DOI 10.1002/pmic.201400462 *Proteomics* 2015, *15*, 2756–2765

RESEARCH ARTICLE

Lysine conjugation properties in human IgGs studied by integrating high-resolution native mass spectrometry and bottom-up proteomics

Violette Gautier[1,2*], *Anja J. Boumeester*[1,3*], *Philip Lössl*[1,2*] *and Albert J. R. Heck*[1,2]

[1] Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Utrecht, The Netherlands
[2] Netherlands Proteomics Center, University of Utrecht, Utrecht, The Netherlands
[3] AbLab, University of Utrecht, Utrecht, The Netherlands

Antibody–drug conjugates (ADCs) are a novel class of biopharmaceuticals several of which are now being investigated in clinical studies. In ADCs, potent cytotoxic drugs are coupled via a linker to reactive residues in IgG monoclonal antibodies. Linkage to lysine residues in the IgGs, using *N*-hydroxysuccinimide ester based chemistry, is one of the possible options. To control drug load and specificity, proper knowledge is required about which lysine residues are most accessible and reactive. Here, we combine native MS and bottom-up proteomics to monitor the overall drug load and site-specific lysine reactivity, using *N*-hydroxysuccinimide-based tandem mass tags. High-resolution Orbitrap native MS enables us to monitor and quantify, due to the achieved baseline resolution, the sequential incorporation of up to 69 tandem mass tag molecules into human IgGs. Complementary, bottom-up proteomics facilitates the identification of some very reactive "hot-spot" conjugation sites. However, we also identify lysine residues that are highly resistant to chemical labeling. Our integrated approach gives insight into the conjugation properties of IgGs at both the intact protein and residue levels, providing fundamental information for controlling drug load and specificity in lysine-linked ADCs.

Received: September 30, 2014
Revised: December 2, 2014
Accepted: January 13, 2015

Keywords:
Antibody–drug conjugates / Covalent chemical labeling / IgGs / Lysine conjugation / Native mass spectrometry / Technology

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Monoclonal antibodies (mAbs) and, more recently, also antibody–drug conjugates (ADCs) are a rapidly emerging class of biopharmaceuticals with more than 30 novel therapeutics currently being investigated in clinical studies [1, 2]. ADCs consist of three components: a mAb that is specific to a particular antigen expressed on the surface of malignant cells, a potent cytotoxic agent, and a linker that enables the covalent attachment of the drug to the antibody carrier [3, 4]. This design facilitates the delivery of a small drug molecule to a specific cellular target, allowing targeted treatment of

Correspondence: Professor Albert J. R. Heck, Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Kruytgebouw Room O603, Padualaan 8, 3584 CH Utrecht, The Netherlands
E-mail: a.j.r.heck@uu.nl
Fax: +31-0-30-253-6919

Abbreviations: **ADC**, antibody–drug conjugate; **G0F**, glycan chain with three mannoses, four *N*-acetylglucosamines, and one fucose; **G1F**, glycan chain with three mannoses, four *N*-acetylglucosamines, one fucose, and one galactose; **IgG4Δhinge**, IgG4 mutant antibody lacking the hinge region; **NHS**, *N*-hydroxysuccinimide; **SASA**, solvent-accessible surface area; **TMT**, tandem mass tag

*These authors contributed equally to this work.
Colour Online: See the article online to view Fig. 3 in colour.

cancer or other diseases with reduced systemic toxicity compared to traditional chemotherapy. Typically, drugs are conjugated to IgG, which exists as a dimer of heavy chain/light chain pairs that are connected by disulfide linkages and noncovalent interactions between their CH3 domains. IgG1 is the most common subclass employed in ADCs and also IgG4 and, to a lesser extent, IgG2 are utilized since they exhibit different structures and functions [5]. For instance, IgG4 has the unique capacity to form half-bodies (one heavy/light chain pair) and undergo half-body exchange [6], a property influencing its immunological functions and therefore the development of IgG4-based therapeutics [7, 8]. An array of strategies has been developed to conjugate cytotoxic molecules to the mAb. Originally, the linkers were attached to either the lysine amino groups, typically using *N*-hydroxysuccinimide (NHS) ester derivatives, or the thiol groups of interchain cysteines [9, 10]. More recently, cleverly engineered mAbs that allow site-specific drug conjugation are gaining popularity in ADC development [11]. However, most ADCs in clinical development as well as the recently US Food and Drug Administration approved ADCs brentuximab vedotin (Adcetris) [12] and trastuzumab emtansine (Kadclya) [13] use the traditional linking chemistry [11, 14]. The resulting ADC populations are generally heterogeneous with respect to payload and conjugation sites of the drug [11, 15], leading to different pharmacokinetics, efficiency, or toxicological properties [16]. This is particularly the case for lysine-directed ADCs, as all lysine residues present in the IgGs may react with the linker and site-specificity is directly governed by the reactivity of the lysine residues, which depends on their local environment, structural flexibility, and solvent accessibility. IgGs contain on average 80 lysine residues, almost half of which are partially modified in the lysine-linked ADC huN901-DM1, as shown by mass spectrometric peptide mapping [15]. To better control the amount of drug molecules attached to an IgG, and the sites they occupy, a sound understanding of global and site-specific conjugation properties of these molecules is essential. Additionally, such information may open up new opportunities to further improve existing ADCs, as for instance highly reactive lysine residues in unfavorable positions (e.g., in the complementary determining region of the antibody) could selectively be mutated to nonreactive residues.

Owing to its sensitivity and versatility, MS-based approaches have emerged as powerful tools in structural biology [17]. Among them, covalent labeling techniques are being used to exploit the reactivity differences between solvent exposed and buried amino acids to map the solvent-accessible surface area (SASA) of proteins [18]. Applied strategies range from nonspecific labeling—for example, hydroxyl radical protein labeling [19, 20]—to the modification of specific residues. Lysine is one of the most commonly targeted amino acids because it is prevalently located on the protein surface and bears a highly reactive primary amino group. The most popular amine-reactive labeling reagents are *N*-acetylsuccinimide [21] and NHS ester derivatives, such as NHS-biotin [22]. Tandem mass tags (TMT) [23], which are often used as isobaric labeling reagents in quantitative proteomics, rely on the same NHS ester chemistry. Next to the NHS group, TMT reagents contain a spacer arm and an isobaric MS/MS reporter. These isotopically labeled reporter groups facilitate multiplex MS/MS experiments, allowing, for example, the determination of labeling kinetics in a single experiment [24].

Recently, we described a method integrating high-resolution native MS and bottom-up proteomics to simultaneously probe kinetics and site-specificity of protein phosphorylation [25]. Here, we explore this strategy further and combine it with TMT labeling to assess the reactivity of lysine residues on two antibodies—that is, IgG1 and IgG4, both of which are of interest for ADC development. Since both the TMT reagent and the lysine-specific ADC linkers (e.g., succinimidyl-4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate in trastuzumab emtansine) represent NHS esters of similar size, TMT labeling should closely resemble the conjugation properties of IgGs during ADC synthesis. At the intact protein level, the incorporation of a TMT molecule can be identified by a mass shift of nominally 229 Da, with regard to the IgG mass of around 150 kDa. The excellent resolving power of the Orbitrap mass analyzer with extended mass range [26] enabled us to monitor sequential conjugation of up to 69 TMT molecules. Moreover, lysine conjugation could be qualitatively and quantitatively assessed for both the deglycosylated and glycosylated IgG1 in a glycoform-specific manner. We next studied an IgG4 mutant antibody lacking the hinge region (IgG4Δhinge) that naturally harbors the interchain disulfide bonds. Dimerization of IgG4Δhinge solely depends on noncovalent interactions and can be considered a model for IgG4 with reduced disulfides, which has been suggested to represent an intermediate during half-body exchange [7]. The native-like setting during MS analysis allowed us to probe label incorporation into half and full IgG4Δhinge antibodies concurrently and study the dynamic equilibrium between these two forms. We complemented the analysis with a bottom-up approach using pepsin instead of trypsin as protease to avoid complication with modified lysine residues during digestion. By applying 5-plex stable isotope labeled TMT, we could quantify site-specific label incorporation on IgG1 at five different TMT concentrations in parallel. This facilitated the identification of highly reactive lysine residues, which represent the preferential conjugations sites within IgG1-based ADCs, as well as several lysine residues that are essentially resistant to conjugation under native conditions.

2 Materials and methods

2.1 Sample preparation

IgG1 and IgG4Δhinge, both produced in HEK293T cells, were kindly provided by Genmab and expressed and purified as described previously [7, 27]. The corresponding amino acid sequences are shown in Supporting Information Fig. 1. In some cases, the IgG antibodies were deglycosylated prior to TMT labeling by incubating 25 μg of protein with one unit of

N-glycosidase F (PNGaseF; Roche Diagnostics, Mannheim) at 37°C overnight. TMT labeling was performed by incubating varying concentrations of the TMT^6 reagents (12 μM of TMT^6–127, 60 μM of TMT^6–128, 120 μM of TMT^6–129, 450 μM of TMT^6–130, 900 μM of TMT^6–131) with 25 μg of IgG in 20 mM triethylammonium bicarbonate at room temperature. The reactions were quenched after 1 min by adding hydroxylamine (5 wt% in 20 mM triethylammonium bicarbonate). Subsequently, the samples were buffer exchanged to either 50 mM ammonium bicarbonate with 6 M urea for bottom-up proteomics analysis or 150 mM ammonium acetate for native MS analysis.

2.2 Native MS

Samples were analyzed using an Exactive Plus extended mass range Orbitrap instrument (Thermo Fisher Scientific, Bremen, Germany), operating in positive-ion mode. Protein solutions were electrosprayed at an estimated concentration of 4 μM using a nanospray ionization (term used by Thermo Fischer Scientific, equivalent to nano-ESI) source and gold-coated borosilicate glass needles. MS settings were kept similar for all measurements: 1.2–1.4 V spray voltage, 50–100 V source fragmentation, 250°C source temperature, 10–30 ms injection time, 50–100 V collision energy, and 17 500–70 000 resolution at m/z 200. Nitrogen gas pressure in the higher energy collisional dissociation (HCD) cell was $5–8 \times 10^{-10}$ bar. Voltage offsets on flatapoles, transport multipole, and ion lenses were manually tuned. The instrument was calibrated by using CsI clusters.

The accurate proteoform masses were obtained by averaging the masses of the charge state envelopes, taking only charge states with >1% relative signal intensity. Three major glycoforms were identified for glycosylated IgG1: G0F/G0F with both glycan chains containing three mannoses, four *N*-acetylglucosamines, and one fucose; G1F/G0F with one additional galactose on one chain; G1F/G1F with additional galactose on both chains. The differentially TMT-labeled forms of the antibodies were manually assigned in the native mass spectra by comparing them to the theoretical masses of deglycosylated IgG1 (145 135.3 Da), glycosylated IgG1 (G0F/G0F 148 026.2 Da, G1F/G0F 148 188.4 Da, G1F/G1F 148 350.6 Da), and deglycosylated IgG4Δhinge (143 423.16 Da), considering a TMT-induced mass shift of 229.263 Da. Proteoform abundances were quantified by adding up the signal intensities of the corresponding charge states and dividing them by the total intensity of all proteoforms. This allowed calculation of the weighted average of TMT incorporations. Relative abundances and experimental masses of the proteoforms were taken as input to generate zero-charge state spectra with Gaussian-distributed peaks using *R*.

2.3 Bottom-up proteomics analysis

IgG1 samples labeled with increasing concentrations of TMT were mixed in a 1:1:1:1:1 ratio and, additionally, a fraction of the sample incubated with the highest concentration of TMT (TMT-131) was kept apart to be analyzed individually. The samples were reduced (2 mM DTT), alkylated (4 mM iodoacetamide), diluted to 2 M urea, and acidified with 0.08 M HCl to reach a pH of 1.5–2 before being digested with pepsin for 2 h at 37°C (pepsin:protein ratio 1:20). After desalting, the samples were analyzed by nanoLC–MS/MS on an Orbitrap Q-Exactive essentially as described previously [28]. MS raw data were analyzed with Proteome Discoverer version 2.0 beta (Thermo Fisher Scientific). Further details about nanoLC–MS/MS analysis and data analysis are provided in the Supporting Information.

2.4 Comparative modeling of IgG1

Comparative structural models of the IgG1 were generated based on the Protein Data Bank (PDB) entry 1HZH using the Rosetta version 3.4 software. The amino acid sequences were aligned with Clustal Omega [29]. Threading of the primary sequence onto the template crystal structure and modeling of missing loop regions were performed as described previously [30]. In total, 1000 structural models were constructed and the best-scoring 10% were clustered with automated detection of the clustering radius by Rosetta. NACCESS [31] was used for SASA calculations of the best-scoring structures representing clusters with more than one member.

3 Results and discussion

3.1 Lysine modification on IgG1 is glycosylation dependent

To test our approach, we first incubated deglycosylated IgG1 with varying concentrations of the lysine residue targeting TMT reagent, allowing the reaction to proceed for 1 min (see Section 2). Representative native mass spectra of nonmodified and TMT-modified IgG1 are shown in Fig. 1, illustrating that sequential incorporations of TMT molecules can be detected by their mass shifts, whereby the ion signals are completely baseline resolved (Fig. 1B, inset). IgG1 contains 88 potentially reactive sites, that is, 84 lysine residues and the four N-termini of the heavy and light chains (Supporting Information Fig. 1). At the highest TMT concentration, a maximum number of 69 TMT molecules are attached to IgG1 (Fig. 1C, Supporting Information Table 1). This number substantially exceeds the reported number of modified lysine residues in the lysine-linked IgG1-based ADC huN901-DM1, as identified by bottom-up proteomics [15]. Likely, there are two main reasons for this discrepancy. First, the linker concentration used during the synthesis of huN901-DM1 might have been different from the TMT concentration used here. Second, it is conceivable that not all modified sites were identified in this early study due to insufficient sensitivity of the mass spectrometric method. Our native MS data show that, in principle,

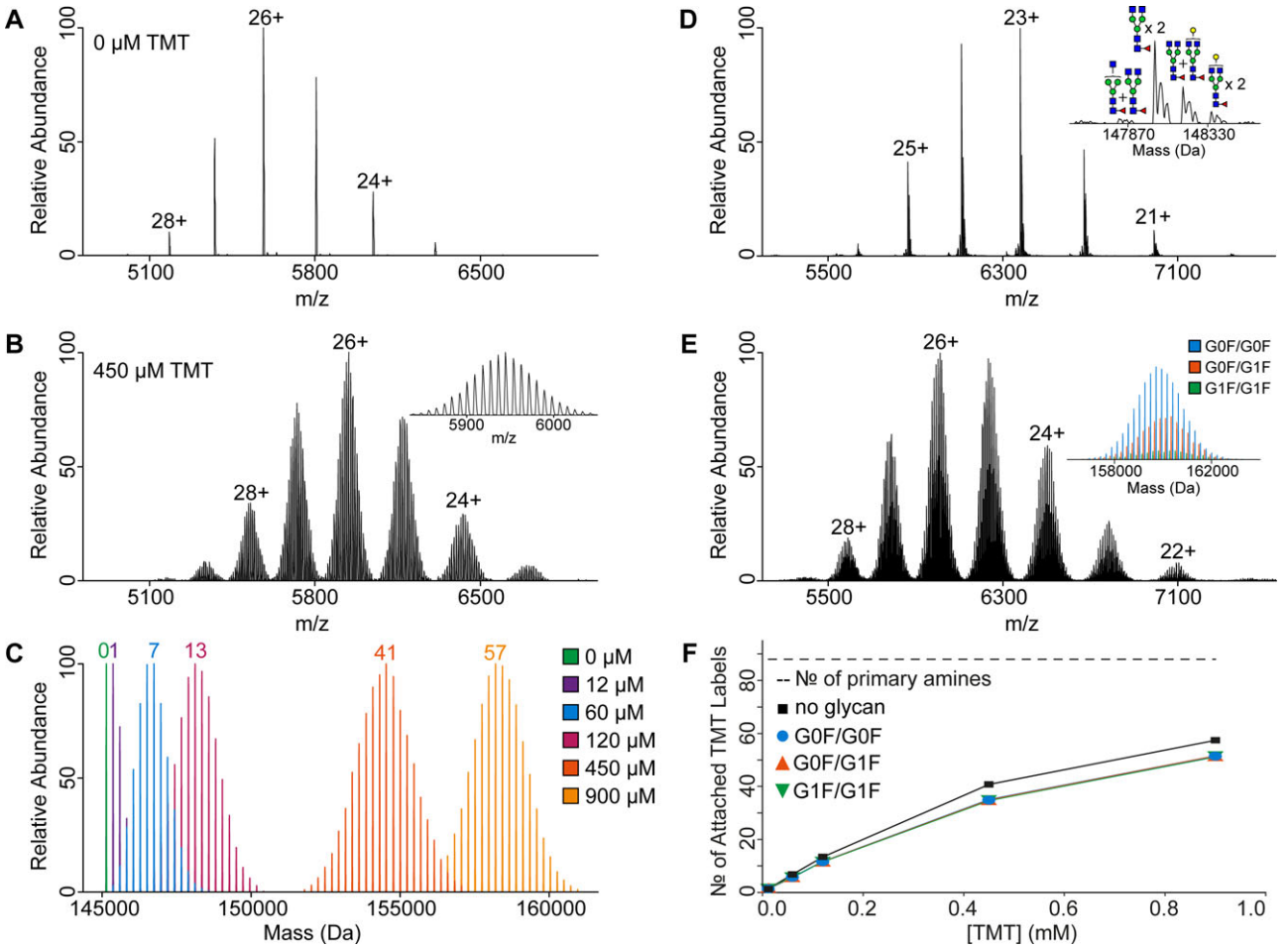

Figure 1. Qualitative and quantitative native MS analysis of the TMT conjugation to intact IgG1. (A, B) Broad-range native mass spectra of deglycosylated IgG1 before (A) and after (B) incubation with TMT. Differentially TMT-modified proteoforms can be baseline resolved (inset in B), facilitating precise relative quantification. (C) Abundance versus mass plot of all proteoforms identified after incubation for 1 min with different TMT concentrations. The number of labels conjugated to the highest abundant proteoform is indicated. (D, E) Native MS analysis of intact glycosylated IgG1 before (D) and after (E) incubation with TMT. In the nonmodified sample, four glycoforms are identified, the glycan structures of which are indicated (D, inset). The TMT incorporation profiles of the three most abundant glycoforms can be unambiguously determined (E, inset), allowing us to derive, for each glycoform, individual kinetics of TMT incorporation. (F) Weighted average number of conjugated labels at different TMT concentrations, revealing the similarity in the three glycoforms, but also the reduced reactivity observed for the glycosylated IgG1 compared to the deglycosylated IgG1.

the vast majority of lysine residues in IgG1 are susceptible to conjugation, highlighting the added value of analyzing the mass of the intact antibody next to the proteolytic digest.

Next, we performed this analysis on the "wild-type" glycosylated IgG1. Wild-type mAbs produced in human cells are typically N-glycosylated at position N297 in both the CH2 domains [32] and thus exist as a mixture of different glycoforms with mass differences in the range of 150–300 Da, complicating the characterization by MS. As shown previously [33], the different IgG1 glycoforms can be baseline resolved by native Orbitrap MS (Fig. 1D). Here, we identify four abundant glycoforms in the IgG1. Strikingly, the obtained mass resolution of ~7800 at m/z ~6200 (FWHM, see Supporting Information Fig. 2) even enables us to distinguish the sequentially TMT-modified variants of the different glycoforms (Fig. 1E, inset), some of which show mass differences of only 66 Da (Supporting Information Table 2). Hence, subtle relative mass shifts of 0.04% are baseline resolved (Supporting Information Fig. 2). Three glycoforms, which can be assigned as IgG1 carrying combinations of the well-characterized glycans G0F and G1F [34], are sufficiently abundant to allow the unambiguous determination of the average amount of incorporated TMT molecules at all used concentrations. Figure 1F, in which the weighted average of incorporated labels is plotted against the used TMT concentration, shows that lysine modification of the three glycosylated proteins proceeds almost at identical rates. In contrast, deglycosylated IgG1 incorporates approximately five additional TMT molecules at concentrations

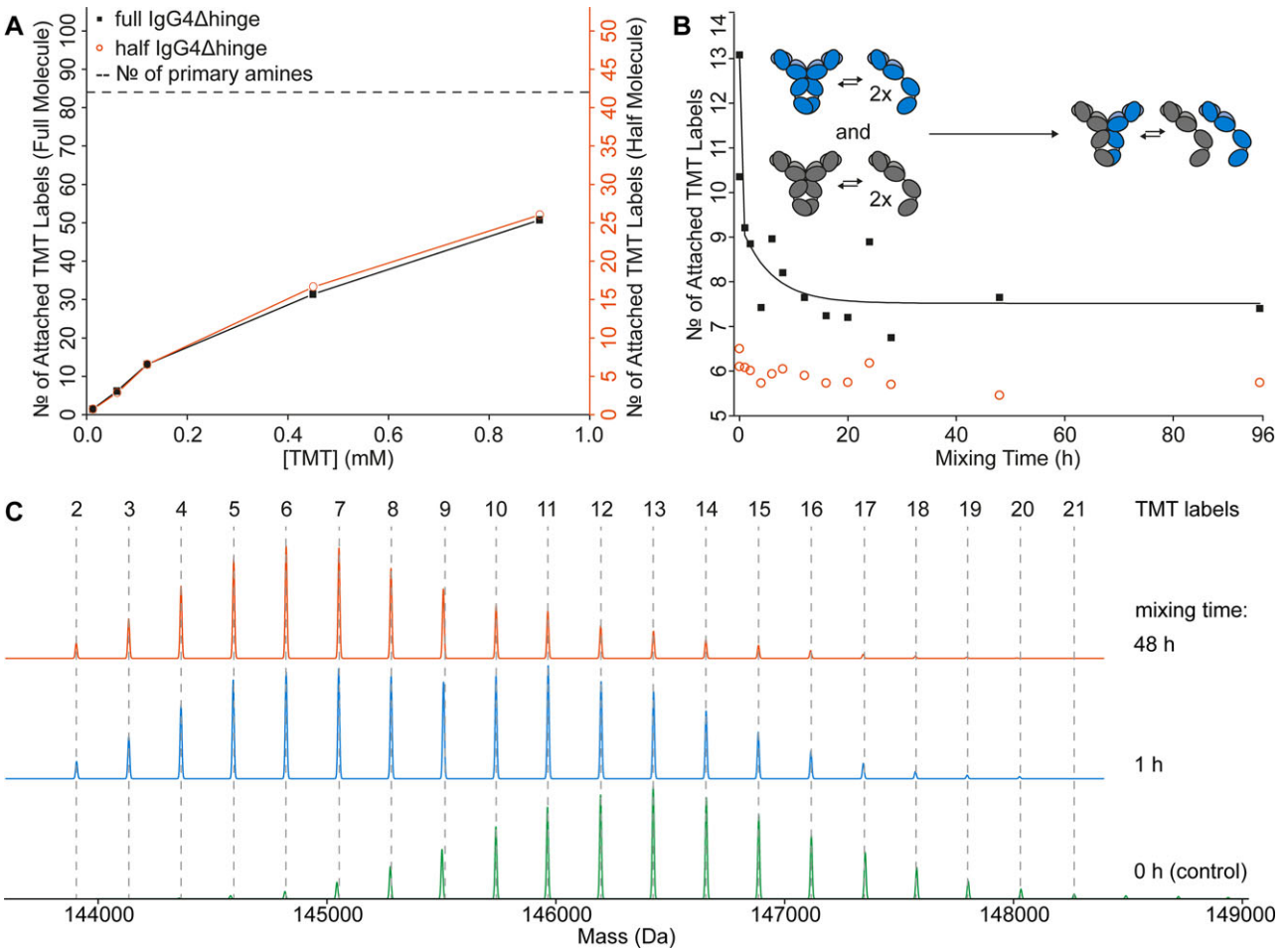

Figure 2. TMT incorporation in IgG4Δhinge half-bodies and full antibodies monitored by native MS. IgG4Δhinge full antibodies are represented by filled black squares, half-bodies are represented by open orange circles. (A) Half and full antibodies follow the same trend of TMT incorporation with exactly half as many lysine residues modified on the monomer. (B) Evolution of the weighted average of attached TMT molecules after mixing highly labeled IgG4Δhinge (obtained after 1 min of incubation with 120 μM TMT concentration) and nonlabeled IgG4Δhinge. The amount of labels decreases for the full antibodies but remains constant for the half-bodies. (C) Abundance distribution of TMT-labeled isoforms of IgG4Δhinge full antibody after 0, 1, and 48 h of mixing time. The apparent number of attached labels decreases, indicating that labeled and nonlabeled IgG4Δhinge exchange half-bodies.

above 400 μM, indicating that some lysine residues become less reactive in the presence of the glycan chain. Indeed, Rose et al. reported that protease cleavage is reduced in glycosylated IgG1, likely as a result of steric protection by the glycan [27]. Furthermore, a recent small-angle X-ray scattering study of IgG1 suggests a conformational opening of the CH2 domains upon deglycosylation [35]. The fact that increased labeling of deglycosylated IgG1 is only observed at relatively high TMT concentrations may imply that the most reactive lysine residues, at which the early conjugation events occur, are not protected, while lysine residues in the vicinity of the glycan chain are less reactive and, therefore, targeted only when TMT is present in high molar excess.

3.2 TMT-labeled IgG4Δhinge shows decreased monomer/dimer interconversion

Next to IgG1, IgG4 represents another mAb employed for ADCs. In our studies, we used a hinge-deleted variant (IgG4Δhinge) that has been shown to exist, in solution, in an equilibrium of monomeric half-body and dimeric antibody [7]. Since noncovalent interactions are retained during native MS analysis, linkages of TMT molecules to IgG4Δhinge monomers and dimers can be probed concurrently (Supporting Information Table 3). Figure 2A illustrates that the kinetics for TMT incorporation are nearly identical for both species, with the number of modifications being doubled for the full antibody compared to the half-body (see also

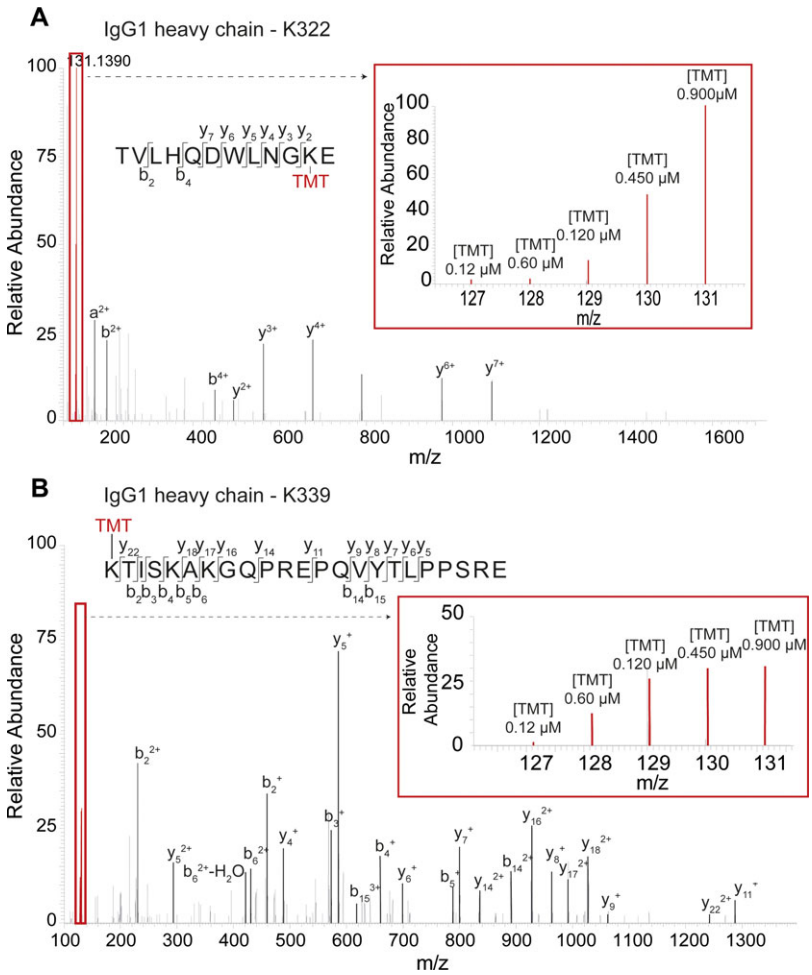

Figure 3. MS/MS spectra of IgG1 heavy chain peptides. The insets show a zoomed-in view of the reporter ions region, which were used for site-specific quantification of TMT incorporation into IgG1. The five isobaric TMT labels were applied at increasing concentrations and samples were mixed after labeling. (A) MS/MS spectrum of the $[M + 2H + TMT]^{2+}$ ion of peptide TVLHQDWLNGKE, bearing the lysine site K322. (B) MS/MS spectrum of the $[M + 4H + TMT]^{4+}$ ion of the IgG1 heavy-chain peptide KTISKAKGQPREPQVYTLPPSRE, bearing the lysine residues K339 (TMT modified), K343, and K345 (unmodified). The *y*- and *b*-ion series unambiguously localize TMT incorporation at K339.

Supporting Information Fig. 3). This allows two conclusions: either dimerization does not lead to additional protection of lysine residues or monomers and dimers remain in a dynamic equilibrium during the chemical labeling. To obtain more-conclusive results, we incubated nonlabeled and highly labeled IgG4Δhinge antibodies, the latter of which had been preincubated with 120 μM of TMT reagent for 1 min, leading to the attachment of on average 13 labels, and monitored their mixing kinetics (Fig. 2B and C). Evidently, the average number of TMT molecules attached to the highly labeled IgG4Δhinge decreases, whereas, for the half-bodies, the degree of modification remains constant (Fig. 2B). This implies that labeled and nonlabeled full bodies and half-bodies interconvert, leading to the formation of mixed species. Although equilibrium is reached after ~20 h, this does not represent complete mixing, since the average number of modifications on the full antibody is still higher than on half-bodies (Fig. 2B). In contrast, complete mixing of nonlabeled IgG4Δhinge was reached after 60 min in previous native MS studies [7]. Therefore, the monomer/dimer interconversion seems to be slowed down or even partially impaired by conjugating TMT to the lysine groups. Considering that IgG4 half-body exchange has been reported to occur also in vivo [6], this effect could be of consequence for IgG4-based ADCs and should be thoroughly studied during drug development.

3.3 IgG1 lysine residues show a high variability in reactivity

To highlight the potential of our TMT-based strategy, we complemented the native MS analysis with quantitative bottom-up proteomics, making use of the multiplexing capacities of TMT reagents. As a proof of concept, we incubated deglycosylated IgG1 with five different concentrations of five isobaric TMT reagents and mixed the labeled samples in a 1:1:1:1:1 ratio. We used pepsin to digest the TMT-conjugated IgG1, since tryptic cleavage becomes highly inefficient when the lysine residues are modified. By analyzing the reporter ion intensities, we can follow the entire course of progressive labeling in a single LC–MS/MS run (Fig. 3A) and map the conjugation site unambiguously, even if peptides contain multiple lysine residues (Fig. 3B). To facilitate comparisons between different lysine residues, we also determined the extent of

A

Light chain IgG1

```
1   EIVLTQSPAT LSLSPGERAT LSCRASQSVS SYLAWYQQKP GQAPRLLIYD ASNRATGIPA RFSGSGSGTD FTLTISSLEP
81  EDFAVYYCQQ RSNWPPTFGQ GTKVEIKRTV AAPSVFIFPP SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ
161 ESVTEQDSKD STYSLSSTLT LSKADYEKHK VYACEVTHQG LSSPVTKSFN RGEC
```

Heavy chain IgG1

```
1   EVQLLESGGG LVQPGGSLRL SCAVSGFTFN SFAMSWVRQA PGKGLEWVSA ISGSGGGTYY ADSVKGRFTI SRDNSKNTLY
81  LQMNSLRAED TAVYFCAKDK ILWFGEPVFD YWGQGTLVTV SSASTKGPSV FPLAPSSKST SGGTAALGCL VKDYFPEPVT
161 VSWNSGALTS GVHTFPAVLQ SSGLYSLSSV VTVPSSSLGT QTYICNVNHK PSNTKVDKRV EPKSCDKTHT CPPCPAPELL
240 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN
321 GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP
401 PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP G(K)
```

B

		Percentage of TMT incorporation				
		[TMT] = 12 µM	[TMT] = 60 µM	[TMT] = 120 µM	[TMT] = 450 µM	[TMT] = 900 µM
Light chain	N-terminus	-	-	-	-	-
	K39	0	0	4	11	25
	K103	0	2	11	36	75
	K107	0	1	6	34	86
	K126	0	6	37	74	100
	K145	1	19	32	63	97
	K149	0	2	8	37	75
	K169	0	5	16	51	97
	K183	-	-	-	-	-
	K188	-	-	-	-	-
	K190	-	-	-	-	-
	K207	1	7	12	29	41
Heavy chain	N-terminus	-	-	-	-	-
	K43	0	0	0	1	1
	K65	0	13	28	77	92
	K76	-	-	-	-	-
	K98	Modified				
	K100	Modified				
	K126	Modified				
	K138	Modified				
	K152	0	0	0	0	1
	K210	-	-	-	-	-
	K215	-	-	-	-	-
	K218	-	-	-	-	-
	K223	-	-	-	-	-
	K227	-	-	-	-	-
	K251	-	-	-	-	-
	K253	-	-	-	-	-
	K279	0	11	20	56	100
	K293	-	-	-	-	-
	K295	-	-	-	-	-
	K322	0	3	12	46	100
	K325	2	12	28	81	99
	K327	0	0	1	26	46
	K331	0	0	0	1	2
	K339	3	16	37	83	96
	K343	0	0	0	0	0
	K345	-	-	-	-	-
	K365	-	-	-	-	-
	K375	0	0	0	0	0
	K397	-	-	-	-	-
	K414	-	-	-	-	-
	K419	-	-	-	-	-
	K444	0	0	0	1	2

Figure 4. Percentage of site-specific TMT modification in IgG1 at different TMT concentrations. (A) Sequences of the light and heavy chains of the IgG1 antibody. Potential reactive sites available for TMT incorporation are highlighted in green. The light chain contains 24 potential reactive sites (12 per monomer) and the heavy chain 64 (32 per monomer), which represents, in total, 88 sites on the IgG1 antibody. (B) Percentage of site-specific TMT incorporation in IgG1 at different TMT concentrations. The percentages of TMT occupancy were determined as follows: First, the percentage of TMT incorporation at the highest TMT concentration (900 µM) was defined in a label-free quantification manner by comparing MS1 areas of the modified and unmodified isoforms of the peptides (from an independently analyzed sample). Second, the relative increase of TMT incorporation at each TMT concentration was defined for each site according to the relative increase of TMT reporter ion intensities from the mixed sample. Finally, the percentage of TMT modification at 900 µM TMT, as determined in the first step, was used as a reference, correlating to the TMT^6–131 reporter ion intensity obtained in the second step. This allowed us to convert the relative TMT incorporation into the percentages of TMT modification at each TMT concentration. Unidentified sites are highlighted in gray and sites, which were identified as TMT modified but for which no percentage could be calculated (no MS1 data available), are highlighted in pink.

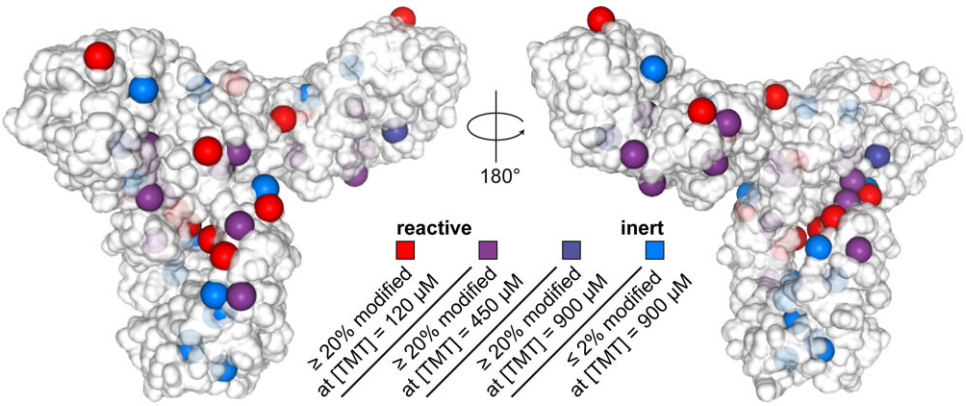

Figure 5. Site-specific conjugation properties of IgG1 lysine residues. The ε-amino groups of the lysine residues, that were covered by our quantitative bottom-up analysis, are shown as spheres, colored according to their experimentally determined reactivity and mapped on the best-scoring comparative model of IgG1, which is represented by its solvent-accessible surface.

lysine modification at the highest TMT concentration (900 µM) by label-free quantification, comparing MS1 areas of TMT-modified peptides and their unmodified counterparts. This was used as reference to define the TMT occupancy for all the identified lysine residues and N-termini at each TMT concentration. Of the 88 potential reactive sites contained in IgG1 (i.e., 84 lysine residues, 4 N-termini), peptides representing 46 TMT-modified sites are identified, 42 of which could be quantified based on their MS1 areas (Fig. 4). After incubation with 120 µM TMT for 1 min, only six lysine residues (K126, K145 on the IgG1 light chain and K65, K279, K325, and K339 on the IgG1 heavy chain) show more than 20% TMT incorporation (Fig. 4), representing readily modified sites (Fig. 5). Applying low concentrations of the conjugating reagent for a short time, evidently, allows selective modification of IgG1 at the most reactive lysine residues. These lysine residues therefore represent hot-spot residues. This finding suggests that, even for lysine-linked ADCs, drug payload and conjugation site can be, to some extent, controlled by careful optimization of the reaction conditions. Furthermore, K169 on the light chain and K322 on the heavy chain represent more modestly reactive lysine residues. In total, at the highest TMT concentration, up to 28 lysine residues of the antibody molecule incorporate TMT molecules (Fig. 4B), whereas a maximum of 69 attached TMT molecules are detected with native MS. In part, this discrepancy can be attributed to the fact that a number of peptides bearing reactive sites are not detected in the bottom-up experiment. The detectability of peptides in LC-MS/MS experiments is governed by several factors such as peptide length, hydrophobicity, retention time, isoelectric point, and mass. Conceivably, the size and amino acid composition (especially, the proportion of basic residues) of some peptic peptides may impair their ionization efficiency or result in the generation of ions beyond the covered m/z range, making them incompatible with LC-MS/MS analysis. Notably, the proteomics data suggest that 14 of the detected lysine residues are inert toward TMT labeling (Fig. 4B). Thus, our data give evidence for a high variability in reactivity of the lysine residues in IgG1.

Since solvent accessibility is one of the principal determinants of reactivity we sought to predict the SASA of the IgG1 used for our analyses. Sequentially, this IgG1 is highly similar to IgG1 b12 [36] (86% sequence identity), a crystal structure of which is available (PDB entry 1HZH). However, the lysine positions within these two IgG1 variants are not identical and some lysine-containing loop regions are not represented by the crystal structure. To obtain a structural model with correctly positioned lysine residues, we used the Rosetta software suite to create comparative models of our IgG1, taking the crystal structure of IgG1 b12 as a template (Supporting Information Table 4). We used these models to calculate the solvent accessibility of the reactive lysine amino groups and N-termini (Supporting Information Table 5). While, for the IgG1 light chain, the SASA predictions are in good agreement with the experimentally determined degree of modification, the values do not correlate that well for the heavy-chain lysine residues (Supporting Information Fig. 4). Reassuringly, those lysine residues, which we identify as highly reactive "hot spots," are also predicted to be solvent exposed. Figure 5 and Supporting Information Fig. S4 illustrate, however, that some of the unreactive lysine residues seem to exhibit a large SASA as well. Evidently, there are other factors influencing lysine reactivity. Supporting our findings, Guo et al. reported that lysine conjugation to Sulfo-NHS-acetate is better correlated with predicted local p*K*a values than with SASA calculations [37]. Furthermore, small-angle X-ray and neutron scattering studies suggest that IgGs are highly flexible in solution [38, 39]. This flexibility will influence the accessibility of lysine residues but cannot be reflected in the comparative models, as they are based on an IgG1 crystal structure. Consequently, experimental probing is vital to determine the actual conjugation properties of a protein.

4 Concluding remarks

The combination of native MS and bottom-up proteomics enabled us to probe the conjugation properties of lysine residues within human IgGs using TMT as labeling reagent. By means of high-resolution native Orbitrap MS, we followed the progressive labeling of individual IgG1 glycoforms suggesting that several lysine residues are protected by the glycan chain. Furthermore, we monitored the TMT labeling of IgG4Δhinge half and full antibodies. Mixing labeled and nonlabeled

IgG4Δhinge antibodies indicated that monomer/dimer interconversion is somewhat impaired when lysine residues are modified. Finally, we used five isobaric TMT reagents to quantify the progressive modification of IgG1 lysine residues site-specifically. Label-free quantification of TMT incorporation at the highest label concentration allowed us to compare the reactivity of different sites and identify hot-spot lysine residues. The conjugation of drugs to mAbs during ADC development relies on the same chemistry as TMT labeling. Therefore, we believe that this strategy presents a fast and straightforward approach to assess the suitability and guide the optimization of specific mAbs for the development of new lysine-linked ADCs.

The authors are indebted to Genmab for providing the IgG1 and IgG4 antibody samples and are grateful for the long-term collaboration. This work was supported by the Netherlands Proteomics Center and the Netherlands Organization for Scientific Research (NWO) funded proteomics facility Proteins@Work (project 184.032.201). Moreover, the ManiFold project (grant agreement number 317371) and the PRIME-XS project (grant agreement number 262067), both funded by the European Union 7th Framework Programme, were used to support the work.

The authors have declared no conflict of interest.

5 References

[1] Mullard, A., Maturing antibody-drug conjugate pipeline hits 30. *Nat. Rev. Drug Discov.* 2013, *12*, 329–332.

[2] Beck, A., Reichert, J. M., Antibody-drug conjugates: present and future. *MAbs* 2014, *6*, 15–17.

[3] Casi, G., Neri, D., Antibody-drug conjugates: basic concepts, examples and future perspectives. *J. Control Release.* 2012, *161*, 422–428.

[4] Teicher, B. A., Chari, R. V., Antibody conjugate therapeutics: challenges and potential. *Clin. Cancer Res.* 2011, *17*, 6389–6397.

[5] Trail, P., Antibody drug conjugates as cancer therapeutics. *Antibodies* 2013, *2*, 113–129.

[6] van der Neut Kolfschoten, M., Schuurman, J., Losen, M., Bleeker, W. K. et al., Anti-inflammatory activity of human IgG4 antibodies by dynamic Fab arm exchange. *Science* 2007, *317*, 1554–1557.

[7] Rose, R. J., Labrijn, A. F., van den Bremer, E. T. J., Loverix, S. et al., Quantitative analysis of the interaction strength and dynamics of human IgG4 half molecules by native mass spectrometry. *Structure* 2011, *19*, 1274–1282.

[8] Labrijn, A. F., Buijsse, A. O., van den Bremer, E. T. J., Verwilligen, A. Y. W. et al., Therapeutic IgG4 antibodies engage in Fab-arm exchange with endogenous human IgG4 in vivo. *Nat. Biotechnol.* 2009, *27*, 767–771.

[9] Ducry, L., Stump, B., Antibody-drug conjugates: linking cytotoxic payloads to monoclonal antibodies. *Bioconjug. Chem.* 2010, *21*, 5–13.

[10] Flygare, J. A., Pillow, T. H., Aristoff, P., Antibody-drug conjugates for the treatment of cancer. *Chem. Biol. Drug. Des.* 2013, *81*, 113–121.

[11] Behrens, C. R., Liu, B., Methods for site-specific drug conjugation to antibodies. *MAbs* 2014, *6*, 46–53.

[12] Francisco, J. A., Cerveny, C. G., Meyer, D. L., Mixan, B. J. et al., cAC10-vcMMAE, an anti-CD30-monomethyl auristatin E conjugate with potent and selective antitumor activity. *Blood* 2003, *102*, 1458–1465.

[13] Lewis, P., Gail, D., Li, G., Dugger, D. L. et al., Targeting HER2-positive breast cancer with trastuzumab-DM1, an antibody-cytotoxic drug conjugate. *Cancer Res.* 2008, *68*, 9280–9290.

[14] Feng, Y., Zhu, Z., Chen, W., Prabakaran, P. et al., Conjugates of small molecule drugs with antibodies and other proteins. *Biomedicines* 2014, *2*, 1–13.

[15] Wang, L., Amphlett, G., Blättler, W. A., Lambert, J. M., Zhang, W., Structural characterization of the maytansinoid-monoclonal antibody immunoconjugate, huN901-DM1, by mass spectrometry. *Protein Sci.* 2005, *14*, 2436–2446.

[16] Hamblett, K. J., Senter, P. D., Chace, D. F., Sun, M. M. C. et al., Effects of drug loading on the antitumor activity of a monoclonal antibody drug conjugate. *Clin. Cancer Res.* 2004, *10*, 7063–7070.

[17] Konermann, L., Vahidi, S., Sowole, M. A., Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Anal. Chem.* 2014, *86*, 213–232.

[18] Mendoza, V. L., Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* 2009, *28*, 785–815.

[19] Watson, C., Sharp, J. S., Conformational analysis of therapeutic proteins by hydroxyl radical protein footprinting. *AAPS J.* 2012, *14*, 206–217.

[20] Deperalta, G., Alvarez, M., Bechtel, C., Dong, K. et al., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *MAbs* 2013, *5*, 86–101.

[21] Scholten, A., Visser, N. F. C., van den Heuvel, R. H. H., Heck, A. J. R., Analysis of protein-protein interaction surfaces using a combination of efficient lysine acetylation and nanoLC-MALDI-MS/MS applied to the E9:Im9 bacteriotoxin–immunity protein complex. *J. Am. Soc. Mass Spectrom.* 2006, *17*, 983–994.

[22] Kvaratskhelia, M., Miller, J. T., Budihas, S. R., Pannell, L. K. et al., Identification of specific HIV-1 reverse transcriptase contacts to the viral RNA:tRNA complex by mass spectrometry and a primary amine selective reagent. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 15988–15993.

[23] Thompson, A., Schäfer, J., Kuhn, K., Kienle, S. et al., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 2003, *75*, 1895–1904.

[24] Zhou, Y., Vachet, R. W., Covalent labeling with isotopically encoded reagents for faster structural analysis of proteins by mass spectrometry. *Anal. Chem.* 2013, *85*, 9664–9670.

[25] van de Waterbeemd, M., Lössl, P., Gautier, V., Marino, F. et al., Simultaneous assessment of kinetic, site-specific, and structural aspects of enzymatic protein phosphorylation. *Angew. Chem. Int. Ed. Engl.* 2014, *53*, 9660–9664.

[26] Rose, R. J., Damoc, E., Denisov, E., Makarov, A., Heck, A. J. R., High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* 2012, *9*, 1084–1086.

[27] Rose, R. J., van Berkel, P. H., van den Bremer, E. T., Labrijn, A. F. et al., Mutation of Y407 in the CH3 domain dramatically alters glycosylation and structure of human IgG. *MAbs* 2013, *5*, 219–228.

[28] Binai, N. A., Bisschops, M. M., van Breukelen, B., Mohammed, S. et al., Proteome adaptation of *Saccharomyces cerevisiae* to severe calorie restriction in Retentostat cultures. *J. Proteome Res.* 2014, *13*, 3542–3553.

[29] Sievers, F., Wilm, A., Dineen, D., Gibson, T. J. et al., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011, *7*, 539.

[30] Combs, S. A., Deluca, S. L., Deluca, S. H., Lemmon, G. H. et al., Small-molecule ligand docking into comparative models with Rosetta. *Nat. Protoc.* 2013, *8*, 1277–1298.

[31] Hubbard, S. J., Thornton, J. M., NACCESS. Computer program. *Department of Biochemistry and Molecular Biology*, University College London 1993. http://www.bioinf.manchester.ac.uk/naccess/.

[32] Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M., Dwek, R. A., The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* 2007, *25*, 21–50.

[33] Rosati, S., Rose, R. J., Thompson, N. J., van Duijn, E. et al., Exploring an Orbitrap analyzer for the characterization of intact antibodies by native mass spectrometry. *Angew. Chem. Int. Ed. Engl.* 2012, *51*, 12992–12996.

[34] Chen, X., Flynn, G. C., Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry. *Anal. Biochem.* 2007, *370*, 147–161.

[35] Borrok, M. J., Jung, S. T., Kang, T. H., Monzingo, A. F., Georgiou, G., Revisiting the role of glycosylation in the structure of human IgG Fc. *ACS Chem. Biol.* 2012, *7*, 1596–1602.

[36] Saphire, E. O., Parren, P. W., Pantophlet, R., Zwick, M. B. et al., Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. *Science* 2001, *293*, 1155–1159.

[37] Guo, X., Bandyopadhyay, P., Schilling, B., Young, M. M. et al., Partial acetylation of lysine residues improves intraprotein cross-linking. *Anal. Chem.* 2008, *80*, 951–960.

[38] Clark, N. J., Zhang, H., Krueger, S., Lee, H. J. et al., Small-angle neutron scattering study of a monoclonal antibody using free-energy constraints. *J. Phys. Chem. B.* 2013, *117*, 14029–14038.

[39] Lilyestrom, W. G., Shire, S. J., Scherer, T. M., Influence of the cosolute environment on IgG solution structure analyzed by small-angle X-ray scattering. *J. Phys. Chem. B.* 2012, *116*, 9611–9618.

