col4A2	(1351)	EQ	F	M
mouse_Col4A2	(1346)	EQ	F	M
		1401		1450
human_Col4A2	(1401)	PG	T	V
mouse_Col4A2	(1396)	PG	T	V
		1451		1500
human_Col4A2	(1451)	FR	G	D
mouse_Col4A2	(1446)	FR	G	D
		1501		1550
human_Col4A2	(1501)	EP	M	C
mouse_Col4A2	(1496)	EP	M	C
		1551		1600
human_Col4A2	(1551)	PG	D	V
mouse_Col4A2	(1546)	PG	D	V
		1601		1650
human_Col4A2	(1601)	AV	H	S
mouse_Col4A2	(1596)	AV	H	S
		1651		1700
human_Col4A2	(1651)	RAT	P	F
mouse_Col4A2	(1646)	RAT	P	F
		1701		1712
human_Col4A2	(1701)	HIS	R	C
mouse_Col4A2	(1696)	HIS	R	C

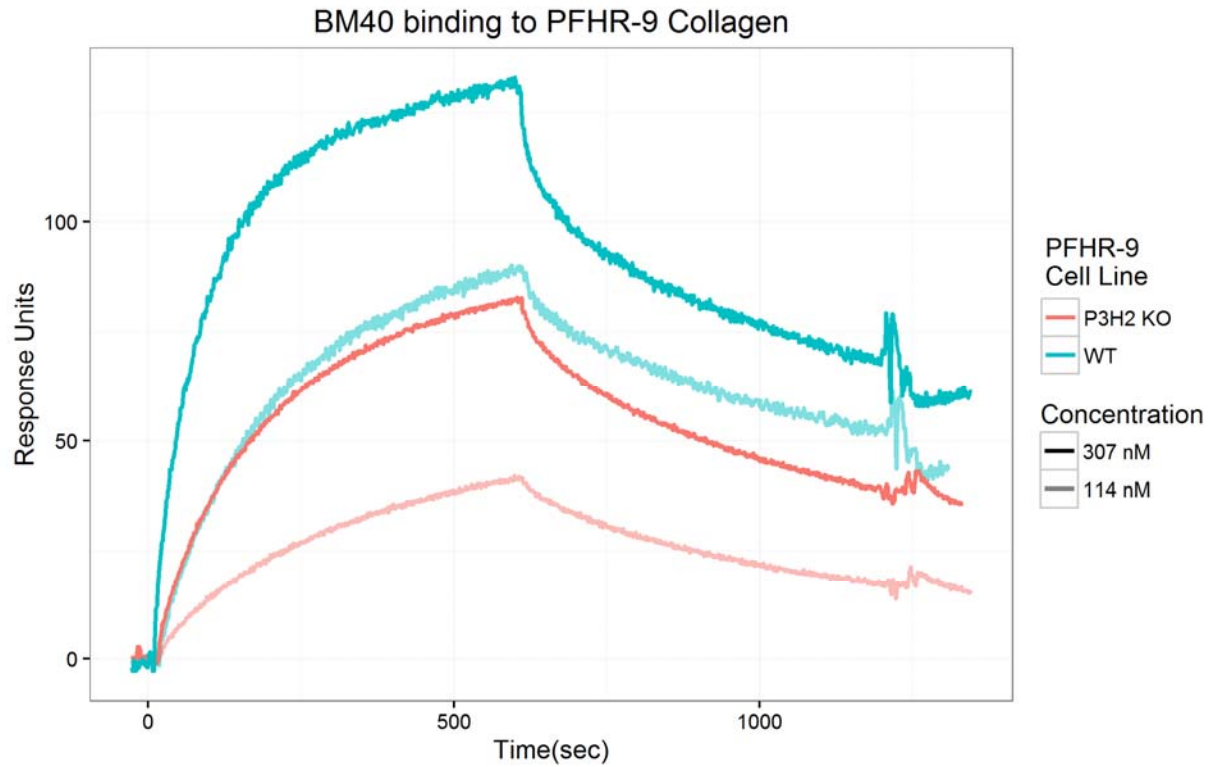
**Supplemental Figure 5.** Sequence alignment of mouse and human Type IV collagen  $\alpha 2$  chain. X-position prolines predicted to be 3-hydroxylated in mouse, based on the mass spectrometry data reported here, are highlighted in blue. Sequences with identity between mouse and human are highlighted in yellow. Similar amino acids (i.e. I/L, T/S) are highlighted in green.



**Supplemental Figure 6.** *Wavelength Scan comparison of wt and P3H2 ko (6A6) PFHR-9 Circular Dichroism Spectra.* Circular dichroism signal is normalized by concentration to molar ellipticity with standard units and are shown on the y-axis. Wavelength scans were acquired at 25 °C from 200 to 300 nm in 0.5 nm steps with 60 second signal averaging per scan. Baseline was set by averaging signal over the range from 240 to 300 nm for each scan. The curve for P3H2 ko type IV collagen scans is shown in red, while the curve for the wt scans is shown in blue.



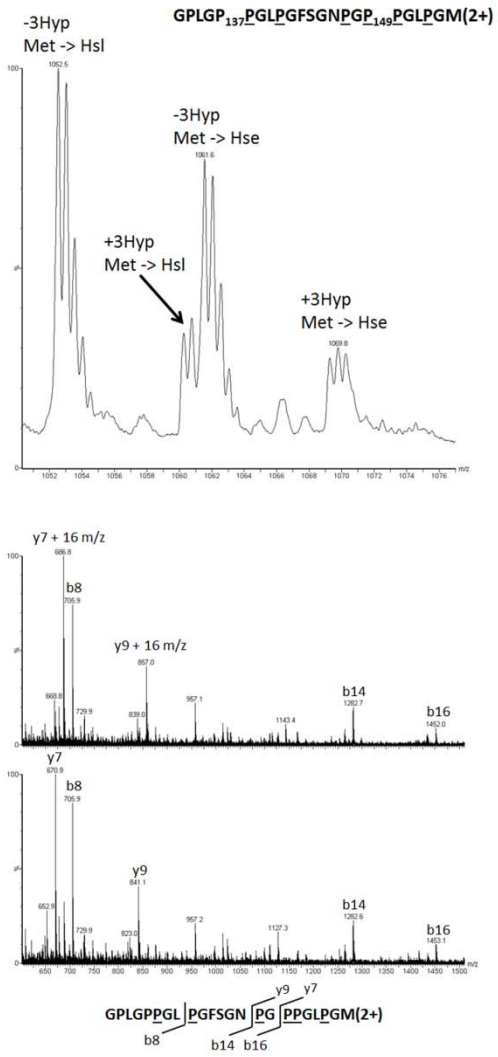
**Supplemental Figure 7.** Comparison of wt (a.) and PFHR-9 P3H2 ko (b.) type IV collagen melting transition. Unfolding curves were modeled using two-transition model with the indicated melting temperatures as input parameters. Data points are shown as red dots, fit as a black line. Absolute CD signal was transformed to fraction of folded collagen IV using a relation found in the methods section of (1).



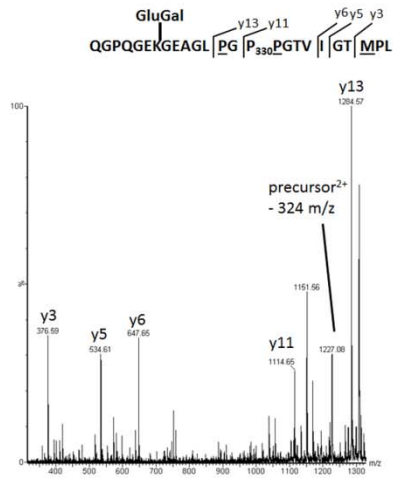
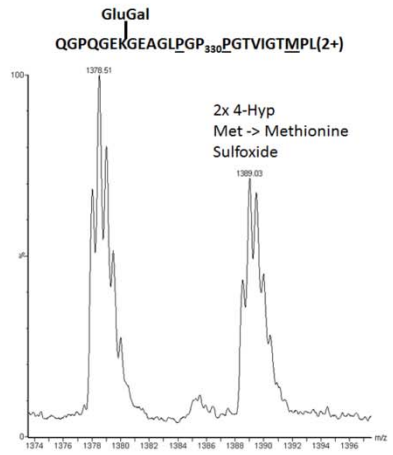
**Supplemental Figure 8.** *Surface Plasmon Resonance Binding of SPARC with Truncation of Inhibitory Domains (BM40- $\Delta I-\alpha C$ ) to type IV collagen from wt or P3H2 ko PFHR-9 cells.* Injection was from 0 to 600 seconds, dissociation was from 600 to 1200 seconds. Lighter curve for each experiment indicates a lower concentration (307 vs 114 nM) Each curve is from an average of 3 injections. Response was normalized for the amount of bound collagen IV.

**Supplemental Figures 9–40.** *Mass Spectra of Diverse type IV collagen Peptides.* See Table 2 for correspondence between Supplemental Figure #, and the relevant site of modification. Top Panels: MS<sup>1</sup> spectra for each peptide. Peaks are labeled to indicate the number of 4-Hyp present, methionine sulfoxide if present, and 3-Hyp together with the residue number containing the putative modification. The sequence of the putative peptide is shown above the spectrum, with residues underlined to indicate oxidation to 4-Hyp, 3-Hyp or methionine sulfoxide, and residues are numbered to indicate modified residues of special interest, including all of the 3-Hyp modifiable residues. Lysine residues modified to galactosylhydroxylysine are also indicated with an -Gal or indicated for glucosylgalactosylhydroxylysine with an adjacent -GluGal. C-terminal methionine residues are present in CNBr digested peptides, where they are modified to homoserine lactone (mass shift = -31 Da) unless otherwise indicated. Bottom Panels: MS<sup>2</sup> spectra for each observed peptide. The predicted b- and y-ions are annotated within the spectrum with their mass and charge state. Multiple related spectra which are differently modified but with the same amino acid sequence are vertically stacked, to easily compare mass differences arising from differential modification. The amino acid sequence is shown above each spectrum, with residue letters separated by lines that indicate the location of b- or y-ion related fragmentation events.

Supplemental Figure 9.

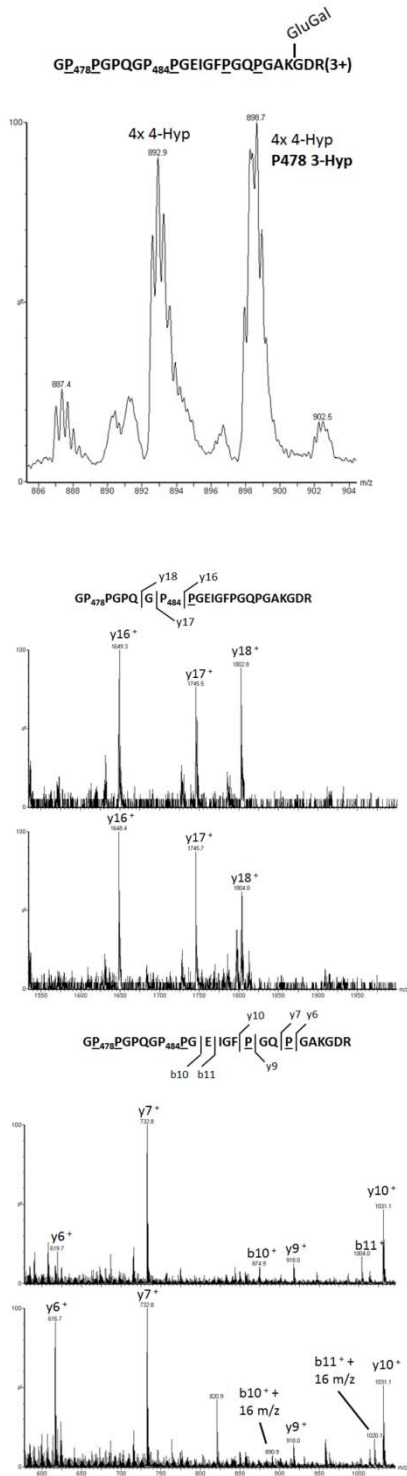


Supplemental Figure 10.

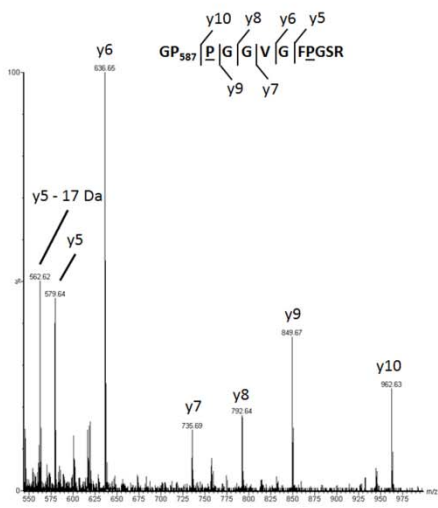
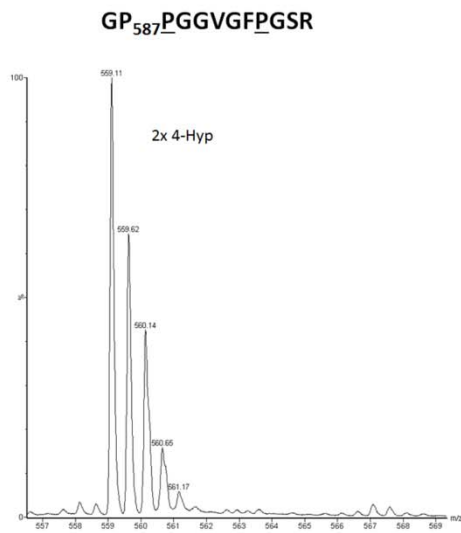




Supplemental Figure 11.

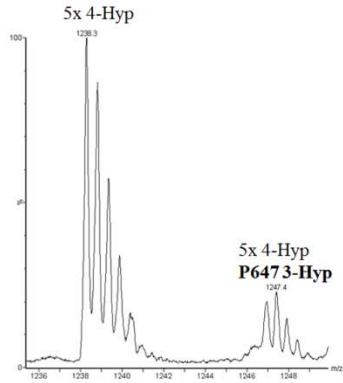


Supplemental Figure 12.

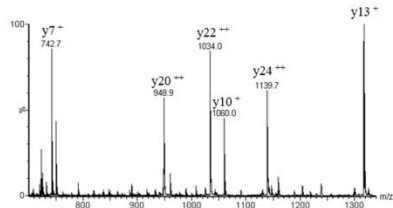
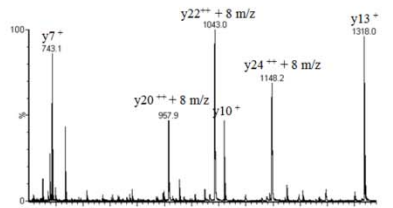


Supplemental Figure 13.

VVPLPGP<sub>647</sub>PGAAGLPGSPGFPGPQGDR(2+)



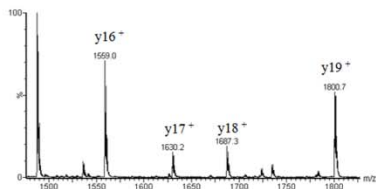
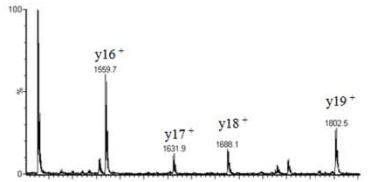
VV|PL|PG|P<sub>647</sub>|PGAAGL|PGS|PGF|PGPQGDR



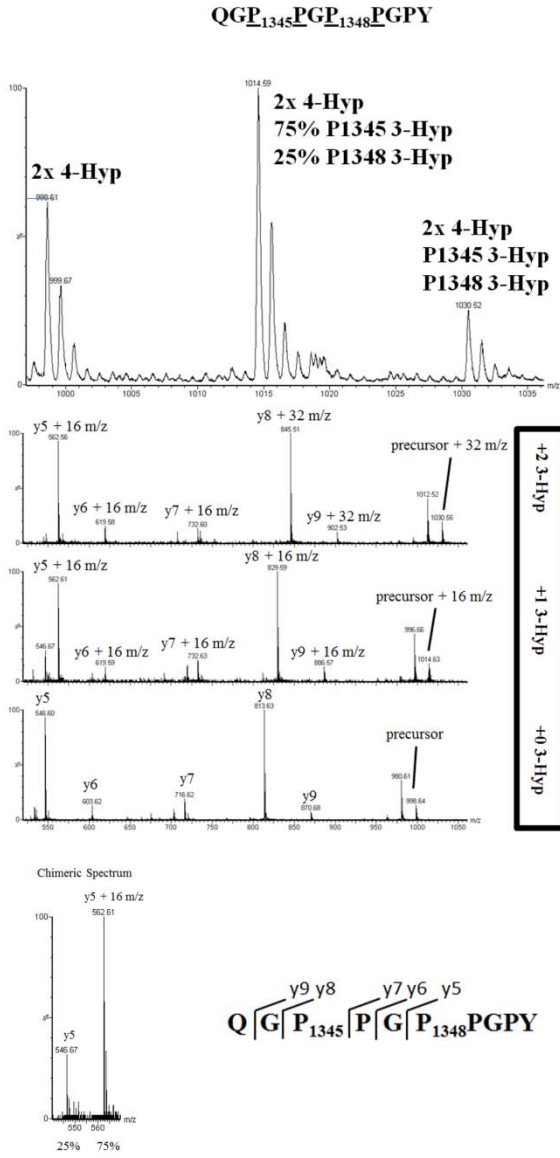
VVPLPGP<sub>647</sub>PG|A|AGLPGSPGFPGPQGDR

y19 y17

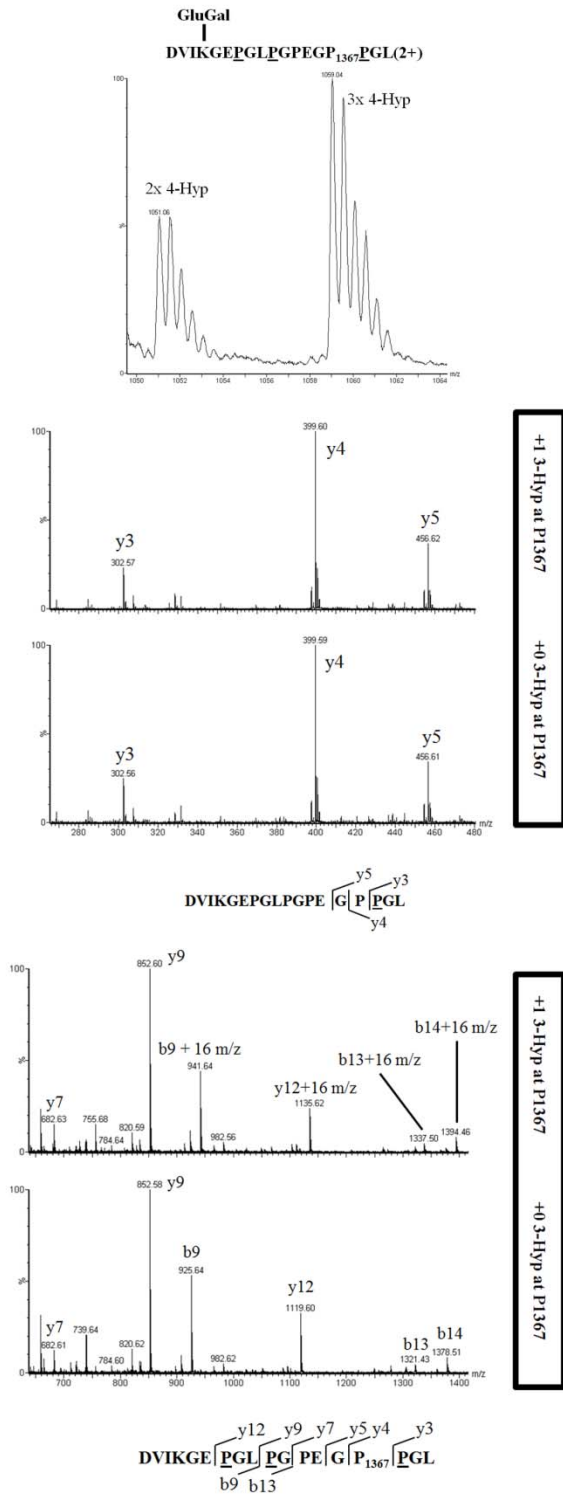
y18 y16



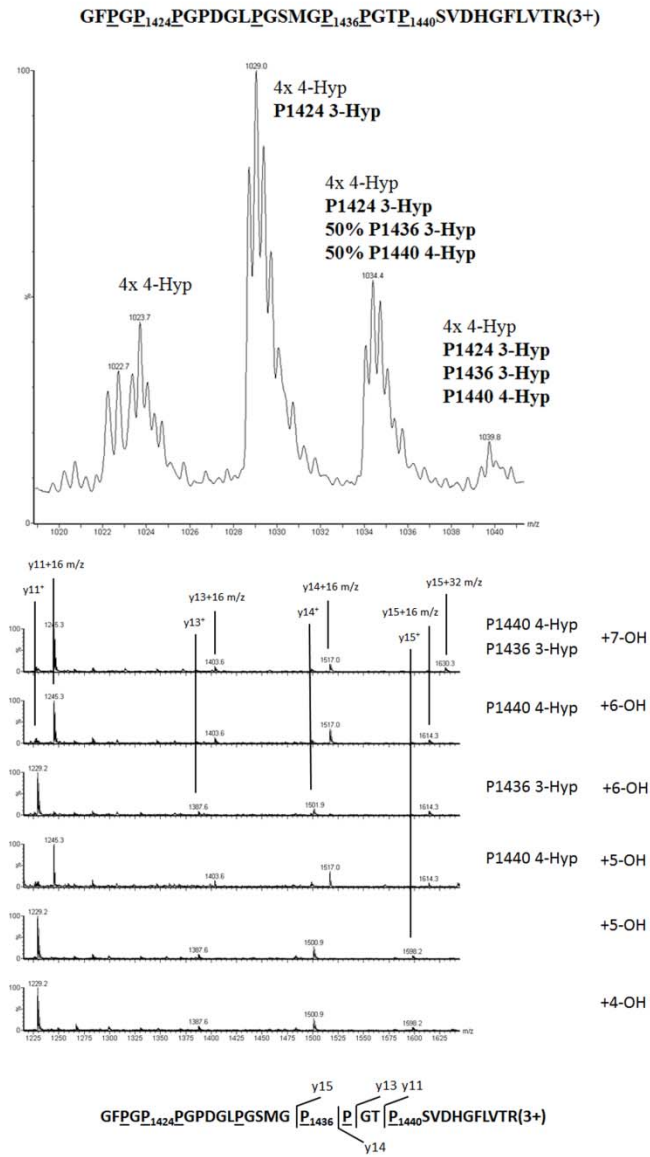
Supplemental Figure 14.



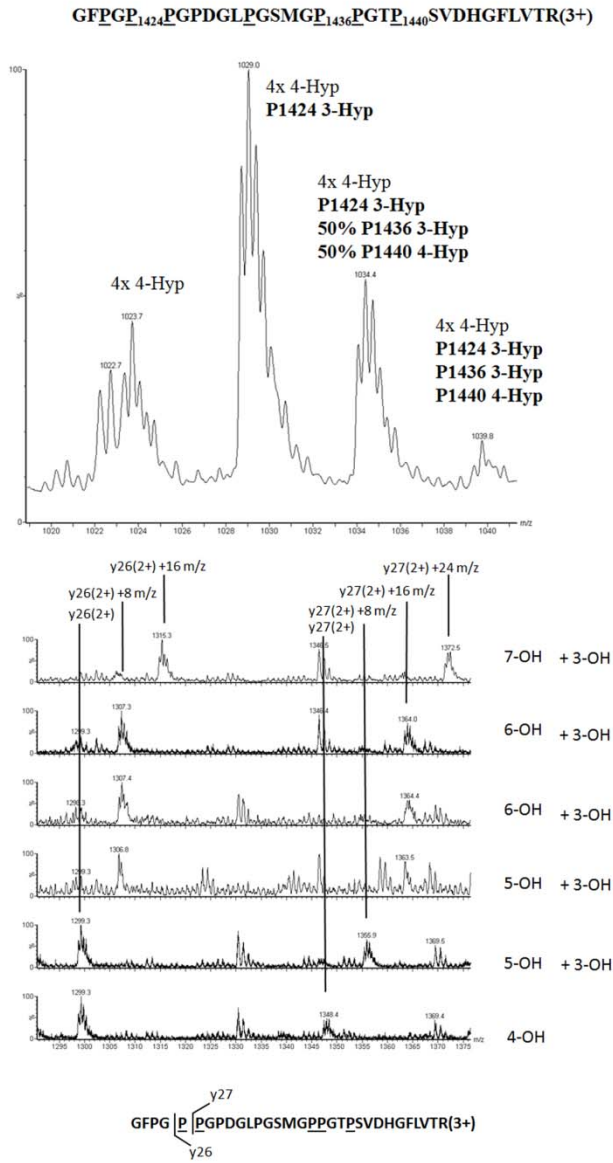
Supplemental Figure 15.



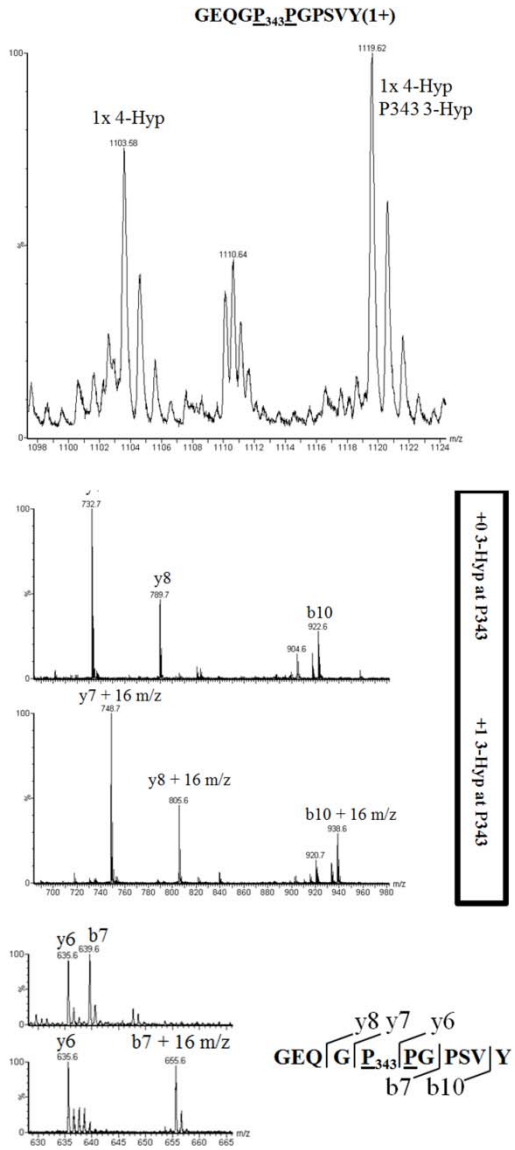
Supplemental Figure 16a.



Supplemental Figure 16b.

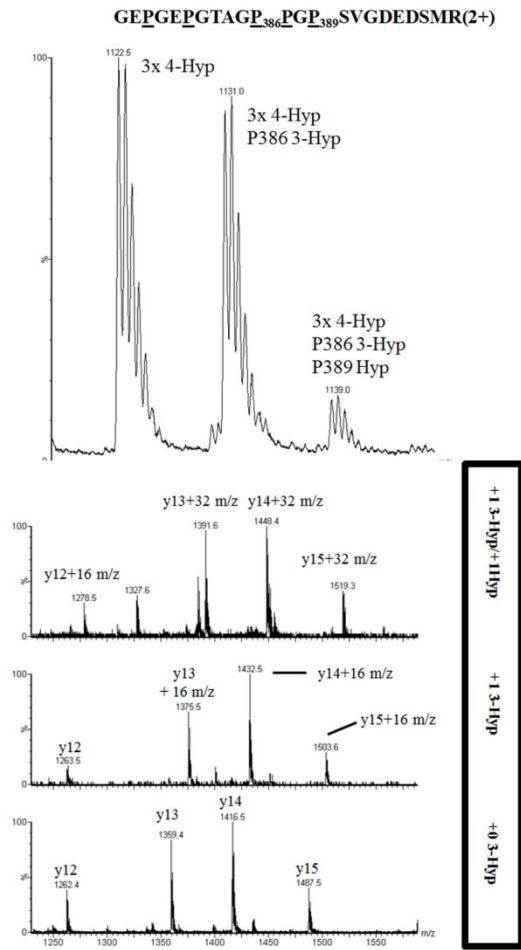


Supplemental Figure 17.

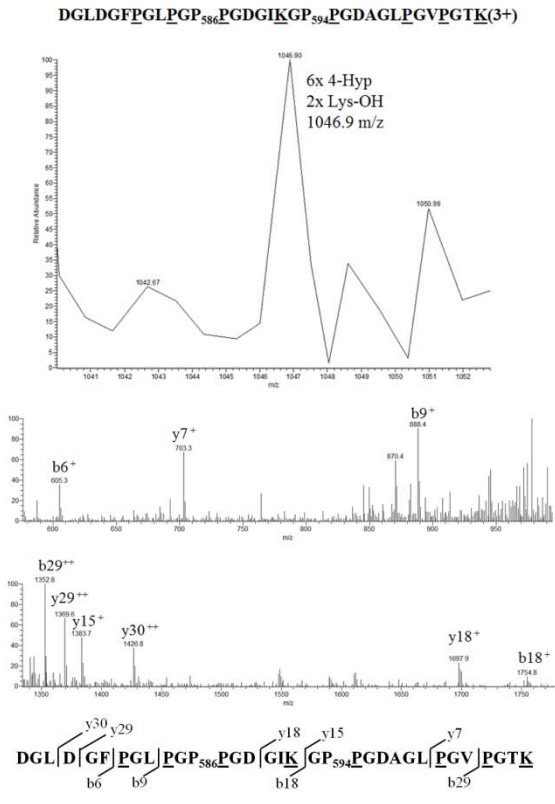




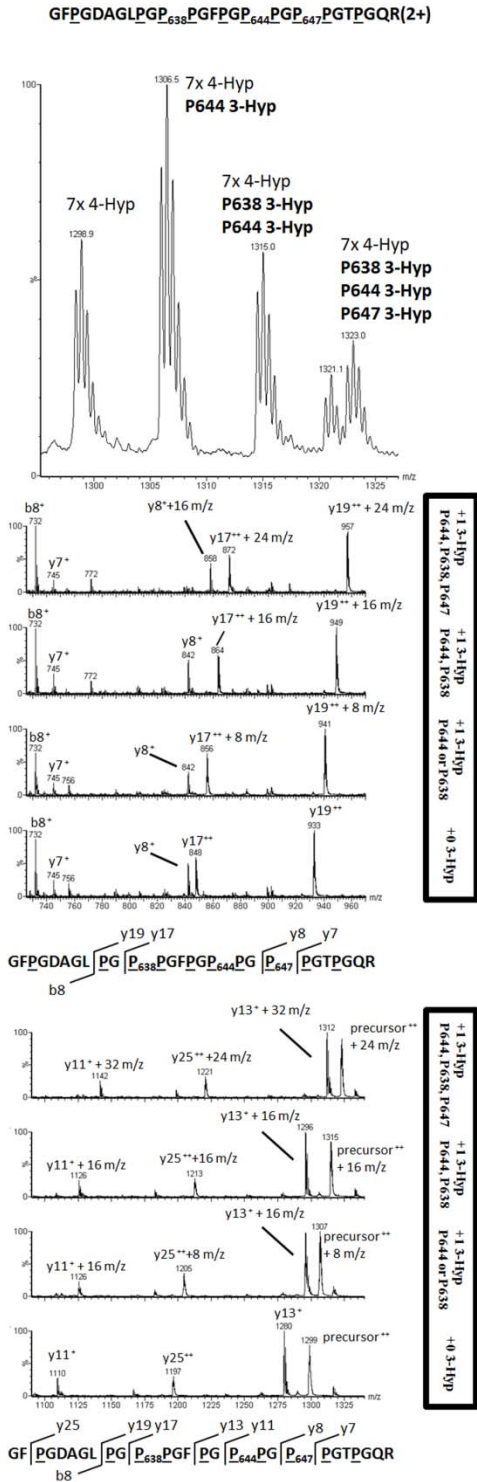
Supplemental Figure 18.



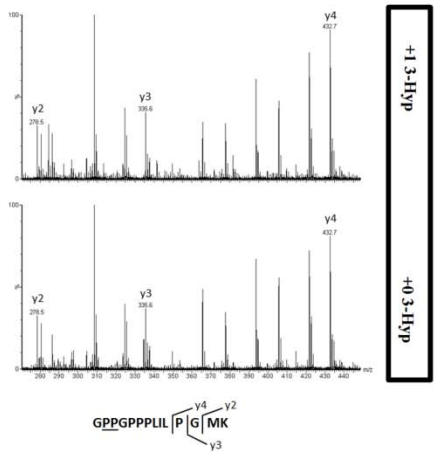
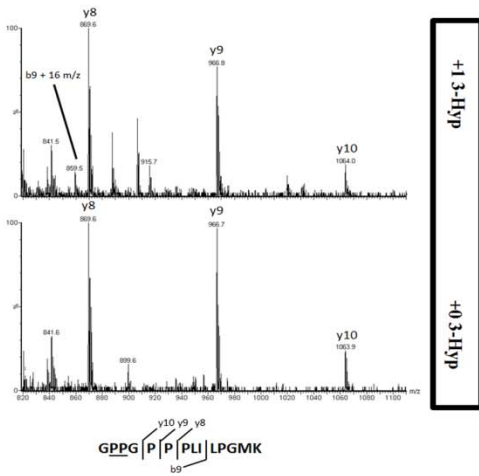
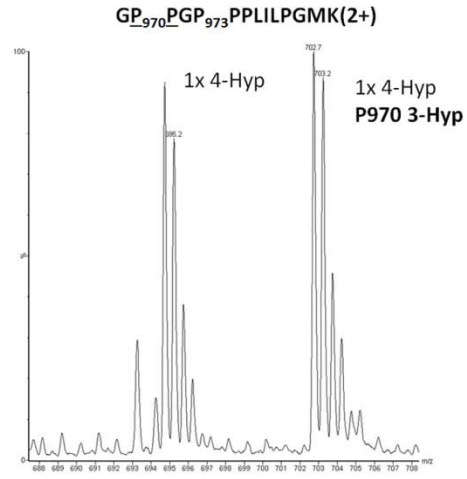
Supplemental Figure 19.



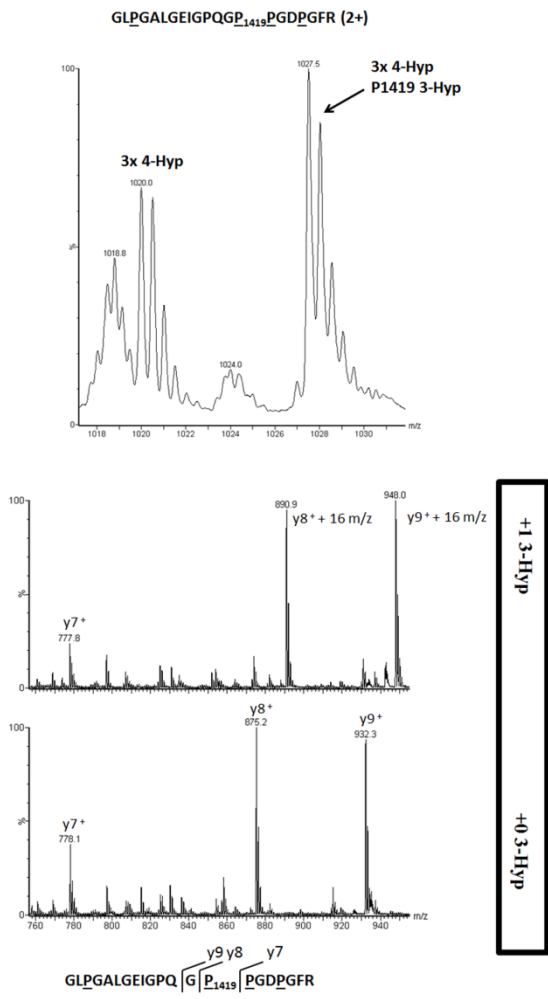
Supplemental Figure 20



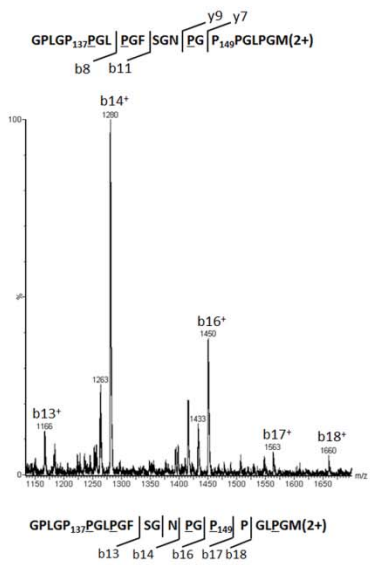
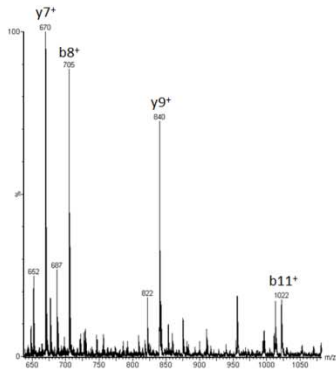
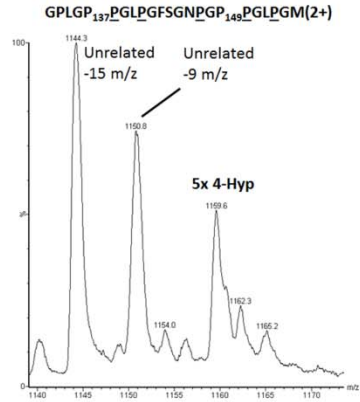
Supplemental Figure 21.



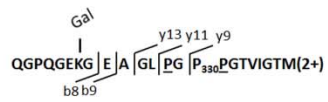
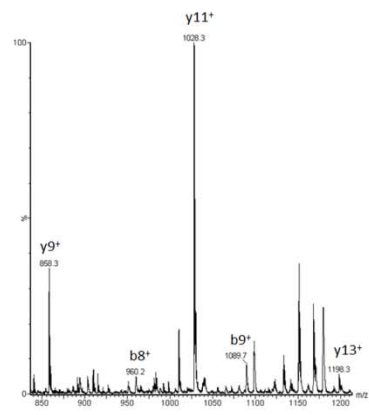
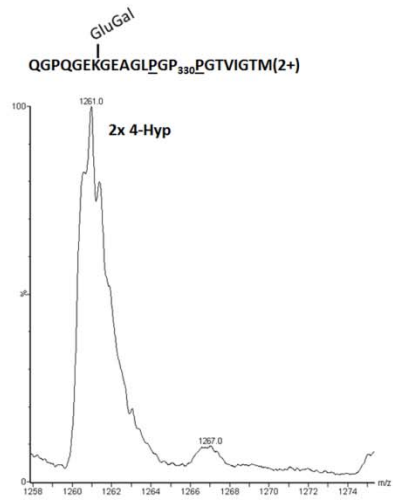
Supplemental Figure 22.



Supplemental Figure 23.

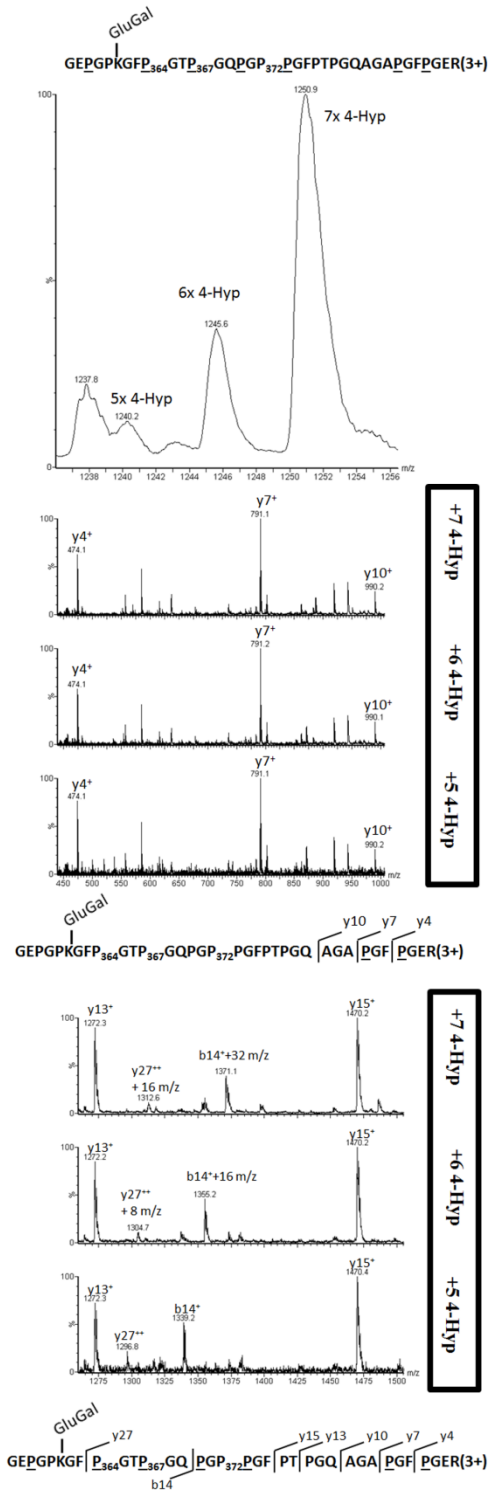


Supplemental Figure 24.



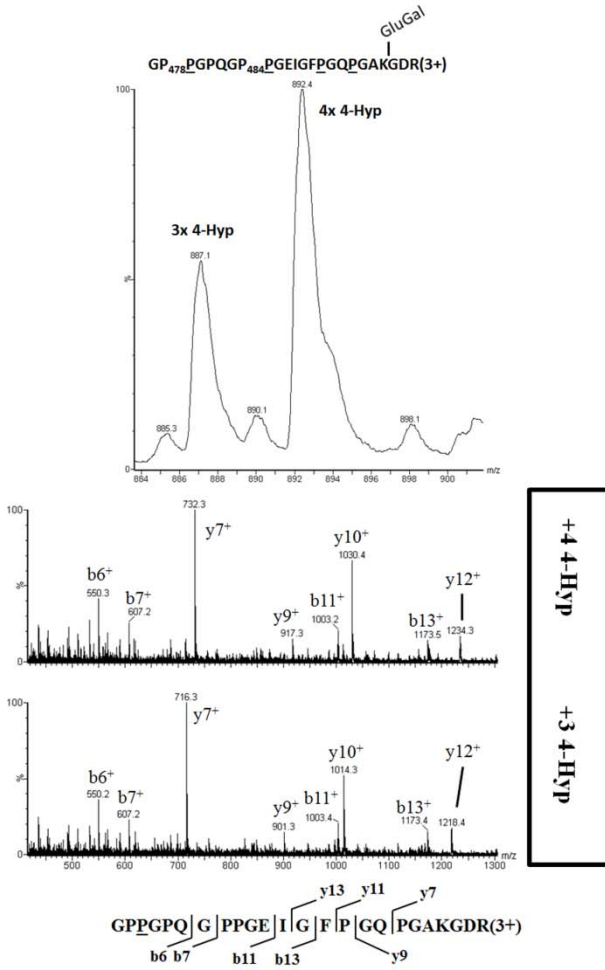
Note: Glu-Gal bond cleaved during MS analysis.  
 B-ion shown above includes K-Gal

Supplemental Figure 25.

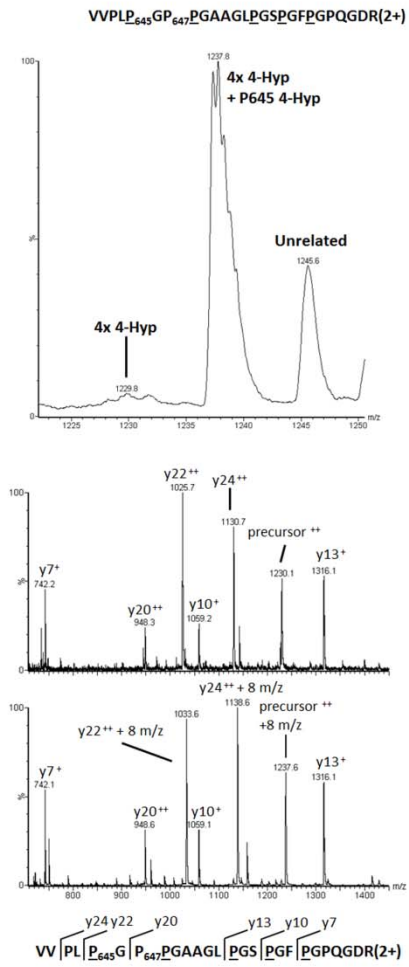




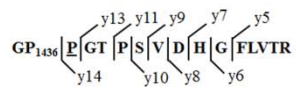
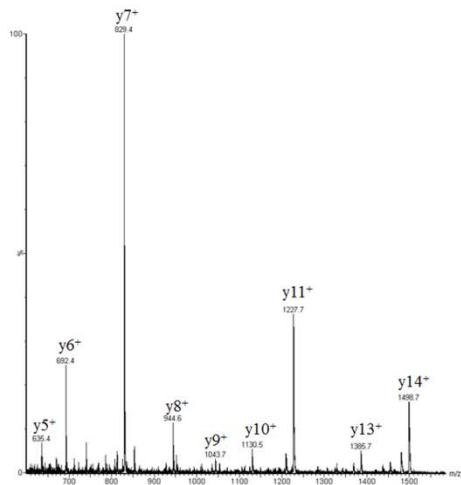
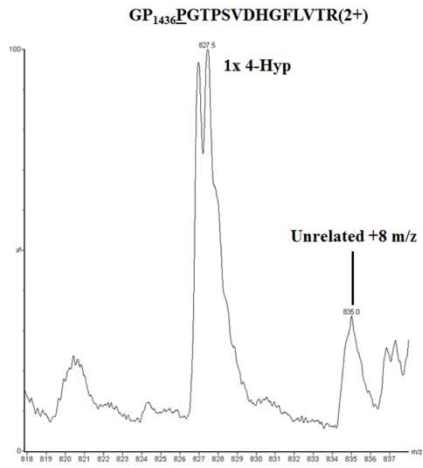
Supplemental Figure 26



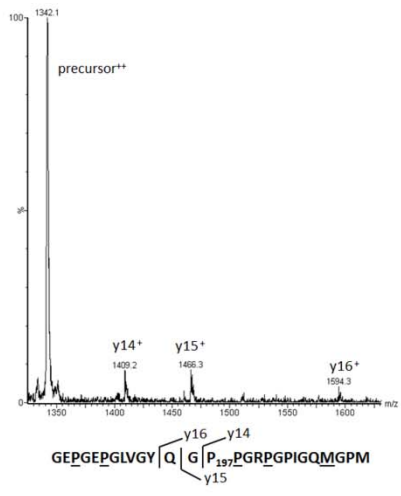
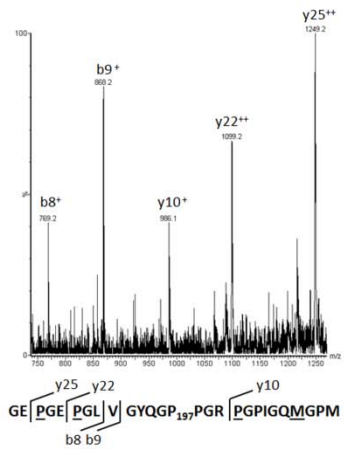
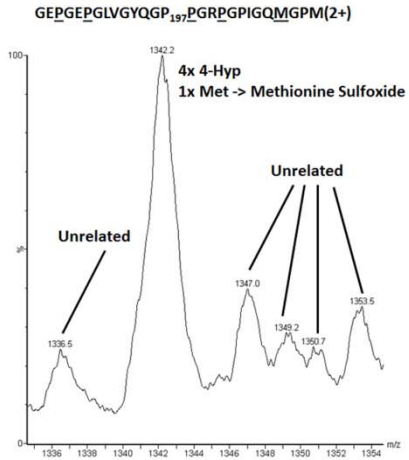
Supplemental Figure 27.



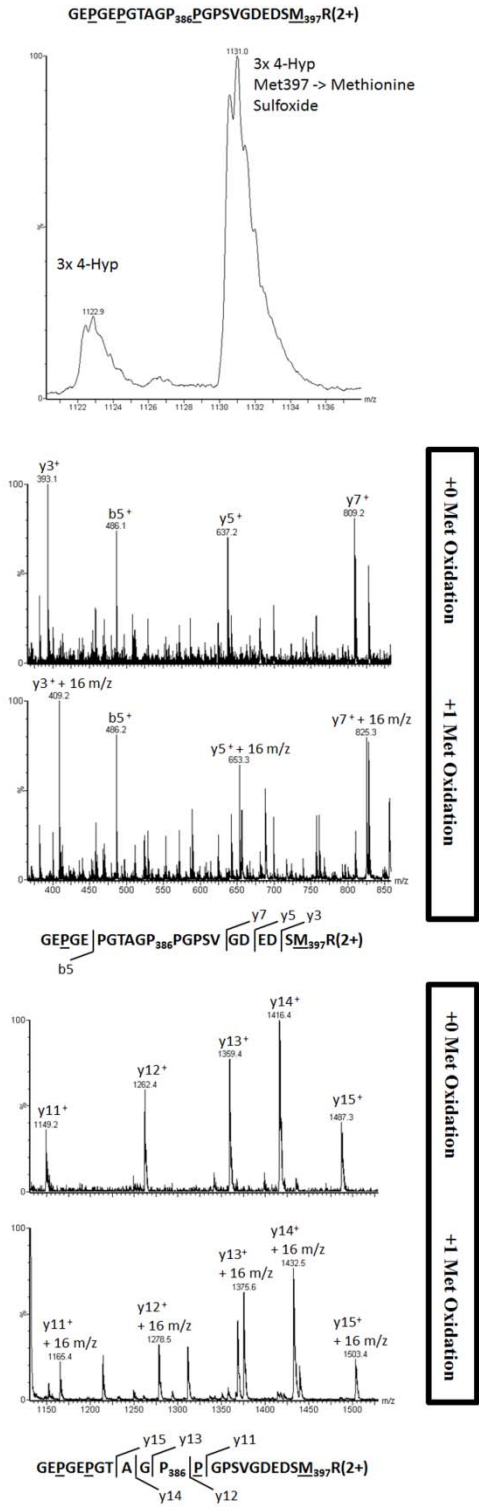
Supplemental Figure 28.



Supplemental Figure 29.



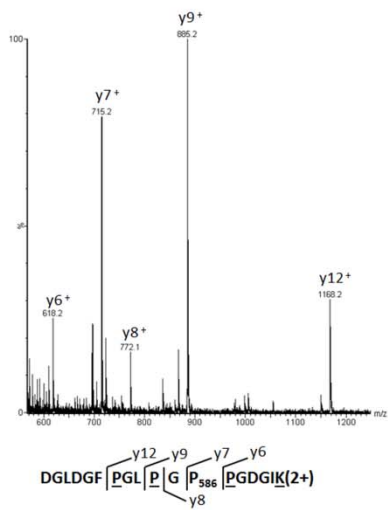
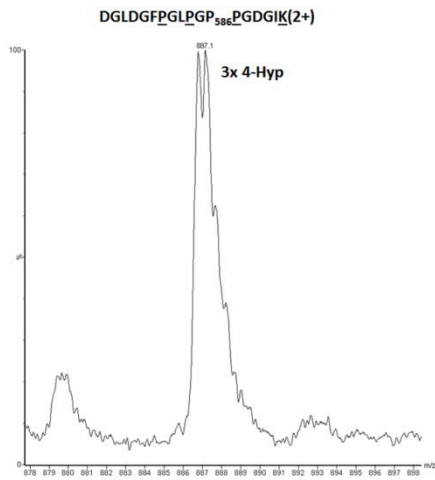
Supplemental Figure 30.



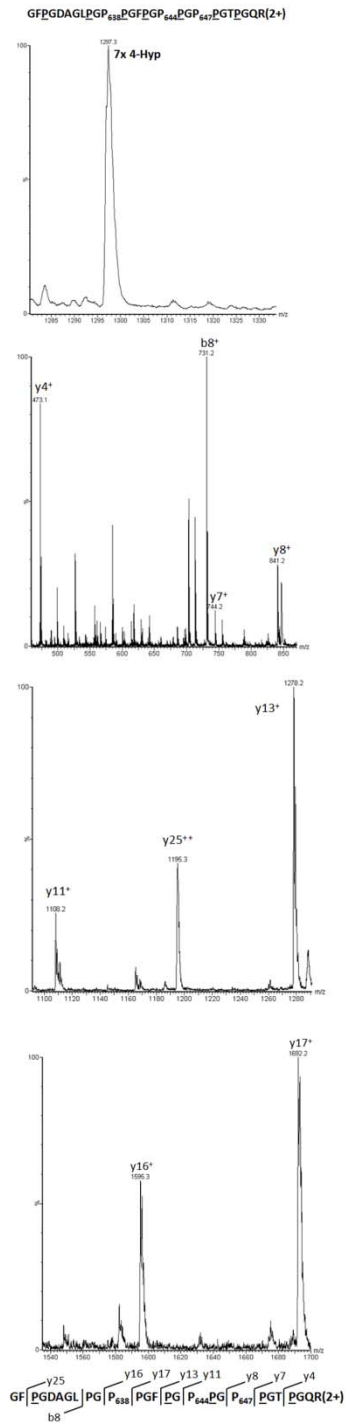
+0 Met Oxidation  
+1 Met Oxidation

+0 Met Oxidation  
+1 Met Oxidation

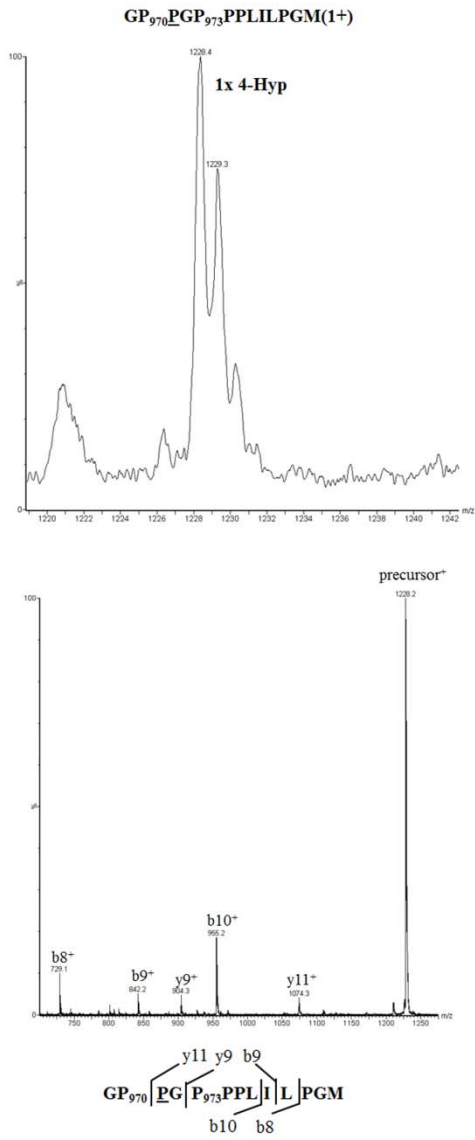
Supplemental Figure 31.



Supplemental Figure 32.

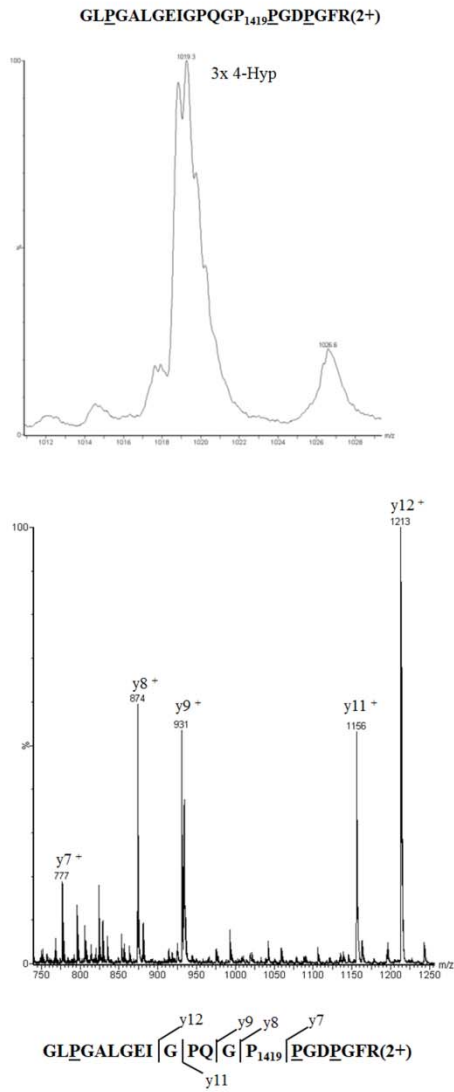


Supplemental Figure 33.

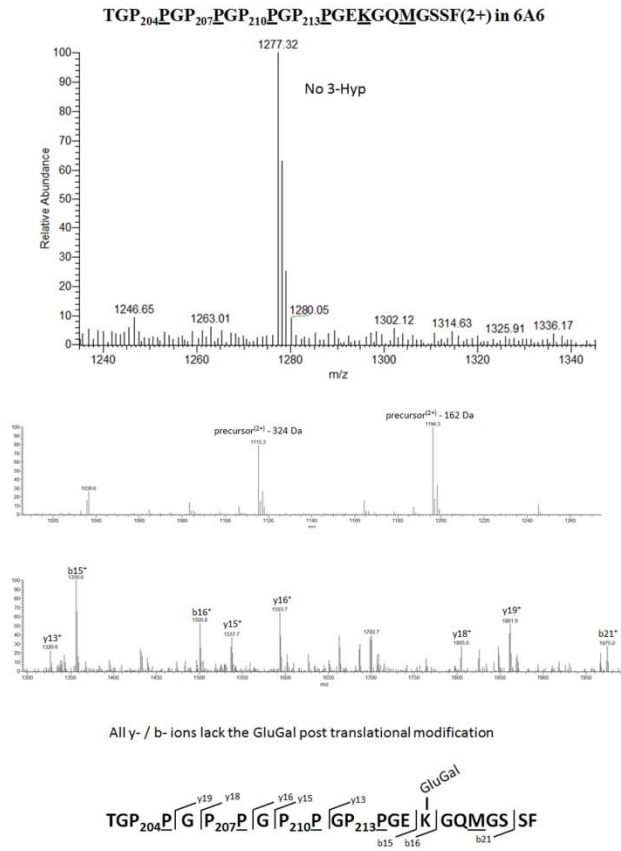




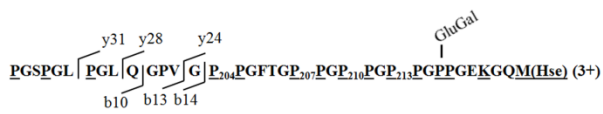
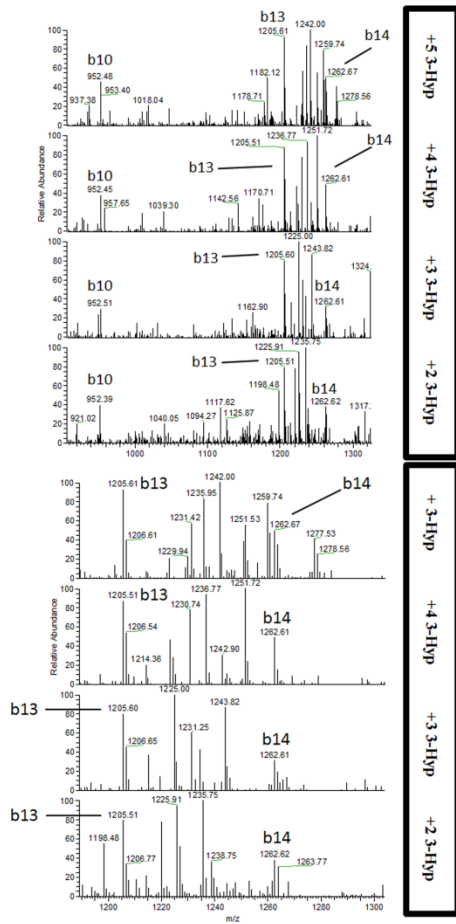
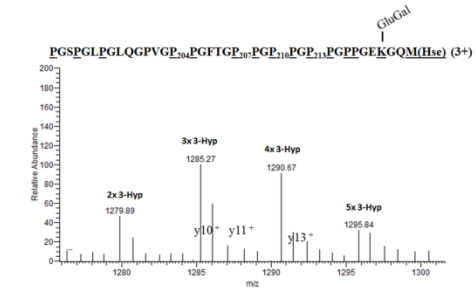
Supplemental Figure 34.



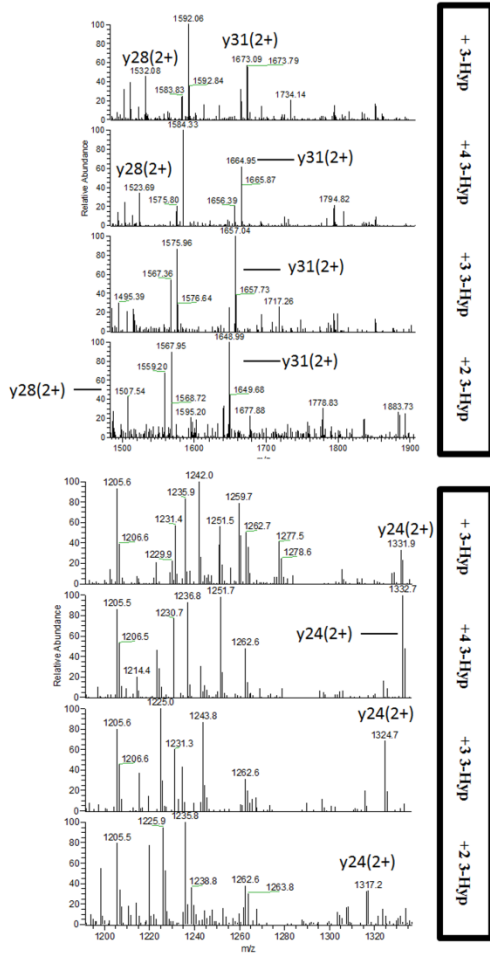
Supplemental Figure 35.



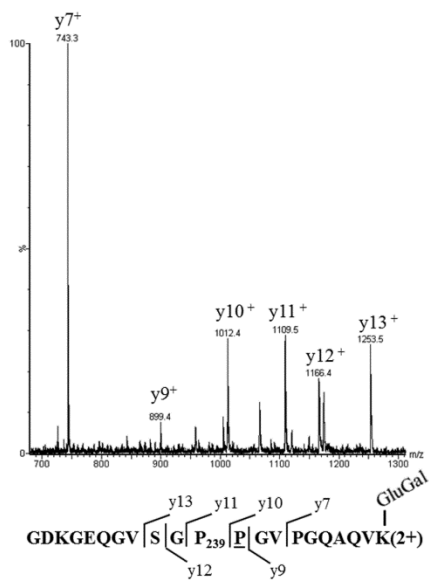
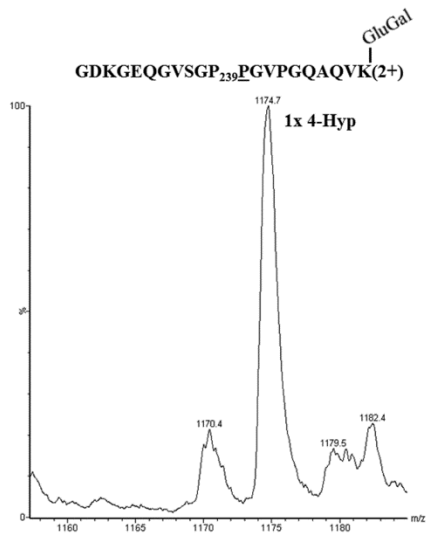
Supplemental Figure 36a



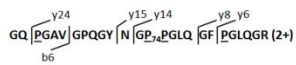
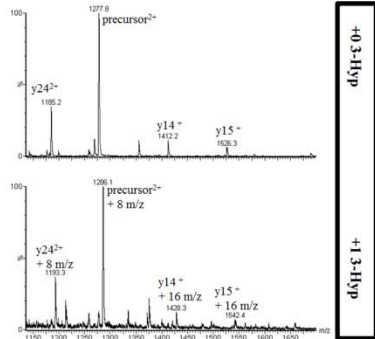
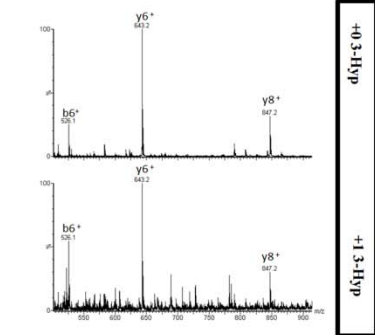
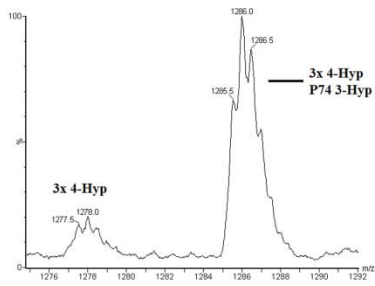
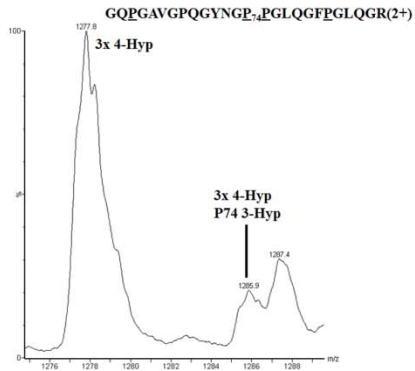
Supplemental Figure 36b



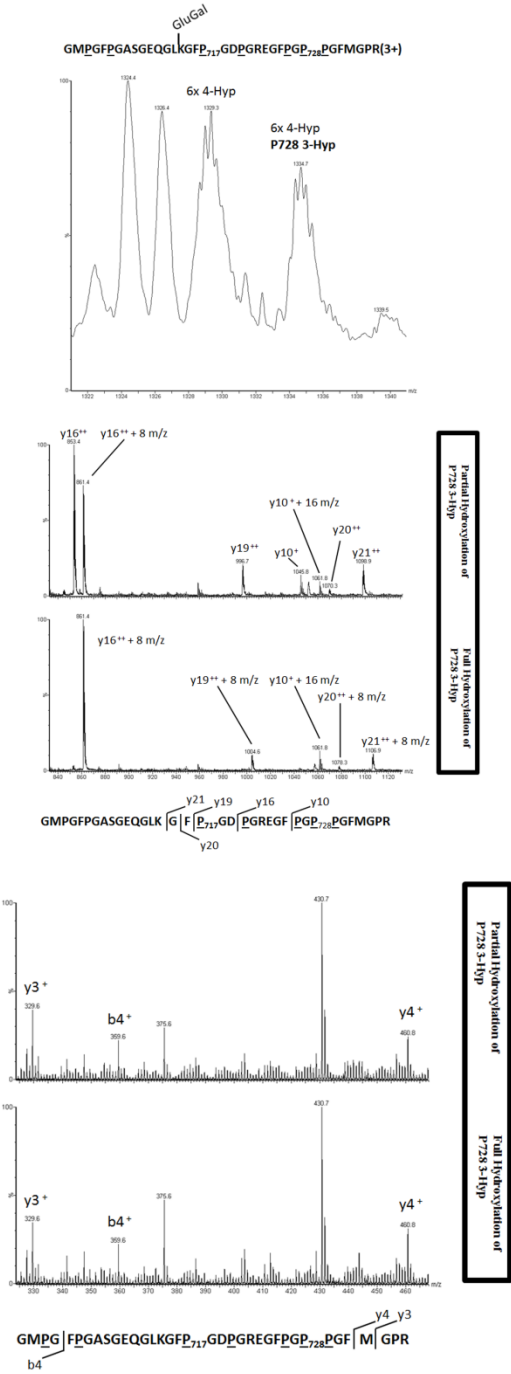
Supplemental Figure 37



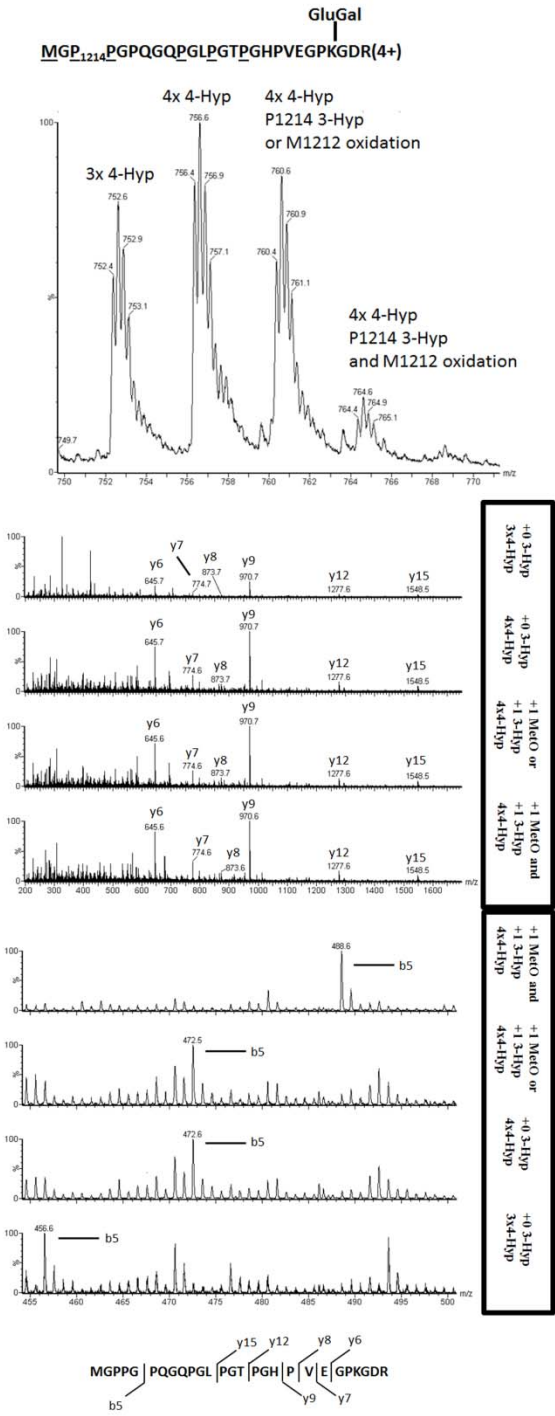
Supplemental Figure 38



Supplemental Figure 39



Supplemental Figure 40





## References

1. Persikov, A.V., Y. Xu, and B. Brodsky, (2004) *Equilibrium thermal transitions of collagen model peptides*. Protein Sci. **13**(4): p. 893-902.