Click chemistry compared to thiol chemistry for the synthesis of site-selective glycoconjugate vaccines using CRM$_{197}$ as carrier protein

G. Stefanetti$^1$ · M. Allan$^2$ · A. Usera$^2$ · F. Micoli$^3$

Received: 13 February 2020 / Revised: 15 April 2020 / Accepted: 30 May 2020 / Published online: 13 June 2020
© The Author(s) 2020

Abstract
Conjugation chemistry is one of the main parameters affecting immunogenicity of glycoconjugate vaccines and a rational approach toward a deeper understanding of their mechanism of action will greatly benefit from highly-defined and well-characterized structures. Herein, different conjugation methods were investigated with the aim of controlling glycosylation site and glycosylation density on the carrier protein. S. Typhimurium lipopolysaccharide O-Antigen and CRM$_{197}$ carrier protein were used as models. In particular, thiol and click chemistry were examined, both involving the linkage of the terminal reducing sugar unit of the O-Antigen chain to different amino acids on the carrier protein. Thiol chemistry allowed O-Antigen conjugation only when the carrier protein was activated on the lysines and with a relative high number of linkers, while click chemistry allowed conjugate generation even when just one position on the protein was activated and to both lysine and tyrosine sites. The study highlights click chemistry as a leading approach for the synthesis of well-defined glycoconjugates, useful to investigate the relationship between conjugate design and immune response.

Keywords Glycoconjugate vaccine · Click chemistry · Thiol chemistry · Conjugation chemistry · Salmonella Typhimurium · O-antigen · CRM$_{197}$

Introduction
Glycoconjugate vaccines are important therapeutics for the prevention of infectious disease from severe pathogens like Neisseria meningitidis, Haemophilus influenza and Streptococcus pneumoniae [1, 2]. They require the covalent linking of a sugar antigen to a carrier protein, which can be obtained by different strategies influencing both the efficiency of conjugation and the structure of the glycoconjugate, with an important impact on immunogenicity [3]. Traditional glycoconjugate approaches involve random linkage of the sugar hapten to a carrier protein, or end-group modification of the saccharide chain to achieve better control and characterization of the resulting vaccine. More defined constructs where the covalent linking between sugar and the protein is limited to well-established attachment sites are highly desirable. They allow a better characterization of the product by physicochemical techniques facilitating the control of manufacturing consistency. Furthermore, the role of glycosylation site on the protein still has to be fully understood and such well-defined products can support studies to further investigate the relationship between glycoconjugate design and immune response. Hence, increasing attention has been made to control the conjugation site [4-6] and provide structurally-defined products not only in their saccharide component but also in the attachment point to the protein [6-9].

We have recently synthesized well-defined O-antigen (OAg)-based glycoconjugate vaccines to protect against nontyphoidal Salmonella (NTS) serovar Typhimurium, with CRM$_{197}$ as carrier protein [6]. NTS is the commonest cause of invasive bacteremia in Africa [10, 11], particularly affecting young children and HIV-infected adults, and OAg-based glycoconjugates represent a leading approach for the development of a vaccine against NTS [12-15].
Importantly, we found that site-selective single or double attachment of OAg to CRM197 was enough to generate levels of functional antibodies in mice, with titers comparable to the IgG induced by more complex random conjugates, and that the attachment site on the carrier protein plays a role on the immunogenicity.

In this study we compared four different orthogonal conjugation chemistries for the generation of selective OAg-based glycoconjugate vaccines against NTS, with the aim to identify efficient strategies to link a different number of sugar chains to defined amino acid sites on the protein. In all the constructs, S. Typhimurium OAg was end-terminally conjugated to CRM197 as carrier protein via the KDO (3-deoxy-D-manno-octulosonic acid) unit. CRM197 is a 58 kDa nontoxic mutant of diphtheria toxin and was selected as carrier protein for this investigation because of its defined structure and extensive use for licensed glycoconjugate vaccines and other vaccines in development [16, 17].

Two methods used refer to thiol chemistry, the thiol-maleimide addition and thioalkylation with halides, while two methods were based on the Huisgen 1,3-dipolar cycloaddition, which was investigated both copper-catalyzed and by the strain promoted variant. Click chemistry has been used for the synthesis of glycoconjugate vaccines only in recent times [6, 8, 9, 18], while the generation of a stable thioether bond by thiol-maleimide addition [19] or thioalkylation with halides [20, 21] are more common conjugation strategies which have been also used for the development of licensed glycoconjugate vaccines against Haemophilus influenza type b [19, 20].

Differently from thiol chemistry, strain promoted copper-free click chemistry resulted in an efficient coupling strategy both in regulating the glycosylation density on the final conjugate and in selectively targeting determined amino acid sites on the carrier protein.

### Materials and methods

#### Reagents

The following chemicals were used in this study: cystamine dihydrochloride, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDAC), (+)-sodium L-ascorbate, 2-(N-morpholino)ethanesulfonic acid (MES), N-hydroxysuccinimide (NHS), tris(2-carboxyethyl)phosphine hydrochloride solution (TECP), propargylamine, copper(II) sulfate pentahydrate (CuSO4·5H2O), tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), phosphate buffer solution (PBS), adipic acid dihydrazide (ADH), sodium cyanoborohydride (NaBH4CN), dimethyl sulfoxide (DMSO), sodium phosphate monobasic (NaH2PO4) [Sigma]; triplex III (EDTA) [Merck]; N-[β-maleimidopropionic acid] hydrazide trifluoroacetic acid salt (BMPH) [Thermo]; N-[ε-maleimidocaproyloxy]succinimide ester (EMCS), succinimidyl 3-(bromoacetamido)propionate (SBAP), NHS-PEG4-N3 [Pierce]. Click-easyTM BCN N-hydroxysuccinimide ester I (BCN NHS I) [Berry & Associates]; DL-1,4-Dithiothreitol (DTT) [Invitrogen]. CRM197 was obtained from Novartis Vaccines and Diagnostics (NV&D).

#### OAg purification and characterization

S. Typhimurium OAg was purified as previously described [22], following fermentation of the animal-derived isolate 2192, obtained from the University of Calgary. OAg resulted pure from proteins (<1% w/w by micro BCA), nucleic acids (<1% w/w by A260) and endotoxins (<0.1 UI/µg by LAL). 2192 OAg was fully characterized [23]: it showed an average molecular weight of 20.5 kDa based on HPLC-SEC analysis dRI profile with dextrans as standard, was 100% O-acetylated on C-2 abequose and 24% glucosylated. Amino groups were detected by TNBS colorimetric method [24, 25] probably as pyrophosphoethanolamine residues in the core region, with a molar ratio of 0.38 respect to N-acetyl glucosamine, unique sugar of the core [23].

#### Synthesis and characterization of derivatized CRM197

**CRM197-BMPH via EDAC chemistry**

CRM197 was solubilized in MES 500 mM pH 6.0 (12 mg/mL); BMPH (42 mg/mL, molar ratio BMPH/COOH groups CRM197 = 10.42) and EDAC (3 mg/mL, molar ratio EDAC/COOH groups CRM197 = 1.15) were added. Mixture was stirred for 1 h at RT, and then purified by desalting against NaH2PO4 100 mM EDTA 10 mM pH 7.0 on a HiPrep 26/10 desalting column 53 mL, prepacked with Sephadex G-25 Superfine (G-25 53 mL) [GE Healthcare].

**CRM197-BMPH via EDAC/NHS chemistry**

CRM197 was solubilized in MES 600 mM pH 6.0 (15.56 mg/mL); NHS (10.8 mg/mL, molar ratio NHS/COOH groups CRM197 = 5.36) and EDAC (6.2 mg/mL, molar ratio EDAC/COOH groups CRM197 = 1.83) were added and the solution mixed at RT for 1 h. After this time, BMPH (2.8 mg/mL, molar ratio BMPH/COOH groups CRM197 = 0.53) was added and the solution stirred for 2 h at RT. The mixture was purified by desalting on a PD-10 desalting column (PD 10) [GE Healthcare] against NaH2PO4 100 mM EDTA 1 mM pH 7.0.

**CRM197-EMCS**

CRM197 was solubilized in NaH2PO4 100 mM EDTA 1 mM pH 8.0 (4.7 mg/mL); EMCS was added (0.19 mg/mL, molar ratio EMCS/Lysine groups on CRM197 = 0.2) after being solubilized in DMSO (final DMSO concentration of 6% v/v).
Mixture was stirred for 2 h at RT, and then purified by desalting (G-25 53 mL column) against NaH$_2$PO$_4$ 100 mM EDTA 1 mM pH 7.0.

**CRM$_{197}$-SBAP**

CRM$_{197}$ was solubilized in NaH$_2$PO$_4$ 100 mM EDTA 1 mM pH 8.0 (4.7 mg/mL); SBAP was added (0.3 mg/mL, molar ratio SBAP/Lysine groups on CRM$_{197}$ = 0.3) after being solubilized in DMSO (final DMSO concentration of 4% v/v). Mixture was stirred for 3 h at RT, and then purified by desalting (G-25 53 mL column) against NaH$_2$PO$_4$ 100 mM EDTA 1 mM pH 7.0.

**CRM$_{197}$-N3** by controlled insertion of linkers on Lys (+3.8, +5.2, +7.1 and +10)

CRM$_{197}$ was solubilized in NaH$_2$PO$_4$ 400 mM pH 7.2 (20 mg/mL) and NHS-PEG4-N$_3$ was added (linker solubilized in DMSO with final DMSO concentration in the mixture of 1% v/v). Different amount of NHS-PEG4-N$_3$ were used to achieve a different degree of derivation on the protein with a molar ratio NHS-PEG4-N$_3$/Lysine groups CRM$_{197}$ of 0.18, 0.26, 0.39, respectively. After mixing at RT for 8 h, the mixture was purified by desalting on a G25 53 mL column against NaH$_2$PO$_4$ 100 mM pH 7.2.

**CRM$_{197}$-maleimide**

4-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)butyl)-3H-1,2,4-triazole-3,5(4H)-dione [26] (20 mM in CH$_3$CN, freshly prepared) was added dropwise, sequentially every minute in 10 copies, to a solution of CRM$_{197}$ at 4 °C (0.85 mg/mL in Tris HCl 500 mM pH 7.4), targeting tyrosine residues with a molar ratio to CRM$_{197}$ of 5:1. The mixture was stirred at 4 °C for 15 min and then desalted on Zeba 7K MWCO spin column [Pierce] with PBS pH 7.4 as the eluting buffer for three times.

**CRM$_{197}$-N3** by selective insertion of linkers on Lys (+1)

CRM$_{197}$-N$_3$ was synthesized as previously reported [6]. In brief, the linker ZQG-NH-PEG3-N$_3$ was first synthesized and characterized [6]. CRM$_{197}$ (32 mg/mL, 32 µL) was then added to ZQG-NH-PEG3-N$_3$ (2 mg/mL, 1 mL) in Tris 100 mM pH 8 and 100 µL of mTGase (stock of 50 mg/mL in PBS prepared from commercial 1% mTGase in maltocyclodextrin) were added. Reaction was incubated at 25°C for 18 h. The mixture was purified by size exclusion chromatography (SEC) on Superdex 200 10/300GL column, with PBS as running buffer. One addition of the linker was observed by Mass Spectrum. LCMS calculated: 58929; observed: [M + 1] 58929.

**Characterization of derivatized CRM$_{197}$**

Protein content was estimated by micro BCA (using BSA as standard and following manufacturer’s instructions [Thermo Scientifics]). HPLC-SEC analysis was used to compare derivatized protein with underivatized CRM$_{197}$. All samples were eluted on TSK gel G3000 PWXL column (30 cm x 7.8 mm; particle size 7 µm; cod. 808021) with TSK gel PWXL guard column (4.0 cm x 6.0 mm; particle size 12 µm; cod.808033) (Tosoh Bioscience). The mobile phase was 0.1 M NaCl, 0.1 M NaH$_2$PO$_4$, 5% CH$_3$CN, pH 7.2 at the flow rate of 0.5 mL/min (isocratic method for 30 min). Void and bed volume calibration was performed with λ-DNA (λ-DNA Molecular Weight Marker III 0.12–21.2 Kbp, Roche) and sodium azide (Na$_3$N, Merck), respectively. Protein peaks were detected at 214 nm and 280 nm (UV detection) and using tryptophan fluorescence (emission spectrum at 336 nm, with excitation wavelength at 280 nm). Linker average loading on CRM$_{197}$ was determined by MALDI-TOF analysis. For MALDI-TOF analysis, the protein was diafiltered, using a Centricon-10 (Millipore), against NaH$_2$PO$_4$ 10 mM pH 7.2. Two microliters of protein (at a concentration of 5 mg/mL) were mixed with 2 µL of a saturated solution of sinapinic acid in 50% acetonitrile solution containing 0.1% TFA. Two microliters of the mix were spotted on a MTP 384 stainless steel target (Bruker Daltonics GmbH, Bremen, Germany) and allowed to air-dry. Measurements were recorded on an Ultraflex III (Bruker GmBH) MALDI-TOF/TOF MS in linear mode. External calibration was performed by spotting 2 µL of protein calibration standard II (Bruker Daltonics) containing the following proteins: trypsinogen (23,982 Da), protein A (44,613 Da) and bovine serum albumin (66,431 Da). All mass spectra were recorded by summing up to 400 laser shots. The Flex Analysis software packages provided by the manufacturer were used for data processing.

**Synthesis of OAg-cystamine-CRM$_{197}$ conjugates**

40 mg of 2192 OAg were solubilized in NaH$_2$PO$_4$ 100 mM pH 7.0 (40 mg/mL) and then cystamine (110 mg/mL, cystamine/OAg (w/w) = 2.75) and NaBH$_3$CN (61 mg/mL, NaBH$_3$CN/OAg (w/w) = 1.53) were added. The mixture was stirred for 3 h at RT and then desalted against water on a G-25 53 mL column. Cystamine disulfide bond was reduced by mixing the OAg at a concentration of 20 mg/mL with DTT 100 mM in NaH$_2$PO$_4$ 100 mM EDTA 5 mM for 1 h at RT. The derivatized OAg was purified by desalting on a G-25 53 mL column against 10 mM NaH$_2$PO$_4$ 5 mM EDTA pH 7.5. The amiation reaction with cystamine was also performed at higher scale on 100 mg of OAg and the mixture purified by tangential flow filtration (10 k 200 cm$^2$ Hydrosart membrane, 10 diafiltration cycles vs. NaCl 1M, followed by 10 cycles against water).
Derivatized OAg intermediates were characterized by phenol sulfuric assay for sugar content [27] and by HPLC-SEC [23] for verifying absence of aggregation or degradation after modification. After the reaction with cystamine, the introduction of NH₂ groups was verified by TNBS colorimetric method [25] using 6-aminohexanoic acid as standard and subtracting the number of NH₂ groups already present on the un-derivatized OAg sample. Percentage of OAg chains activated was calculated as moles of linked cystamine/moles of OAg (calculated by HPLC-SEC/semicarbazide assay [22]) %. After the reduction with DTT, introduction of SH groups was verified by Ellman analysis [28]. Activation on the terminus KDO was calculated as moles of linked SH moles of KDO %, while the ratio % between SH groups/cystamine moles gave the efficiency of reduction reaction with DTT.

Conjugation was performed solubilizing the derivatized OAg in NaH₂PO₄ 100 mM EDTA 1 mM pH 7.2 (10 mg/mL), using a molar ratio of thiol groups to CRM 197 of 30 to 1. After mixing ON at RT, the mixture was purified by size exclusion chromatography with a Sephacryl S-300 column 1.6 cm x 90 cm [GE Healthcare], eluting with PBS pH 7.2 at 0.5 mL/min.

**Synthesis of OAg-propargylamine-CRM₁₉₇**

OAg (40 mg) was solubilized in NaH₂PO₄ pH 7.0 (40 mg/mL), then propargylamine (27.5 mg/mL, propargylamine/OAg (w/w) = 0.67) and NaBH₃CN (62.8 mg/mL, NaBH₃CN/OAg (w/w) = 1.53) were added. The reaction was stirred for 3 h at RT and purified by desalting on a G-25 53 mL column against water. Derivatized OAg intermediates were characterized by phenol sulfuric assay for sugar content [27] and by HPLC-SEC [23] for verifying absence of aggregation or degradation after modification.

**Conjugation was performed adding by OAg-propargylamine to CRM₁₉₇-N₃ (5 mg/mL, average of 5.2 Lysines activated) in NaH₂PO₄ pH 7.2 (40 mM) EDTA 1 mM pH 7.2 with THPTA 25 mM, sodium ascorbate 10 mM and CuSO₄·5H₂O 5 mM, with a molar ratio alkyn/azide = 5. Reaction was stirred at RT for 6 h and conjugate formation followed by HPLC-SEC (TSK gel 6000PW + TSK gel 5000PW).

**Synthesis of OAg-ADH-BCNesterI-CRM₁₉₇**

OAg (40 mg) was solubilized in AcONa 100 mM pH 4.5 (40 mg/mL), then ADH (48 mg/mL, ADH/OAg (w/w) = 1.2) and NaBH₃CN (48 mg/mL, NaBH₃CN/OAg (w/w) = 1.2) were added. The mixture was stirred for 2 h at 30 °C and then purified by desalting on a G-25 53 mL column against water.

For introduction of the second linker, BCN NHS ester I, OAg-ADH was dissolved in water/DMSO 1:9 (v/v) at a concentration of 50 mg/mL. When the derivatized OAg was completely solubilized, TEA was added (molar ratio TEA/total NH₂ groups = 5; total NH₂ groups included both phosphoethanolamine groups on the OAg and the hydrazide groups introduced with the linker ADH) followed by Click easy BCN NHS ester I (molar ratio BCN NHS ester I/total NH₂ groups = 12). The solution was mixed at RT for 3 h. The sample was purified by desalting on a G-25 53 mL column against water.

Derivatized OAg intermediates were characterized by phenol sulfuric assay for sugar content [27] and by HPLC-SEC [23] for verifying absence of aggregation or degradation after modification. Introduction of NH₂ groups was verified by TNBS colorimetric method [24, 25] using ADH as standard and subtracting the number of NH₂ groups already present on the underivatized OAg sample. Free ADH was detected by RP-HPLC [29]. The percentage of OAg chains activated was calculated as moles of linked ADH (NH₂)/moles of KDO % [22].

Total alkyne groups introduced with BCN NHS I were quantified by TNBS considering the residual number of unreacted NH₂ groups remained after this derivatization step. Percentage of derivatization with BCN NHS I was calculated as molar ratio percentage of linked alkyne groups/total NH₂ groups by TNBS before derivatization, indicating the percentage of moles of NH₂ groups activated with this reaction.

For conjugation, CRM₁₉₇-N₃ at a concentration of 10 mg/mL in PBS was added of OAg-ADH-BCNesterI (molar ratio alkyne/azide = 4). Mixture was stirred for 6 h at RT. Conjugate purification was performed by hydrophobic interaction chromatography on a Phenyl HP column [GE Healthcare], loading 500 μg of protein for mL of resin in 50 mM NaH₂PO₄ 3M NaCl pH 7.2. The purified conjugate was eluted in water and the collected fractions were dialysed against 10 mM NaH₂PO₄ pH 7.2.

**OAg-CRM₁₉₇ conjugates characterization**

Total saccharide was quantified by phenol sulfuric assay [27], protein content by micro BCA and the ratio of saccharide to protein calculated. OAg-CRM₁₉₇ conjugates profiles were compared with free CRM₁₉₇ and free OAg by HPLC-SEC. All samples were eluted on a TSK gel 6000PW (30 cm x 7.5 mm) column (particle size 17 μm; Sigma S-05765) connected in series with a TSK gel 5000PW (30 cm x 7.5 mm) column (particle size 17 μm; Sigma S-05764) with TSK gel PWH guard column (7.5 mm ID x 7.5 cm L; particle size 13 μm; Sigma S-06732) (Tosoh Bioscience). The use of the two columns in series gave better separation of conjugate from free saccharide and protein, allowing the conjugate to enter into the column. The mobile phase was 0.1 M NaCl, 0.1 M NaH₂PO₄, 5% CH₃CN, pH 7.2 at the flow rate of 0.5 mL/min (isocratic method for 60 min). OAg peaks were detected by dRI, while UV detection at 214 nm and 280 nm was used for free protein and conjugate detection. Protein and conjugate peaks were also detected using tryptophan fluorescence (emission spectrum at 336 nm, with excitation wavelength at 280 nm). Free
protein was estimated by HPLC-SEC, running a calibration curve of the unconjugated protein in the range 5–50 µg/mL under the same conditions as for the conjugate. The percentage of unconjugated CRM197 was calculated by dividing the amount of free protein detected by HPLC-SEC by the total amount of protein quantified in the sample by micro BCA. Free saccharide was verified to be lower than 20% by comparing eventual peak of free OAg in the conjugate with the peak of an OAg standard injected at the concentration corresponding to 20% of the total sugar amount as estimated by phenol sulfuric assay.

Results

Thiol conjugation chemistry

Thiol-conjugation chemistry was initially tested. The KDO moiety was targeted for introducing cystamine linker at the reducing end of the OAg chain without modifying the repeating units structure. After DTT-reduction, the activated OAg was conjugated to the carrier protein (Fig. 1).

OAg derivatization with cystamine was characterized by sugar recovery higher than 80%. TNBS analysis indicated that 70–80% of OAg chains were activated. After DTT addition, sugar recovery was again higher than 75% and a complete reduction of the thiols was confirmed by the analysis of thiol groups introduced. CRM197 was derivatized using different hetero bi-functional linkers, with a functional group able to react with the thiol unit at the end of the OAg chain. Different amino acids were targeted, also trying to result in variable protein loading (Fig. 2A; Table 1). In all cases, protein recovery was higher than 90%. Overall, derivatization on Lysine residues allow to introduce the highest number of linkers per mole of protein, respectively 8 using SBAP and 11.3 using EMCS (1–2, Table 1). Using controlled conjugation chemistries, an average of 3.1 lysines were instead targeted with EMCS (3, Table 1). Reaction of CRM197 with BMPH by EDAC chemistry resulted in the introduction of an average number of 3.4 linkers per protein (4, Table 1). When NHS was added to EDAC for activating the COOH groups on the protein, trying to increase the activation degree, an average of 5.1 linkers was introduced (5, Table 1). Tyrosines were also modified obtaining an average loading of 4.1 linkers per protein (6, Table 1).

Thiolated OAg was then conjugated to derivatized CRM197 characterized by different amino acids activated (Lys, Tyr or Glu/Asp) and different linker loading. Using this conjugation chemistry, it was not possible to obtain conjugate formation unless the number of linkers per protein was at least 8 (1–2, Table 1; Fig. 1).

The two conjugation reactions that succeeded, both based on lysine chemistry, were characterized by no presence of unreacted protein in the conjugation mixture. Even if the number of linkers per CRM197 was high (8 and 11.3), corresponding conjugates were characterized by a low OAg to protein molar ratio (1.6 and 2.7 respectively) (1–2, Table 1).

Looking at the HPLC-SEC profiles of the conjugation mixtures, even when conjugation was successful, the presence of a sugar population at higher molecular weight was observed, probably deriving from oxidation of the thiolated OAg to generate a sugar dimer (not shown). Different methods were attempted to avoid the generation of the oxidation product, and hence of an unreactive OAg form, to see if this could then result in higher conjugation efficiency. Conjugation of DTT-treated OAg-cystamine with CRM197-BMPH under N2 caused precipitation of the protein; one-pot-two-step OAg-cystamine reduction with TCEP as reducing agent and conjugation to CRM197-BMPH (5.2 linkers) was also not successful. When DTT-treated cystamine was conjugated with CRM197-BMPH in the presence of TCEP, the formation of the disulfide aggregate was avoided, however most of protein did not conjugate.

In summary thiol conjugation strategies did not allow to get all the panel of wanted conjugates and it became needed to identify a different conjugation approach.

Click chemistry

There are two main ways of performing the click 1,3-dipolar reaction: by using metal as a catalyst (usually copper), or alternatively lowering the activation barrier for [3 + 2] cycloaddition by employing intrinsically highly strained cyclic alkenes that readily react with azide groups. For starting, both the approaches were compared for the reaction of the alkyne-derivatized S. Typhimurium OAg with CRM-N3.

For the copper click chemistry, the OAg was derivatized with propargylamine by reductive amination on the terminal KDO (Fig. 3A). For the copper-free approach instead, the alkyne linker was introduced on the OAg-ADH, with activation higher 80% (Fig. 3B).

The inclusion of azido groups on CRM197 was initially performed by targeting lysines using NHS-PEG4-N3 (Fig. 2B, controlled modification). Different loadings were obtained, in the range of 3.8–10 linker introduced per protein, depending on the amount of linker added. In all cases, protein recovery was higher than 90% (Table 2). CRM197 with an average number of 5.2 linkers introduced (CRM197-N3[LY5.2]) was used at 5 mg/mL for comparing these two approaches.

Despite a lower sugar to protein ratio, the copper free conjugation was more efficient (Table 3). In both cases, reaction time did not seem to have a strong impact on conjugate formation (Table 3). Because of the higher conjugation efficiency and to avoid the use of a toxic metal, the copper-free approach was selected for further experiments.
Maintaining protein concentration at 5 mg/mL and increasing the alkyn to azide molar ratio from 1 to 2, conjugation efficiency was not altered after 6 h of reaction, with 78% of CRM\textsubscript{197} conjugated (Table 4). Increasing protein concentration from 5 to 10 mg/mL and the alkyn to azide molar ratio from 2 to 4, conjugation became quantitative. When a protein with a higher average number of linkers was used (from 5.2 to 10), no free CRM\textsubscript{197} was detected in the conjugation mixture (Table 4; Fig. 4).

The optimal conditions identified here for click-chemistry conjugation were applied for the synthesis of a panel of glycoconjugates differing for type of amino acid targeted and number of sites activated on the protein (Table 5; Fig. 2B) [6]. In order to compare the conjugates generated by the free-metal click and the thiol conjugation chemistry with a similar linker loading on the carrier protein, CRM\textsubscript{197} with an average number of 7.1 linkers introduced (CRM\textsubscript{197-N3(LYS7.1)}) was considered (7, Table 5). As previously reported, the conjugation to the alkyn-derivatized OAg was successful with all CRM\textsubscript{197} conjugated and residual free OAg removed by purification through HIC (Table 5) [6]. The conjugate was characterized by a similar OAg to protein molar ratio (2.0) to what previously observed with thiol chemistry on activated lysine groups on CRM\textsubscript{197}. We have determined that click chemistry can be used for the synthesis of conjugate vaccines when the linker loading is low targeting both lysines and tyrosine that were immunogenic in mice (8–10 Table 5; Fig. 2B) [6]. Importantly, the reaction conditions allowed the conjugation with only one linker CRM\textsubscript{197-N3(LYS1)}, where 38% of CRM\textsubscript{197} was conjugated after 6 h using a chemoenzymatic approach (8 Table 5; Fig. 2B selective modification).

As previously observed with CRM\textsubscript{197-N3(LYS5.2)}, the conjugation of CRM\textsubscript{197-N3(LYS1)} is fast as increasing the reaction time from 2 to 6 h did not impact the conjugate formation (Table 6).

**Discussion**

The design of structurally defined glycoconjugates brings advantages in terms of physicochemical characterization of the vaccine and could lead to the selection of candidates with enhanced efficacy. With this regard, we have recently demonstrated that the conjugation site plays a role in determining the immunogenicity in mice and one single attachment point may be sufficient to induce high levels of bactericidal antibodies [6]. The use of structurally defined constructs with epitopes displayed at precise sites is also desirable to better understand the antigen presentation to the immune system and to rationalize the development of efficacious glycoconjugates vaccines. However, selective
Fig. 2 Activation of the carrier protein. (A) Thiol chemistry: CRM$_{197}$ was derivatized on Lys, Tyr or Glu/Asp using different hetero bi-functional linkers able to react with cysteamine-activated OAg. (B) Click Chemistry: CRM$_{197}$ was derivatized on Lys or Tyr with NHS-PEG4-N$_3$ to introduce azido groups. Lysine activation was performed by selective (Lys + 1) or controlled modification (Lys + 1).

Table 1 OAg-CRM$_{197}$ conjugation by thiol chemistry

| Conjugate                  | Chemistry      | Target | Average CRM$_{197}$ labeling | OAg chains per CRM$_{197}$ | Conjugate (%) |
|---------------------------|----------------|--------|-----------------------------|---------------------------|---------------|
| 1 - OAg-cysteamine-SBAP-CRM$_{197}$ | Thiol (Alkylation) | Lys    | +8.0                        | 1.6                        | 100           |
| 2 - OAg-cysteamine-EMCS-CRM$_{197}$ | Thiol (Addition)      | Lys    | +11.3                       | 2.7                        | 100           |
| 3 - OAg-cysteamine-EMCS-CRM$_{197}$ | Thiol (Addition)      | Lys    | +3.1                        | -                          | No            |
| 4 - OAg-cysteamine-BMPH-CRM$_{197}$ | Thiol (Addition)      | Glu/Asp| +3.4                        | -                          | No            |
| 5 - OAg-cysteamine-(NHS)BMPH-CRM$_{197}$ | Thiol (Addition)      | Glu/Asp| +5.1                        | -                          | No            |
| 6 - OAg-cysteamine-maleimide-CRM$_{197}$ | Thiol (Addition)      | Tyr    | +4.1                        | -                          | No            |
chemistries are usually characterized by lower conjugation efficiency than random ones.

The possibility of applying click chemistry [30] in bioconjugation was first demonstrated for the preparation of peptidotriazoles via solid phase synthesis [31]. Recently, it has been applied to the synthesis of glycoconjugate vaccines [5, 6, 18, 32]. One of these reactions, the Huisgen 1,3-dipolar cycloaddition [33], has become the most popular click reaction, especially because it can proceed rapidly at room temperature by employing copper as a catalyst [31, 34]. It is characterized by mild reaction conditions, high yields and simple work-up, selectivity, specificity and can be usually performed in water. One of the most important properties of the click chemistry is its bio-orthogonality. The azide moiety is

![Fig. 3 Click conjugation chemistry. Two conjugation strategies: (A) Copper-free click reaction. (B) Copper-catalyzed click reaction](image)

### Table 2

| Linker:lysine ratio (mol/mol) | Theoretical Lys labelling* | Average Lys labelling |
|------------------------------|---------------------------|----------------------|
| 0.14                         | + 5.5                     | + 3.8                |
| 0.18                         | + 7.0                     | + 5.2                |
| 0.26                         | + 10.1                    | + 7.1                |
| 0.39                         | + 15.2                    | + 10                 |

*Calculated considering 39 lysines on CRM197

### Table 3

| Copper | Alkyne:azide ratio (mol/mol) | Reaction time (h) | Conjugate (%) |
|--------|------------------------------|-------------------|---------------|
| Yes    | 5                            | 2                 | 28            |
|        | 4                            | 4                 | 36            |
|        | 6                            | 6                 | 40            |
| No     | 1                            | 2                 | 69            |
|        | 4                            | 4                 | 75            |
|        | 6                            | 6                 | 78            |

CRM197-N3(LYS5.2) at 5 mg/ml
absent in almost all naturally existing compounds, lacks reactivity with natural biomolecules and undergoes ligation only with a limited set of functionalities, such as alkyne groups. The sulfhydryl group is a popular target in many modification strategies. The frequency of sulfhydryl occurrence in proteins or other molecules is usually very low compared to other groups like amines or carboxylates. Therefore, the use of sulfhydryl-reactive chemistries can also restrict modification to only a limited number of sites within a target molecule.

In this study we have screened different click and thiol conjugation chemistries with the aim to identify a powerful conjugation methodology for the synthesis of site-selective glycoconjugates targeting one single point of the saccharide chain and one-to-few points of attachment on the carrier protein. We have used S. Typhimurium OAg and CRM197 as carbohydrate hapten and carrier protein models. It is worth mentioning that S. Typhimurium OAg and CRM197 are large molecules (20.5 kDa and 58.4 kDa, respectively), therefore the possibility to design site-selective constructs is hardly challenged by steric hindrance factors. Click chemistry has proved to be a more powerful tool in this sense, allowing the synthesis of conjugates even where only one position on the protein was available for linkage to the KDO unit of the sugar chain. When using thiolchemistry, we were not successful in obtaining glycoconjugates unless the number or linkers on the protein was at least 8. Possible reasons to explain this behavior could be related with the instability of the functional groups on the linker in the presence of CRM197 and the verified oxidation of thiolated OAg in the conjugation conditions.

Indeed, insertion of thiol groups in meningococcal C oligomers followed by thiol-maleimide conjugation on CRM197 has been previously shown feasible, although the conjugation with CRM197-EMCS (6.5 linker per protein) reaction linked on average only 1.8 oligosaccharides per protein. In our work we have confirmed that conjugation of S. Typhimurium OAg to CRM197-EMCS is feasible, provided that a high number of linkers per protein are present (+11.3), resulting in a conjugate bearing about 2.7 OAg chain per protein. We did not observe conjugation in the presence of only 3.1 linker per protein. While in our investigation we did not test CRM197-EMCS with an average of 6.5 linkers and we have used a longer sugar model (20.5 vs. 4.5 kDa), we think that both finding underline how, in the presence of thiol chemistry, the synthesis of conjugate vaccine with a controlled number of glycopeptide linkages may present some issues [35]. In addition, work from Nilo et al. confirmed that thiol-maleimide chemistry resulted in lower sugar loading on CRM197 compared to copper-free conjugation [26].

Copper-free chemistry was preferred because it produced better conjugation yields in preliminary tests and avoids the use of a toxic metal. Differently from the thiol chemistry tested, copper-free click chemistry allowed the linkage of the terminal sugar end of the OAg chain to the protein with high conjugation yields (higher than 80%) even when few linkers were present on the protein (average number of 4) and to have conjugate formation with just one linker on the protein (38% of CRM197 conjugated). It is important also to consider that the conditions tested for the thiol chemistry have an intrinsic bias, since they used higher concentration of both OAg and protein, and higher molar ratio of active groups of OAg per linker on the carrier protein. The click conjugation proceeds with a fast rate with the reaction completed in only 2 h even when only

**Table 4** Reaction conditions optimized for the copper-free conjugation

| Azide per CRM197 | [CRM197] (mg/mL) | Alkyne:azide (mol/mol) | Conjugate (%) |
|-----------------|-----------------|------------------------|--------------|
| 5.2             | 5               | 1:1                    | 78           |
| 5.2             | 10              | 2:1                    | 78           |
| 5.2             | 10              | 4:1                    | 90           |

Reaction time is 6 h

![Fig. 4 HPLC-SEC profiles of OAg-ADH-BCNesterI-CRM197 conjugation with CRM197-N4(LYS5.2), green, or CRM197-N4(LYS10), blue, in comparison to free CRM197, black. Conjugation mixture were analyzed after 6 h at RT, using carrier protein at 10 mg/mL and alkyne/azide molar ratio of 4](image-url)
one linker is present on the carrier protein. However, the linkers used for the synthesis of click and thiol conjugates were different, and an impact of the type of linker on conjugate formation cannot be excluded. While it has been previously reported that the linkers used for thiol-maleimide addition induced a lower anti-linker response compared to the cyclooctene ring generated by SPAAC [26], our group has previously shown how conjugate with even only one sugar moiety per protein synthesized by copper-free click reaction can induce strong immunogenicity [6].

In order to generalize the outcome of the comparison between thiol and click chemistries, different sugar and – more importantly - protein models should be investigated. While we would expect different carrier proteins to behave differently, we think that the outcome of the current investigation will likely apply to the diphtheria toxoid (DT), another commonly used carrier protein, structurally related to CRM 197.

In conclusion, we compared thiol and click chemistry for the synthesis of site-selective conjugates on the OAg. Click chemistry allows the synthesis of glycoconjugates where different amino acid can be targeted and in a different number, impacting the number of sugar present per carrier protein and the immunogenicity [6]. Click chemistry is therefore a powerful tool for the synthesis of glycoconjugate vaccines as it allows the investigation of different parameters important for the immunogenicity.

Table 5  OAg-CRM197 conjugation by click chemistry

| Conjugate | Chemistry | Target | Average CRM197 labeling | OAg chains per CRM197 | Conjugate (%) |
|-----------|-----------|--------|------------------------|-----------------------|---------------|
| 7 - OAg-ADH-BCNesterl-CRM197 Click (Copper-free) | Lys | + 7.1 | 2.0 | 100 |
| 8 - OAg-ADH-BCNesterl-CRM197 Click (Copper-free) | Lys | + 1 | 0.7 | 31 |
| 9 - OAg-ADH-BCNesterl-CRM197 Click (Copper-free) | Lys | + 3.8 | 1.5 | 91 |
| 10 - OAg-ADH-BCNesterl-CRM197 Click (Copper-free) | Tyr | + 4.3 | 3.7 | 81 |

*Data from [6]*

Acknowledgements  The authors thank Dr. Qi-Ying Hu and Zachary Robinson (former Novartis Institutes for BioMedical Research) for assistance during the synthesis of the compounds OAg-ADH-BCNesterl-CRM197 (Lys + 1) and OAg-cysteamine-maleimide-CRM197.

Compliance with ethical standards

Conflicts of interest  The authors declare that they have no conflicts of interest.

Ethical approval  This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Avci, F.Y., Kasper, D.L.: How bacterial carbohydrates influence the adaptive immune system. Annu. Rev. Immunol. 28:107–130 (2010) https://doi.org/10.1146/annurev-immunol-030409-101159
2. Pace, D.: Glycoconjugate vaccines. Expert. Opin. Biol. Ther. 13(1): 11–33 (2013) https://doi.org/10.1517/14712598.2012.725718 (fdoi)
3. Costantino, P., Rappuoli, R., Bertil, F.: The design of semi-synthetic and synthetic glycoconjugate vaccines. Expert. Opin. Drug Discov. 6(10):1045–1066 (2011) https://doi.org/10.1517/17460441.2011.609554
4. Bernardes, G.J., Castagner, B., Seeger, R.H.: Combined approaches to the synthesis and study of glycoproteins. ACS Chem. Biol. 4(9):703–713 (2009) https://doi.org/10.1021/cb900014n
5. Adamo, R., Nilo, A., Castagner, B., Bouteira, O., Bertil, F., Bernardes, G.J.L.: synthetically defined glycoprotein vaccines: current status and future directions. Chem. Sci. 4:2995 (2013)
33. Huisgen, R.: Kinetics and Mechanism of 1,3-Dipolar Cycloadditions. Angew. Chem. Int. Ed Engl. 2(11), 633–645 (1963) https://doi.org/10.1002/anie.196306331
34. Rostovtsev, V.V., Green, L.G., Fokin, V.V., Sharpless, K.B.: A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. Angew. Chem. Int Ed Engl. 41(14), 2596–2599 (2002) https://doi.org/10.1002/anie.200200715
35. Adamo, R., Nilo, A., Harfouche, C., Brogioni, B., Pecetta, S., Brogioni, G., Balducci, E., Pinto, V., Filippini, S., Mori, E., Tontini, M., Romano, M.R., Costantino, P., Berti, F.: Investigating the immunodominance of carbohydrate antigens in a bivalent unimolecular glycoconjugate vaccine against serogroup A and C meningococcal disease. Glycoconj. J. 31(9), 637–647 (2014) https://doi.org/10.1007/s10719-014-9559-1

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.