Supporting Information

Expanding the dipeptidyl peptidase 4-regulated peptidome via an optimized peptidomics platform

Arthur D. Tinoco, Debarati M. Tagore, and Alan Saghatelian*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

saghatelian@chemistry.harvard.edu

Table S1. The charge state determines the fraction in which peptides elute from from a PolySULFOETHYL A SCX column at pH 2.6.

Protein (peptide region)	Peptide Sequence	Charge	Salt Fraction (mM KCl)
EF-1α(281-291)	APVNVTTEVKS	2	40
CtsB(74-86)	LPETFDAREQWSN	2	40
Sorbitol dehydrogenase(25-40)	YPIPELGPNDVLLKMH	3	100
Mepβ(21-41)	LPAPEKFVKDIDGGIDQDIFD	3	400
Atp6v1g1(107-118)	RPEIHENYRING	4	400
DBI(92-105)	RPGLLDLKGKAKWD	5	400

Table S2. Absolute fold changes of peptides identified in the salt free SCX fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Clathrin heavy chain(1662-1672)	**PP**YGQPQPGFG	3.88**
Legumain(74-88)	DDIANSEENPTPGVV	5.13**
Low density lipoprotein receptor-related protein 2(4552-4559)	PENVENQN	3.02**
Phosphatidylinositol-binding clathrin assembly protein(645-658)	**PPNP**FGPVSGAQIQ	4.77**
Putative uncharacterized protein OS=Mus musculus GN=Dab2 PE=2 SV=1(270-279)	**APAP**VGPLVG	2.86*
Putative uncharacterized protein OS=Mus musculus GN=Pcdh24 PE=2 SV=1(1300-1308)	**NP**GLDTTDL	4.36**

*, $p<0.05$

**, $p<0.01$

Table S3. Absolute fold changes of peptides identified in the SCX-40 mM KCl fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold. The precursor amino acid for the *DPP4*$^{+/+}$ elevated peptides is shown in bold and parentheses. (M) = oxidized methionine.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
60 S ribosomal protein L6(272-280)	**VP**QLQGYLR	4.11**
ATP synthase a(88-101)	**LP**HTFTPTTQLS(M)N	2.90**
ATP synthase b, mitochondrial(68-80)	**VP**KTGVTGPYVLG	1.97**
ATP synthase b, mitochondrial(68-84)	**VP**KTGVTGPYVLGTGLS	2.13**
ATP synthase e, mitochondrial(3-13)	**PP**VQVSPLIKF	2.91**
ATP synthase coupling factor 6, mitochondrial (90-96)	**FP**TFKFD	3.43*
Brain protein 44(90-96)	**IP**KNWSL	2.59*
Cathepsin B(80-88)	**LP**ETFDARE	4.22**
Cathepsin B(80-89)	**LP**ETFDAREQ	3.02**
Cathepsin B(80-90)	**LP**ETFDAREQW	2.09**
Cathepsin B(80-91)	**LP**ETFDAREQWS	3.05**
Cathepsin S(239-248)	**LP**FGDEDALK	3.28**
Cordon-bleu protein-like 1(1071-1086)	**LP**AETSLPLVFPKPMT	3.45**
Cytochrome b-c1 complex subunit Rieske, mitochondrial(34-47)	**VP**AASEPPVLDVKR	2.51**
Cytochrome c oxidase polypeptide 5A, mitochondrial(79-87)	**VPEP**KIIDA	2.00**
Cytochrome c oxidase polypeptide 6A1, mitochondrial(84-91)	**KPFP**WGDG	6.25**
Cytochrome c oxidase subunit 4 isoform 1(44-56)	**YPLP**VAHVTMLS	3.31**
Cytochrome c oxidase subunit 6B isoform 1(15-26)	**AP**FDSRFPNQNQ	5.42**
Cytoplasmic dynein 1 light intermediate chain 1(505-523)	**KP**ASVSPTTPTSPTEGEAS	2.68**
Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial(186-201)	**PPVPSP**SQPPSSKPVS	4.11**

Disabled homolog 2(625-634)	**GP**LKDIPSDA	2.65**
Glyceraldehyde-3-phosphate-dehydrogenase (219-227)	**IP**ELNGKLT	2.45*
Growth arrest-specific protein 2(219-227)	**AP**SGSFFARDNTANF	3.56**
Heme-binding protein 1(66-84)	VGGTNDKGVGMGMTVPVSF	5.18**
Histidine triad nucleotide-binding protein(27-34)	**IP**AKIIFE	17.96**
Kinesin light chain 4(600-619)	**AP**LQVSRGLSASTVDLSSSS	2.12**
Low-density lipoprotein receptor-related protein(4637-4652)	**TP**GYTATEDTFKDTAN	3.39*
Lysosomal protective protein(346-354)	**IP**ESLPRWD	17.63**
Major Urinary Protein 6(110-117)	**IP**KTDYDN	26.06**
Major Urinary Protein 6(110-118)	**IP**KTDYDNF	28.67**
Meprin A subunit beta(21-28)	**LPAP**EKFV	40.51**
Microtubule-associated protein tau(195-204)	**APVPMP**DLKN	2.40**
MKIAA0248 protein (fragment)(1790-1803)	**TPRP**TDPIPTSEVN	4.37**
Peptide methionine sulfide reductase(187-197)	**GP**ITTDIREGQ	2.76**
Peroxiredoxin-5, mitochondrial(50-62)	**AP**IKVGDAIPSVE	3.69**
Phosphatidylethanolamine-binding protein 1 (177-184)	**VP**KLYEQL	4.92**
Protein kinase C and casein kinase substrate in neurons protein 2(340-358)	**KP**GSNLSVPSNPAQSTQLQ	3.34**
Protein NDRG1(122-129)	**LP**GVLHQF	3.48**
Protein NDRG1(375-393)	**TP**NSGATGNNAGPKSMEVS	3.42**
Putative uncharacterized protein OS=Mus musculus GN=Ubqln1 PE=2 SV=1(495-517)	**AP**STAPSEDTNPQGGTAEPGHQQ	3.19*
Serine/Arginine repetitive matrix protein 2 (2275-2289)	**YP**SSSRTPQAPTPAN	3.04**
Serum Albumin(247-258)	**FP**NADFAEITKL	4.05**
Sorbitol dehydrogenase(25-38)	**YPIP**ELGPNDVLLK	7.49**
Sorbitol dehydrogenase(25-40)	**YPIP**ELGPNDVLLKMH	12.48**
Sulfotransferase 1 C2(244-256)	**AP**KSILDQSISPF	3.60**
Triosephosphate isomerase(238-247)	**KP**EFVDIINA	8.44**
Tripeptidyl-peptidase 1(497-505)	**PP**LGFLNPR	10.41**

Tripeptidyl-peptidase 1(497-506)	**PP**LGFLNPRL	3.19**
Tripeptidyl-peptidase 1(497-507)	**PP**LGFLNPRLY	4.81**
WT Elevated Peptides		$DPP4^{+/+}$/ $DPP4^{-/-}$
Low-density lipoprotein receptor-related protein(4639-4652)	**(P)**GYTATEDTFKDTAN	3.88**
Putative uncharacterized protein OS=Mus musculus GN=Ubqln1 PE=2 SV=1(497-517)	**(P)**STAPSEDTNPQGGTAEPGHQQ	3.19**
Serum Albumin(249-258)	**(P)**NADFAEITKL	4.05**

*, p<0.05

**, p<0.01

Table S4. Absolute fold changes of peptides identified in the SCX-100 mM KCl fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold. The precursor amino acid for the *DPP4*$^{+/+}$ elevated peptides is shown in bold and parentheses.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
40 S ribosomal protein s2(264-275)	**SP**YQEFTDHLVK	13.9*
ATP synthase subunit b,mitochondrial(45-57)	**PPLP**EYGGKVRLG	5.91**
cAMP-regulated phosphoprotein 19(81-97)	**AP**DKTEVTGDHIPTPQD	5.26**
Catalase(23-40)	**RP**DVLTTGGGNPIGDKLN	35.6*
Cordon-bleu protein-like 1(404-420)	**AP**APPSKTPLAQTDERN	5.22**
Cytochrome c oxidase polypeptide 6A1, mitochondrial(84-94)	**KPFP**WGDGNHT	4.14**
Cytochrome c oxidase polypeptide 6A1, mitochondrial(84-95)	**KPFP**WGDGNHTL	5.73*
Legumain(158-167)	**FP**NDDLHVKD	3.59**
Lysosomal alpha-glucosidase(810-820)	**AP**LDTINVHLR	3.34*
Peroxiredoxin-5, mitochondrial(58-73)	**IP**SVEVFEGEPGKKVN	2.40**
Polypyrimidine tract-binding protein 1(5-18)	**VP**DIAVGTKRGSDE	4.98*
Profilin-1(44-58)	**TP**AEVGVLVGKDRSS	5.20*
Triosephosphate isomerase(238-249)	**KP**EFVDIINAKQ	3.43*
Triosephosphate isomerase(43-55)	**PP**TAYIDFARQK	3.94**
WT Elevated Peptides		DPP4$^{+/+}$/DPP4$^{-/-}$
cAMP-regulated phosphoprotein 19(83-97)	**(P)**DKTEVTGDHIPTPQD	5.26**

*, p<0.05

**, p<0.01

Table S5. Absolute fold changes of peptides identified in the SCX-300 mM KCl free fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequence is highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Cytochrome c oxidase subunit 5A, mitochondrial(58-76)	**KP**DIDAWELRKGMNTLVGY	2.79**

*, p<0.05

**, p<0.01

Table S6. Selected peptide fractionation by off-gel electrophoresis (OGE). The peptides are from the *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney peptidome. Fraction A = Lanes 1,2; B = Lanes 3,4; C = Lanes 5,6; D = Lanes 7-12.

Protein (peptide region)	Peptide Sequence	PI	Fraction
Mepβ(21-41)	LPAPEKFVKDIDGGIDQDIFD	3.6	A
Slc9a3r1(275-296)	SPRPALARSASSDTSEELNSQD	4.1	A
EF-1α(281-291)	APVNVTTEVKS	6.9	C, D
Slc22a12 (3-20)	FPELLDRVGGLGRFQ	7	C
Vimentin(440-460)	RTLLIKTVETRDGQVINETSQ	7.1	C
Atp6v1g1(107-118)	RPEIHENYRING	7.8	D

Table S7. Absolute fold changes of peptides identified in the OGE-Lane 1-2 fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Leucine-rich repeat-containing protein 4B(239-248)	**LP**FGDEDALK	2.92**
Low-density lipoprotein receptor-related protein 2 (4637-4652)	**TP**GYTATEDTFKDTAN	3.19*
Major urinary protein 2(110-118)	**IP**KTDYDNF	5.29**
MKIAA0248 protein(1790-1803)	**TPRP**TDPIPTSEVN	2.98*
Peptide Methionine Sulfoxide(238-249)	**KP**EFVDIINAKQ	2.34**

*, $p<0.05$

**, $p<0.01$

Table S8. Absolute fold changes of peptides identified in the OGE-Lane 3-4 fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold. The precursor amino acid for the *DPP4*$^{+/+}$ elevated peptides is shown in bold and parentheses.

Protein (peptide region)	Peptide Sequence	Fold Change
***DPP4*$^{-/-}$ Elevated Peptides**		DPP4$^{-/-}$/ DPP4$^{+/+}$
Peroxiredoxin-5, mitochondrial(58-73)	**IP**SVEVFEGEPGKKVN	5.87**
***DPP4*$^{+/+}$ Elevated Peptides**		DPP4$^{+/+}$/DPP4$^{-/-}$
Putative uncharacterized protein OS=Mus musculus GN=Dab2 PE=2 SV=1(301-323)	**(P)**STTVVPGAIISGQPPSFGQPLVF	7.15**

*, p<0.05

**, p<0.01

Table S9. Absolute fold changes of peptides identified in the OGE-Lane 5-6 fraction of the global peptide profiling experiments (N = 4) with WT and DPP4$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
3-hydroxyanthranilate-3,4-dioxygenase(158-173)	**KPNP**DQLLKELPFPLN	4.13**
Alpha-globin transcription factor CP2(108-115)	**LP**ELNGKL	5.49**
ATP synthase-coupling factor 6, mitochondrial (90-96)	**FP**TFKFD	7.74*
Cytochrome b-c1 complex subunit Rieske, mitochondrial(34-50)	**VP**AASEPPVLDVKRPFL	3.24**
Cytochrome c oxidase polypeptide 6A1(49-57)	**LP**GVGVSML	2.01**
Fructose-bisphosphate aldolase B(5-13)	**FP**ALTPEQK	26.5**
Microtubule-associated protein tau(538-546)	**APVPMP**DLK	6.69**
Non-specific lipid-transfer protein(466-480)	**GP**GGKEATWVVDVKN	3.40**
PDZK1-interacting protein 1(90-113)	FRSSEHKNAYENVLEEEGRVRSTP	7.92**
Putative uncharacterized protein OS=Mus musculus GN=Prodh PE=2 SV=1(22-38)	**KP**QAQEQPPASPEALRG	4.02*
Solute carrier family 22 member 12(3-17)	**FP**ELLDRVGGLGRFQ	4.74**

*, $p<0.05$

**, $p<0.01$

Table S10. Absolute fold changes of peptides identified in the OGE-Lane 7-12 fraction of the global peptide profiling experiments (N = 4) with WT and DPP4$^{-/-}$ kidney samples. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Collectrin(18-29)	**HP**DAENAFKVRL	2.89*
Cytochrome c oxidase polypeptide 6A1(84-95)	**KPFP**WGDGNHTL	3.65**
Protein kinase c and kinase substrate in neurons protein 2(340-358)	**KP**GSNLSVPSNPAQSTQLQ	10.9**

*, $p<0.05$

**, $p<0.01$

Table S11. *DPP4*$^{-/-}$ elevated kidney peptides detected by application of either SCX or OGE electrophoresis fractionation during 2D fractionation, which were also detected by standard 1D RP-HPLC MS methods. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

SCX Fractionation	**OGE Electrophoresis Fractionation**
Atp6v1g1(107-118) **RP**EIHENYRING	Atp6v1g1(107-118) **RP**EIHENYRING
CtsB(74-86) **LP**ETFDAREQWSN	Mepβ(21-41) **LPAP**EKFVKDIDGGIDQDIFD
EF-1α(281-291) **AP**VNVTTEVKS	Slc9a3r1(275-296) **SPRP**ALARSASSDTSEELNSQD
Mepβ(21-28) **LPAP**EKFV*	Slc22a12(3-20) **FP**ELLDRVGGLGRFQ

* = Partial Match

Table S12. Absolute fold changes of peptides identified in the salt free SCX fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ gut samples applying SCX-RP-peptidomics workflow. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Chymotrypsinogen B(21-28)	**VP**AIQPVL	5.86**
Chymotrypsinogen B(21-28)	**VP**AIQPVLTG	9.83*
VIP Peptides(26-33)	**WP**LFGPPS	1.91*

*, $p<0.05$

**, $p<0.01$

Table S13. Absolute fold changes of peptides identified in the SCX-40 mM KCl free fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ mice guts applying SCX-RP-peptidomics workflow. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
40 S ribosomal protein S20(9-19)	**TP**VEPEVAIHR	4.93*
Actin, cytoplasmic 2(31-38)	**FP**SIVGRP	5.68*
Actin, cytoplasmic 2(242-254)	**LP**DGQVITIGNER	6.30*
Alpha-actinin-4(885-901)	**AP**YQGPDAAPGALDYKS	4.94*
Alpha-actinin-4(885-902)	**AP**YQGPDAAPGALDYKSF	6.03*
Cathepsin B(80-89)	**LP**ETFDAREQ	1.64**
Ezrin-radixin-moesin-binding phosphoprotein 50 (240-254)	**IP**SQEHLDGPLPEPF	6.39*
Glyceraldehyde-3-phosphate dehydrogenase (219-228)	**IP**ELNGKTG	14.95*
Inorganic pyrophosphatase(276-283)	**LP**TDVDKW	4.32*
Junction plakoglobin(709-717)	**VP**LDPLDMH	11.13**
Profilin-1(44-54)	**TP**AEVGVLVGK	6.45*
Putative uncharacterized protein GN=Pdlim1, PE=1, SV=1(130-147)	**SPAP**STRVITNQYNSPTG	2.29*
Tripeptidyl-peptidase 1(497-507)	**PP**LGFLNPRLY	2.31*
Tubulin alpha-1B chain(71-83)	**EP**TVIDEVRTGTY	1.80*
Uncharacterized protein C19orf21(479-490)	**VP**DVPQGTETPH	3.04*

*, $p<0.05$

**, $p<0.01$

Table S14. Absolute fold changes of peptides identified in the SCX-100 mM KCl free fraction of the global peptide profiling experiments (N = 4) with *DPP4*$^{+/+}$ and *DPP4*$^{-/-}$ gut samples applying SCX-RP-peptidomics workflow. The preferred DPP4 truncation sites in the *DPP4*$^{-/-}$ elevated peptide sequences are highlighted in bold.

Protein (peptide region)	Peptide Sequence	Fold Change
DPP4$^{-/-}$ Elevated Peptides		DPP4$^{-/-}$/ DPP4$^{+/+}$
Defensin-related cryptdin-5(20-33)	**DP**IHKTDEETNTEE	6.27*
Histone H2B type 1-M(110-122)	**HA**VSEGTKAVTKY	10.55**
Plasminogen activator inhibitor 1 RNA-binding protein(112-126)	**RP**DQQLQGDGKLIDR	7.94*
Putative uncharacterized protein GN=Tf, PE=2, SV=1(94-114)	GREEKPAASDSSGKQSTQVMA	18.86**
Secretogranin-1(21-35)	**AP**VDNRDHNEEMVTR	29.85*

*, p<0.05

**, p<0.01

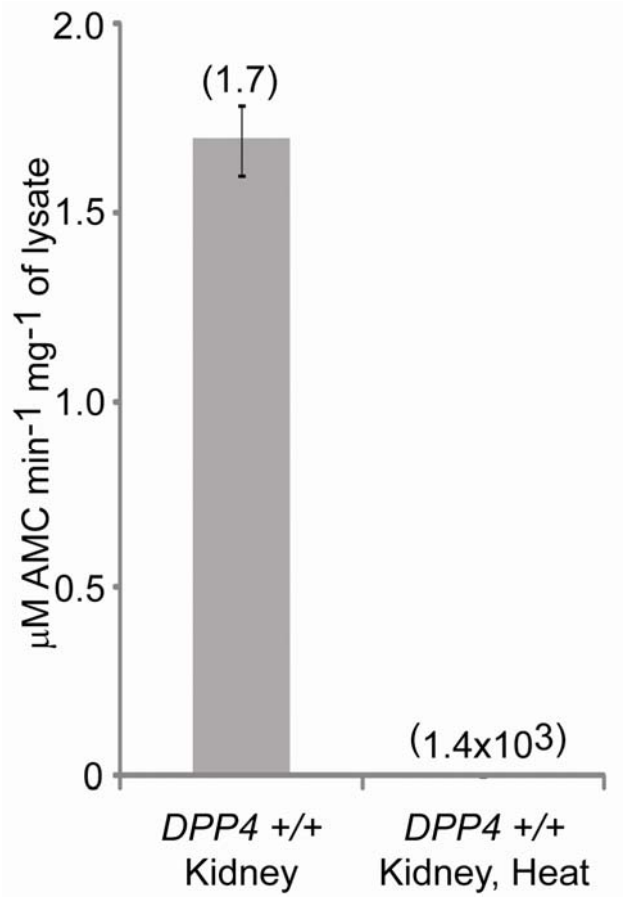

Figure S1. Measurement of DPP4 activity from kidney lysates. Kidney samples were dounced homogenized in 25 mM Tris-HCl, 140 mM NaCl, 10 mM KCl, pH 7.5, 0.1% BSA followed by centrifugation at 20,800 x g for 20 min at 4 °C. Heat treated samples were microwaved in water at high power for 2 min prior to homogenization. The lysate concentrations were obtained using a Bradford assay and the activity assay was run in the same buffer that was used to dounce the samples using 50 μg of total protein (100 μL of a 0.5 mg mL^{-1} lysate) and 22.5 μM of the substrate, H-GlyPro-AMC (aminomethylcoumarin), in a final volume of 100 μL. The reaction was monitored by measuring the fluorescence increase associated with the liberation of the AMC group at 360 nm using a Spectramax plate reader. A standard curve was run with each sample and used to determine the absolute amount of AMC generated during the reaction. These results show that pre-heating samples is able to inactivate DPP4. Actual values for peak heights are given in parenthesis above each bar.

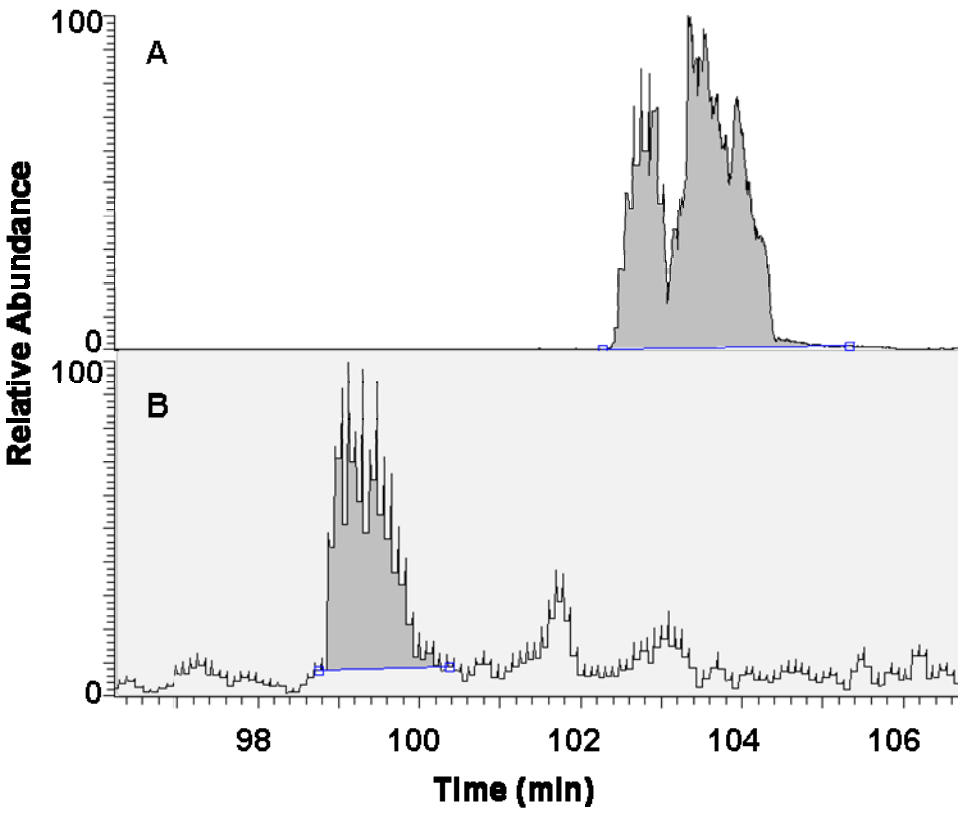

Figure S2. Peak integration of the leucine heavy-label peptide standard, RPGL*L*DL*KGKAKWD, in the presence (area: 5.6±0.4 x 10^6) and absence (area: 66.5±4.2 x 10^6) of *DPP4*$^{-/-}$ sample background.

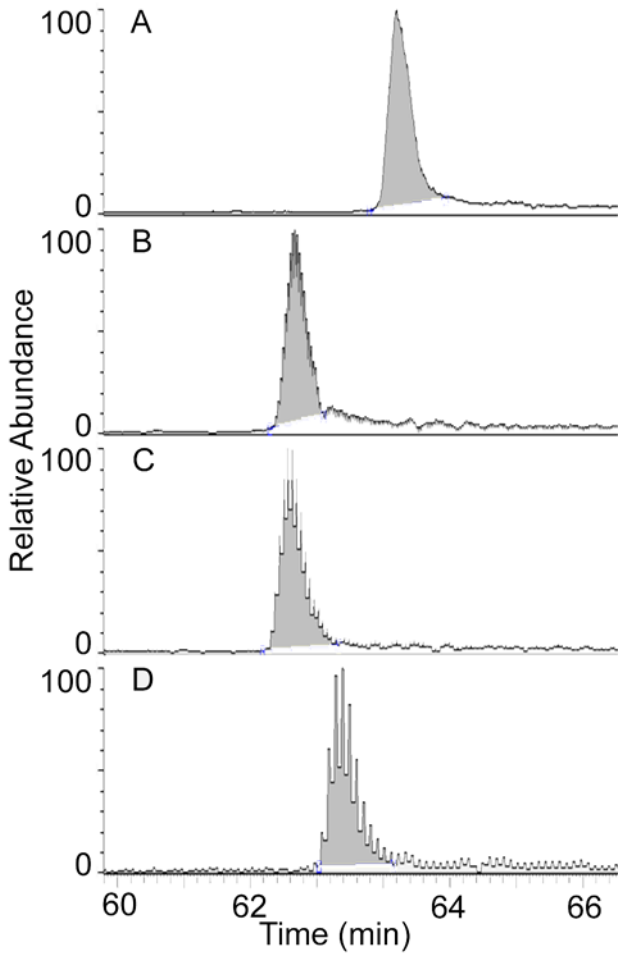

Figure S3. Comparison of the shape and areas of the MS peak for the vacuolar H+ ATPase, G1 subunit peptide RPEIHENYRING (z: +3; m/z: 499.92) collected in- A) Full MS (Most Gaussian peak shape; Area: $3.3x10^7$), B) Top 3 tandem MS (Area: $7.9x10^6$), C) Top 6 tandem MS (Area: $6.3x10^6$), and Top 10 tandem MS (Least Gaussian peak shape; Area: $3.2x10^6$) modes. As peak shape becomes more jagged (i.e., top-to-bottom) XCMS cannot accurately identify and integrate peaks. Therefore, samples are run using full MS (top) for accurate ion quantification by XCMS.

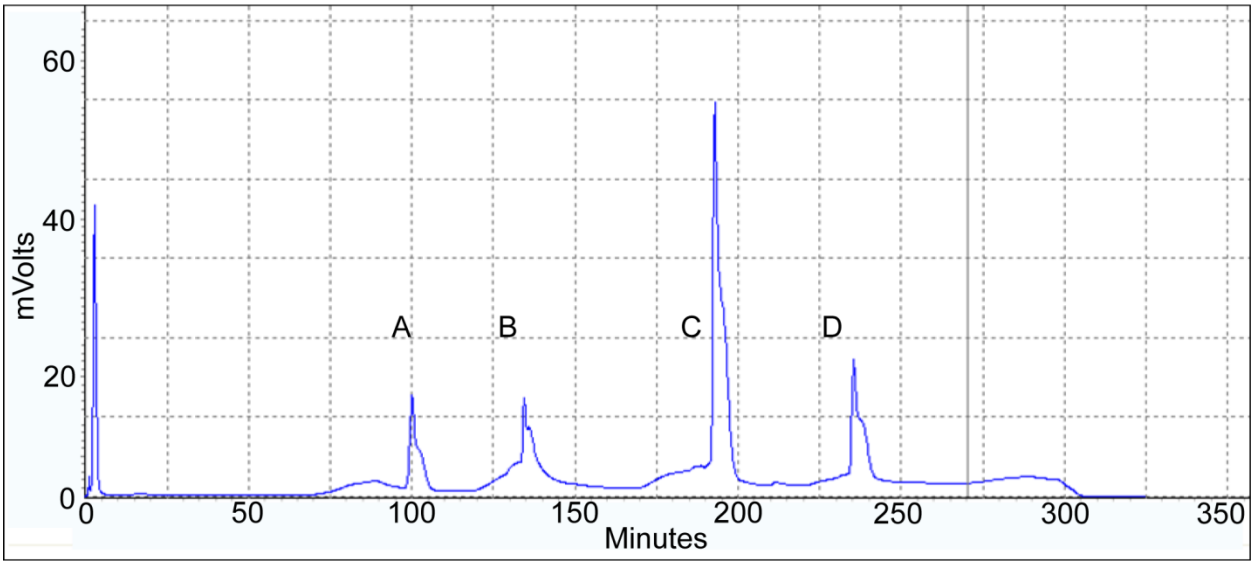

Figure S4. Order of peptide elution from a PolySULFOETHYL A™ SCX column- Fraction A (40 mM KCl): LPLFDRVLVE, +2; Fraction B (100 mM KCl): LPAPEKFVKDIDGGIDQDIFD, +3; Fraction C (200 mM KCl) GLLDLKGKAKWD, +4; and Fraction D (400 mM KCl): RPGLLDLKGKAKWD, +5.