

## Supporting Information;

Site-Specific Intact N-Linked Glycopeptide Characterization of Prostate-Specific Membrane Antigen from Metastatic Prostate Cancer Cells

Stephen Mackay<sup>†,‡,§</sup>, Naomi L. Hitefield<sup>†,‡,||</sup>, Ian O. Oduor<sup>†</sup>, Autumn B. Roberts<sup>†,‡</sup>, Tanya C. Burch<sup>†,‡</sup>, Raymond S. Lance<sup>†,⊥</sup>, Tina D. Cunningham<sup>#</sup>, Dean A. Troyer<sup>†,‡</sup>, Oliver J. Semmes<sup>†,‡</sup>, Julius O. Nyalwidhe<sup>†,‡,\*</sup>

<sup>†</sup>Leroy T. Canoles Jr. Cancer Research Center, Eastern Virginia Medical School, Norfolk, Virginia 23507, United States

<sup>‡</sup>Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia 23507, United States

<sup>§</sup>University of North Carolina, Chapel Hill, North Carolina 27516, United States

<sup>||</sup>University of Georgia, Athens, Georgia 30602, United States

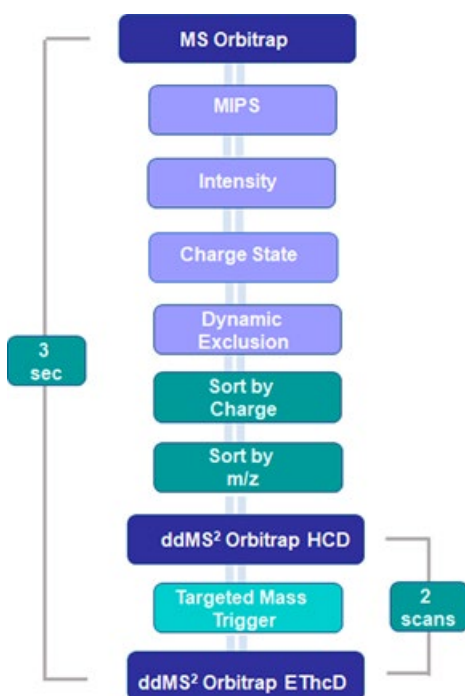
<sup>⊥</sup>Spokane Urology, Spokane, Washington 99202, United States

<sup>#</sup>School of Health Professions, Eastern Virginia Medical School, Norfolk, Virginia 23507, United States

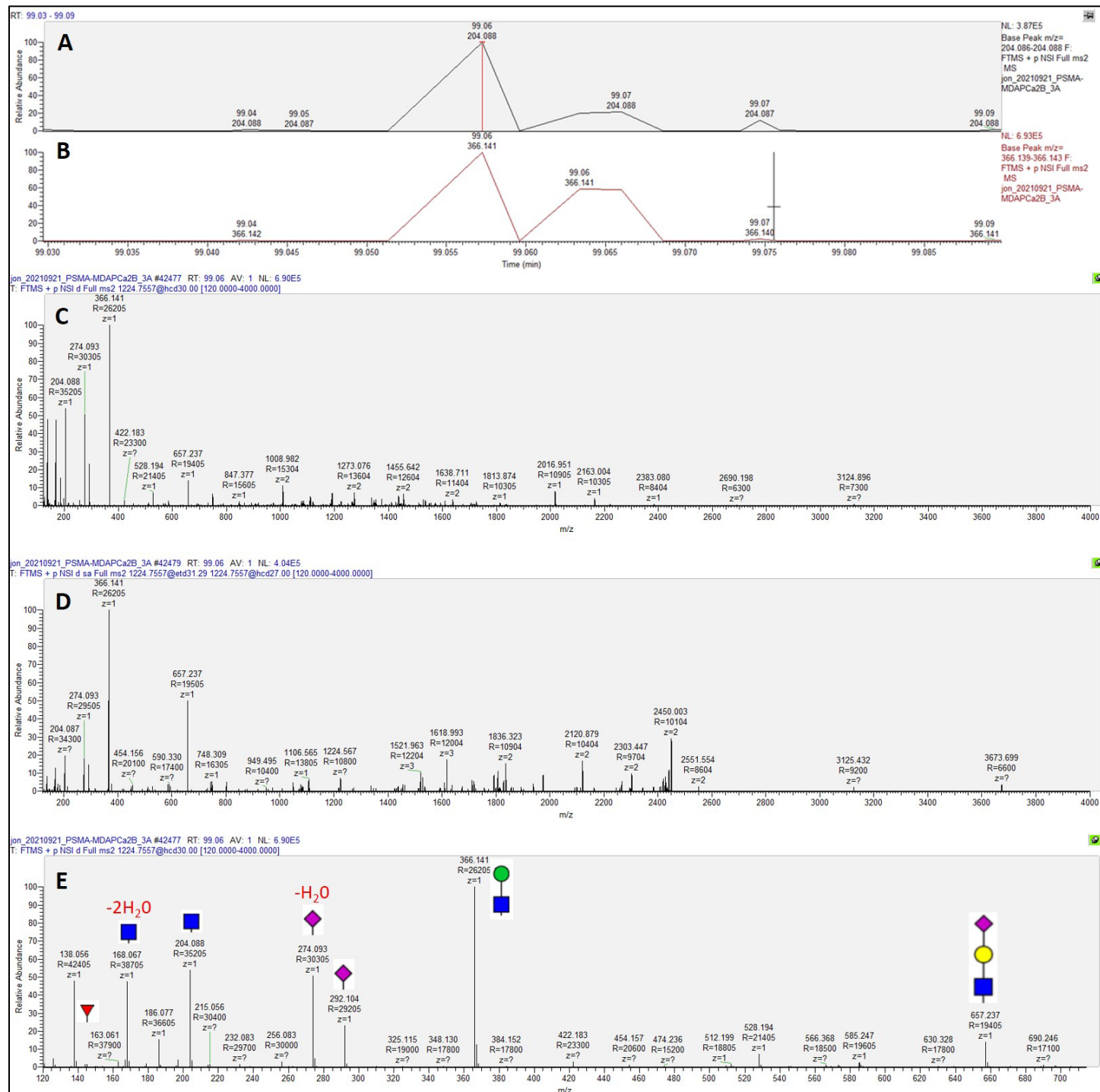
\*Email: [nyalwijo@evms.edu](mailto:nyalwijo@evms.edu). Phone: 757-446-5682

## SUPPLEMENTARY FIGURES

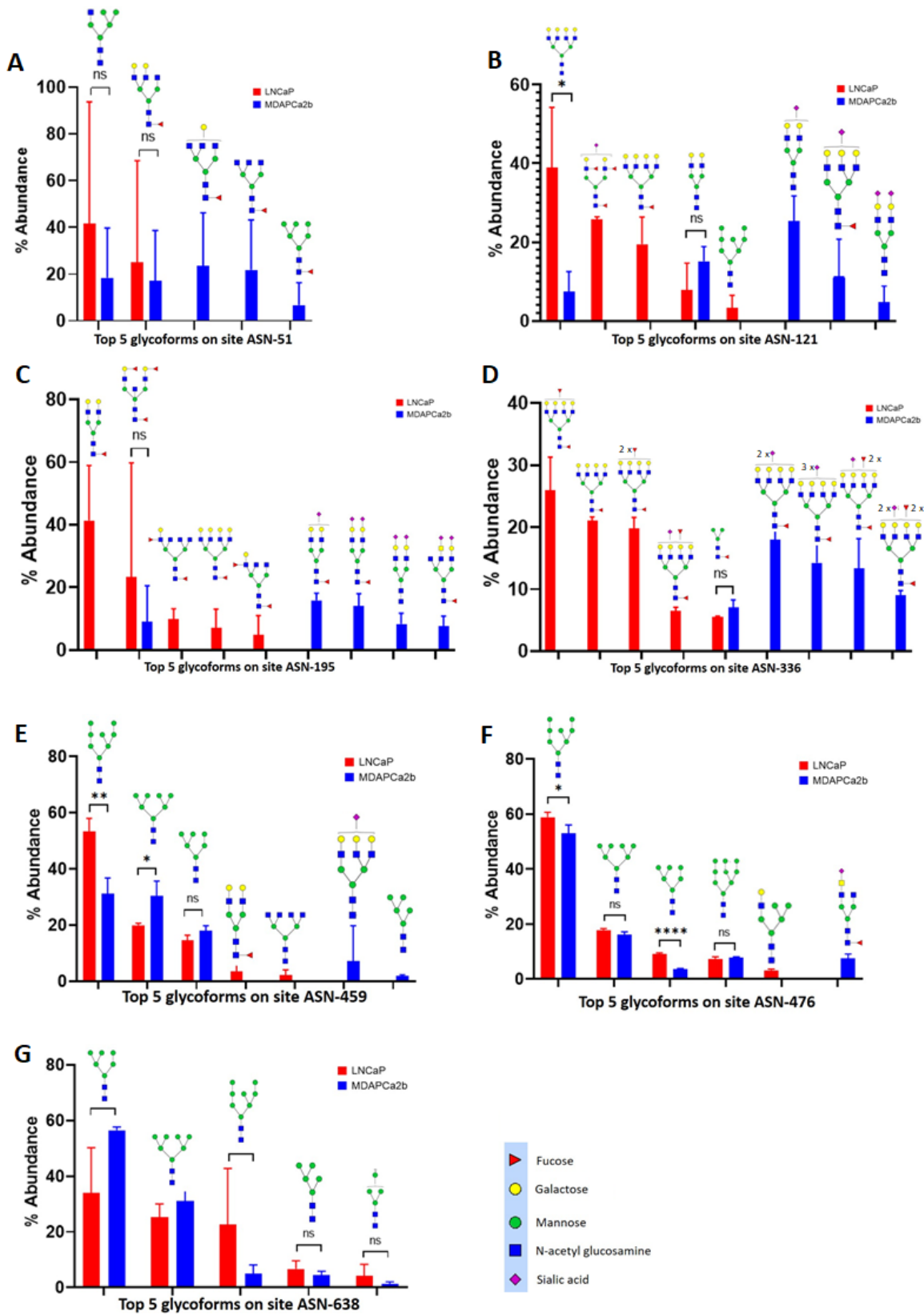
**Figure S1.** Instrument methods and scan parameters that were used are summarized in the schematic below. To eliminate potential sources of experimental bias, immunoprecipitated PSMA from different cell line cohorts were processed for LC/MS/MS at the same time for sample preparation consistency. The samples were run in a block randomized manner with two blank injections of 0.1% formic acid (buffer A) to minimize carry-over between samples. The LC/MS/MS runs were performed immediately after implementation of standard instrument maintenance and QC protocols to ensure optimal performance of the instruments in September 2021. A full-scan MS spectrum acquired in the Orbitrap at 60,000, resolution and scanned at  $m/z$  375-2000 under standard automatic gain control and a maximum injection time of 50 ms. Monoisotopic peak determination was performed for peptides to include charge states 2-7. Undetermined charge states were also included in the analyses. ddMS<sup>2</sup> scans were isolated in the quadrupole with an isolation window of 1.6  $m/z$  and an HCD activation energy of 30%. Ions were detected in the Orbitrap at a resolution of 30,000 and a scan range of  $m/z$  of 120-4000 with standard automatic gain control and maximum injection time of 75 ms. Oxonium ion based mass list consisting of HexNAc 204.0867, HexNAc-fragment 138.0545, HexNAcHex 366.1396, Hex-2H<sub>2</sub>O 127.0389, Hex 163.0601, HexNAc-2H<sub>2</sub>O 168.0655, HexNAc-H<sub>2</sub>O 186.076, NeuAc-H<sub>2</sub>O 274.09, NeuAc 292.1024 and HexNAcHexNeuAc ( $m/z$  657.23) were used to trigger ETD scans. EThcD scans were acquired in the Orbitrap with supplemental collision energy of 27%, resolution of 30000, a scan range of 120-4000  $m/z$  and maximum injection time of 200 ms. All the spectral data were acquired in the profile mode.



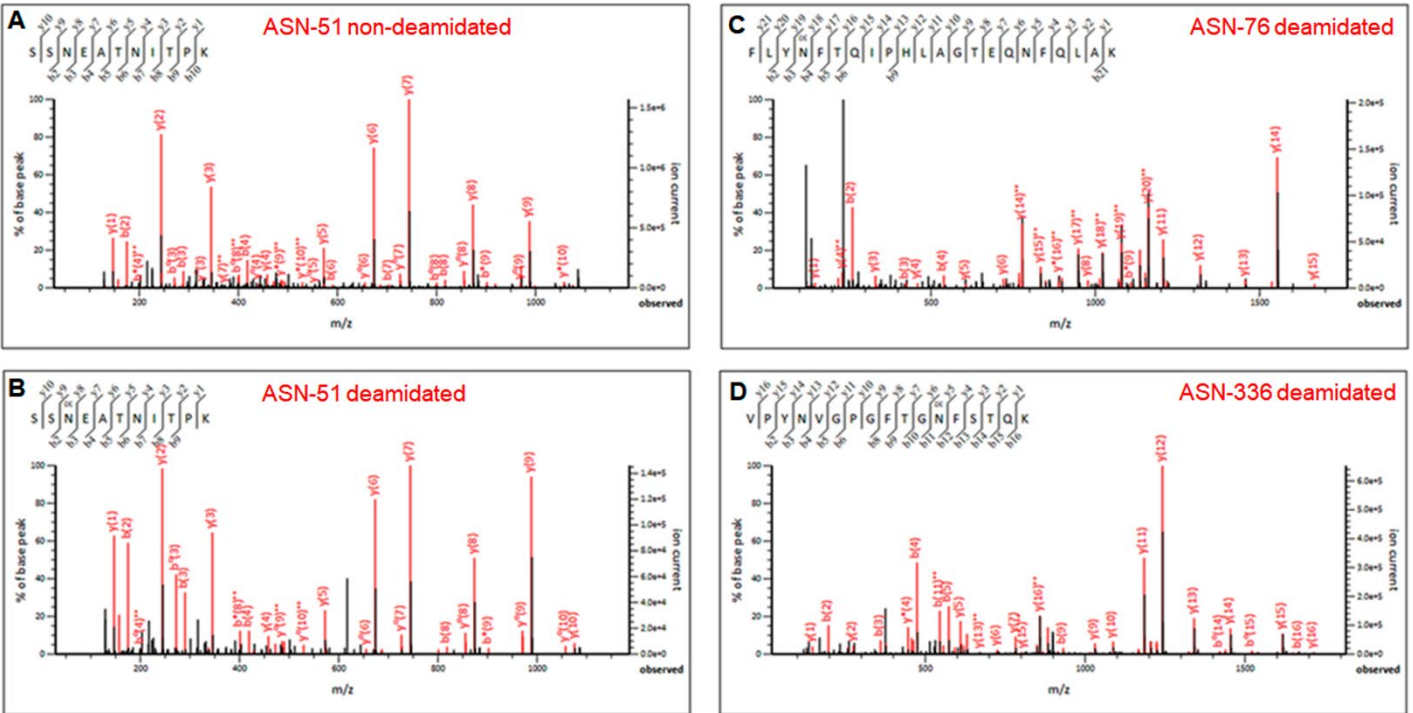
**Figure S2.** Representative extracted ion chromatograms of tandem mass spectra of the HexNAc oxonium ion (HexNAc<sup>+</sup>) with an *m/z* range 204.086-204.088 (**A**) and HexHexNAc with the *m/z* range 366.139-366.143 (**B**) and retention time 99.03 and 99.09 minutes corresponding to the ASN-336 glycopeptide with sequence VPYNVGGFTGNFSTQK conjugated with the glycan FA4G4S2 (*m/z* 1224.7557 4+). (**C**) Annotated HCD MS/MS spectrum and (**D**) EThcD MS/MS spectra contain oxonium ions from the constituent carbohydrates with [HexNAc]<sup>+</sup> 204.087, [Neu5Ac]<sup>+</sup> 292.1027, [Neu5Ac-H<sub>2</sub>O]<sup>+</sup> 274.093, [HexHexNAc]<sup>+</sup> 366.1395, [HexNAcHex Neu5Ac]<sup>+</sup> 657.237 and [dHex]<sup>+</sup>147.0652. (**E**). Annotated spectra showing some of the oxonium ions.



**Figure S3.** Glycans in seven N-linked glycosylation sites in PSMA. Site-specific comparative quantitative glycan analysis for the relative abundance glycan abundances of the five most abundant glycopeptides on seven PSMA in the ASN-51, ASN-195, ASN-121, ASN-336, ASN-459, ASN-476 and ASN-638 glycosylation sites in LNCaP (red bars) and MDAPCa2b (blue bars), respectively. The glycan abundance between the cell lines were compared using multiple unpaired t-tests. The error bars show standard deviation across triplicate analyses from each cell line. **(A)** At the ASN-51 glycosite, M5G0N1 and FA3G2 were detected in both cell lines and there are no significant differences in relative abundance between of the glycopeptides between them as determined by multiple unpaired t-tests. Three other glycopeptides conjugated with FA3G1, FA3 and FM6 are detected in MDAPCa2b cells. **(B)** At the ASN-121 glycosite, two glycans A3G3 and A2G2 ) are detected in both cells. There is a significant difference in the abundance of A3G3 between them (\* P=0.027817) but not in the abundance of A2G2. Two other glycopeptides conjugated with A2G2S1 and FA4G4 are detected in LNCaP cells but not in MDAPCa2b, whereas FA2G2, A3G3S1 and A2G2S2 are detected in MDAPCa2b but not in LNCaP. **(C)** At the ASN-195 glycosite, one glycan, FA3BG2F2 is detected in both cells without significant expression differences. Four glycans, FA2G2, FA4G1F1, FA4G3, and FA3G1F1 are detected in LNCaP cells but not in MDAPCa2b cells and four glycans; A2G2S1, A2G2S2, A2G2S2 and A3G3S2, are detected in MDAPCa2b but not in LNCaP cells. **(D)** At the ASN-336 glycosite, one glycan, FA is detected in both cells without significant expression differences. Four glycans, FA4G4F, FA4G4, FA2G4F2, and FA4G4F1S1 are detected in LNCaP cells but not in MDAPCa2b cells. Four glycans, FA4G4F1S1, FA4G4S3 , FA4G4F2S1 and FA4G4F2S2 are identified in MDAPCa2b but not LNCaP cells. **(E)** At the ASN-459 glycosite, three glycans Man8, Man7 and Man6 are detected in both cells. There is a significant difference in the abundance of Man8 (\* P= 0.006154) and Man7 (\* P= 0.025249), but not in the abundance of Man6 between the two cells. Two glycans, FA3G1 and A4 are detected in LNCaP but not in MDAPCa2b. Two glycans, A3G3S2 and Man6 detected in MDAPCa2b but not LNCaP. **(F)** At the ASN-476 glycosite, four glycans Man8, Man7, A2G2 and Man9, are detected in both cells. There is a significant difference in the abundance of Man8 (\* P= 0.046446) and Man6 (\*\*\*\* P= 0.000022), but not in the abundance of Man7 and Man9. One glycan, M5G0N1 is detected in LNCaP but not in MDAPCa2b cells. One glycan FA2G1S1 is detected in MDAPCa2b but not LNCaP. **(G)** At the ASN-638 glycosite, five glycans Man6, Man7, Man8, Man5 and Man4, are detected in both cells and there are no significant differences in the abundances.



**Figure S4.** MS/MS spectra for trypsin derived and PNGase-F treated PSMA peptides corresponding to three N-linked glycosylation sites. **(A)** The lack of deamidation upon PNGase-F treatment in ASN-51 as shown by annotated y9 and b3 fragments demonstrates lack of N-linked glycosylation at the site. **(B)** The presence of deamidation of ASN-51 upon PNGase-F treatment as shown by annotated y9 and b3 fragments se presence. **(C)** In the case of ASN-76 only the deamidated peptide species were identified with y19 and b4 ions which confirms that this residue is fully N-linked glycosylated. **(D)** In the case of ASN-336 only the deamidated peptide species were identified with y6 and b12 ions which confirms that this residue is fully N-linked glycosylated.



**Figure S5.** Tandem mass spectra for ASN-76 glycopeptides with high mannose structures including Man7, Man6, and Man5 respectively from LNCaP and MDAPCa2b cells. MS/MS spectra of an intact PSMA N-glycopeptides with sequence FLYN\*FTQIPHLAGTEQNFQLAK and conjugated with high mannose glycans including **(A)**. Tandem mass spectrum identifying a Man7 glycan conjugated to ASN-76. **(B)** Tandem mass spectrum identifying a Man6 glycan conjugated to ASN-76. **(C)** Tandem mass spectrum identifying a Man5 glycan conjugated to ASN-76. Both b and y ions peptide fragment ions and oxonium ions, that include HexNAcHex (366.1396) and the intact peptide with the sugar moiety with HexNAcHex sugar moieties are identified. The *m/z* error plots demonstrate the accuracy of the assignments for both the peptide and oxonium ions with most of the ions matching exactly between the expected and observed mass values.

**Figure S5A**

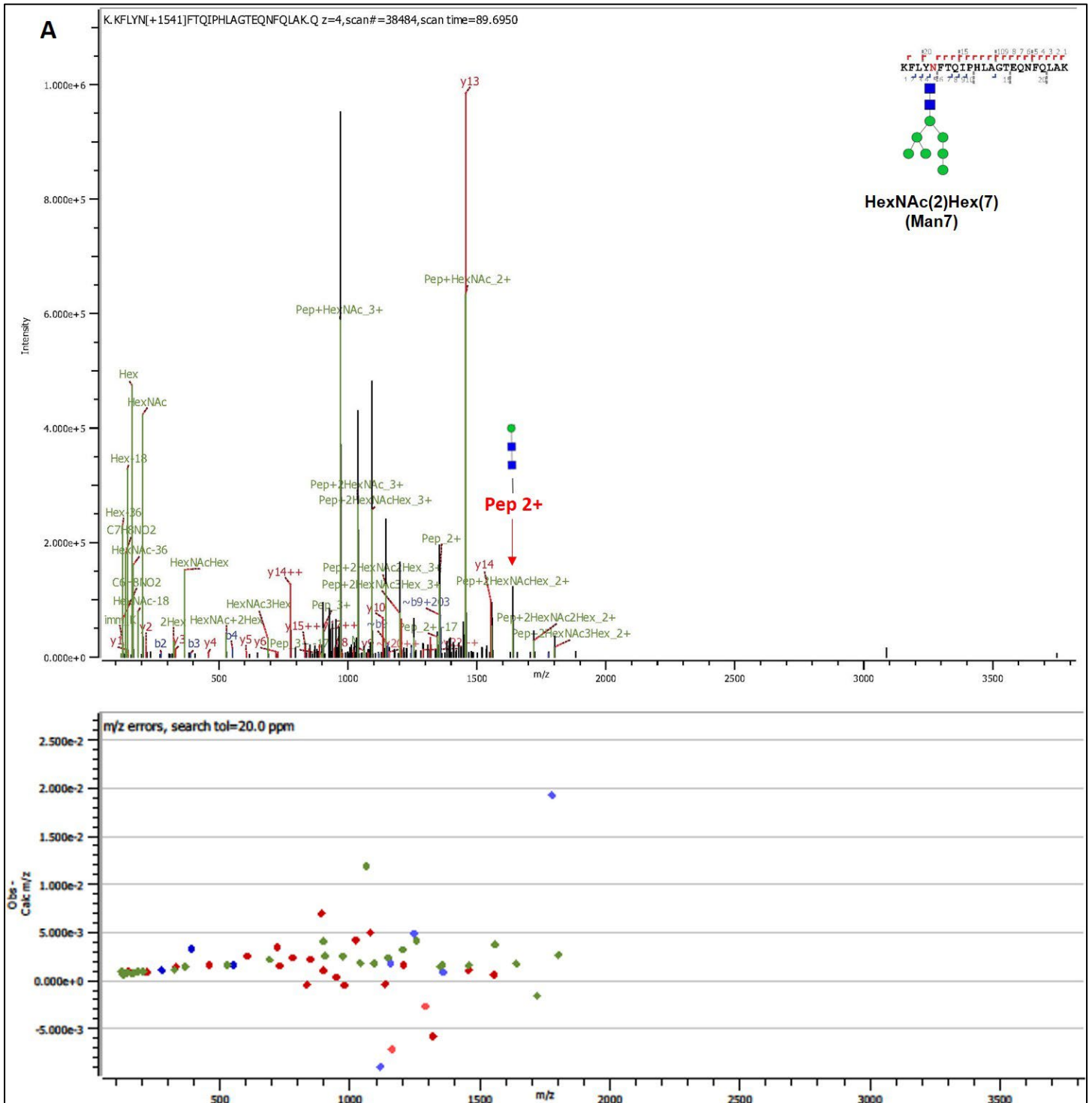
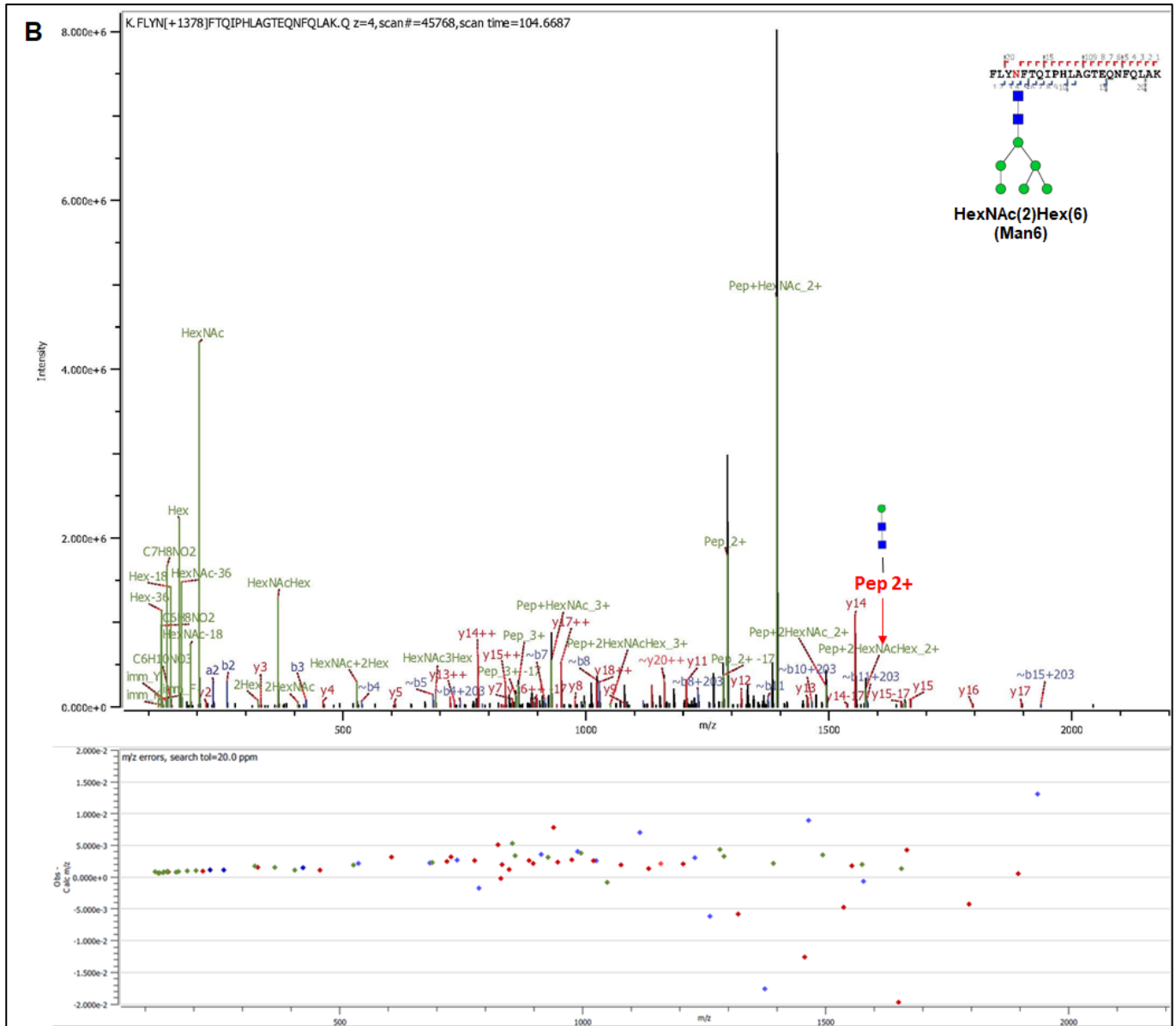


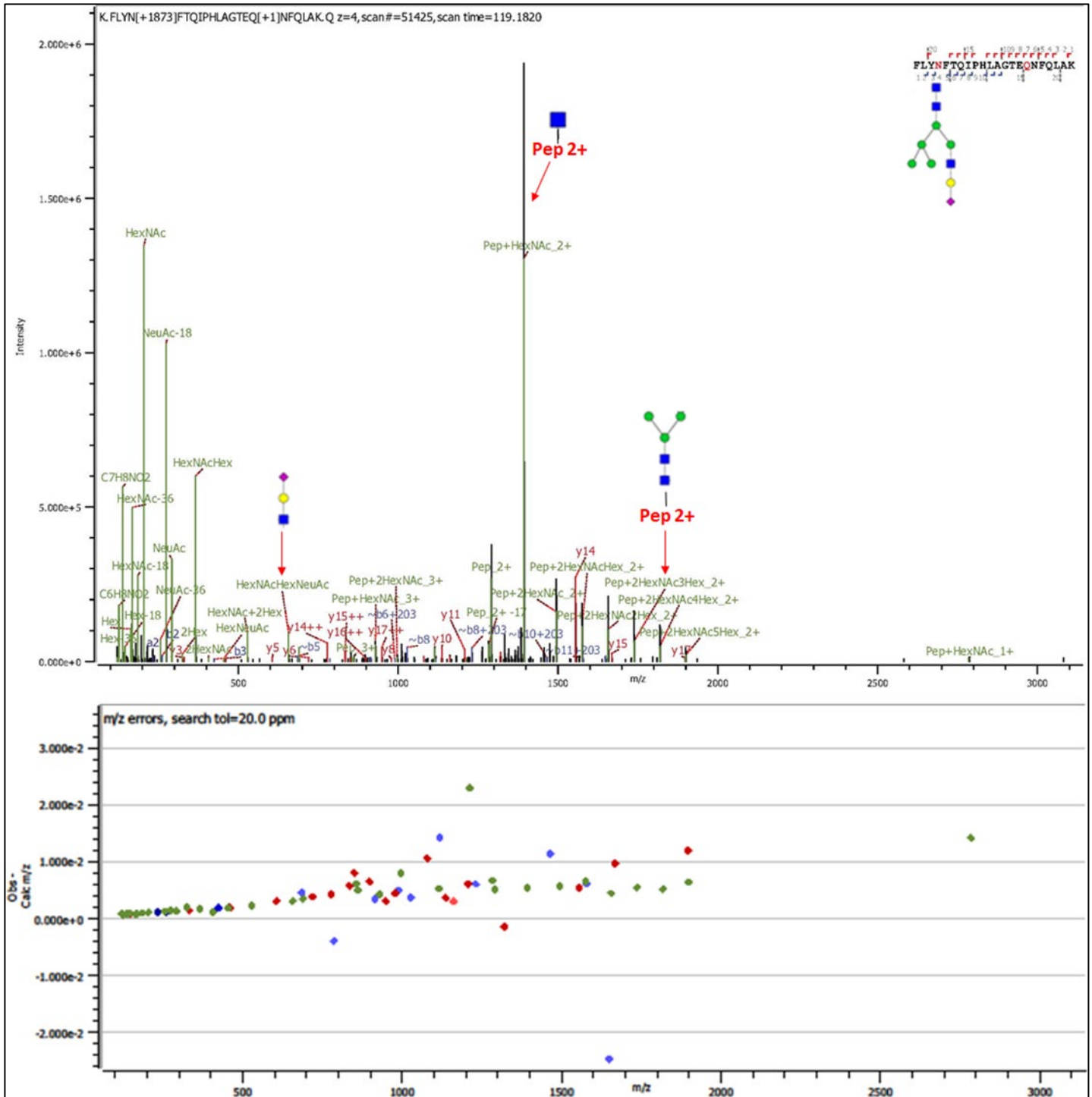
Figure S5B



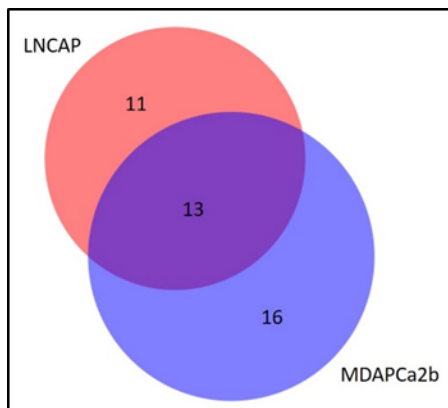




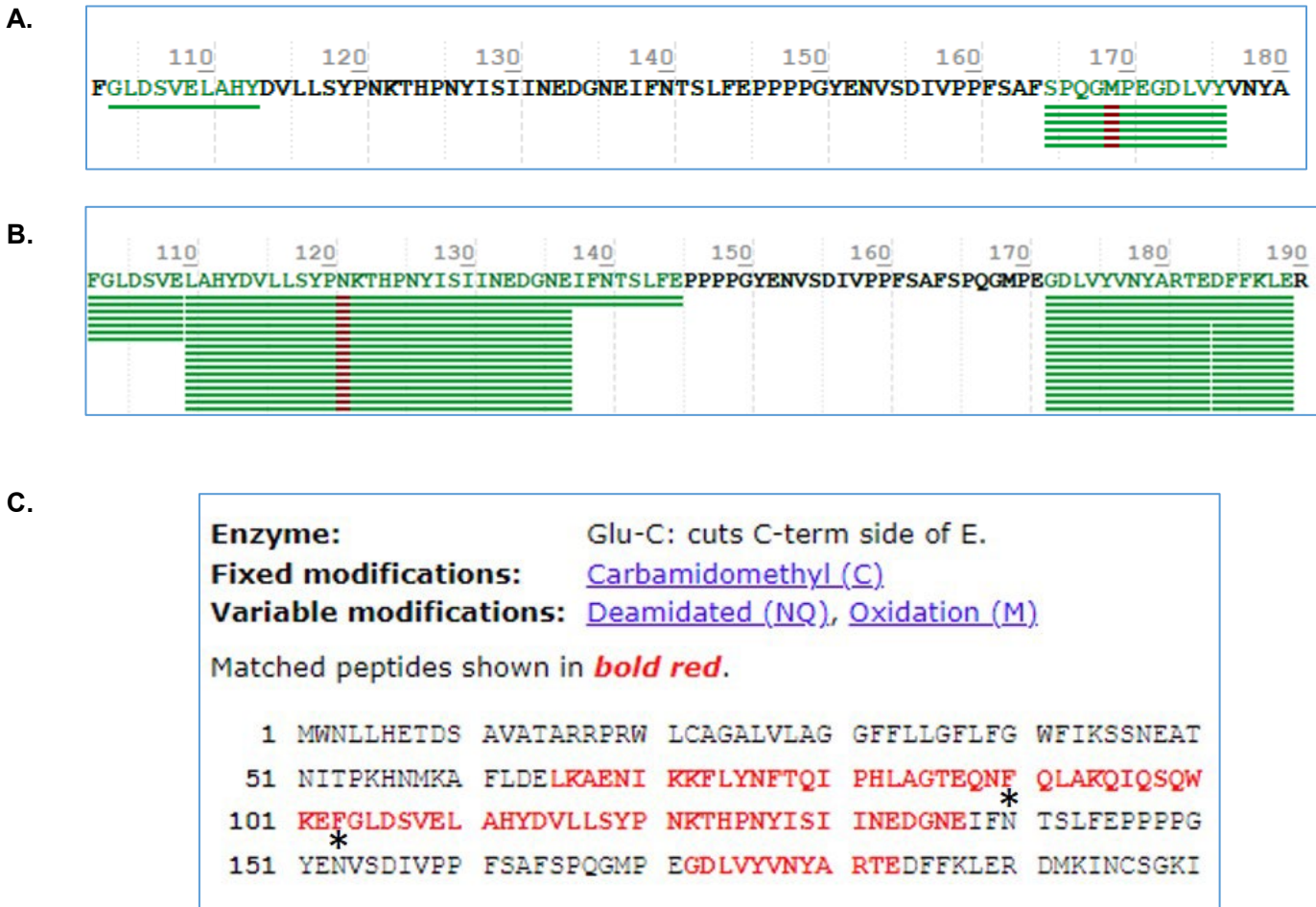
**Figure S6.** MS/MS spectra for an ASN-76 sialylated glycopeptide. MS/MS spectra of an intact PSMA N-glycopeptide with sequence FLYN\*FTQIPHLAGTEQNFQLAK and conjugated with triantennary complex glycans with a terminal sialic acid residue M5G1S1. Both b and y ions peptide fragment ions and oxonium ions, that include HexNAcHex (366.1396), neutral loss sialic acid and the terminal sialic acid containing oxonium ion HexNAcHexNeuAc are identified. The intact peptide with the sugar moieties including HexNAc2Hex3  $m/z$  1736.8245 (+2) are identified. The  $m/z$  error plots demonstrated the accuracy of the assignments for both the peptide and oxonium ions with most of the ions matching exactly between the expected and observed mass values.



**Figure S7.** Venn diagram indicating the number of top five most abundant glycans detected by in the LNCaP and MDAPCa2b cell lines. Out of the 40 glycans, 11 are exclusively identified in LNCaP cells and 16 are exclusively identified in MDAPCa2b cells. Thirteen glycans are identified in both cell lines.



**Figure S8.** Non-identification of ASN-140 and ASN-153 containing peptides after chymotrypsin and Glu-C digestion. Sequence coverage of peptides identified after digestion of immunoprecipitated PSMA with chymotrypsin panel **A** and GluC panels **B** and **C**. The expected peptides the NTSLF and ENVSDIVPPF for chymotrypsin and IFNTSLFE and PPPPGYENVSDIVPPFSAFSPQGMPE for GluC are not detected using intact glycopeptides with and without HILIC purification and after PNGase-F cleavage.








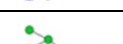








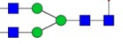






## SUPPLEMENTARY TABLES

**Table S1.** The origins and malignant characteristics of four prostate cancer cell lines that were used in the study are summarized below. Western blot analysis was performed to determine the expression levels of PSMA using all the four cell lines. PSMA glycosylation profiles were determined in two cell lines, LNCaP and MDAPCa2b.

<b>Cell Line</b>	<b>Origin</b>	<b>Characteristics</b>
RWPE-1	Immortalized (HPV) prostate epithelial cell line from a normal 54 year old Caucasian male	Non-malignant, non-metastatic, does not form tumors, androgen sensitive, expresses PSA
LNCaP	Adenocarcinoma supraclavicular metastatic lymph node tumor from a 50 year old Caucasian male	Low metastatic potential develops tumors in mice, hormone sensitive, expresses PSA
PC3-ML2	PC3 subline. PC3 is an Adenocarcinoma bone metastatic tumor from a 62 year old male Caucasian	More aggressive compared to PC3-N2, hormone refractory, form metastases when injected in SCID mice
MDAPCa2b	Adenocarcinoma metastatic bone tumor from a 63 year old African American male	Adenocarcinoma metastatic bone tumor, hormone refractory

**Table S2.** Summary of the top five most abundant glycans identified in eight N-linked glycosylation sites in PSMA glycopeptides from LNCaP and MDAPCa2b cells. The Venn diagram below shows the distribution of the identified glycans in the two cell lines.

#	Glycan Composition	Oxford Nomenclature	CFG Structure
1	HexNAc(2)Hex(4)	M4	
2	HexNAc(2)Hex(5)	M5	
3	HexNAc(2)Hex(6)	M6	
4	HexNAc(2)Hex(7)	M7	
5	HexNAc(2)Hex(8)	M8	
6	HexNAc(2)Hex(9)	M9	
7	HexNAc(3)Hex(5)	M5G0N1	
8	HexNAc(3)Hex(6)	M5G1	
9	HexNAc(4)Hex(5)	A2G2	
10	HexNAc(5)Hex(6)	A3G3	
11	HexNAc(6)Hex(3)	A4	
12	HexNAc(2)Hex(3)Fuc(1)	FA	
13	HexNAc(2)Hex(6)Fuc(1)	M6F	
14	HexNAc(3)Hex(5)Fuc(1)	FM5G1	
15	HexNAc(4)Hex(5)Fuc(1)	FA2G2	
16	HexNAc(5)Hex(3)Fuc(1)	FA3	
17	HexNAc(5)Hex(4)Fuc(1)	FA3G1	
18	HexNAc(5)Hex(5)Fuc(1)	FA3G2	
19	HexNAc(5)Hex(4)Fuc(2)	FA3G1F1	
20	HexNAc(5)Hex(5)Fuc(3)	FA3BG2F2	
21	HexNAc(5)Hex(6)Fuc(1)	FA4G3	

22	HexNAc(6)Hex(7)Fuc(1)	FA4G4	
23	HexNAc(6)Hex(7)Fuc(2)	FA4G4F	
24	HexNAc(6)Hex(7)Fuc(3)	FA2G4F2	
25	HexNAc(6)Hex(4)Fuc(2)	FA4G1F1	
26	HexNAc(3)Hex(6)NeuAc(1)	M5G1S1	
27	HexNAc(4)Hex(5)NeuAc(1)	A2G2S1	
28	HexNAc(4)Hex(5)NeuAc(2)	A2G2S2	
29	HexNAc(5)Hex(6)NeuAc(1)	A3G3S1	
30	HexNAc(5)Hex(6)NeuAc(2)	A3G3S2	
31	HexNAc(3)Hex(5)Fuc(1)NeuAc(1)	FA2G1S1	
32	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	FA2G2S1	
33	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	FA2G2S2	
34	HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	FA2G2F2S1	
35	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	FA3G3S2	
36	HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	FA4G4S2	
37	HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	FA4G4S3	
38	HexNAc(6)Hex(7)Fuc(2)NeuAc(1)	FA4G4F1S1	
39	HexNAc(6)Hex(7)Fuc(3)NeuAc(1)	FA4G4F2S1	
40	HexNAc(6)Hex(7)Fuc(3)NeuAc(2)	FA4G4F2S2	