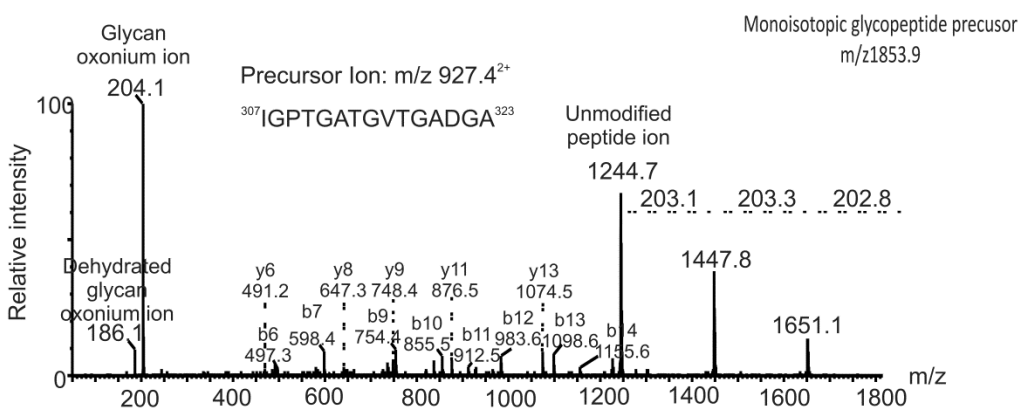
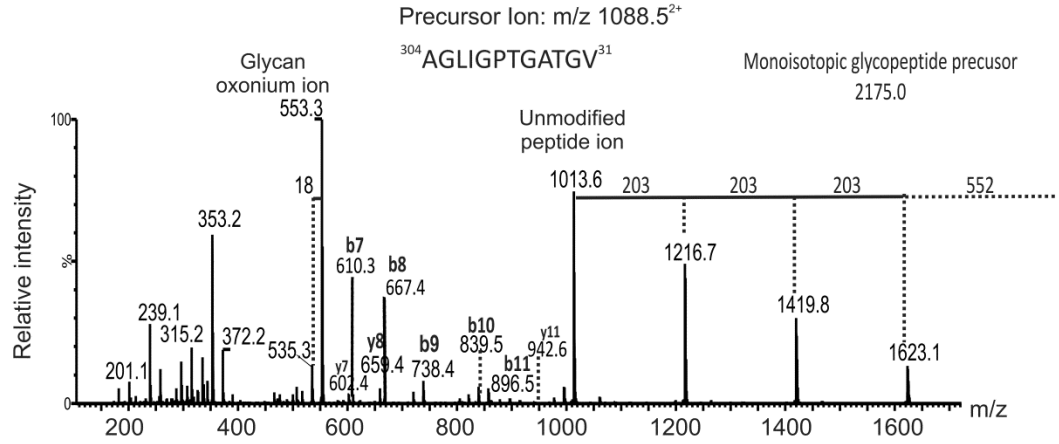
Supplementary Data - Strong et al 2014

a)



b)

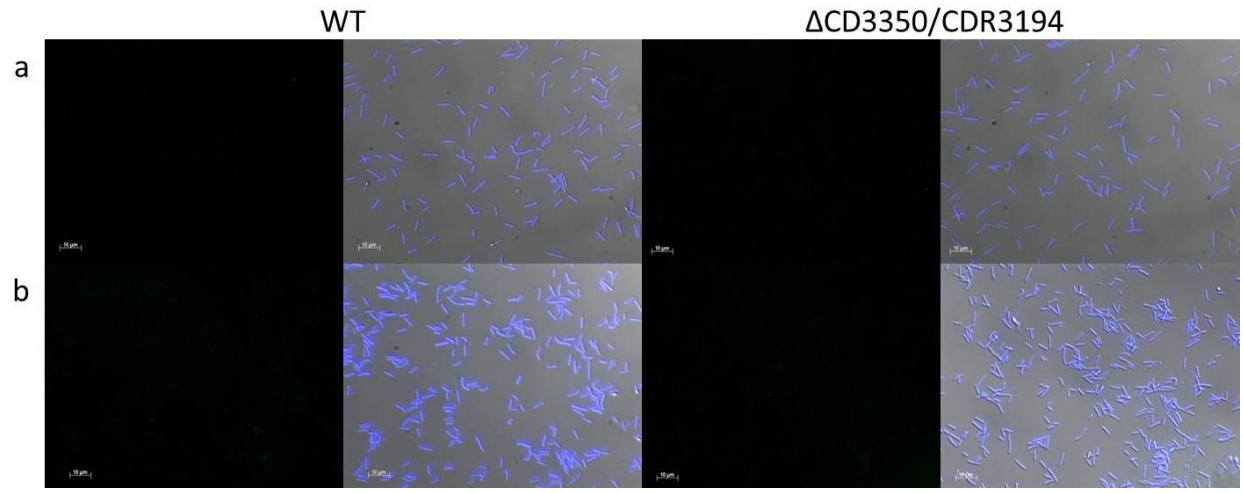


c)

```
1   MSRNKYFGPF DDNDYNNGYD KYDDCNNGRD DYNSCDCHHC CPPSCVGPTG PMGPRGRTGP
61  TGPTGPTGPG VGGTGPTGPT GPTGPTGNTG NTGATGLRGP TGATGGTGPT GATGAIGFGV
121 TGPTGPTGAT GATGADGVTG PTGPTGATGA DGITGPTGAT GATGFGVTGP TGPTGATGVG
181 VTGATGLIGP TGATGTPGAT GPTGAIGATG IGITGPTGAT GATGADGATG VTGPTGPTGA
241 TGADGVTGPT GATGATGIGI TGPTGATGAT GIGITGATGL IGPTGATGAT GATGPTGVTG
301 ATGAAGLIGP TGATGVTGAD GATGATGATG ATGPTGADGL VGPTGATGAT GADGLVGPTG
361 PTGATGVGIT GATGATGATG PTGADGLVGP TGATGATGAD GVAGPTGATG ATGNTGADGA
421 TGPTGATGPT GADGLVGPTG ATGATGLAGA TGATGPIGAT GPTGADGATG ATGATGPTGA
```

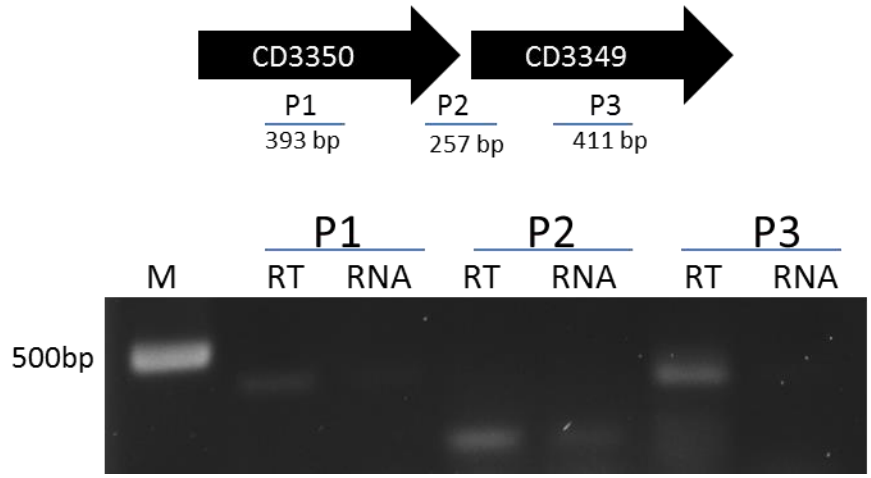
Supplementary Figure S1 Gel electrophoresis and mass spectrometry analyses of *C. difficile* QCD-32g58 endospore cell surface protein extract. Strain QCD-32g58. the MSMS spectrum of the doubly charged precursor ion at m/z 927.4. The y an b ion sequence corresponded to the peptide sequence ^{307}IGPTGATGVTGADGA^{323} from putative exosporium protein

(CdifQ_040500019311) with modification with three putative HexNAc residues. (c) The MSMS spectrum of the doubly charged peptide precursor on at m.z 1088.5 gave a series of peptide y and b ions, corresponding to the putative exosporial peptide ^{304}AGLIGPTGATGV^{317} . Neutral losses corresponding to three HexNac moieties and an unknown glycan of 552 Da were observed in the high m/z region of the spectrum. This gave a total mass excess of 1162 Da. De novo sequencing of the resulting MSMS spectra showed peptides corresponding to a putative exosporium glycoprotein (CdifQ_040500019311). c) A total of nine glycopeptides were identified, corresponding to 17-21 % sequence coverage.

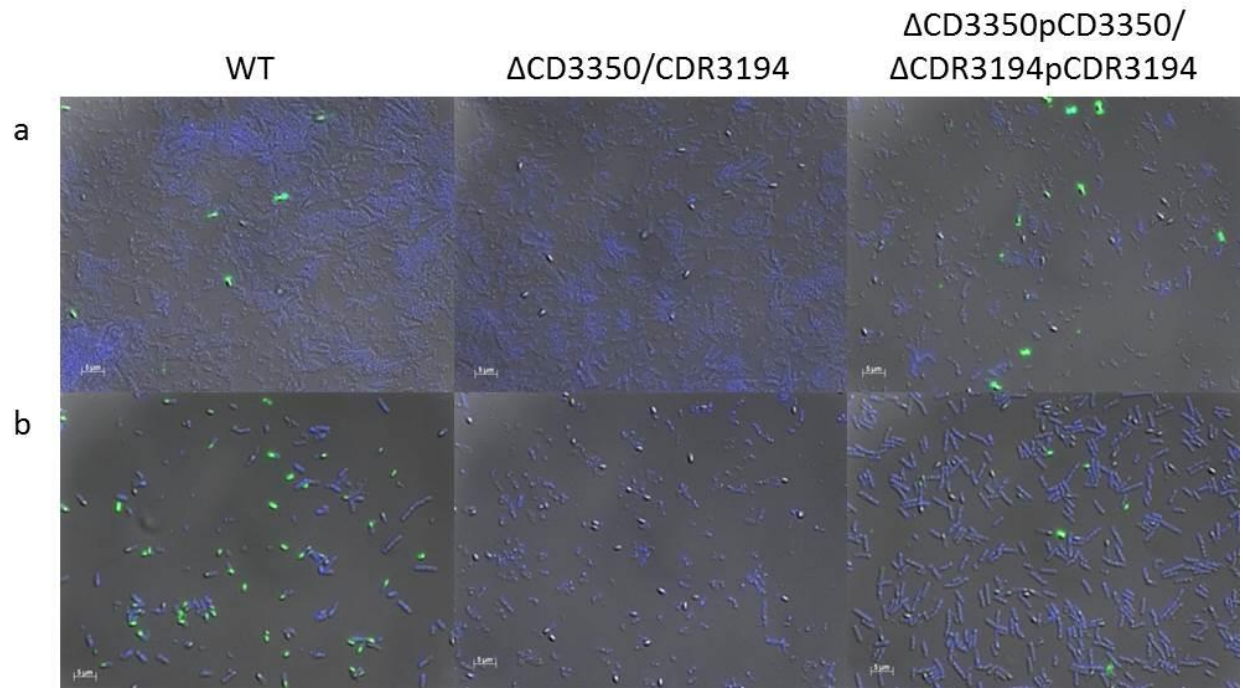


Supplementary figure S2 Immunofluorescence of anti-β-O-GlcNAc binding to vegetative cells.

(a) 630Δerm, (b) R20291, comparing wild type to respective mutant strains (a) ΔCD3350 (b) ΔCDR3194. Left hand column FITC labelling only, right hand column shows merged images of FITC, DAPI and transmitted light channels. GlcNAc visualised with mouse anti-β-O-GlcNAc and anti-mouse IgM-FITC conjugate.



Supplementary figure S3 RT-PCR analysis demonstrating co-transcription of CD3350 and CD3349. Upper panel shows expected size of each product with primers pairs P1 (CD3350) P2 (intergenic region) and P3(CD3349); lower panel shows agarose gel analysis of products. **RT lanes**, RT-PCR was performed using total RNA from *C. difficile* 630 cells. **RNA lanes**, standard PCR reaction with same primers using total RNA to demonstrate no contaminating DNA in RNA samples. **M**, DNA marker 500bp.



Supplementary figure S4 Restoration of anti-GlcNAc reactivity through complementation. 72 hour plate grown cultures of (a) 630Δerm, (b) R20291, comparing wild type to ΔCD3350/ΔCDR3194 and ΔCDR3194p3350; complements were induced with 500 ng anyhydrotetracycline. Merged images of FITC, DAPI and transmitted light channels. GlcNAc visualised with mouse anti-β-O-GlcNAc and anti-mouse IgM-FITC conjugate.

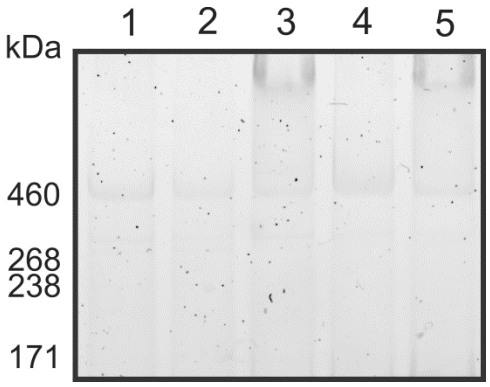
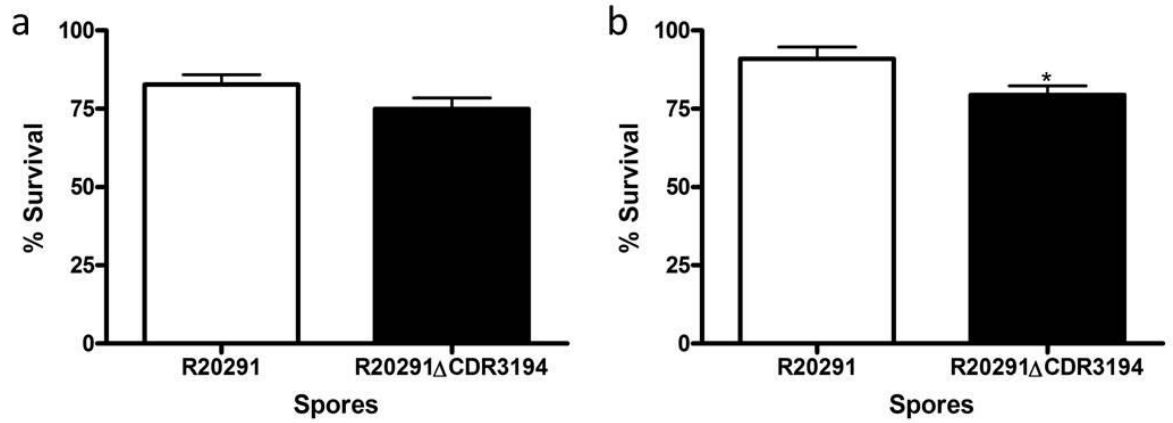


Figure S5 Glycostaining of *C. difficile* spore surface extracts. Surface extracts were run on 3-8% Tris-Acetate NuPAGE prior to glycostaining with Pro-Emerald Q. Lane 1 630Δerm; Lane 2 630ΔsgtA; Lane 3 R20291; Lane 4 R20291ΔsgtA; Lane 5 QCD-32g58



Supplemental figure S6. Resistance assays (a) lysozyme and (b) ethanol. a) R20291 WT and ΔsgtA spores were incubated with 250 µg/ml lysozyme 1 hour 37°C then percentage survival was calculated. b) R20291 WT and ΔsgtA spores were incubated in 70% ethanol 20 minutes room temperature then percentage survival was calculated. Assays performed in triplicate on three independent occasions. Statistical analysis is t-test with Welch's correction (* $p = <0.05$).