Supporting information

Posttranslational modifications of the photoreceptor-specific ABC transporter ABCA4[†]

Yaroslav Tsybovsky[‡1], *Benlian Wang*[§1], *Faraz Quazi*[@], *Robert S. Molday*[@], *Krzysztof Palczewski*[‡*]

[‡]Department of Pharmacology and [§]Center for Proteomics and Mass Spectrometry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA

[@]Department of Biochemistry and Molecular Biology, Centre for Macular Research, University of British Columbia, Vancouver, BC, V6T 1Z3 Canada

[1]Equal contributions

*To whom correspondence should be addressed: E-mail: Krzysztof Palczewski (kxp65@cwru.edu). Phone: (216) 368-4631. Fax: (216) 368-1300.

Table S1. Combined trypsin/chymotrypsin sequence coverage of native bovine ABCA4[1].

Region	Position in sequence	Glycosylated ABCA4	Deglycosylated ABCA4
ECD1	46 - 651	82%	85%
ECD2	1392 – 1677	85%	90%
Cyto1	856 – 1370	88%	88%
Cyto2	1897 – 2281	97%	91%
Transmembrane	24 – 45	69%	52%
	652 – 855		
	1371 – 1391		
	1678 - 1896		
Total		84%	81%

[1]Mass spectrometry was performed as described in Experimental Procedures. ECD1, ECD2, Cyto1, Cyto2: exocytoplasmic domains 1 and 2, cytoplasmic domains 1 and 2, respectively.

Table S2. Glycosylated peptides and composition of oligosaccharides in native bovine ABCA4[1].

N-glycosylation site	Peptide sequence	Observed N-glycopeptide m/z	Glycan composition
N415	NA**N**STFEELER	1344.54(2+)	$Hex6(HexNAc)_2$
		1425.58(2+)	$Hex7(HexNAc)_2$
		1506.60(2+)	$Hex8(HexNAc)_2$
		1587.63(2+)	$Hex9(HexNAc)_2$
N504	DIF**N**ITDR	1206.99(2+)	$Hex2(HexNAc)1+Hex3(HexNAc)_2$
		1288.04(2+)	$Hex3(HexNAc)1+Hex3(HexNAc)_2$
		1352.56(2+)	$Hex2(HexNAc)1(NeuAc)1+Hex3(HexNAc)_2$
		1433.59(2+)	$Hex3(HexNAc)1(NeuAc)1+Hex3(HexNAc)_2$
		1328.03(2+)	$Hex3(HexNAc)1(phos)1+Hex3(HexNAc)_2$
N1455	CLKEEWLPEFPCG**N**SSPWK	1153.83(3+)	$(HexNAc)1+Hex_3(HexNAc)_2$
		1221.53(3+)	$(HexNAc)2+Hex_3(HexNAc)_2$
N1527	STEILQDLTDR**N**VSDFLVK	1137.20(3+)	$Hex_5(HexNAc)_2$
		1191.22(3+)	$Hex_6(HexNAc)_2$
		1245.24(3+)	$Hex_7(HexNAc)_2$
		1299.26(3+)	$Hex_8(HexNAc)_2$
		1353.28(3+)	$Hex_9(HexNAc)_2$
N1586	LPAPPFTGEALVGFLSDLGQLM**N**VS GGPMTR	1065.51(4+)	$Hex_4(HexNAc)_2$
		1106.02(4+)	$Hex_5(HexNAc)_2$
		1146.53(4+)	$Hex_6(HexNAc)_2$
N1660	DKNPEEYGITVISQPLNLTK	1213.24(3+)	$Hex_6(HexNAc)_2$
		1267.26(3+)	$Hex_7(HexNAc)_2$
		1321.28(3+)	$Hex_8(HexNAc)_2$
		1375.30(3+)	$Hex_9(HexNAc)_2$

[1]Mass spectrometry was performed as described in Experimental Procedures. Hex, hexose; HexNAc, N-acetylhexosamine; NeuAc, N-acetylneuraminic acid; phos, phosphate.

Table S3. Specificities of lectins used in this study[1].

Abbreviation	Full name	Specificity	Binding detected?
ConA	Concanavalin A	α-linked mannose	yes
LCA	*Lens culinaris* agglutinin	α-linked mannose	yes
PSA	*Pisum sativum* agglutinin	α-linked mannose-containing oligosaccharides	yes
WGA	Wheat germ agglutinin	N-acetylglucosamine and sialic acid	yes
Suc. WGA	Succinylated WGA	N-acetylglucosamine	yes
MAL	*Maackia amurensis* lectin II	(α-2,3) linked sialic acid	yes
UEA-I	*Ulex europaeus* agglutinin I	α-linked fucose	no
RCA120	*Ricinus communis* agglutinin I	oligosaccharides ending in galactose; N-acetylgalactosamine	no
DBA	*Dolichos biflorus* agglutinin	α-linked N-acetylgalactosamine	no
GSL-I	*Griffonia (Bandeiraea) simplicifolia* lectin I	α-linked N-acetylgalactoeamine	no
SBA	Soybean agglutinin	terminal α- or β-linked N-acetylgalactosamine, galactose	no
PNA	Peanut agglutinin	galactosyl (β-1,3) N-acetylgalactosamine	no
PHA-E, PHA-L	*Phaseolus vulgaris* agglutinin E and L	terminal galactose, N-acetylglucosamine and mannose residues of complex glycans on mammalian glycoproteins	no

[1]Details of the lectin binding assays are explained in Experimental Procedures.

Fig. S1. Sequence coverage of glycosylated native bovine ABCA4. Peptides detected after digestion by trypsin are indicated in red, and those by chymotrypsin in blue. Protein purification, enzymatic digestion and mass spectrometry experiments are described in Experimental Procedures.

```
   1  MGFARQIKLL LWKNWTLRKR QKIRFVVELV WPLSLFLVLI WLRNVNPLYS KHECHFPNKA MPSAGMLPWL QGIFCNVNNP CFQSPTAGES PGIVSNYNNS
 101  ILARVYRDFQ ELLMDAPESQ HLGQVWRELR TLSQLMNTLR MHPERIAGRG IRIREVLKDD EMLTLFLVKN IGLSDSVVYL LVNSQVRPEQ FARGVPDLML
 201  KDIACSEALL ERFLIFPQRR AAQTVRGSLC SLSQGTLQWM EDTLYANVDF FKLFHVFPRL LDSRSQGMNL RSWGRILSDM SPRIQEFIHR PSVQDLLWVT
 301  RPLVQTGGPE TFTQLMGILS DLLCGYPEGG GSRVFSFNWY EDNNYKAFLG IDSTRKDPIY SYDERTTTFC NALIQSLESN PLTKIAWRAA KPLLMGKILF
 401  TPDSPATRRI LKNANSTFEE LERVRKLVKV WEEVGPQIWY FFDKSTQMSM IRDTLENPTV KAFWNRQLGE EGITAEAVLN FLYNGPREGQ ADDVDNFNWR
 501  DIFNITDRAL RLANQYLECL ILDKFESYDD EFQLTQRALS LLEENRFWAG VVFPDMHPWT SSLPPHVKYK IRMDIDVVEK TNKIKDRYWD SGPRADPVED
 601  FRYIWGGFAY LQDMVEHGIT RSQAQEEVPV GIYLQQMPYP CFVDDSFMII LNRCFPIFMV LAWIYSVSMT VKSIVLEKEL RLKETLKNQG VSNRVIWCTW
 701  FLDSFSIMSM SICLLTIFIM HGRILHYSNP FILFLFLLAF SIATIMQCFL LSTFFSRASL AAACSGVIYF TLYLPHILCF AWQDRITADM KMAVSLLSPV
 801  AFGFGTEYLA RFEEQGVGLQ WSNIGNSPME GDEFSFLMSM KMMLLDAALY GLLAWYLDQV FPGDYGTPLP WYFLLQESYW LGGEGCSTRE ERALEKTEPI
 901  TEEMEDPEYP EGINDCFFER ELPGLVPGVC VKNLVKIFEP YGRPAVDRLN ITFYESQITA FLGHNGAGKT TTLSIMTGLL PPTSGTVLVG GKDIETNLDA
1001 IRQSLGMCPQ HNILFHHLTV AEHILFYAQL KGRSWDEAQL EMEAMLEDTG LHHKRNEEAR DLSGGVQRKL SVAIAFVGDA KVVVLDEPTS GVDPYSRRSI
1101 WDLLLKYRSG RTIIMSTHHM DEADILGDRI AIISQGRLYC SGTPLFLKNC FGTGFYLTLV RRMKTIQSQG RGREATCSCA SKGFSVRCPA CAEAITPEQV
1201 LDGDVNELTD MVHHHVPEAK LVECIGQELI FLLPNKNFKQ RAYASLFREL EETLADLGLS SFGISDTPLE EIFLKVTEDL DSGHLFAGGT QQKRENINLR
1301 HPCSGPSEKA GQTPQGSSSH PGEPAAHPEG QPPPEREGHS RLNSGARLIV QHVQALLVKR FQHTIRSHKD FLAQIVLPAT FVFLALMLSL IIPPFGEYPA
1401 LTLHPWMYGQ QYTFFSMDQP DSEWLSALAD VLVNKPGFGN RCLKEEWLPE FPCGNSSPWK TPSVSPDVTH LLQQQKWTAD QPSPSCRCST REKLTMLPEC
1501 PEGAGGLPPP QRIQRSTEIL QDLTDRNVSD FLVKTYPALI RSSLKSKFWV NEQRYGGISV GGKLPAPPFT GEALVGFLSD LGQLMNVSGG PMTREAAKEM
1601 PAFLKQLETE DNIKVWFNNK GWHALVSFLN VAHNAILRAS LHKDKNPEEY GITVISQPLN LTKEQLSEIT VLTTSVDAVV AICVIFAMSF VPASFVLYLI
1701 QERVNKAKHL QFVSGVSPTT YWLTNFLWDI MNYTVSAALV VGIFIGFQKK AYTSSENLPA LVALLMLYGW AVIPMMYPAS FLFDIPSTAY VALSCANLFI
1801 GINSSAITFV LELFENNRTL LRINAMLRKL LIIFPHFCLG RGLIDLALSQ AVTDVYAQFG EAHSSNPFQW DLIGKNLAAM AVEGVVYFLL TLLIQYQFFF
1901 SRWTTEPAKE PITDEDDDVA EERQRIISGG NKTDILRLNE LTKVYSGTSS PAVDRLCVGV RPGECFGLLG VNGAGKTTTF KMLTGDTAVT SGDATVAGKS
2001 ILTNISDVHQ SMGYCPQFDA IDDLLTGREH LYLYARLRGV PAEEIERVTN WSIQSLGLSL YADRLAGTYS GGNKRKLSTA IALIGCPPLV LLDEPTTGMD
2101 PQARRMLWNT IMGIIREERA VVLTSHSMEE CEALCTRLAI MVKGAFQCLG TIQHLKSKFG DGYIVTMKIR SPKDDLLPDL GPVEQFFQGN FPGSVQRERH
2201 YNTLQFQVSS SSLARIFRLL VSHKDSLLIE EYSVTQTTLD QVFVNFAKQQ NETYDLPLHP RTAGASRQAK EVDKGNSAPQ G
```

Fig. S2. Sequence coverage of ABCA4 deglycosylated with PNGas F. Seven peptides containing the N-glycosylation motif N-X-S/T (where X is any amino acid except proline) were detected based on conversion of Asn residues to Asp by PNGase F. Peptides detected after digestion by trypsin are indicated in red, and those by chymotrypsin in blue.

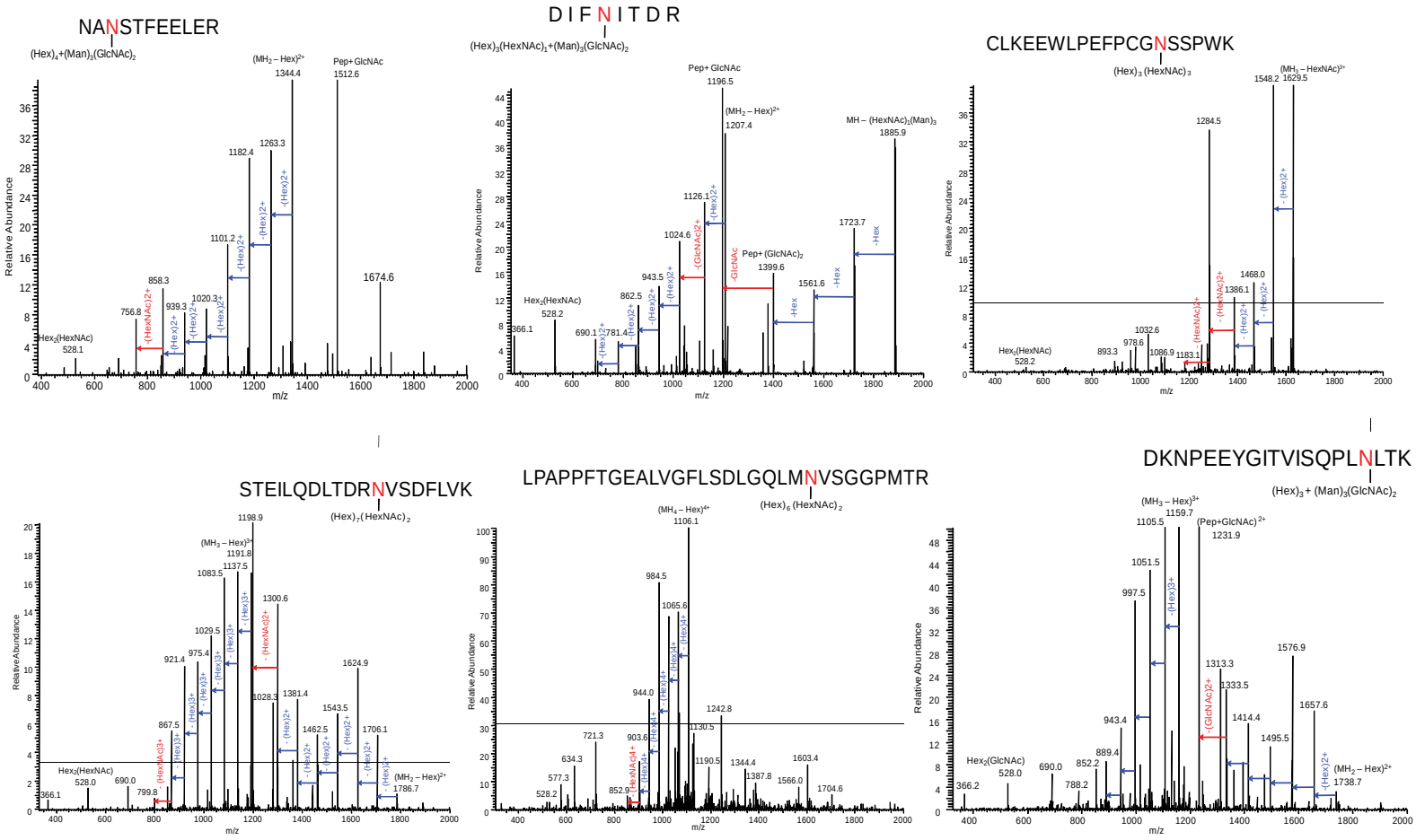

Fig. S3. Determination of glycan compositions in native bovine ABCA4. Representative tandem mass spectra of glycosylated peptides listed in **Table S2** reflect collision-induced dissociation of attached oligosaccharides. See Experimental Procedures and Results for details.

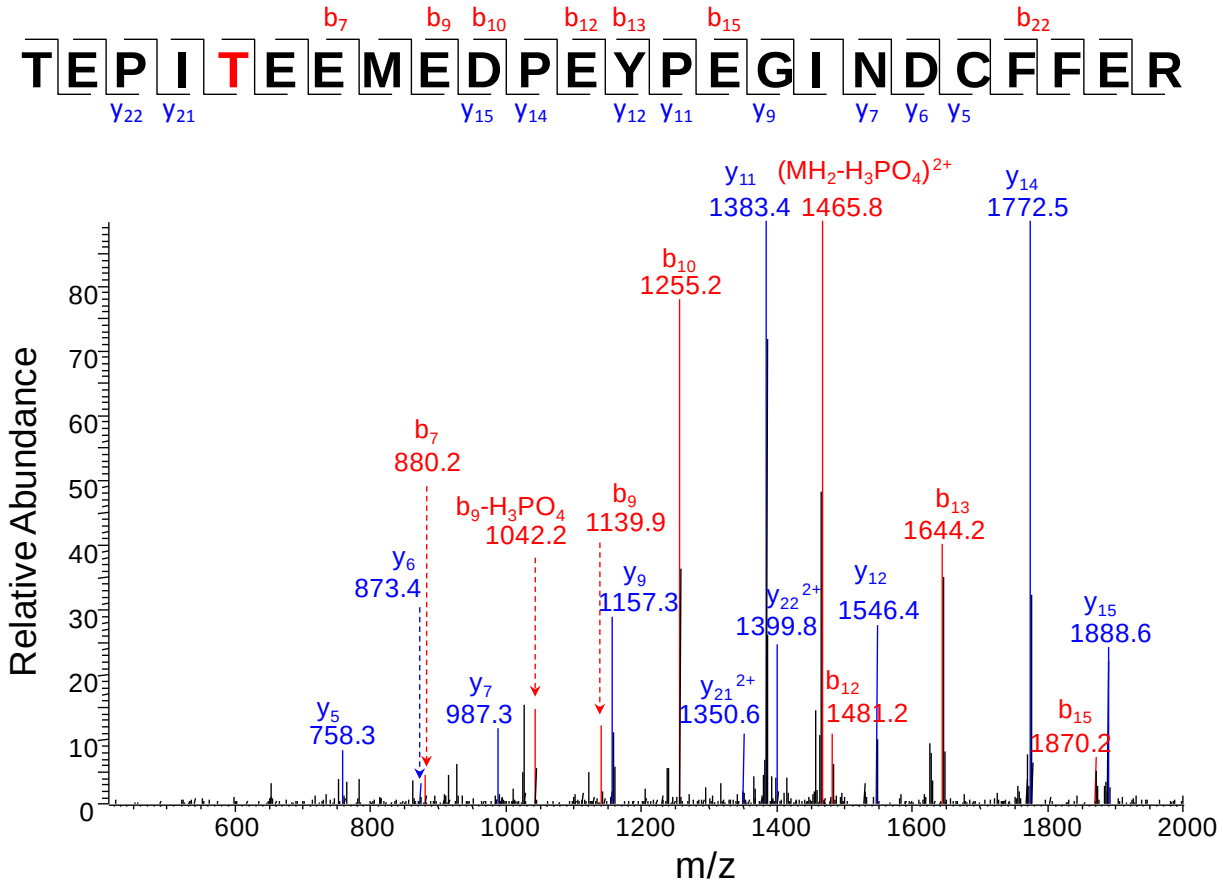

Fig. S4. Tandem mass spectrum of **TEPITEEMEDPEYPEGINDCFFER,** peptide 897-920, with doubly charged precursor ion m/z of 1514.09 suggests phosphorylation of T901 (shown in red).

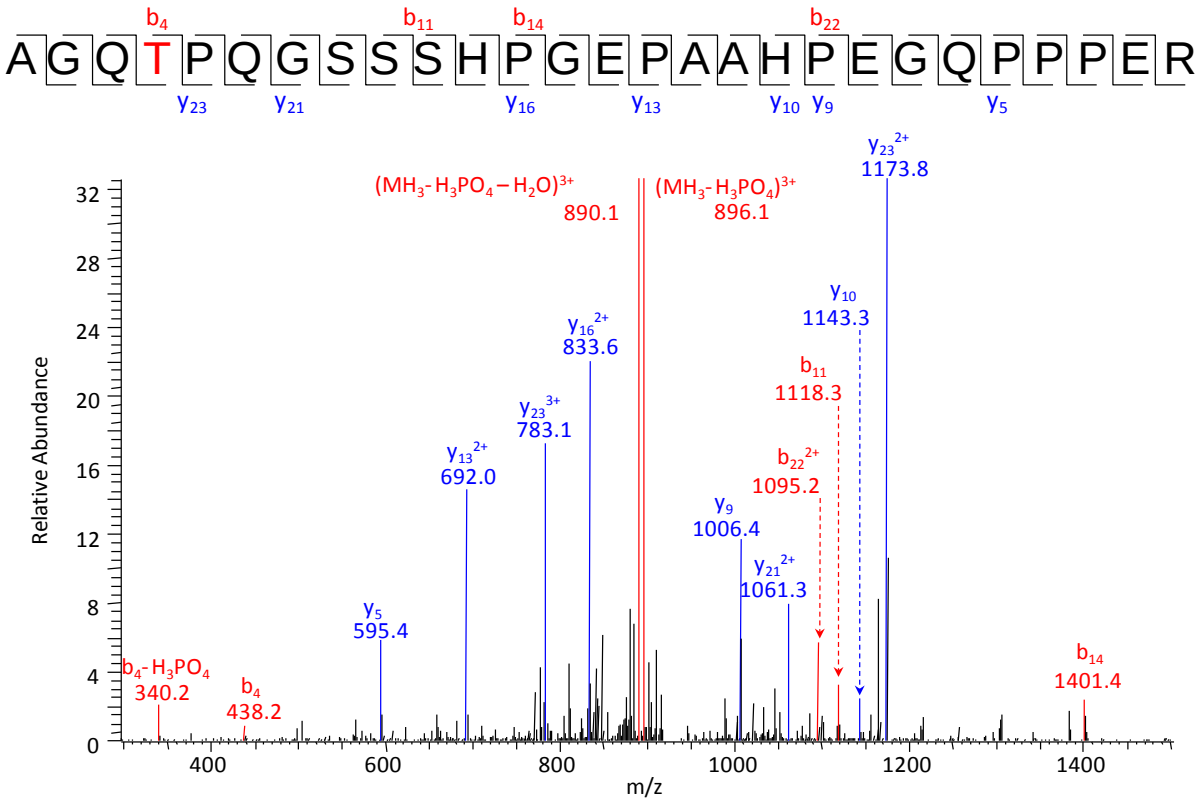

Fig. S5. Tandem mass spectrum of **AGQTPQGSSSHPGEPAAHPEGQPPPER,** peptide 1310-1336, with triply charged precursor ion m/z of 928.42. Mass shift on precursor ion of +80 Da and loss of H_3PO_4 on precursor ion and b_4 ions suggest phosphorylation of T1313 (shown in red).

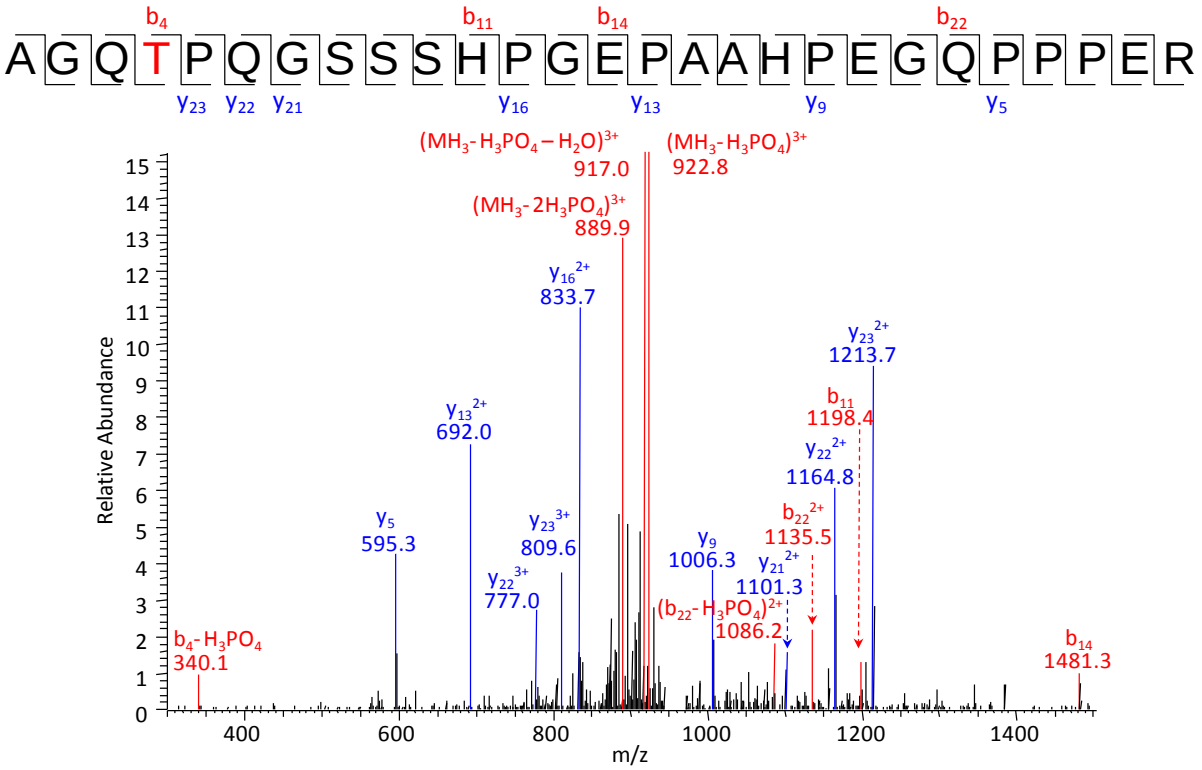

Fig. S6. Tandem mass spectrum of **AGQTPQGSSSHPGEPAAHPEGQPPPER,** peptide 1310-1336, with triply charged precursor ion m/z of 955.07. Mass shift of +160 Da and the neutral loss of $2H_3PO_4$ on the precursor ion suggest two phosphorylation sites on this peptide. One site is on T1313 (shown in red), the other is among S1317 -1319.

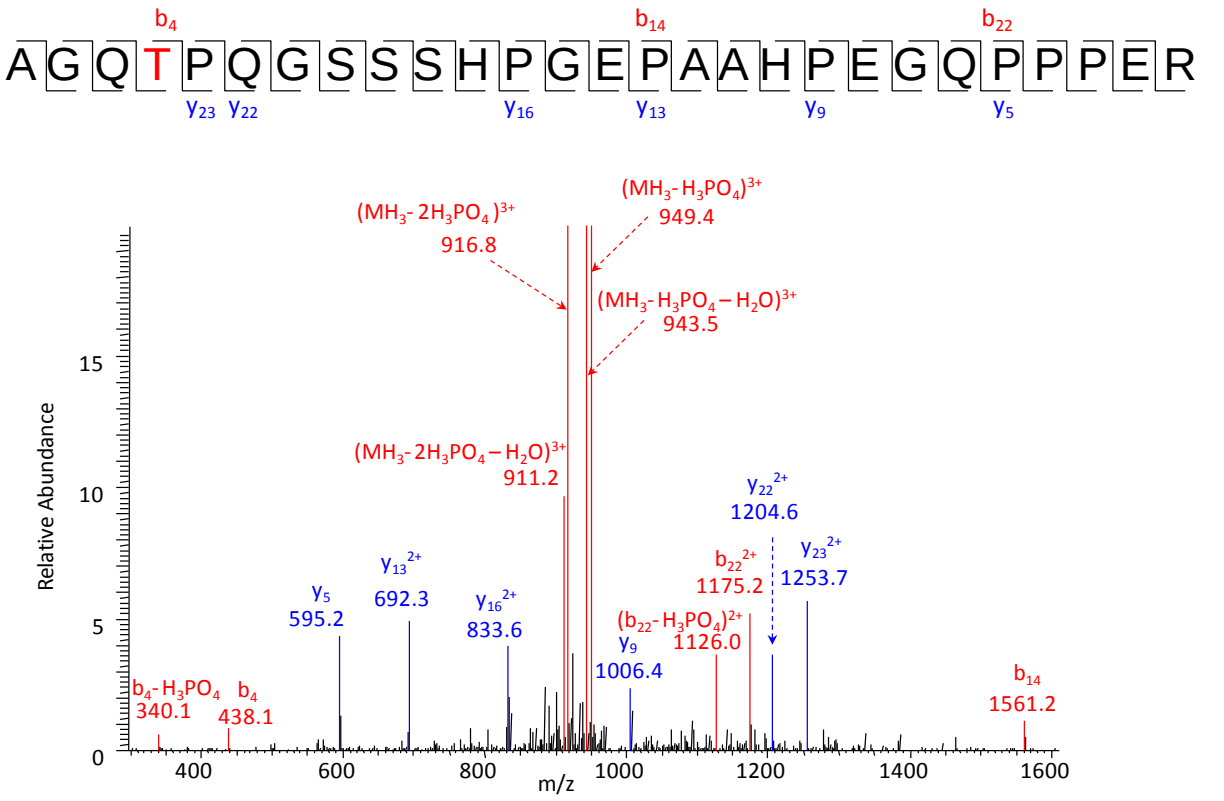

Fig. S7. Tandem mass spectrum of peptide **AGQTPQGSSSHPGEPAAHPEGQPPPER**, peptide 1310-1336, with the triply charged precursor ion m/z of 981.73. Mass shift of +240 Da suggests three phosphorylation sites on this peptide. One site is on T1313 (shown in red), the two others are among S1317 -1319.

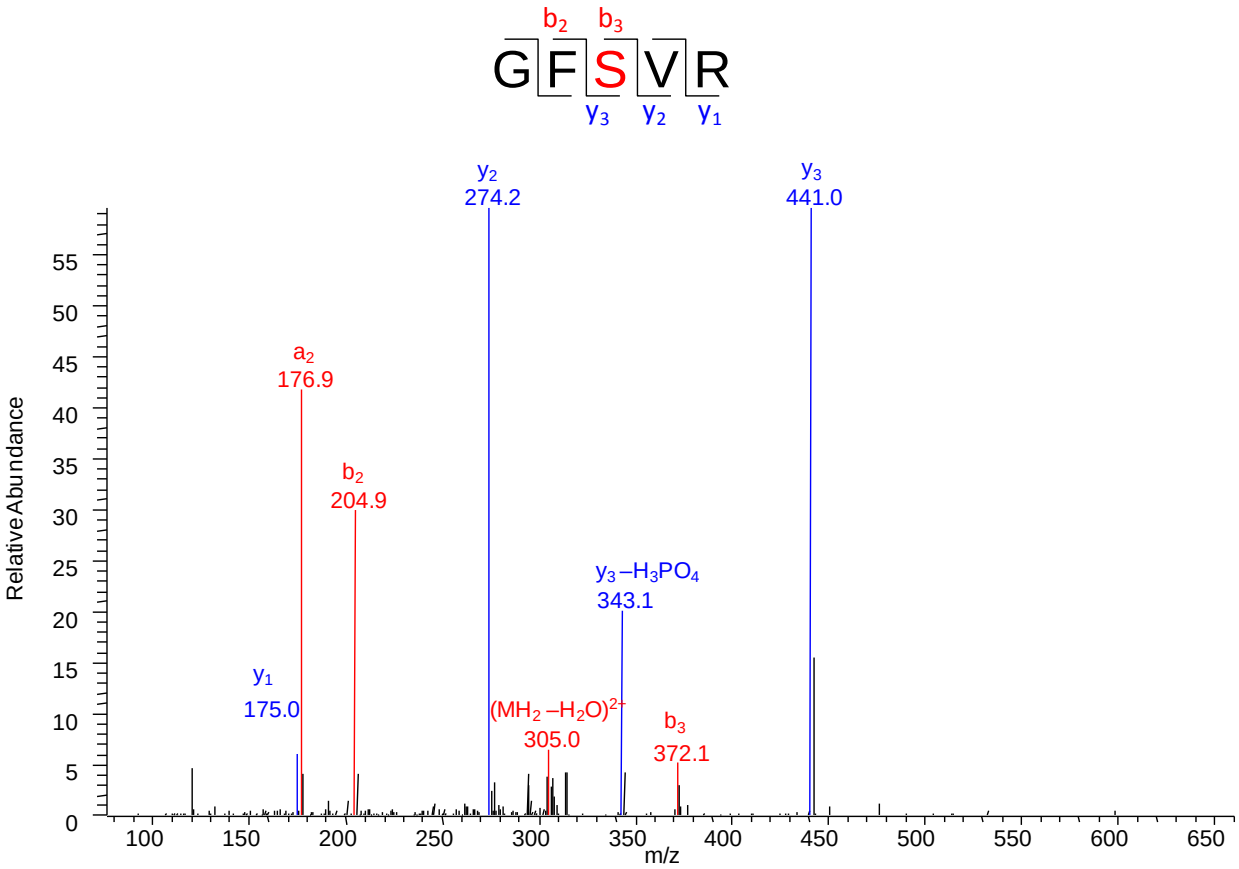

Fig. S8. Tandem mass spectrum of **GFSVR**, peptide 1183-1187, with doubly charged precursor ion m/z of 323.24 suggests a phosphorylation site on S1185 (shown in red).

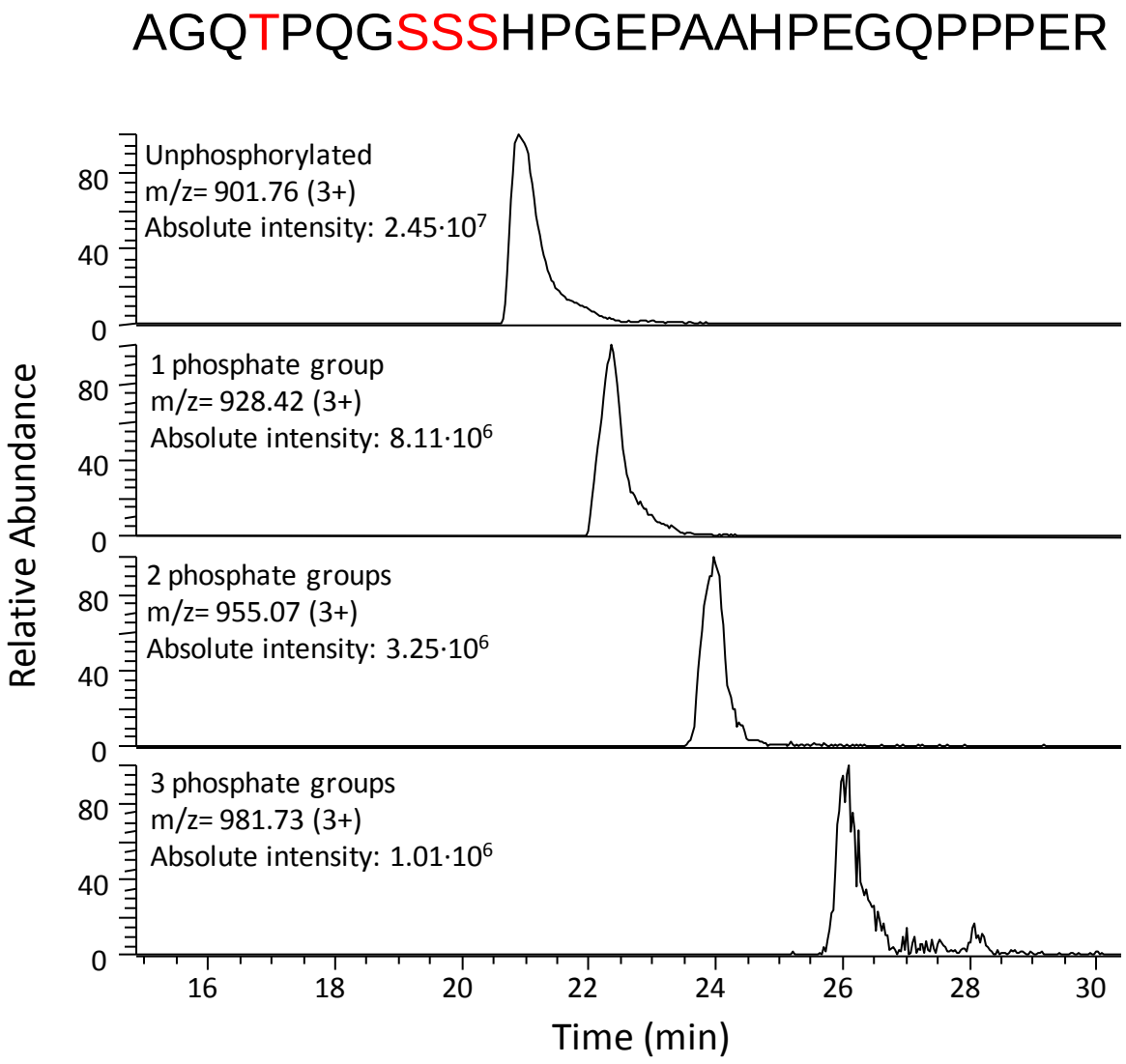

Fig. S9. Chromatograms showing base peaks corresponding to the **AGQTPQGSSSHPGEPAAHPEGQPPPER** peptide (residues 1310-1326) phosphorylated at 0, 1, 2 and 3 sites (top to bottom). Phosphorylation sites described in the current paper are shown in red (only two of the three consecutive Ser residues became phosphorylated). Calculated from the corresponded peak areas, the ratios of the unphosphorylated peptide to the peptides carrying 1, 2 and 3 phosphate groups are 1:0.29, 1:0.11 and 1:0.03, respectively.

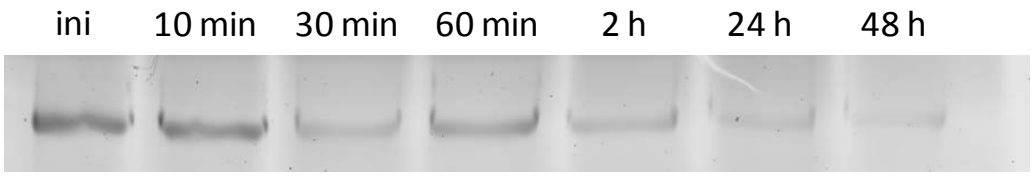

Fig. S10. Dephosphorylation of native bovine ABCA4 with PP2A at 4°C. PP2A treated samples of ABCA4 were resolved on a SDS-PAGE gel and stained with ProQ Diamond as described in Experimental Procedures. ini: ABCA4 before addition of PP2A.