Identification of novel glycosylation events on human serum-derived Factor IX

[†]School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD 4072, Australia.

[‡]Centre for Biopharmaceutical Innovation, Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, St Lucia, QLD 4072, Australia.

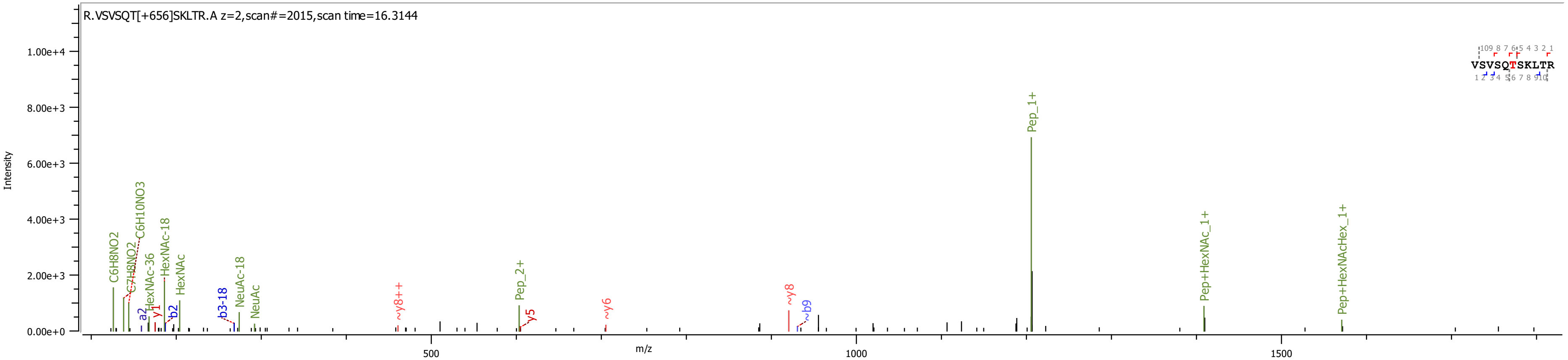
Cassandra L. Pegg,[†] Lucia F. Zacchi,[‡] Dinora Roche Recinos[‡] Christopher B. Howard,[‡] and Benjamin L. Schulz[†‡*]

[*]Corresponding author, Email: b.schulz@uq.edu.au

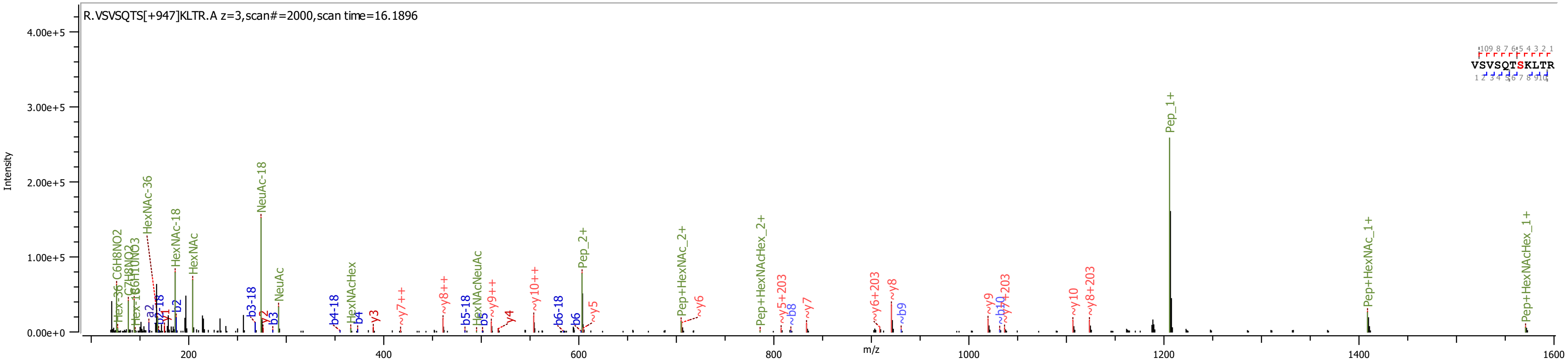
Fig. S-2a Annotated spectra of the novel *O*-linked site of human serum Factor IX from Byonic searches. Spectra have been included from the trypsin, chymotrypsin/PNGase and Glu-C digests of the sample.

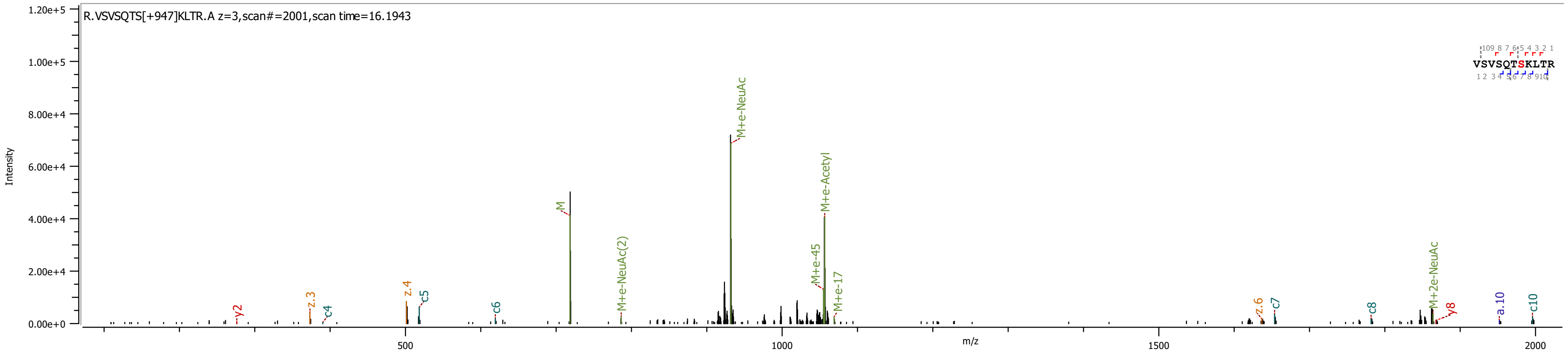
Trypsin digested sample:

HCD



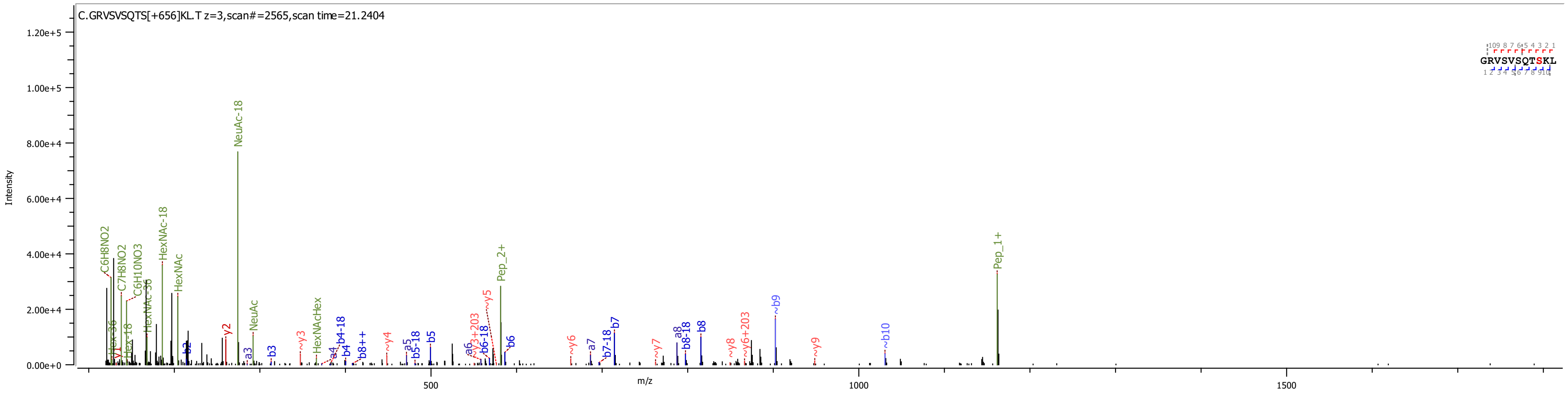
HCD-pd-ETciD

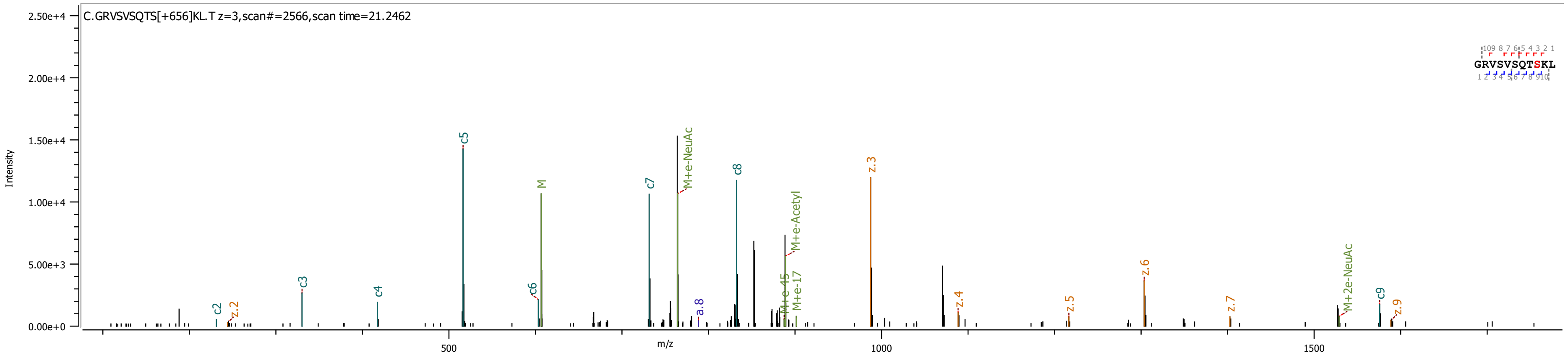




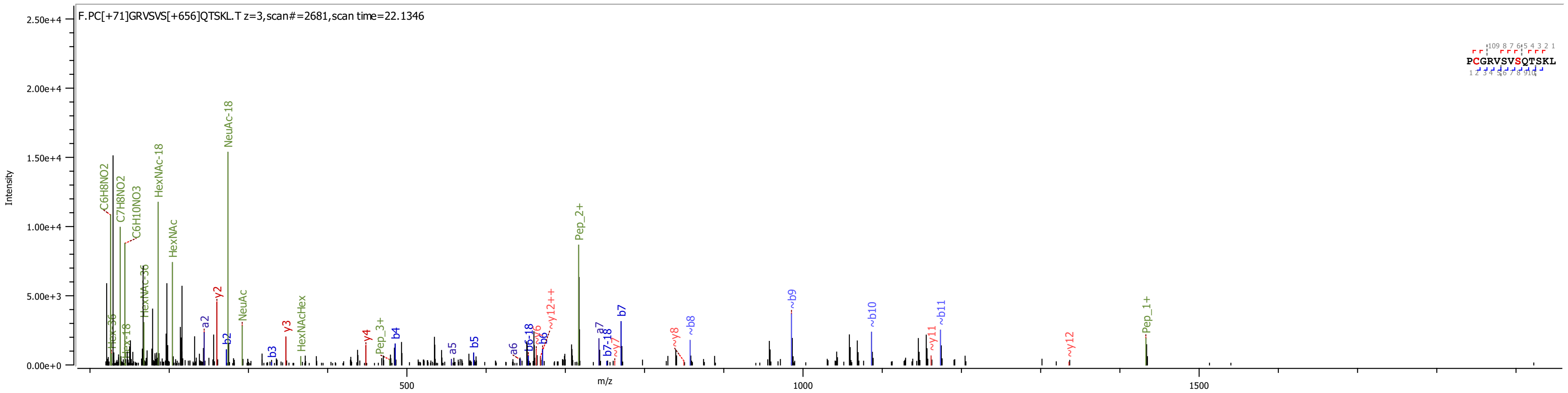
Chymotrypsin/PNGaseF digested sample:

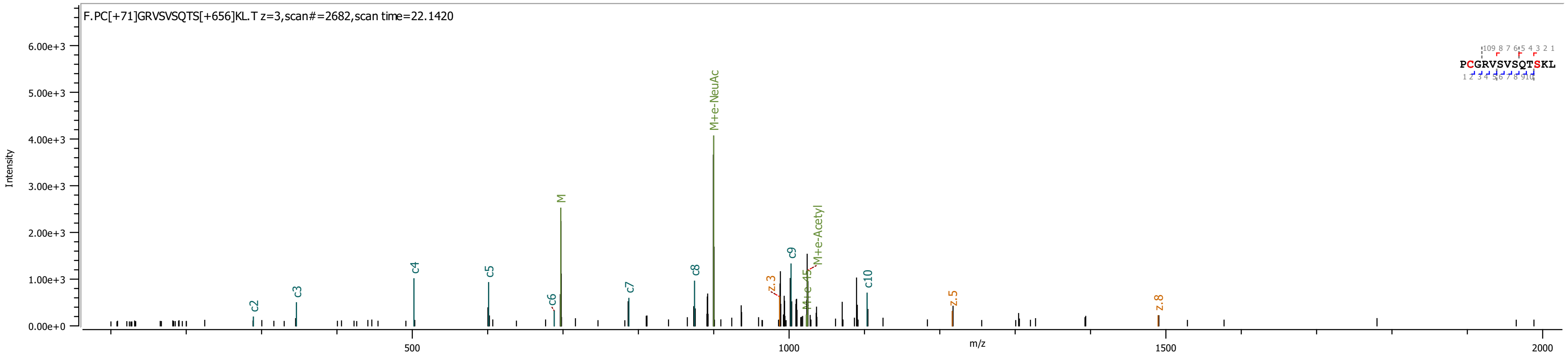
HCD-pd-ETciD




C.GRVSVSQTS[+656]KL.T z=3,scan#=2566,scan time=21.2462
Intensity
2.50e+4
2.00e+4
1.50e+4
1.00e+4
5.00e+3
0.00e+0
m/z
500
1000
1500
c2
z.2
c3
c4
c5
c6
M
c7
M+e-NeuAc
a.8
c8
M+e-45
M+e-17
M+e-Acetyl
z.3
z.4
z.5
z.6
z.7
M+2e-NeuAc
c9
z.9
GRVSVSQTSKL

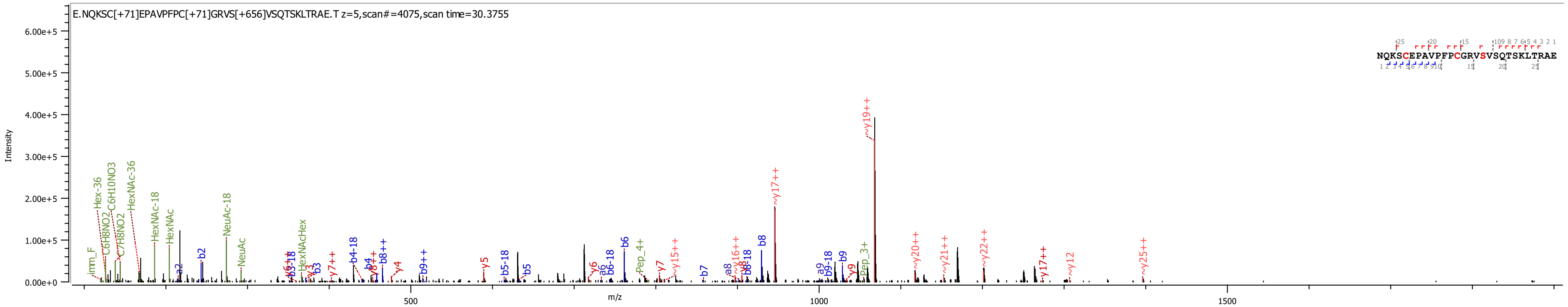
HCD-pd-ETciD


F.PC[+71]GRVSVS[+656]QTSKL.T z=3,scan#=2681,scan time=22.1346
Intensity
2.50e+4
2.00e+4
1.50e+4
1.00e+4
5.00e+3
0.00e+0
m/z
500
1000
1500
C6H8NO2
Hex-36
Hex-18
C7H8NO2
C6H10NO3
HexNAc-36
HexNAc-18
HexNAc
a2
y2
b2
NeuAc-18
NeuAc
b3
y3
HexNAcHex
y4
Pep_3+
b4
a5
b5
a6
b6-18
y6
b6
~y12++
Pep_2+
a7
b7-18
~y7
b7
~y8
~b8
~b9
~b10
~y11
~b11
~y12
Pep_1+
PCGRVSVSQTSKL



Glu-C digested sample:

HCD



HCD

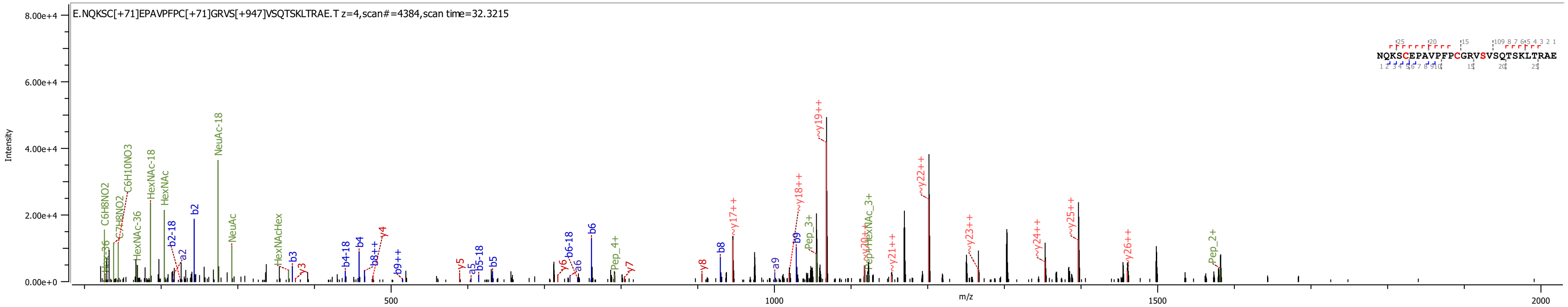
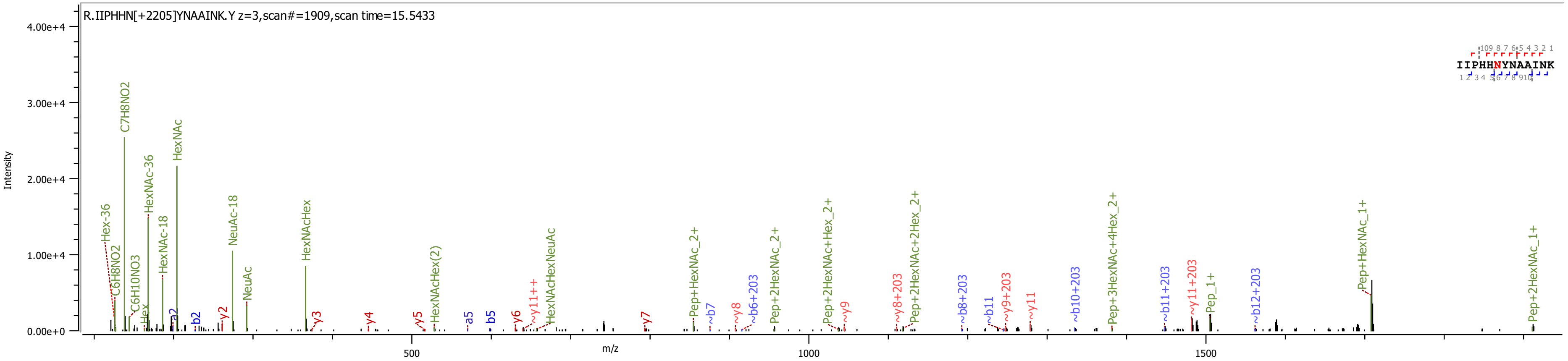

E.NQKSC[+71]EPAVPFPC[+71]GRVS[+947]VSQTSKLTRAE.T z=4,scan#=4384,scan time=32.3215
Intensity
8.00e+4
6.00e+4
4.00e+4
2.00e+4
0.00e+0
m/z
500
1000
1500
2000
Hex-36
C6H8NO2
C7H8NO2
C6H10NO3
HexNAc-36
HexNAc-18
HexNAc
b2-18
a2
b2
NeuAc-18
NeuAc
HexNAcHex
b3
y3
b4-18
b4
b8++
y4
b9++
y5
a5
b5-18
b5
y6
b6-18
a6
b6
Pep_4+
y7
y8
b8
~y17++
a9
b9
~y18++
Pep_3+
~y19++
~y20++
Pep+HexNAc_3+
~y21++
~y22++
~y23++
~y24++
~y25++
~y26++
Pep_2+
25 20 15 10 9 8 7 6 5 4 3 2 1
NQKSCEPAVPFPCGRVSVSQTSKLTRAE
1 2 3 4 5 6 7 8 9 10 15 20 25

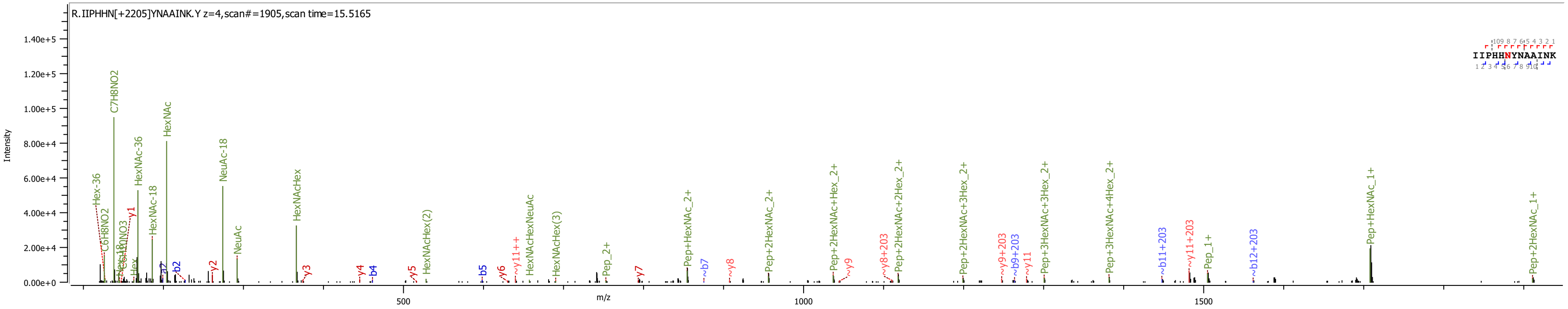
Fig. S-2b Annotated spectra of the novel *N*-linked site of human serum Factor IX from Byonic searches. Spectra have been included from the trypsin and Glu-C digests of the sample.

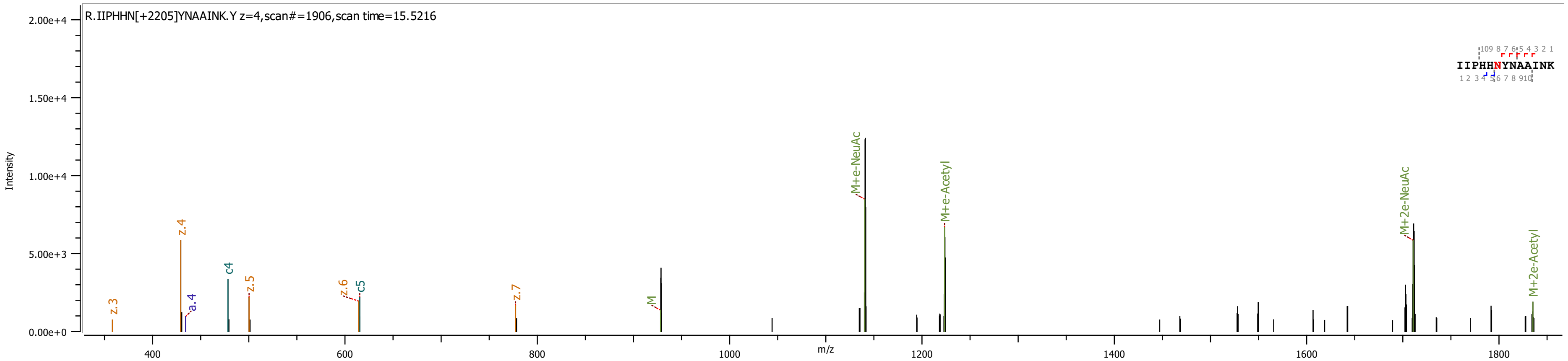
Trypsin digested sample:

HCD



HCD-pd_ETciD





Glu-C digested sample:

HCD

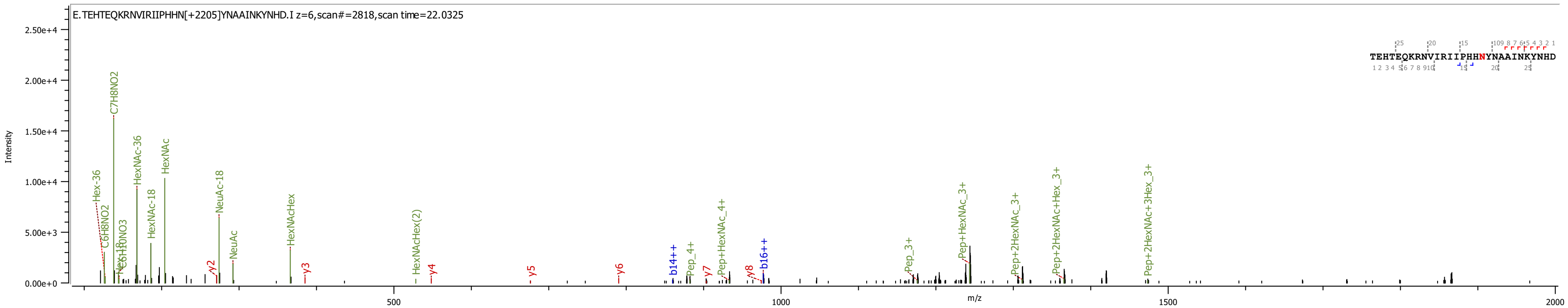
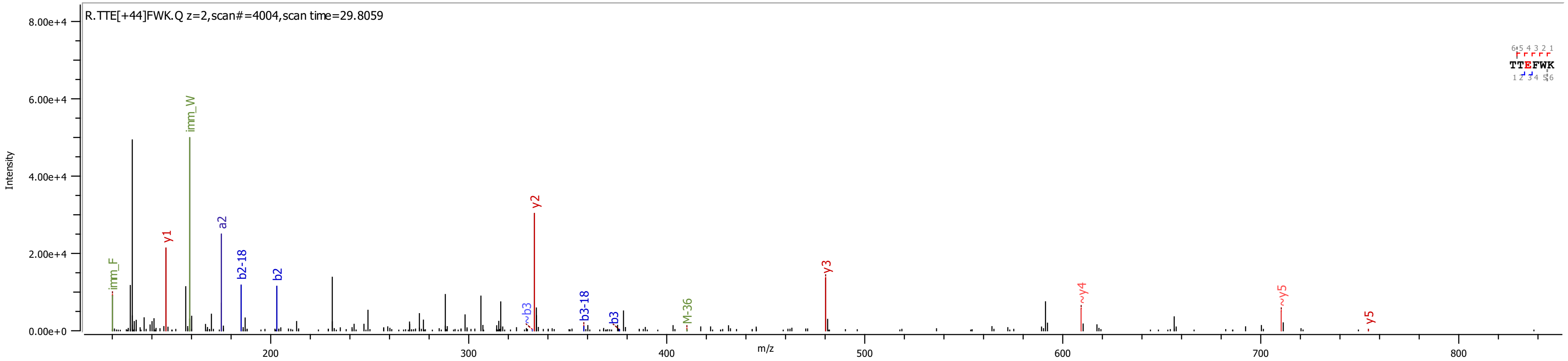


Fig. S-2c Annotated spectra of reported post-translational modifications (PTMs) of Factor IX from Byonic searches. Spectra have been listed in the order the PTMs are observed from the N-terminal region of Factor IX.

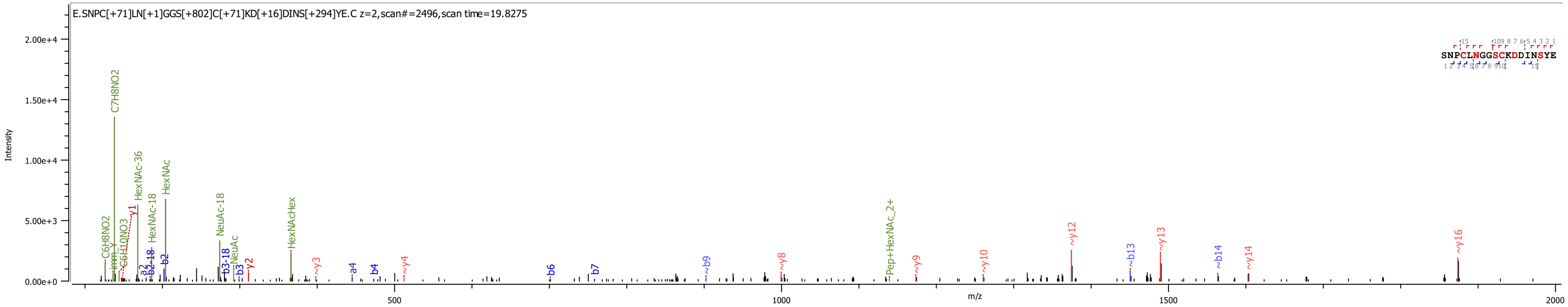
Reported modification of γ-carboxylation at E40[11-12] from the trypsin digested sample:

HCD



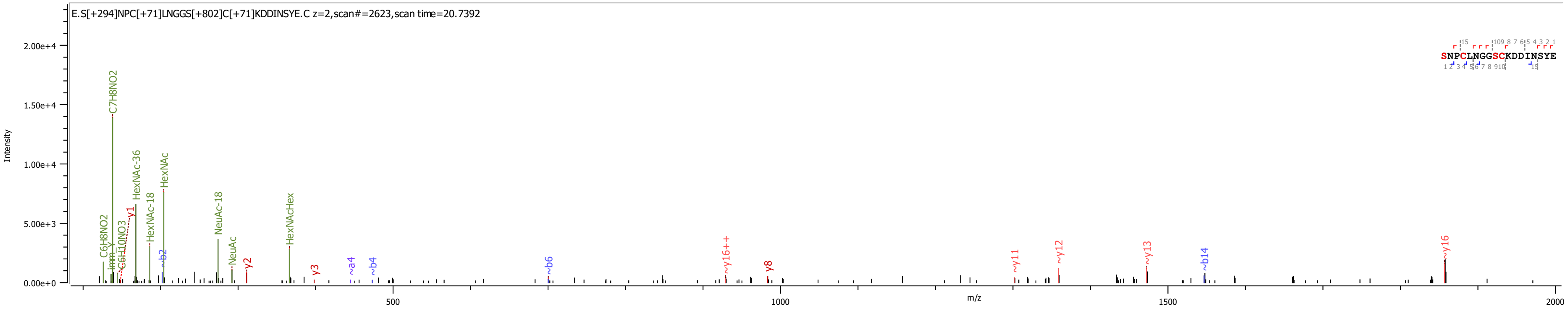
Reported modifications of *O*-linked glucose at Ser53[17], *O*-linked fucose at Ser6[18-19] and β-hydroxylation of Asp64[15-16] from the Glu-C digested sample. Please note Byonic has assigned the glycans in different positions to those reported in the literature.

HCD



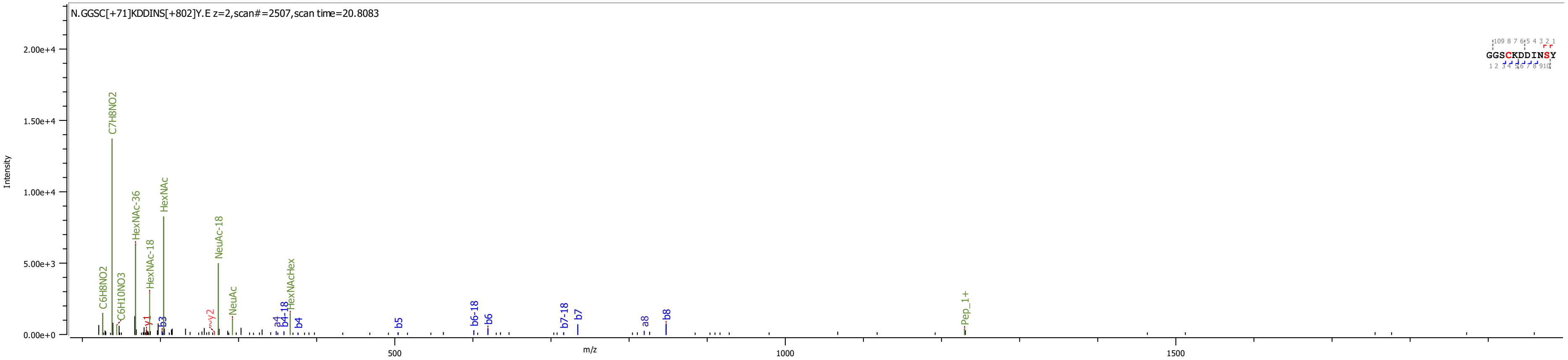
Reported modifications of *O*-linked glucose at Ser53[17] and *O*-linked fucose at Ser6[18-19] from the Glu-C digested sample.

HCD



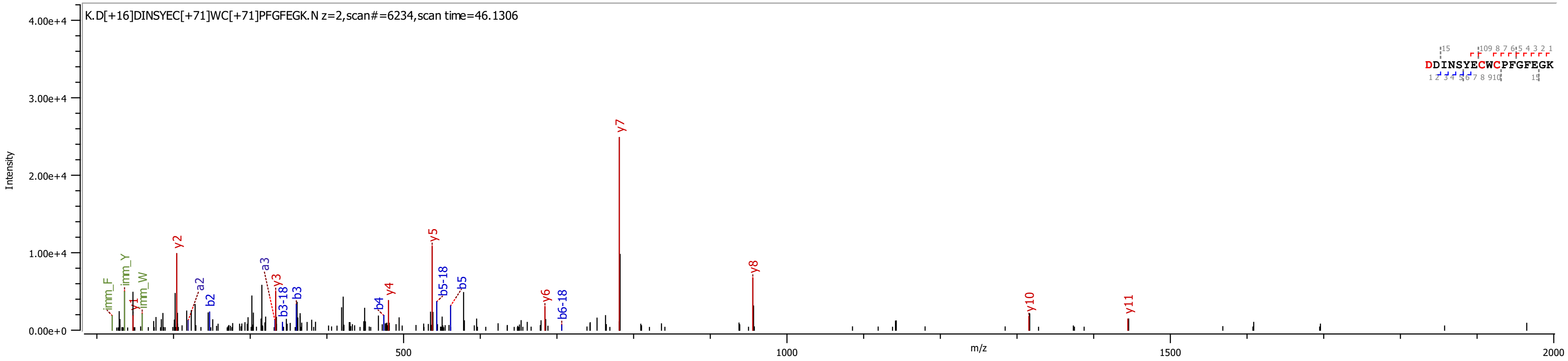
Reported modification of *O*-linked glucose at Ser6[18-19] from the chymotrypsin/PNGaseF digested sample.

HCD



Reported modification of β-hydroxylation of Asp64[15-16] from the trypsin digested sample.

HCD



Reported modification of phosphorylation at S68 (Noted in UniProt identifier P00740) from the trypsin digested sample.

HCD

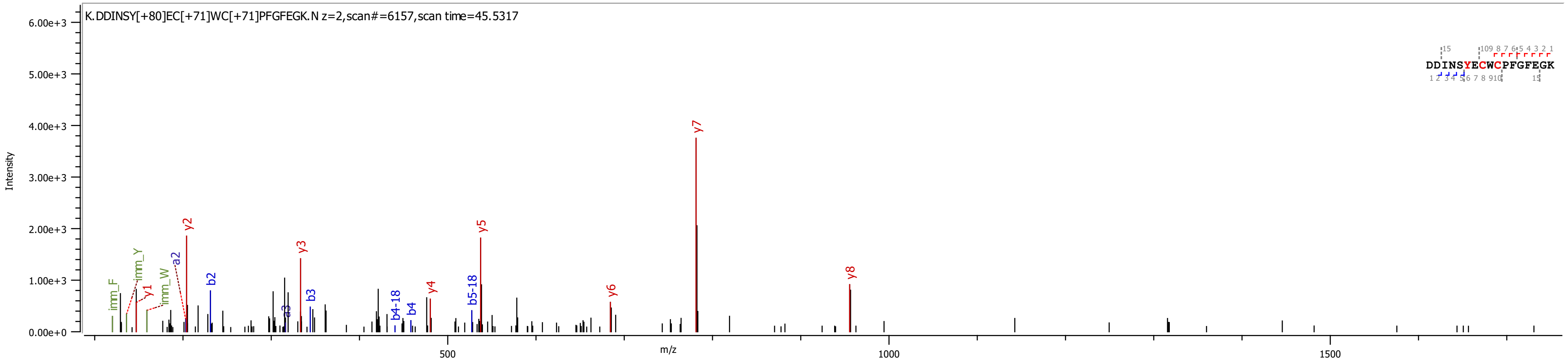


Fig. S-3 Multiple sequence alignment of human Factor IX, human protein C, human Factor X and human Factor VII and the sequences of Factor IX from Bos taurus (bovine), Pan troglodytes (chimpanzee), Mus musculus (mouse), Canis lupus familiaris (dog), Rattus norvegicus (rat) and Sus scrofa (pig). UniProt accession numbers are listed for each protein and amino acid numbering is based on the full protein sequences.

```
CLUSTAL O(1.2.4) multiple sequence alignment

SP|P00740|FA9_HUMAN  MQRVNMIMAESPGLITICL-L---GYL------------------LSAECTVFLDHENAN 38
SP|P04070|PROC_HUMAN MWQLTS-------LLLFVATWGISGT-------------------PAPLDSVFSSSERAH 34
SP|P00742|FA10_HUMAN MGRPLH-----LVLLSASL-A---GLL------------------L-LGESLFIRREQAN 32
SP|P08709|FA7_HUMAN  ------MVSQ--ALRLLCLLLGLQGCLAAGGVAKASGGETRDMPWKPGPHRVFVTQEEAH 52
SP|P00741|FA9_BOVIN  MWCLNMIMAESPGLVTICL-L---GYL------------------LSAECTVFLDRENAT 38
SP|Q95ND7|FA9_PANTR  MQRVNMIMAESPGLITICL-L---GYL------------------LSAECTVFLDHENAN 38
SP|P16294|FA9_MOUSE  MKHLNTVMAESPALITIFL-L---GYL------------------LSTECAVFLDRENAT 38
SP|P19540|FA9_CANLF  -------MAEASGLVTVCL-L---GYL------------------LSAECAVFLDRENAT 31
SP|P16296|FA9_RAT    -------MADAPGLIPIFL-L---GYL------------------LSTECAVFLDRENAT 31
SP|P16293|FA9_PIG    ------------------------------------------------------------


SP|P00740|FA9_HUMAN  KILNRPKRYNSGKLE-EFVQGNLERECMEEKCSFEEAREVFENTERTTEFWKQYVDGDQC 97
SP|P04070|PROC_HUMAN QVLRIRKRANS--FLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKHVDGDQC 92
SP|P00742|FA10_HUMAN NILARVTRANS--FLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQC 90
SP|P08709|FA7_HUMAN  GVLHRRRRANA--FLEELRPGSLERECKEEQCSFEEAREIFKDAERTKLFWISYSDGDQC 110
SP|P00741|FA9_BOVIN  KILHRPKRYNSGKLE-EFVRGNLERECKEEKCSFEEAREVFENTEKTTEFWKQYVDGDQC 97
SP|Q95ND7|FA9_PANTR  KILNRPKRYNSGKLE-EFVQGNLERECMEEKCSFEEAREVFENTERTTEFWKQYVDGDQC 97
SP|P16294|FA9_MOUSE  KILTRPKRYNSGKLE-EFVRGNLERECIEERCSFEEAREVFENTEKTTEFWKQYVDGDQC 97
SP|P19540|FA9_CANLF  KILSRPKRYNSGKLE-EFVRGNLERECIEEKCSFEEAREVFENTEKTTEFWKQYVDGDQC 90
SP|P16296|FA9_RAT    KILTRPKRYNSGKLE-EFVQGNLERECIEERCSFEEAREVFENTEKTTEFWKQYVDGDQC 90
SP|P16293|FA9_PIG    --------YNSGKLEESFVRGNLERECIEEKCSFEEAREVFENTEKTNEFWKQYVDGDQC 52
                              *:  :  .:  . ***** ** *.:***:*:*:: : *  ** .: *****

SP|P00740|FA9_HUMAN  ESNP--------CLNGGSCKDDINSYECWCPFGFEGKNCELD----VTCNIKNGRCEQFC 145
SP|P04070|PROC_HUMAN LVLPLEHPCASLCCGHGTCIDGIGSFSCDCRSGWEGRFCQRE-VSFLNCSLDNGGCTHYC 151
SP|P00742|FA10_HUMAN ETSP--------CQNQGKCKDGLGEYTCTCLEGFEGKNCELF--TRKLCSLDNGDCDQFC 140
SP|P08709|FA7_HUMAN  ASSP--------CQNGGSCKDQLQSYICFCLPAFEGRNCETHKDDQLICVNENGGCEQYC 162
SP|P00741|FA9_BOVIN  ESNP--------CLNGGMCKDDINSYECWCQAGFEGTNCELD----ATCSIKNGRCKQFC 145
SP|Q95ND7|FA9_PANTR  ESNP--------CLNGGSCKDDINSYECWCPFGFEGKNCELD----VTCNIKNGRCEQFC 145
SP|P16294|FA9_MOUSE  ESNP--------CLNGGICKDDISSYECWCQVGFEGRNCELD----ATCNIKNGRCKQFC 145
SP|P19540|FA9_CANLF  ESNP--------CLNDGVCKDDINSYECWCRAGFEGKNCELD----VTCNIKNGRCKQFC 138
SP|P16296|FA9_RAT    ESNP--------CLNGGICKDDINSYECWCQAGFEGRNCELD----ATCSIKNGRCKQFC 138
SP|P16293|FA9_PIG    EPNP--------CLNGGLCKDDINSYECWCQVGFEGKNCELD----ATCNIKNGRCKQFC 100
                        *        * . * * * : .: * *  .:**  *:        *  .** * ::*

SP|P00740|FA9_HUMAN  KNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTS-KLTRAETVFPD----VD 200
SP|P04070|PROC_HUMAN LEEVG-WRRCSCAPGYKLGDDLLQCHPAVKFPCGRPWKRMEKKRSHLK------------ 198
SP|P00742|FA10_HUMAN HEE-QNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSIT 199
SP|P08709|FA7_HUMAN  SDHTGTKRSCRCHEGYSLLADGVSCTPTVEYPCGKIPILEKRN----------------- 205
SP|P00741|FA9_BOVIN  KRDTDNKVVCSCTDGYRLAEDQKSCEPAVPFPCGRVSVSHISKKLTRAETIFSN----TN 201
SP|Q95ND7|FA9_PANTR  KNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTS-KLTRAETVFPD----VD 200
SP|P16294|FA9_MOUSE  KNSPDNKVICSCTEGYQLAEDQKSCEPTVPFPCGRASISYSSKKITRAETVFSN----MD 201
SP|P19540|FA9_CANLF  KLGPDNKVVCSCTTGYQLAEDQRSCEPAVPFPCGRVSVPHISMTRTRAETLFSN----MD 194
SP|P16296|FA9_RAT    KNSPDNKIICSCTEGYQLAEDQKSCEPAVPFPCGRVSVAYNSKKITRAETVFSN----TD 194
SP|P16293|FA9_PIG    KTGADSKVLCSCTTGYRLAPDQKSCKPAVPFPCGRVSVSHSPTTLTRAEIIFSN----MD 156
                              * *  ** *  :   * *:  :***:

SP|P00740|FA9_HUMAN  YVNSTEAET---------ILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVL-NGKVDA 250
SP|P04070|PROC_HUMAN ----------------------RDTEDQEDQVDPRLIDGKMTRRGDSPWQVVLLDSKKKL 236
SP|P00742|FA10_HUMAN WKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEG 259
SP|P08709|FA7_HUMAN  ----------------------------ASKPQGRIVGGKVCPKGECPWQVLLL-VNGAQ 236
SP|P00741|FA9_BOVIN  YENSSEAEI---------IWDNVTQSNQSFDEFSRVVGGEDAERGQFPWQVLL-HGEIAA 251
SP|Q95ND7|FA9_PANTR  YVNSTEAET---------ILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVL-NGKVDA 250
SP|P16294|FA9_MOUSE  YENSTEAVFIQDDITDGAILNNVTESSESLNDFTRVVGGENAKPGQIPWQVIL-NGEIEA 260
SP|P19540|FA9_CANLF  YENSTEVEK---------ILDNVTQ---PLNDFTRVVGGKDAKPGQFPWQVLL-NGKVDA 241
SP|P16296|FA9_RAT    YGNSTEL--ILDDITNSTILDNLTENSEPINDFTRVVGGENAKPGQIPWQVIL-NGEIEA 251
SP|P16293|FA9_PIG    YENSTEVEP---------ILDSLTESNQSSDDFIRIVGGENAKPGQFPWQVLL-NGKIDA 206
                                                   .   *::.*:    *: ***.:*   :

SP|P00740|FA9_HUMAN  FCGGSIVNEKWIVTAAHCVETGVK---ITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNA 307
SP|P04070|PROC_HUMAN ACGAVLIHPSWVLTAAHCMDESKK---LLVRLGEYDLRRWEKWELDLDIKEVFVHPNYS- 292
SP|P00742|FA10_HUMAN FCGGTILSEFYILTAAHCLYQAKR---FKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFT- 315
```

```
SP|P08709|FA7_HUMAN  LCGGTLINTIWVVSAAHCFDKIKNWRNLIAVLGEHDLSEHDGDEQSRRVAQVIIPSTYV- 295
SP|P00741|FA9_BOVIN  FCGGSIVNEKWVVTAAHCIKPGVK---ITVVAGEHNTEKPEPTEQKRNVIRAIPYHSYNA 308
SP|Q95ND7|FA9_PANTR  FCGGSIVNEKWIVTAAHCVDTGVK---ITVVAGEHNIEETEHTEQKRNVIRIIPHHNYNA 307
SP|P16294|FA9_MOUSE  FCGGAIINEKWIVTAAHCLKPGDK---IEVVAGEYNIDKKEDTEQRRNVIRTIPHHQYNA 317
SP|P19540|FA9_CANLF  FCGGSIINEKWVVTAAHCIEPDVK---ITIVAGEHNTEKREHTEQKRNVIRTILHHSYNA 298
SP|P16296|FA9_RAT    FCGGAIINEKWIVTAAHCLKPGDK---IEVVAGEHNIDEKEDTEQRRNVIRTIPHHQYNA 308
SP|P16293|FA9_PIG    FCGGSIINEKWVVTAAHCIEPGVK---ITVVAGEYNTEETEPTEQRRNVIRAIPHHSYNA 263
                      **. ::   ::::****.    .   :    *: :  . :  *    :   :    :

SP|P00740|FA9_HUMAN  AINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNI-FLKFG-SGYVSGWGRVFHKGR- 364
SP|P04070|PROC_HUMAN -KSTTDNDIALLHLAQPATLSQTIVPICLPDSGLAERELNQAGQETLVTGWGYHSSREKE 351
SP|P00742|FA10_HUMAN -KETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQK-TGIVSGFGRTHEKGR- 372
SP|P08709|FA7_HUMAN  -PGTTNHDIALLRLHQPVVLTDHVVPLCLPERTFSERTLAFVR-FSLVSGWGQLLDRGA- 352
SP|P00741|FA9_BOVIN  SINKYSHDIALLELDEPLELNSYVTPICIADRDYTNI-FLKFG-YGYVSGWGKVFNRGR- 365
SP|Q95ND7|FA9_PANTR  AINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNI-FLKFG-SGYVSGWGRVFHKGR- 364
SP|P16294|FA9_MOUSE  TINKYSHDIALLELDKPLILNSYVTPICVANREYTNI-FLKFG-SGYVSGWGKVFNKGR- 374
SP|P19540|FA9_CANLF  TINKYNHDIALLELDEPLTLNSYVTPICIADREYSNI-FLKFG-SGYVSGWGRVFNKGR- 355
SP|P16296|FA9_RAT    TINKYSHDIALLELDKPLILNSYVTPICVANKEYTNI-FLKFG-SGYVSGWGKVFNKGR- 365
SP|P16293|FA9_PIG    TVNKYSHDIALLELDEPLTLNSYVTPICIADKEYTNI-FLKFG-SGYVSGWGRVFNRGR- 320
                        . . ***:*.*  *  :   :.* *: :   ::  :        *:*:*    :

SP|P00740|FA9_HUMAN  ----SALVLQYLRVPLVDRATCLRST-----KFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415
SP|P04070|PROC_HUMAN AKRNRTFVLNFIKIPVVPHNECSEVM-----SNMVSENMLCAGILGDRQDACEGDSGGPM 406
SP|P00742|FA10_HUMAN ----QSTRLKMLEVPYVDRNSCKLSS-----SFIITQNMFCAGYDTKQEDACQGDSGGPH 423
SP|P08709|FA7_HUMAN  ----TALELMVLNVPRLMTQDCLQQSRKVGDSPNITEYMFCAGYSDGSKDSCKGDSGGPH 408
SP|P00741|FA9_BOVIN  ----SASILQYLKVPLVDRATCLRST-----KFSIYSHMFCAGYHEGGKDSCQGDSGGPH 416
SP|Q95ND7|FA9_PANTR  ----SALVLQYLRVPLVDRATCLRST-----KFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415
SP|P16294|FA9_MOUSE  ----QASILQYLRVPLVDRATCLRST-----TFTIYNNMFCAGYREGGKDSCEGDSGGPH 425
SP|P19540|FA9_CANLF  ----SASILQYLKVPLVDRATCLRST-----KFTIYNNMFCAGFHEGGKDSCQGDSGGPH 406
SP|P16296|FA9_RAT    ----QASILQYLRVPLVDRATCLRST-----KFSIYNNMFCAGYREGGKDSCEGDSGGPH 416
SP|P16293|FA9_PIG    ----SATILQYLKVPLVDRATCLRST-----KVTIYSNMFCAGFHEGGKDSCLGDSGGPH 371
                          :  *  :.:* :    *         .  : . *:***     .*:* ******

SP|P00740|FA9_HUMAN  VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 461
SP|P04070|PROC_HUMAN VASFHGTWFLVGLVSWGEGCGLLHNYGVYTKVSRYLDWIHGHIRDKEAPQKSW-AP---- 461
SP|P00742|FA10_HUMAN VTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVIT 483
SP|P08709|FA7_HUMAN  ATHYRGTWYLTGIVSWGQGCATVGHFGVYTRVSQYIEWLQKLMRSEPRPGVLLRAPFP-- 466
SP|P00741|FA9_BOVIN  VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 462
SP|Q95ND7|FA9_PANTR  VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 461
SP|P16294|FA9_MOUSE  VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 471
SP|P19540|FA9_CANLF  VTEVEGISFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 452
SP|P16296|FA9_RAT    VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTKLT-------------- 462
SP|P16293|FA9_PIG    VTEVEGTSFLTGIISWGEECAVKGKYGIYTKVSRYVNW---------------------- 409
                     .:  ..  ::.*::***: *.   ::*:**:*: ::.*

SP|P00740|FA9_HUMAN  -----
SP|P04070|PROC_HUMAN -----
SP|P00742|FA10_HUMAN SSPLK 488
SP|P08709|FA7_HUMAN  -----
SP|P00741|FA9_BOVIN  -----
SP|Q95ND7|FA9_PANTR  -----
SP|P16294|FA9_MOUSE  -----
SP|P19540|FA9_CANLF  -----
SP|P16296|FA9_RAT    -----
SP|P16293|FA9_PIG    -----
```