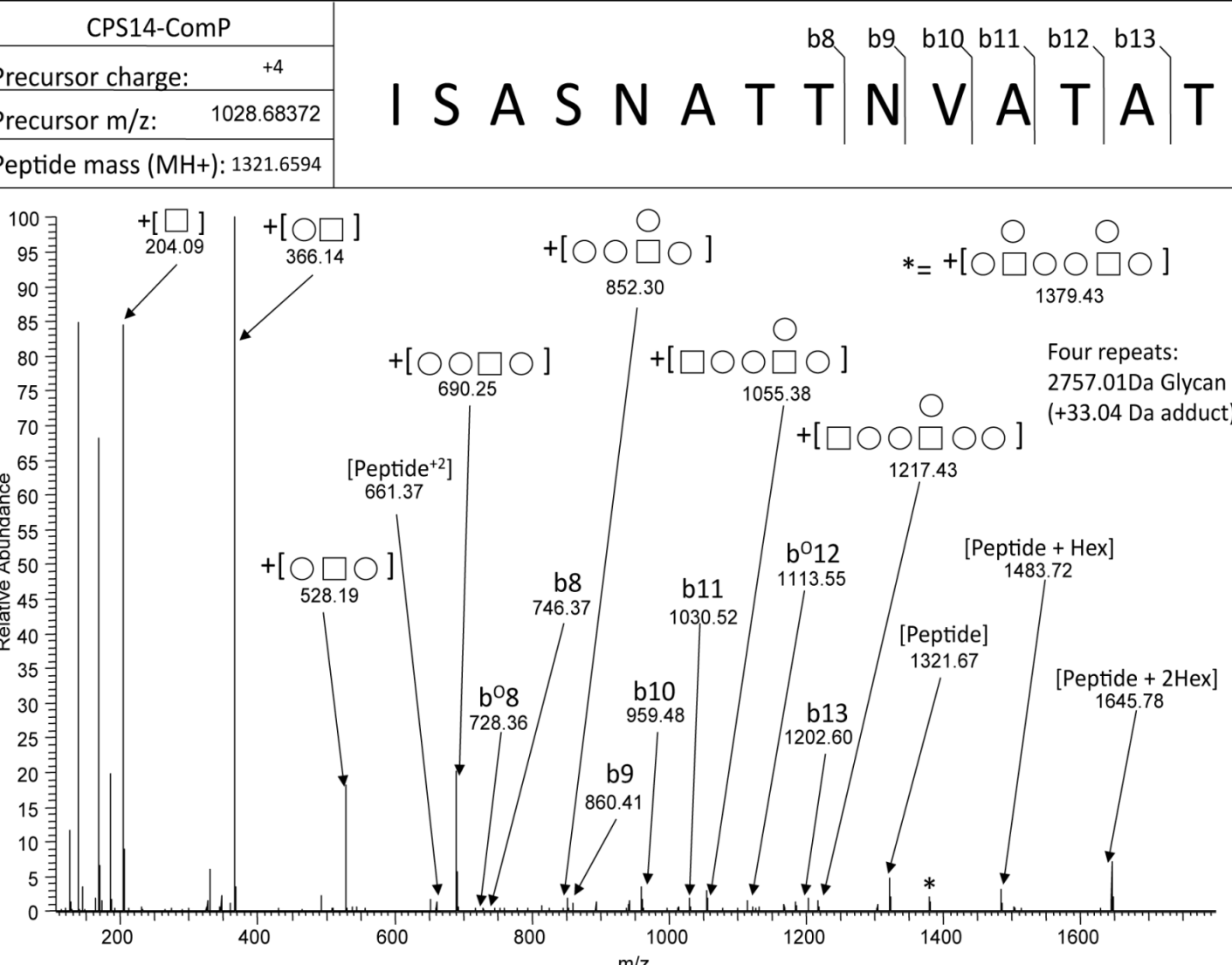


## **Supplemental Files**

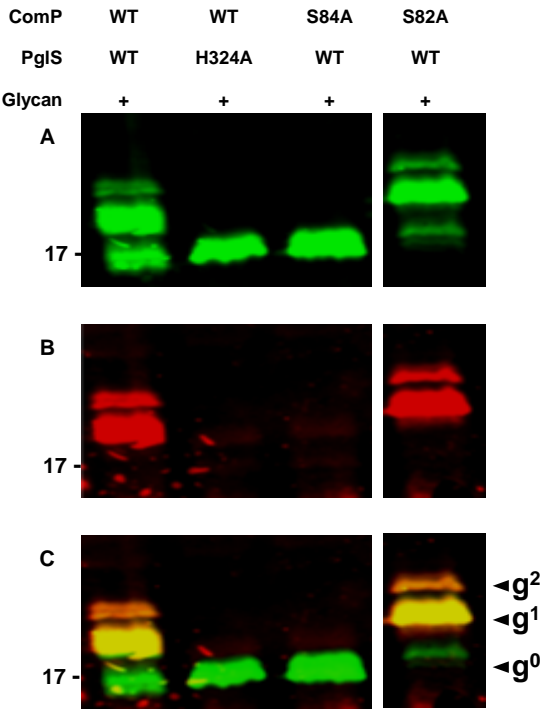
**A platform for glycoengineering a polyvalent pneumococcal bioconjugate vaccine using *E. coli* as a host**

Harding et al.



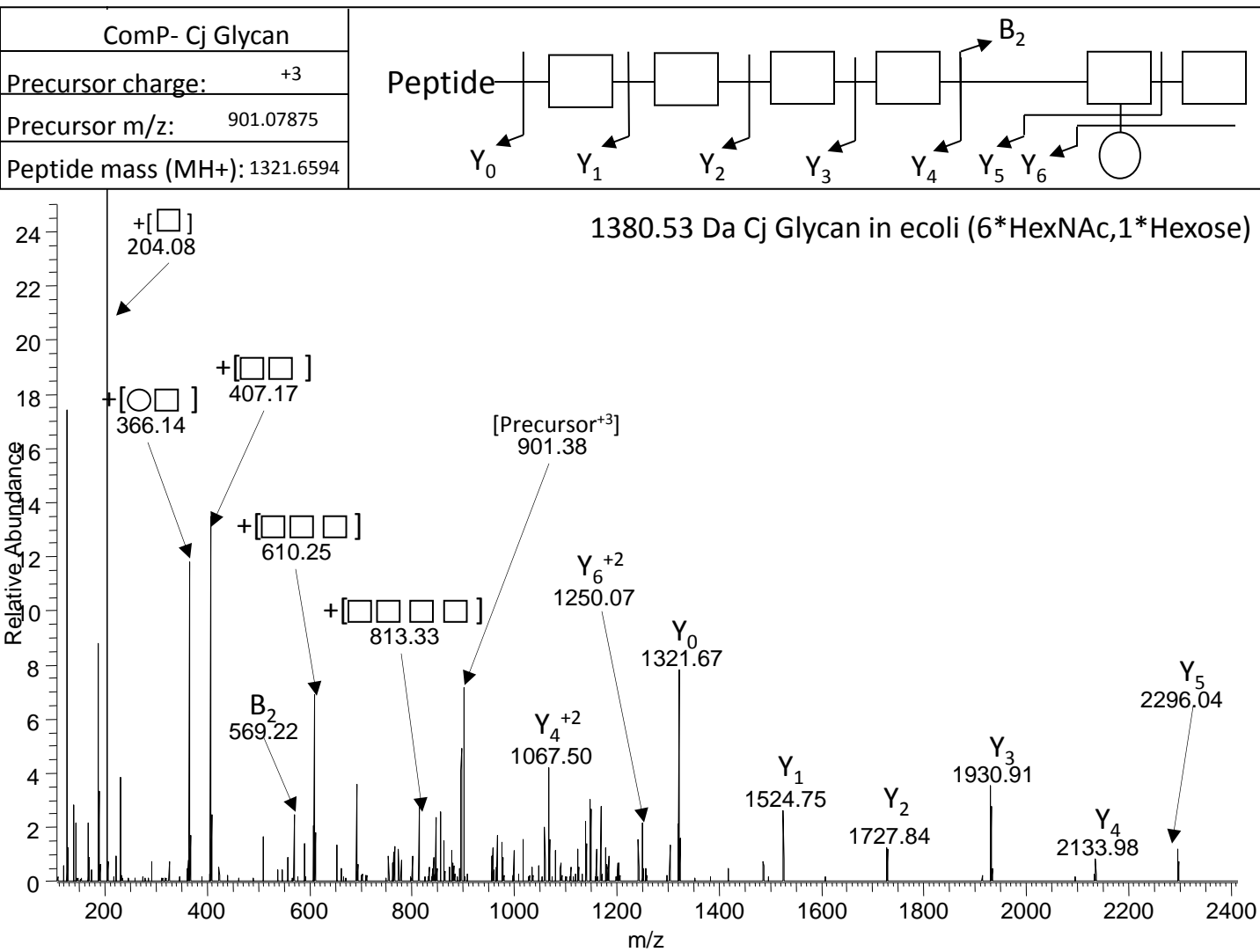
**Higher energy collisional dissociation (HCD) fragmentation spectra of GluC digested CPS14-ComP bioconjugates.** GluC digested CPS14-ComP was subjected to HCD fragmentation enabling the confirmation of a semi-GluC derived single peptide attached to a glycan with the CPS14 repeating subunit. Additional glycopeptides were also observed decorated with extended glycans corresponding to up to four tetrasaccharide repeat units.

Supplementary Figure 2



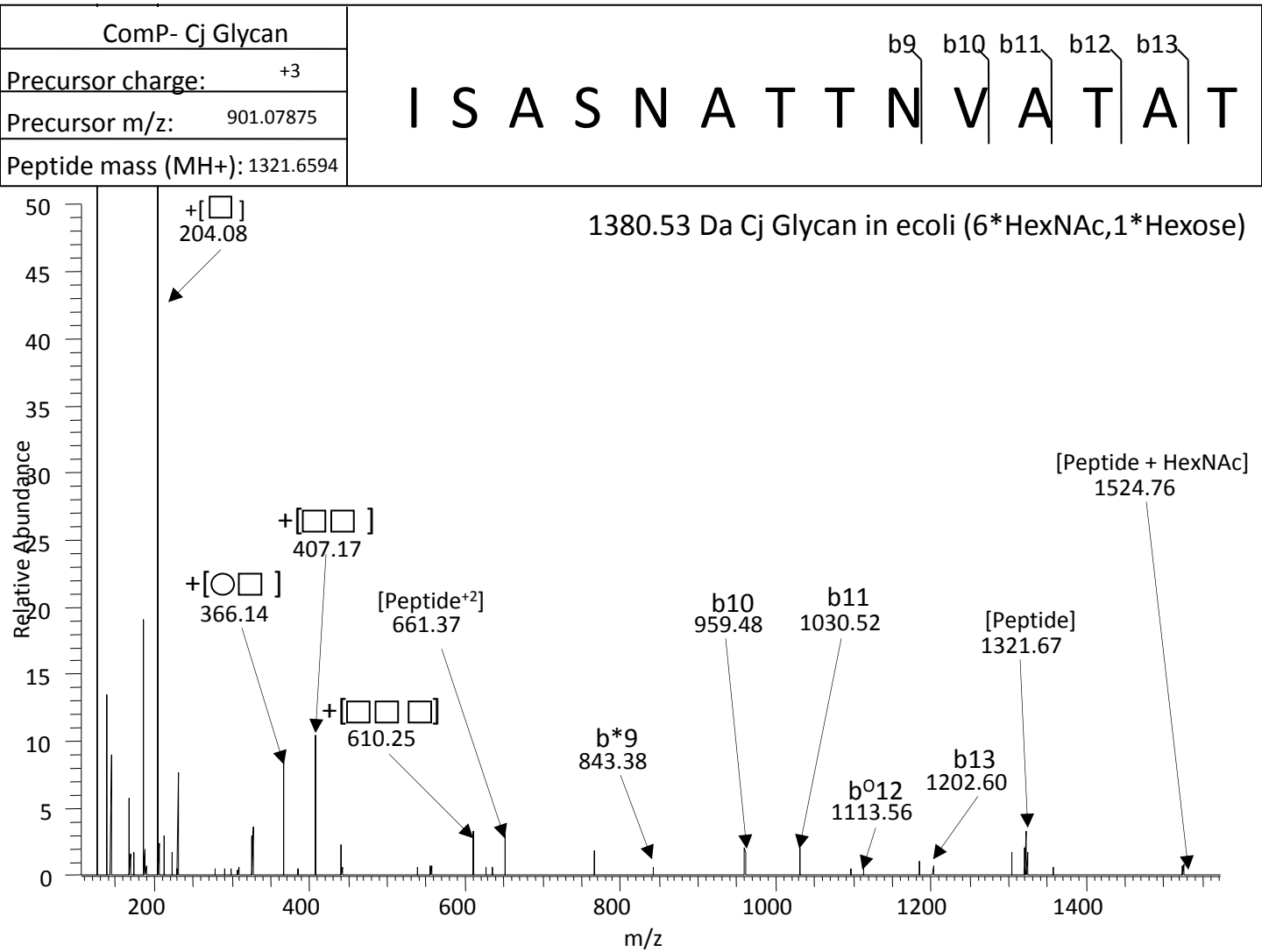
**Point mutational analysis of ComP confirms serine 84 is the likely site of glycosylation.** Serine 84 of ComP was mutated to an alanine and probed for glycosylation. (A-C) SDB1 cells expressing either WT ComP or ComP[S84A] in the presence of PglS and the *C. jejuni* heptasaccharide were probed via western blotting for protein glycosylation. As a negative control, the inactive PglS[H324A] variant was co-expressed with WT ComP and the *C. jejuni* heptasaccharide. (A) Anti-His channel probing for ComP expression and glycosylation. (B) Anti-glycan channel probing for the *C. jejuni* heptasaccharide. (C) Merged image for panels A and B. Source data are provided as a Source Data file.

Supplementary Figure 3



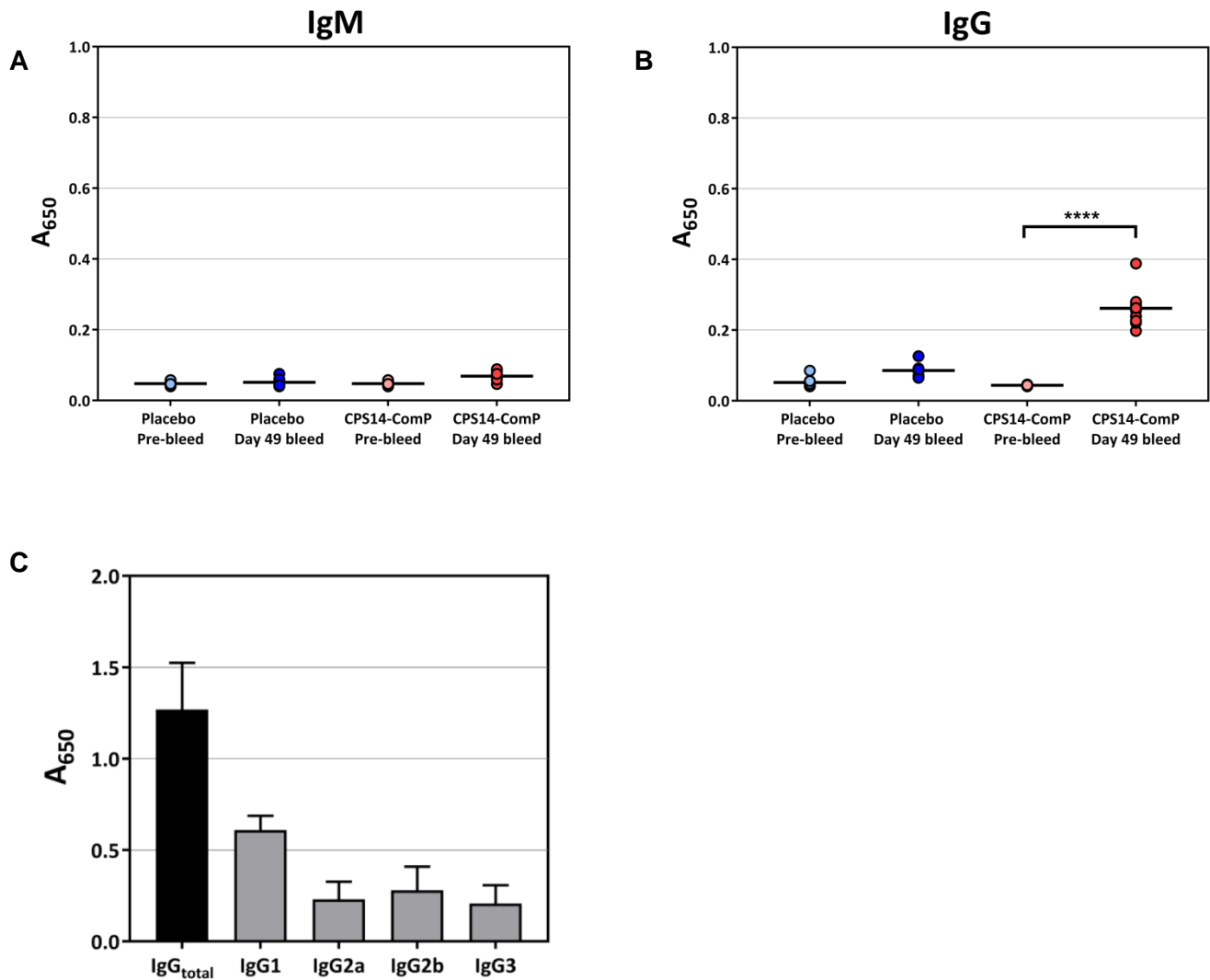
**Higher energy collisional dissociation (HCD) fragmentation spectra of GluC digested Comp glycosylated with the *C. jejuni* heptasaccharide (Comp-Glycan<sub>Cj</sub>).** GluC digested Comp-Glycan<sub>Cj</sub> was subjected to HCD fragmentation enabling the confirmation of a single peptide attached to a glycan with the CPS14 repeating subunit. Low collision energies regimes were undertaken to confirm the glycosylation of the peptide ISASNATTNVATAT with a 1380.53 Da glycan corresponding to 6\*HexNAc,1\*Hexose.

Supplementary Figure 4



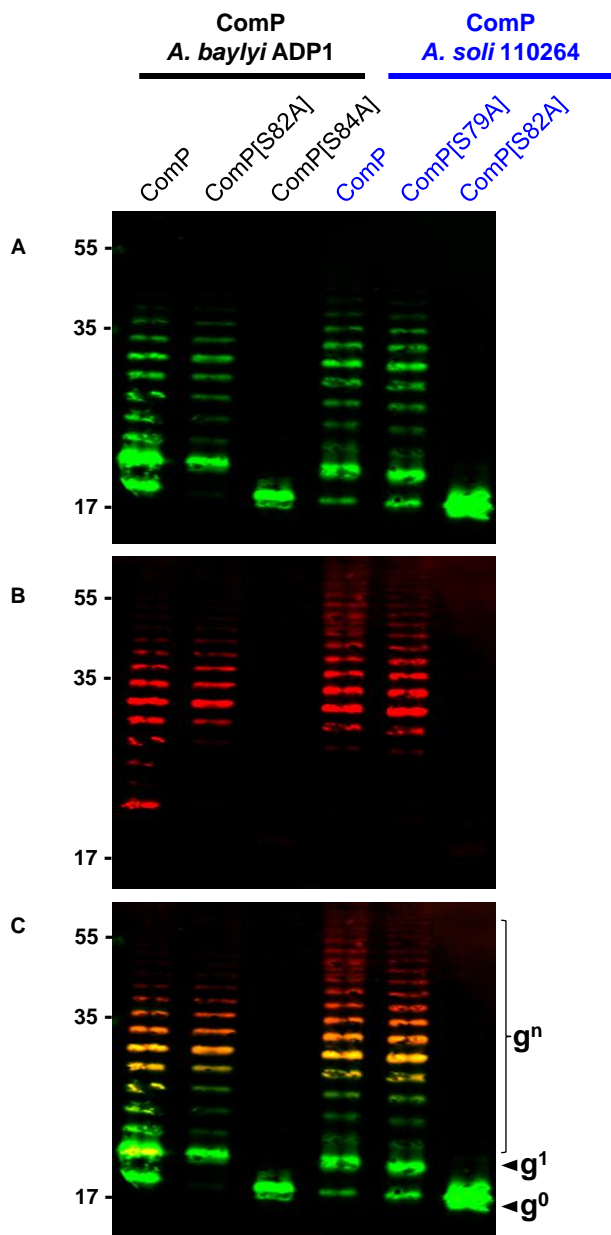
**Higher energy collisional dissociation (HCD) fragmentation spectra of GluC digested ComP glycosylated with the *C. jejuni* heptasaccharide (ComP-Glycan<sub>Cj</sub>).** GluC digested ComP-Glycan<sub>Cj</sub> was subjected to HCD fragmentation enabling the confirmation of a single peptide attached to a glycan with the CPS14 repeating subunit. High collision energies regimes were undertaken to confirm the glycosylation of the peptide ISASNATTNVATAT with a 1380.53 Da glycan corresponding to 6\*HexNAc,1\*Hexose.

Supplementary Figure 5



**Antibody responses towards monovalent CPS14-ComP bioconjugate vaccination.** Sera from placebo vaccinated and CPS14-ComP vaccinated mice were used to probe for IgM (A) and IgG (B) antibody responses measured against heat killed whole cell *S. pneumoniae* serotype 14. The Rout method for identifying outliers was performed and used to remove two data points from the IgG Placebo Day 49 bleed group as well as two points from the IgG CPS14-ComP Pre-bleed group. A Kruskal-Wallis test was subsequently performed and demonstrated that a statistically significant increase in serotype 14 specific IgG titers was observed (\*\*\*\* p= 0.0001). (C) IgG subclass responses in mice vaccinated with the CPS14-ComP bioconjugate. Sera from CPS14-ComP vaccinated mice were used for an ELISA probing for IgG subclasses against heat killed whole cell *S. pneumoniae* serotype 14. Kruskal-Wallis analysis a statistically significant increase of IgG1 subtype over IgG2a (\* p=0.0166) and IgG3 (\*\* p=0.0080). Each dot represents a single vaccinated mouse. Error bars indicate the standard deviation of the mean. Source data are available as a Source Data file.

Supplementary Figure 6



**A conserved and homologous serine is the likely site of glycosylation in ComP proteins from *A. baylyi* ADP1 and *A. soli* 110264.** Serines 82 and 84 of ComP<sub>ADP1</sub> and the homologous serines 79 and 82 of ComP<sub>110264</sub> were mutated to an alanine and probed for glycosylation in the presence of PglS and the serotype 8 capsular polysaccharide. (A-C) SDB1 cells expressing ComP variants in the presence of PglS and CPS8 were probed via western blotting for protein glycosylation. (A) Anti-His channel probing for ComP expression and glycosylation. (B) Anti-glycan channel probing for CPS8. (C) Merged image for panels A and B.

Supplementary Table 1

Strains and Plasmids	Description	Reference/ Source
Strains		
<i>E. coli</i> SDB1	W3110, $\Delta$ waal ligase, $\Delta$ wecA glycosyltransferase	<sup>1</sup>
<i>E. coli</i> DH5 $\alpha$	General cloning strain	Invitrogen
<i>S. pneumoniae</i> serotype 8, 9V, and 14	Wild type pneumococci strains expressing either the serotype 8, 9V, or 14 capsular polysaccharides	Statens Serum Institut
Plasmids		
pEXT20	Cloning vector, Amp <sup>R</sup> , IPTG inducible	<sup>2</sup>
pMN1	6X His-tagged Comp <sub>ADP1</sub> cloned in BamHI and Sall sites of pEXT20, Amp <sup>R</sup> , IPTG inducible	This work
pMN2	Non-coding region and PglS <sub>ADP1</sub> cloned in Sall and PstI sites of pMN1, Amp <sup>R</sup> , IPTG inducible	This work
pMN4	PglS <sub>ADP1</sub> [H324A] in pMN2 background	This work
pMN9	Comp <sub>ADP1</sub> [S82A] mutant of pMN2	This work
pMN10	Comp <sub>ADP1</sub> [S84A] mutant of pMN2	This work
pMAF10	HA-tagged PglB cloned in pMLBAD, TpR, Arabinose inducible	<sup>3</sup>
pAMF10	C-10 $\times$ His-tagged NmPglL cloned into pEXT20, AmpR, IPTG inducible	<sup>4</sup>
pIH18	C-6X His-tagged AcrA from <i>C. jejuni</i> cloned into pEXT21, SpR, IPTG inducible	<sup>5</sup>
pAMF22	C-6X His-tagged dsbA1 from <i>N. meningitidis</i> MC58 cloned into pMLBAD, Tp <sup>R</sup> Arabinose inducible	<sup>4</sup>
pACYCpglBmut	pACYC184-based plasmid encoding the <i>C. jejuni</i> pgl locus with mutations W458A and D459A in PglB. Cm <sup>R</sup> , IPTG inducible.	<sup>3</sup>
pNLP80	<i>S. pneumoniae</i> CPS14 cluster on pWSK129, Kan <sup>R</sup>	<sup>6</sup>
pB-8	<i>S. pneumoniae</i> CPS8 cluster on pBBR1MCS-3, Tc <sup>R</sup>	<sup>7</sup>
pWKS130-9V	<i>S. pneumoniae</i> CPS9V cluster on pWKS130, Kan <sup>R</sup>	This work
pCH1	6X His-tagged Comp <sub>110264</sub> cloned into pEXT20, Amp <sup>R</sup> , IPTG inducible	This work
pCH2	Comp <sub>110264</sub> [S79A] mutant of pCH1	This work
pCH3	Comp <sub>110264</sub> [S82A] mutant of pCH1	This work
pCH4	pEXT20 containing the $\Delta$ E553 variant of exotoxin A of <i>P. aeruginosa</i> (EPA) with an N-terminal DsbA signal sequence and a C-terminal glycotag consisting of Comp <sub>110264</sub> without the first 28 amino acids. A peptide linker (GGGS) fuses EPA to the Comp fragment.	This work

<sup>1</sup>Garcia-Quintanilla, F., Iwashkiw, J. A., Price, N. L., Stratilo, C. & Feldman, M. F. Production of a recombinant vaccine candidate against *Burkholderia pseudomallei* exploiting the bacterial *N*-glycosylation machinery. *Front Microbiol* **5**, 381, doi:10.3389/fmicb.2014.00381 (2014).  
<sup>2</sup>Dykxhoorn, D. M., St Pierre, R. & Linn, T. A set of compatible tac promoter expression vectors. *Gene* **177**, 133-136 (1996).  
<sup>3</sup>Feldman, M. F. *et al.* Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci U S A* **102**, 3016-3021, doi:10.1073/pnas.0500044102 (2005).  
<sup>4</sup>Faridmoayer, A., Fentabil, M. A., Mills, D. C., Klassen, J. S. & Feldman, M. F. Functional characterization of bacterial oligosaccharyltransferases involved in *O*-linked protein glycosylation. *J Bacteriol* **189**, 8088-8098, doi:10.1128/JB.01318-07 (2007).  
<sup>5</sup>Hug, I. *et al.* Exploiting bacterial glycosylation machineries for the synthesis of a Lewis antigen-containing glycoprotein. *J Biol Chem* **286**, 37887-37894, doi:10.1074/jbc.M111.287755 (2011).  
<sup>6</sup>Price, N. L. *et al.* Glycoengineered Outer Membrane Vesicles: A Novel Platform for Bacterial Vaccines. *Sci Rep* **6**, 24931, doi:10.1038/srep24931 (2016).  
<sup>7</sup>Kay, E. J., Yates, L. E., Terra, V. S., Cuccui, J. & Wren, B. W. Recombinant expression of *Streptococcus pneumoniae* capsular polysaccharides in *Escherichia coli*. *Open Biol* **6**, 150243, doi:10.1098/rsob.150243 (2016).



Supplementary Table 2

Primer	Sequence
igr F	5'ACTGGTCGACTAGTAGTACTATATGGCTT TAAA
igr R	5'ACTGCTGCAGTTAATATTCTATTGAACAAA ATTTTAAC
H325A F	5'GAGAATGGTTTACATACTCAGCGAATTTG TTCTTAGATTTAATG
H325A R	5'CATTAAATCTAAGAACAAATTCGCTGAGTA TGTAACCACTTCTC
S82A F – ADP1	5'GGAGTCCAAGAAATTGCGGCAAGTAATG CCA
S82A R– ADP1	5'GTGGCATTACTTGCCGCAATTTCTTGGAC TCC
S84A F– ADP1	5'CAAGAAATTTTCAGCAGCGAATGCCACTAC GAAC
S84A R– ADP1	5'GTTCTAGTGGCATTCTGCTGCTGAAATTT CTTG
S82A 110254 F	5'ACAGATCGCGTCCGGCGCCgCAGCAGC GACAACAAATGTAGCGT
S82A 110254 R	5'ACGCTACATTTGTTGTCGCTGCTGcGGC GCCGGACGCGATCTGT
S79A 110254 F	5'CGGGCGTCACACAGATCGCGgCCGGCG CCTCAGCAGCGACAACA
S79A 110254 R	5'TGTTGTCGCTGCTGAGGCGCCGGcCGC GATCTGTGTGACGCCCCG