Supplemental Information

Retaining the structural integrity of protein disulfide bonds is crucial for effectiveness of glycoconjugate vaccine candidates

Filippo Carboni[a,+], Annabel Kitowski[b,+], Charlotte Sorieul[a], Daniele Veggi[a], Marta C. Marques[b], Davide Oldrini[a], Evita Balducci[a], Barbara Brogioni[a], Linda Del Bino[a], Francesca Angiolini[a], Lucia Dello Iacono[a], Immaculada Margarit[a], Maria Rosaria Romano[a], Gonçalo J. L. Bernardes[b,c,*], Roberto Adamo[a,*]

[a]*GSK, Via Fiorentina 10, 53100 Siena, Italy*
[b]*Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal*
[c]*Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*
[+]*F.C. and A.K. contributed equally to this work*
[*]*To whom correspondence should be addressed. E-mail: gbernardes@medicina.ulisboa.pt, roberto.x.adamo@gsk.com*

Methods

1. Procedures for protein modification and conjugation

Site-selective introduction of DCA

CRM_{197} were diluted in sodium phosphate 50 mM pH 7.5 to a final concentration of 20 mg/mL. 12 equivalents of tris(2-carboxyethyl) phosphine (0.05 M solution) were added and the solution was kept gently moving at room temperature for 3 hours. A solution of 1,3-dichloroaceton (10 equivalents, 20 mg/mL in DMSO) was added and the reaction was performed at room temperature in the dark for 3.5 hours. The modified protein was purified on a G25 column equilibrated with 100 mM sodium phosphate pH 6.3. The obtained fractions were concentrated with Amicon ultracentrifuge filters 10 kDa MWCO.

Introduction of the azido-containing linker in CRM-DCA

CRM_{197}-DCA in 100 mM sodium phosphate pH 6.3 (5-8 mg/mL) was mixed with 600 equivalents of Aminoxy-peg5-azide. For the linker solution, 21 mg of aminoxy-peg5-azide were dissolved in 70 µL DMSO, resulting in a concentration of 300 mg/mL. The reaction performed at room temperature over the weekend while moving gently. The reaction was purified on a G25 column which was equilibrated with PBS 1x. The modified protein was eluted with PBS 1x and concentrated with Amicon ultracentrifuge filters 10 kDa MWCO.

Site-selective introduction of DHA

CRM_{197} were diluted to a final concentration of 5 mg/mL with 50 mM sodium phosphate pH 11. 12 equivalents of tris(2-carboxyethyl) phosphine (0.05M solution) were added and the reaction mixture was kept in the dark gently moving at room temperature for 3 hours. Methyl-2,5-dibromopentanoate (500 equivalents, from 2 M stock in DMSO) was added, the mixture was vortex for 10 seconds and the reaction was completed after 5 hours at room temperature. CRM-DHA was purified using ZebaSpin Desalting columns (7 kDa MWCO), which were previously equilibrated with 100 mM sodium phosphate pH 6.3. The final concentration of the obtained protein solution was determined by bicinchoninic acid assay.

Introduction of the azide-containing linker on CRM-DHA

For the introduction of the linker bis(11-azidoundecyl)disulfide into the DHA residues of CRM-DHA, 100 equivalents of the linker (from 43.8 mM stock in DMSO) were mixed with 100 equivalents of tris(2-carboxyethyl) phosphine (from 0.05 M solution) and incubated at room temperature for 2 hours. The

protein CRM-DHA (in 100 mM sodium phosphate pH 6.3) was added directly to the linker solution and the reaction was keep at room temperature gently moving for 4 hours. The reaction was purified using ZebaSpin Desalting columns (7 kDa MWCO) which were previously equilibrated with 12 mM sodium phosphate pH 6.3. The protein concentration was determined by bicinchoninic acid assay and should be above 2 mg/mL for the following glycoconjugation reaction.

Derivatization of polysaccharides from Group B *Streptococcus* serotype Ia and III and *Streptococcus pneumoniae* serotype 14

For reaction with the linker Dibenzocyclooctyne-N-hydroxysuccinimidyl ester (DBCO-NHS, Click Chemistry Tools), the polysaccharides (10 mg/mL) were diluted with 1.25 M NaOH to reach a final concentration of polysaccharide of 2 mg/mL and NaOH concentration 0,5M for Group B *Streptococcus* serotype III and *Streptococcus pneumoniae* serotype 14, or 1M for Group B *Streptococcus* serotype Ia. The reaction was incubated at 70 °C for 0.5 hours for Group B *Streptococcus* serotype III and *Streptococcus pneumoniae* serotype 14, or 4,5 hours for Group B *Streptococcus* serotype Ia before being neutralized with concentrated acetic acid and purified on G25 columns equilibrated with water. The solvent was evaporated and the residue was dissolved in DMSO and water (1:1 ratio 400 µl DMSO/H_2O per 5 mg polysaccharide). A dilution of Triethylamine (2 µl TEA in 18 µl H_2O, 5 µl per 5 mg polysaccharide) and a solution of the linker DBCO-NHS (1 mg in 50 µl DMSO, 1 mg linker per 5 mg polysaccharide) were added and the mixture was kept gently moving at room temperature for 3-4 hours. The reaction was purified on G25 columns, which were previously equilibrated with water. The solvent was evaporated and the ratio of linker to polysaccharide was determined using NMR spectroscopy.

Glycoconjugation between GBS PSIII-DBCO or Sp14-DBCO and CRM-DCA-N_3

A dried sample of GBS PSIII-DBCO or Sp14-DBCO was solubilized in H_2O and mixed with a sample of CRM-DCA-N_3. The final ratio protein/polysaccharide should be 1:4. A suspension of dehydroascorbic acid (36 mg/mL, 2.5 µg per µg protein) in H_2O was added and the reaction was performed at room temperature for 3hours.

For the purification, a column packed with CMT Ceramic Hydroxyapatite type I, 80µm, was equilibrated with 2 mM sodium phosphate 550 mM, sodium chloride pH 7.2. The reaction mixture (in PBS 1x) was diluted to a final concentration of 2 mM sodium phosphate and sodium chloride was added to reach a final concentration of 550 mM. The free polysaccharide was eluted with 2 mM sodium phosphate 550 mM NaCl, followed by the elution of the conjugate in 35 mM sodium phosphate pH 7.2 and elution of the free protein in 550 mM sodium phosphate pH 7.2. The conjugate was concentrated and stored in 10 mM sodium phosphate pH 7.2.

Glycoconjugation between GBS PSIa-DBCO and CRM-DHA-N_3

To generate a glycoconjugate freshly generated CRM-DHA-N_3 and polysaccharide GBS PSIa-DBCO, were mixed in a ratio of 1:2 protein/polysaccharide, at protein concentrations of ~2 mg/mL.

A suspension of dehydro-ascorbic acid (36 mg/mL in H_2O) was added with a concentration of 2.5 µg per 1 µg protein. The reaction was kept gently moving at room temperature overnight.

For the purification, unreacted polysaccharide was separated from the glycoconjugate and free protein by ammonium sulfate precipitation. For this purification the crude reaction mixture was saturated with ammonium sulfate (500 mg/mL) and incubated on ice for 15 min. The precipitated protein was obtained after centrifugation. A second precipitation of the supernatant during this step, did not result in any further protein precipitation. The protein pellet was dissolved in 1 mL 10 mM sodium phosphate pH 7 and saturated again with ammonium sulfate. This precipitation procedure was repeated for six times and the final protein pellet was dissolved in 800 µl 10 mM sodium phosphate pH 7. The protein concentration was determined using bicinchoninic acid assay.

2. Protein structural characterization

ESI MS spectra of modified CRM_{197}

Exact masses were measured by electron spray ionization cut-off spectroscopy, using a Q-Tof microMacromass (Waters) mass spectrometer calibrated with myoglobin (Sigma Aldrich) in 50% acetonitrile and 0.2% formic acid. Direct infusion of reference and modified CRM_{197}, previously dialyzed against 5 mM ammonium acetate on Amicon with 10kDa cutoff membrane for desalting, was performed with samples diluted in 25% acetonitrile and 0.2% formic acid and at a flow rate of 10µL/min. The electrospray source was operated with a capillary voltage of 3.0 kV, and a cone voltage of 30 V. Nitrogen was used as the desolvation gas at a total flow of 500 L/h, and Argon as the collision gas at a total flow of 50 L/h. Total ion chromatogram, combined ion series and deconvoluted mass spectrum were reconstructed using the MaxEnt algorithm installed on the MassLynx software (Waters).

Dynamic light scattering studies

Dynamic light scattering (DLS) measurements were conducted on a Malvern Zetasizer Nano ZS (Malvern, UK) with a backscattering detection at 173°, equipped with a He-Ne laser (λ = 632.8 nm), at 25 °C, using low-volume 0.3 cm quartz cells (Hellma GmbH & Co, Müllheim, Germany). CRM_{197} protein variants were diluted to a concentration of 5 µM in the same buffer as used for the CD spectroscopy experiments. The samples were left equilibrating for 5 min at 25 °C before each

measurement set (10 measurements of 10 seconds). All buffers were filtered through a sterile 0.45μm pore size filter (Whatman, Florham Park, NJ) and protein samples were centrifuged before dilution. Normalized intensity autocorrelation functions were analyzed using the CONTIN method [ref?], yielding a distribution of hydrodynamic radius (RH) and diameter (DH). For each measurement, RH and DH values were obtained from the peak of the particle number distribution.

Circular dichroism spectroscopy

Far-UV circular dichroism (CD) data were acquired on a Jasco J-815 CD spectropolarimeter (Hachioji, Tokyo, Japan) using quartz cuvettes of 1.0 mm (Hellma GmbH & Co, Müllheim, Germany). All spectra were collected at 25ºC, between 195 and 300 nm, with sampling velocity of 200 nm min-1, data pitch of 0.5 nm, data integration time of 1 s, and 1 nm bandwidth. Each spectrum represents an average of 5 scans. Protein stock solutions were diluted to a final peptide concentration of 10 μM in 10 mM sodium phosphate buffer (pH 7.4) with 150 mM NaCl. Absorbance spectra were also monitored to control the light scattering and the signal detection saturation. In addition to blank subtraction, experimental instrument-related baseline drift was corrected by subtracting the average of the signal between 250 and 260 nm to each spectrum. The raw ellipticity, θ, was subsequently converted into the mean residue molar ellipticity ([θ]; in deg cm2 dmol-1 Res-1). All conditions were measured independently and in duplicate.

Protein crystallization, diffraction, data collection and processing

The CRM_{197}-DCA was concentrated to 20 mg/ml in 50mM Tris-HCl, 200 mM NaCl, pH8 and was screened in 480 different crystallization conditions. Using a Crystal Gryphon robot (Art Robbins Instruments), each experiment was prepared using 200nl reservoir and 200nl protein sample. The best crystal was grown in buffer containing 0.1 sodium cacodylate, 0.2 M magnesium acetate. 20% w/v PEG 8 K as precipitant at pH 6.5. Crystals were soaked in the original mother liquor supplemented with 15% ethylene glycol prior to cryo-cooling in liquid nitrogen. Structure determination and refinement Diffraction of the crystals was performed at beamline ID23-1 of the European Synchrotron Radiation Facility (ESRF) and several full datasets were collected at 100K, at wavelength λ =0.976Å, on a Pilatus 6M detector. Diffraction datasets were indexed and integrated using iMOSFLM and reduced using Aimless, via the CCP4 suite[2].

Crystals of CRM_{197}-DCA belonged to space group P31 1 2. The structure of the CRM_{197}-DCA was solved at 2.03 Å resolution by molecular replacement with Phaser[3] using as model the diphtheria toxin mutant CRM_{197}, protein data bank (PDB) code 4AE0. Initial molecular replacement solutions were subjected to subsequent cycles of manual building in Coot with Phenix.refine[4].

3. In vivo testing of synthesized glycoconjugates

Assessment of immune responses

Animal studies were conducted according to experimental guidelines from Animal Welfare Body of the Animal Resource Centre of our Company. Groups of eight CD1 or BALB/c mice were immunized on days 1, 21 and 35 by intraperitoneal injection of 0.5µg/1µg (saccharide content) of each glycoconjugate formulated with Alum Hydroxide. Sera were bled at days 1, 35 and 49. IgG titers in collected sera were estimated by ELISA. Microtiter plates (96 wells, NUNC, Maxisorp) were coated with 100ng of HSA (Sigma Aldrich)-conjugated PSIa or PSIII or 2µg of Pn14. Plates were incubated overnight at 2-8°C, washed three times with PBST (0.05% Tween-20 in PBS pH 7.4) and saturated with 250 µL/well of PBST-B (2% Bovine Serum Albumin-BSA in PBST) for 90 min at 37°C. Two-fold serial dilutions of test and standard sera in PBST-B were added to each well. Plates were then incubated at 37°C for 1h, washed with PBST, and then incubated for 90 min at 37°C with anti-mouse IgG-alkaline phosphatase (Sigma) diluted 1:2000 in PBST-B. After washing, the plates were developed with a 4 mg/mL solution of p-Nitrophenyl Phosphate (pNPP) in 1 M diethanolamine (DEA) pH 9.8, at room temperature for 30 min. The absorbance was measured using a SPECTRAmax plate reader with wavelength set at 405 nm.IgG titers were calculated by the reciprocal serum dilution giving Optical density (OD) equal to 1. Functional activity of anti-GBS antibodies was evaluated by opsonophagocytic killing assay (OPKA) using differentiated HL-60 cells and GBS COH1 and 515[1]. OPKA titers were expressed as the reciprocal serum dilution mediating 50% bacterial killing, estimated through piecewise linear interpolation of OPKA data.

4. *Supplemental Figures*

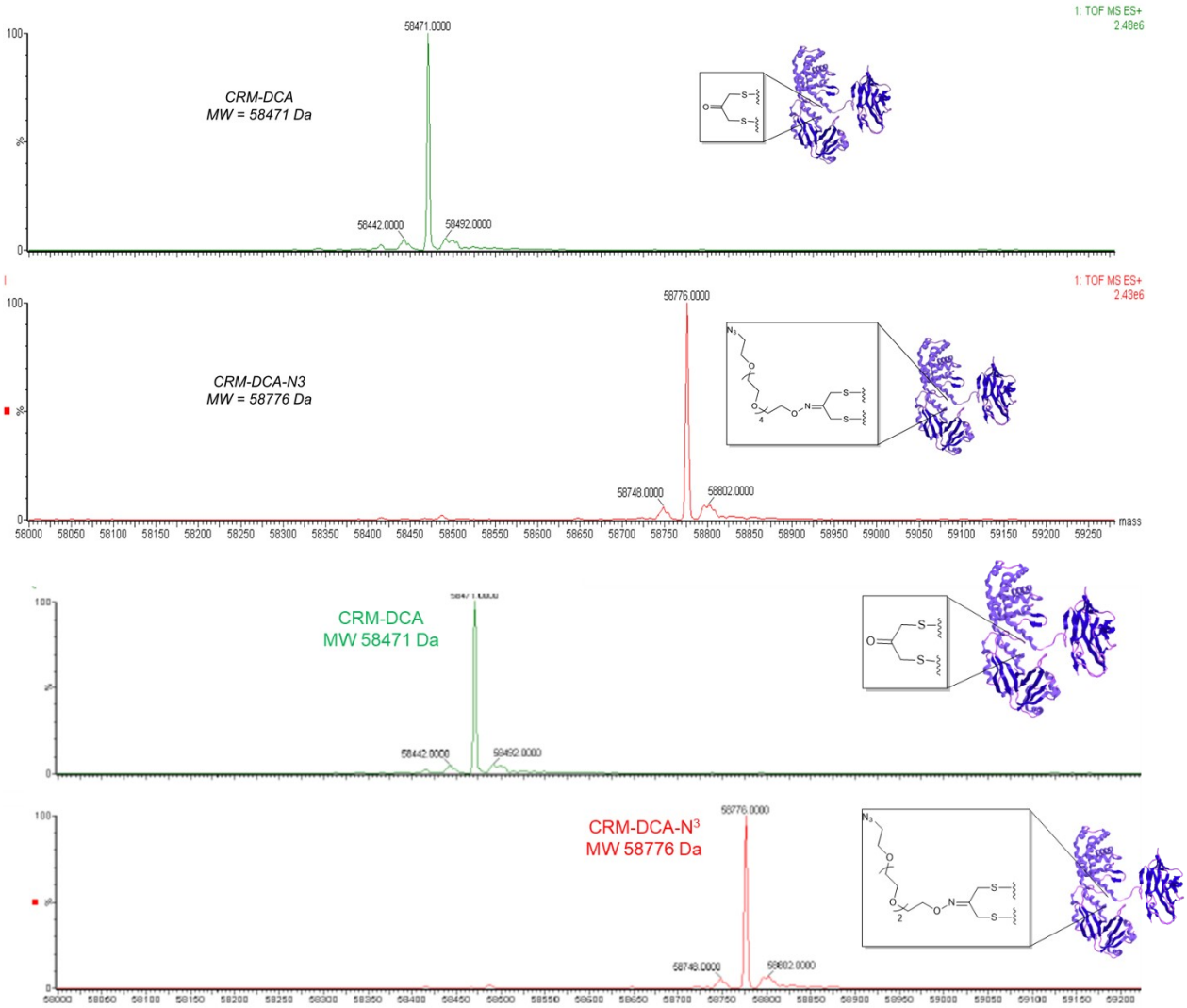

Figure S1. LC-MS spectra of native CRM-DCA and CRM-DCA-N3.

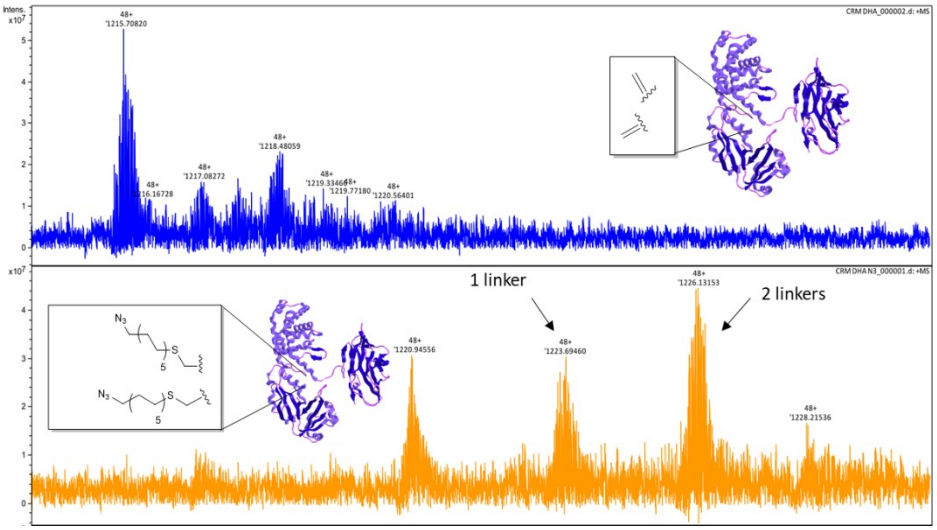

Figure S2. Comparison of protein species present in charge state 48+ of CRM-DHA (blue) and CRM-DHA-N3 (yellow),deconvolution revealed the major species as CRM-DHA-N3 containing 2 linker moieties.

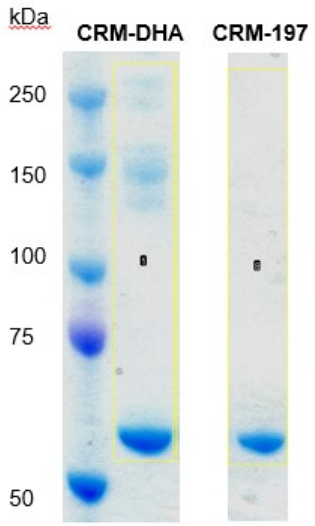

Figure S3. SDS Page of CRM-Dha showing minor level of aggregation as compared to native CRM_{197}. Samples were freshly prepared and used for conjugation.

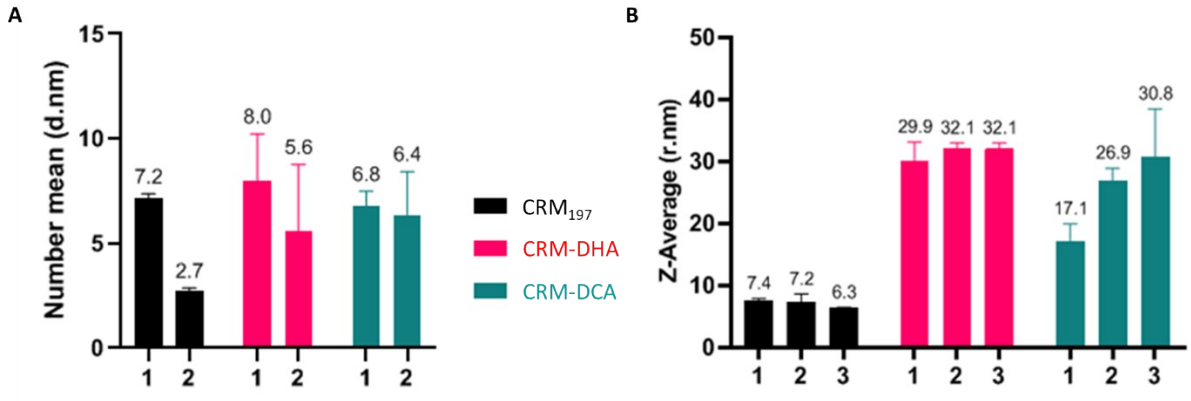

Figure S4. DLS measurements of CRM_{197}, CRM_{197}-DHA and CRM_{197}-DCA A) Number-averaged size distribution profiles of each sample (10 µM) were used to determine the respective average DH. B) Z-Average radius profiles for each sample. The data is the average of at least three independent replicates. Error bars correspond to the standard deviation of the mean.

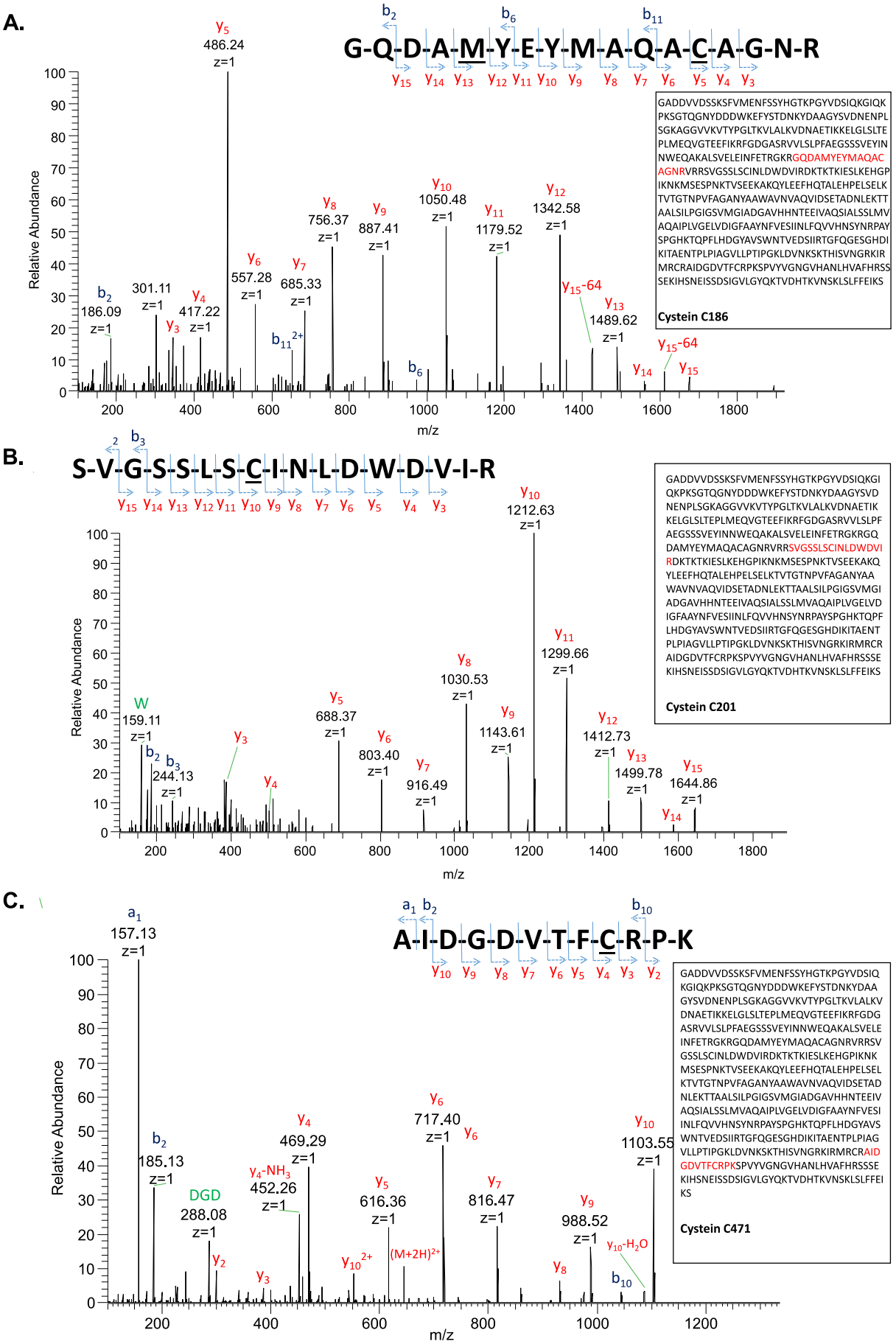

Figure S5. A) MS/MS spectrum of the m/z 930.884 doubly charged ion of the tryptic peptide *GQDAMYEYMAQACAGNR* from CRM_{197}, containing the **DHA** modification at the cysteine residue, **B)** MS/MS

spectrum of the m/z 915.47 doubly charged ion of the tryptic peptide *SVGSSLSCINLDWDVIR* from CRM_{197}, containing the **DHA** modification at the cysteine residue, **C)** MS/MS spectrum of the m/z 644.338 doubly charged ion of the tryptic peptide *AIDGDVTFCRPK* from CRM_{197}, containing minor DHA modifications at the cysteine residue**.**

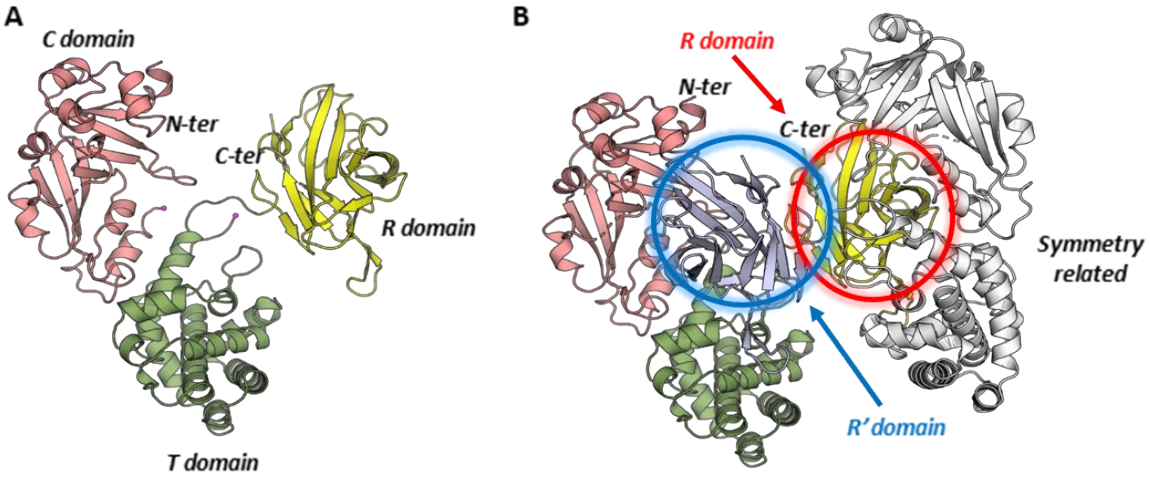

Figure S6. A) The X-ray structure of the monomeric form of CRM_{197}-DCA. Catalytic (C), transmembrane (T), and receptor (R) domains are colored in pink, green, and yellow, respectively.The Cα atoms of the last visible residues of the disordered loop (186-199), connecting the C and T domains, are represented as magenta spheres. B) Organization of the domain-swapped dimer of CRM_{197}-DCA. The symmetry-related molecule (white cartoon) shows the R domain-swapped molecule that forms the dimer of the CRM_{197}-DCA. The red circle points out the R domain of the CRM_{197}-DCA molecule (yellow), while the blue circle indicates the R' domain (light blue) belonging to the symmetry-related molecule.

Table S1. Data collection and refinement statistics.

	CRM_{197}-DCA (PDB ID 7O4W)
Crystal	
Space group	P 3_1 1 2
Cell dimensions	
a, *b*, *c* (Å)	98.00, 98.00, 98.27
α, β, γ (°)	90, 90, 120
Data collection	
Beamline	ESRF ID23-1
Wavelength (Å)	0.976254
Resolution (Å)	49.1-2.03 (2.10-2.03)
Total reflections	393242 (38576)
Unique reflections	35072 (3440)
R_{merge}	0.051 (1.225)
R_{meas}	0.056 (1.349)
$I/\sigma(I)$	26.2 (2.1)
$CC_{1/2}$	0.99 (0.71)
Completeness (%)	100 (100)
Redundancy	11.2 (11.2)
Wilson B-factor (Å)	49.92
Refinement	
Resolution (Å)	49.1-2.03
No. reflections	35037
R_{work} / R_{free}	20.4/23.7
No. atoms	
Protein	3940
Ligands	40
Water	133
B factors	
Protein	54.66
Ligands	62.26
Water	50.11
R.m.s. deviations	
Bond lengths (Å)	0.020
Bond angles (°)	1.211
Clash scores	11.32
Ramachandran[#]	
Favored (%)	96.50
Allowed (%)	3.11

[a] Values in parentheses are for highest-resolution shell

[#] Measured using Molprobity

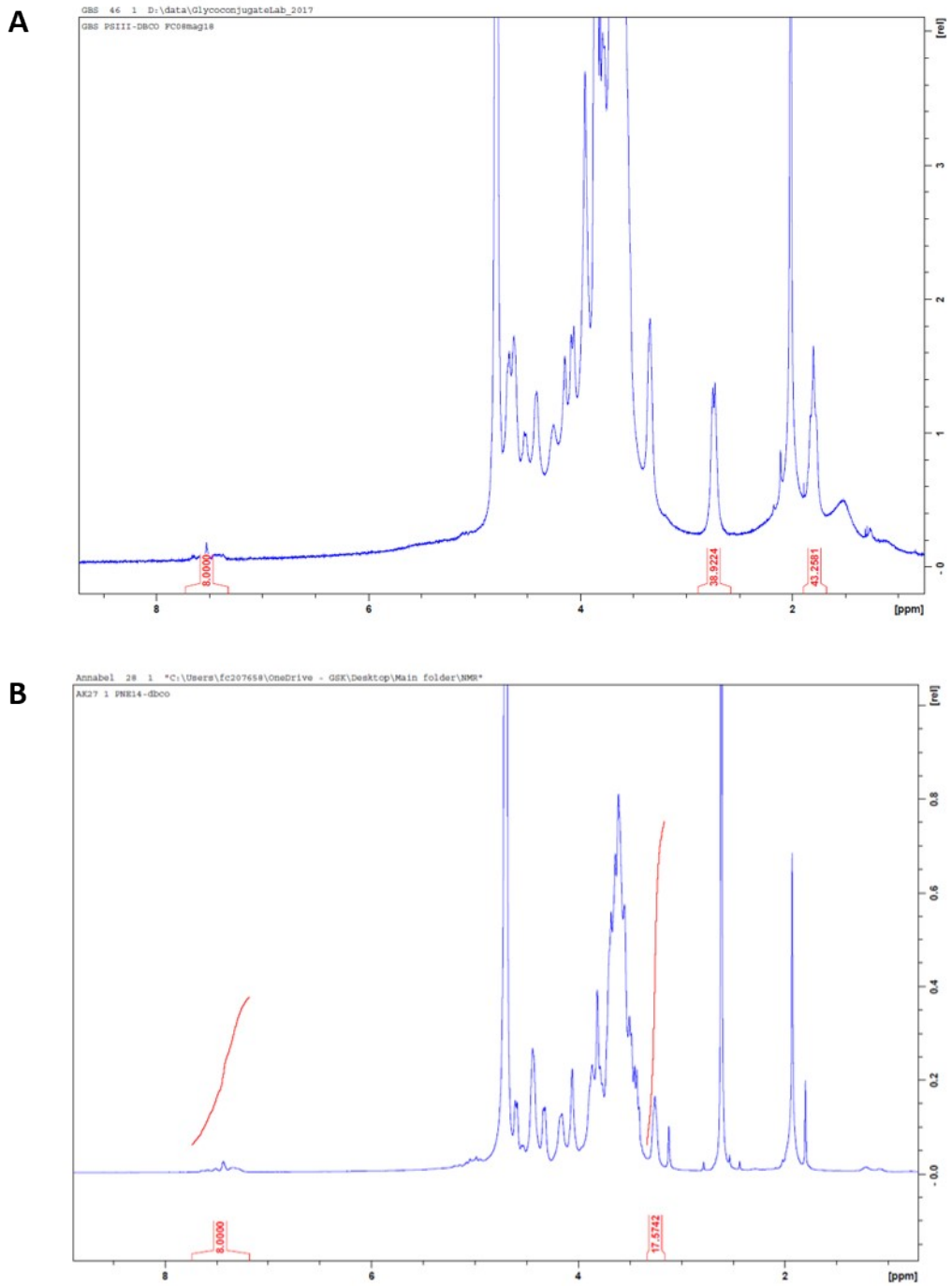

Figure S7. ^{1}H NMR spectra of modified GBS PSIII (A) and Sp14 PS (B). Peak intensities of aromatic DABCO and H-3equatorial from Neu5Ac or H-2 from Glc residue were used to calculate level of linker incorporation for GBS PS and Sp14 PS, respectively.

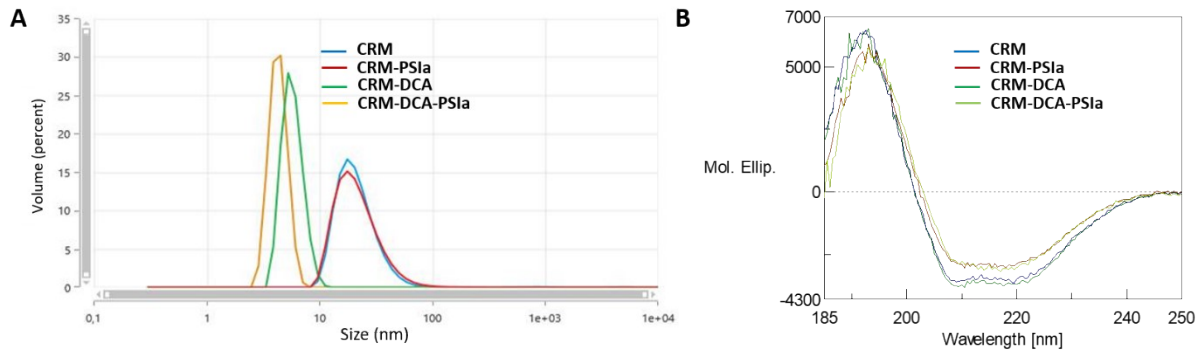

Figure S8. A) DLS profile of conjugates and related unconjugated proteins. B) CD spectra of CRM and CRM-DCA prior and after conjugation to PSIa.

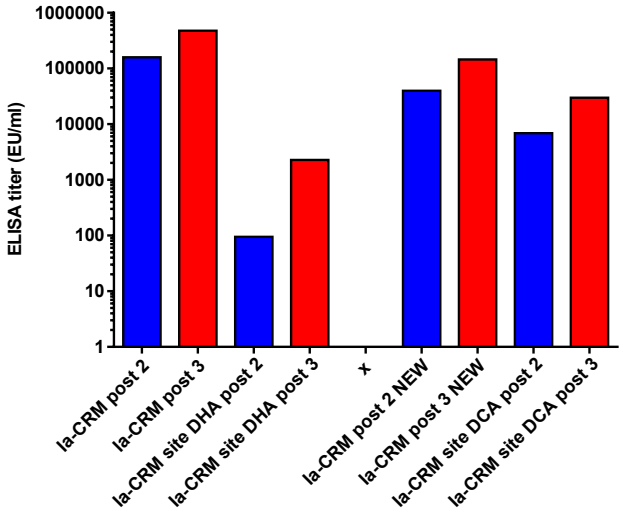

Figure S9. Anti-CRM_{197} antibodies as measure by ELISA in pooled sera after second and third administration.

References

1. F. Carboni, R. Adamo, M. Fabbrini, R. De Ricco, V. Cattaneo, B. Brogioni, D. Veggi, V. Pinto, I. Passalacqua, D. Oldrini, R. Rappuoli, E. Malito, I. Y. R. Margarit and F. Berti, *Proc Natl Acad Sci U S A*, 2017, **114**, 5017-5022.
2. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr D Biol Crystallogr*, 2011, **67**, 235-242.
3. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J Appl Crystallogr*, 2007, **40**, 658-674.
4. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 213-221.