

REPORT

 OPEN ACCESS

## Detection of a phosphorylated glycine-serine linker in an IgG-based fusion protein

Oksana Tyshchuk<sup>a</sup>, Hans Rainer Völger<sup>a</sup>, Claudia Ferrara <sup>b</sup>, Patrick Bulau<sup>c</sup>, Hans Koll<sup>a</sup>, and Michael Mølhøj<sup>a</sup>

<sup>a</sup>Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Roche Diagnostics GmbH, Penzberg, Germany; <sup>b</sup>Oncology Discovery & Translational Area, Roche Innovation Center Zurich, Schlieren, Switzerland; <sup>c</sup>Roche Pharma Technical Development Penzberg, Roche Diagnostics GmbH, Penzberg, Germany

### ABSTRACT

Molecular mass determination by electrospray ionization mass spectrometry of a recombinant IgG-based fusion protein (mAb1-F) produced in human embryonic kidney (HEK) cells demonstrated the presence of a dominant +79 Da product variant. Using LC-MS tryptic peptide mapping analysis and collision-induced dissociation (CID) and electron-transfer/higher-energy collision dissociation fragmentations, the modification was localized to the C-terminal serine residue of a glycine-serine linker [(G<sub>4</sub>S)<sub>2</sub>] of a fused heavy chain containing in total 2 (G<sub>4</sub>S)<sub>2</sub>-linkers. The modification was identified as a phosphorylation (+79.97 Da) by the presence of a 98 Da neutral loss reaction with CID, by spiking a synthetic phosphoserine peptide, and by dephosphorylation with alkaline phosphatase. A thermolysin digest combined with higher-energy collision dissociation (HCD) positioned the phosphoserine to one specific glycine-serine linker of the fused heavy chain, and the relative level of phosphorylated linker was determined to be 11.3% and 0.4% by LC-MS when the fusion protein was transiently expressed in HEK or in stably transformed Chinese hamster ovary cells, respectively. This observation demonstrates that fusions with glycine-serine linker sequences should be carefully evaluated during drug development to prevent the introduction of a phosphorylation site in therapeutic fusion proteins.

**Abbreviations:** Fc, fragment crystallizable region; IgG, immunoglobulin G; LC, liquid chromatography; pS, pSer, phosphoserine; Xyl, xylose

### ARTICLE HISTORY

Received 5 August 2016  
Revised 5 September 2016  
Accepted 8 September 2016

### KEYWORDS

Alkaline phosphatase; CID; EThcD; fusion protein; Glycine-serine linker; HCD; mass spectrometry; phosphorylation; phosphoserine; post-translational modification



## Introduction


Engineered fusion proteins with dual or multifunctional specificities or activities, e.g., appended IgGs, bispecific antibodies, have become a focus for targeted immune therapy. These versatile molecules can be used for many purposes, such as targeting various malignancies or effective transport of protein drug across cell membranes or other biological barriers.<sup>1–5</sup> Fusion proteins not only consist of protein domains, but often also include one or more suitable linkers or spacers to successfully fuse the domains and allow the protein to fold properly. Fusions without linkers may result in low yield, low potency, or misfolding.<sup>6–8</sup> Depending on the functionality required, flexible, rigid or cleavable linkers may be incorporated in fusion proteins. Flexible linkers are rich in small and hydrophilic amino acids such as glycine, serine or threonine. Rigid linkers often have a helical conformation or are rich in proline residues, which ensures effective spatial separation of the domains and reduced their interference. Cleavable linkers may contain protease cleavage sites for the release of a domain specifically where the protease is located *in vivo*.<sup>9</sup>

Glycine-serine (GS) linkers have no ordered secondary structure and are rich in small size glycine for flexibility, as well as polar serine residues to ensure solubility.<sup>9–11</sup> A widely applied protease-

resistant GS-linker has the sequence of (G<sub>4</sub>S)<sub>n</sub> as proposed by a study of naturally occurring linkers in multi-domain proteins.<sup>12</sup> To ensure optimal flexibility or separation of the adjacent domains and to promote intermolecular interactions, the length of the linker can be adjusted by the number (n) of the G<sub>4</sub>S-units. (G<sub>4</sub>S)<sub>n</sub>-linkers are frequently used in recombinant fusion proteins by generating loops that connect domains or multiple target specificities. A prominent example of the application of a GS-linker is the single-chain variable fragment (scFv), which is produced as a single polypeptide linking the antibody heavy chain variable domain (V<sub>H</sub>) and the light chain variable domain (V<sub>L</sub>) and hereby facilitating the association of the 2 domains.<sup>13–15</sup> The flexibility of the (G<sub>4</sub>S)<sub>3</sub>-linker allows the correct orientation of the V<sub>H</sub> and V<sub>L</sub> domains and does not interfere with their folding.<sup>13</sup> Other antibody formats and fusions involving GS-linkers have been reported and are currently in clinical trials.<sup>16–28</sup>

The development of therapeutic proteins, including antibodies, antibody derivatives and fusion proteins, typically involves extensive characterization of the integrity, homogeneity, and the presence of chemical and post-translational modifications (PTMs) to assess critical quality attributes.<sup>29–30</sup> Generally, unexpected heterogeneities are unwanted in biotherapeutic proteins.

**CONTACT** Michael Mølhøj  [michael.molhoj@roche.com](mailto:michael.molhoj@roche.com)  Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany.

 Supplemental data for this article can be accessed on the [publisher's website](#).

Published with license by Taylor and Francis Group, LLC © Roche Diagnostic GmbH

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

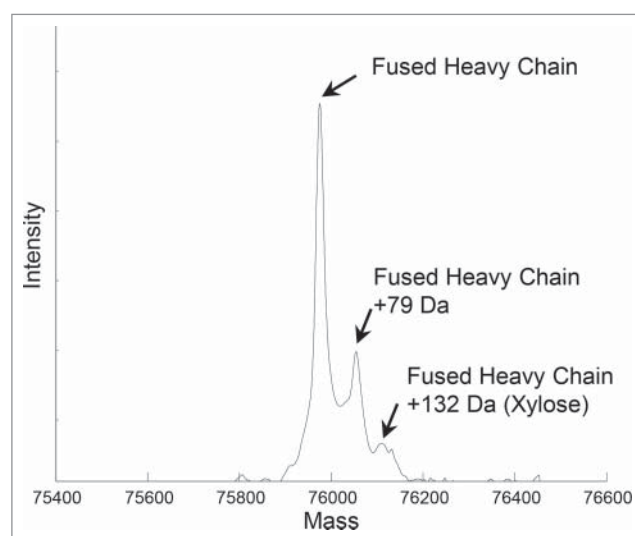
If identified, they need to be diminished or eliminated by either manufacturing optimizations or further protein engineering. At a minimum, PTMs need to be monitored to ensure consistency between batches. Failure to detect PTMs represents a safety risk due to potential immunogenicity. Nevertheless, a complex PTM has been reported to be common for serine residues in the GSG sequence motifs of GS-linkers due to a series of xylose-containing *O*-glycans, with *O*-Xyl alone being the most abundant glycosidic substituent.<sup>18,21,31-32</sup> The relative level of *O*-xylosylated GS-linkers has been reported to differ between human embryonic kidney (HEK) cells and Chinese hamster ovary (CHO) cells, with approximately 30% present in fusion proteins expressed in HEK cells, and about 3% in the case of CHO cell expression.<sup>31</sup>

Here, we report the detection of a +79 Da product variant of an IgG-based fusion protein (mAb1-F) that could be identified as a posttranslational phosphorylation of the C-terminal serine of a (G<sub>4</sub>S)<sub>2</sub>-linker. Neighboring C-terminal residues relative to the affected serine residue likely determine the susceptibility to phosphorylation. Our results suggest that unwanted phosphorylation of GS-linkers can be avoided during early drug development by carefully preventing the introduction of a potential phosphorylation site upon fusion to peptides and proteins.

## Results

### UHR-ESI-QTOF-MS analysis of the IgG-based fusion protein

An engineered bispecific IgG-based fusion protein (mAb1-F) was transiently expressed in HEK cells and the molecular mass determined by ultra-high resolution electrospray ionization quadrupole time-of-flight mass spectrometry (UHR-ESI-QTOF-MS). Beforehand, the fusion protein was deglycosylated with PNGase F to remove the *N*-glycan heterogeneity of the Fc-region and reduced with Tris(2-carboxyethyl)phosphine (TCEP). In addition to the expected fused heavy chain molecular mass of mAb1-F, the presence of an unknown variant with



**Figure 1.** Total mass determination of mAb1-F transiently expressed in human embryonic kidney cells. Deconvoluted mass spectrum of the deglycosylated and reduced fused heavy chain. In addition to the expected fused heavy chain, another peak demonstrated the presence of an intense +79 Da product variant of the fused heavy chain. A less intense signal due to a +132 Da xylose product variant was also present.

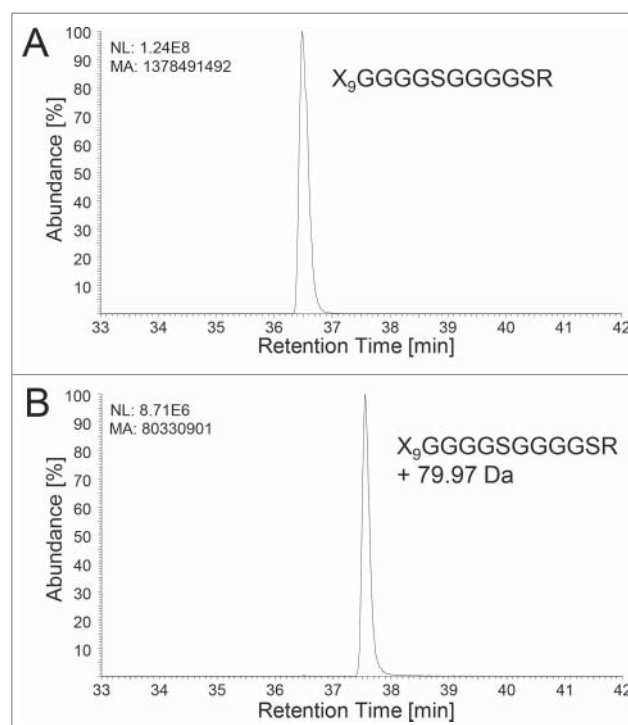


**Figure 2.** Schematic representation of the fused heavy chain structure of mAb1-F. The fused heavy chain consists of 2 >10 kDa non-IgG proteins linked by 2 glycine-serine [(G<sub>4</sub>S)<sub>2</sub>] linkers I and II to an IgG heavy chain constant domain.

an additional mass of +79 Da was verified (Fig. 1). A +132 Da modification (potentially *O*-xylose) could also be detected. A schematic representation of the fused heavy chain consisting of 2 >10 kDa non-IgG proteins linked by 2 GS-linkers I and II [(G<sub>4</sub>S)<sub>2</sub>] to a heavy chain constant domain is shown in Fig. 2.

### +79 Da modification of a tryptic glycine-serine linker peptide

To elucidate the identity and position of the heterogeneity responsible for the +79 Da variant, the fusion protein from HEK cells was further characterized by liquid chromatography (LC)-mass spectrometry (MS)/MS tryptic peptide mapping. Evaluation of the MS data proved the presence of the expected tryptic peptide of the fused heavy chain with the sequence X<sub>9</sub>GGGGSGGGGS covering a GS-linker [(G<sub>4</sub>S)<sub>2</sub>] and the presence of the same peptide with a mass corresponding to a 79.97 Da modification

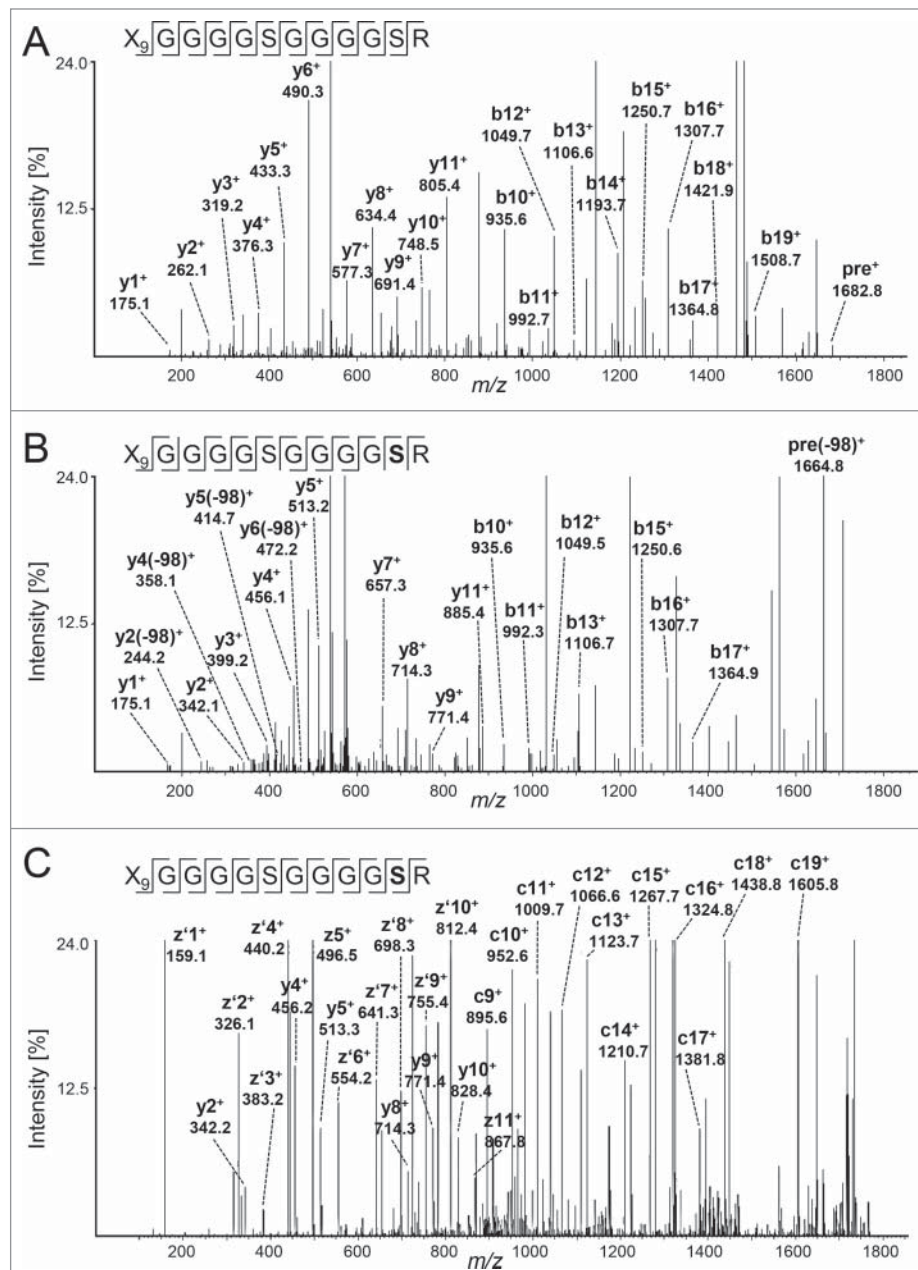


**Figure 3.** Extracted ion current (EIC) chromatograms ( $z = 2$  and  $3$ ) of (A) the unmodified (elution time 36.5 min) and (B) the +79.97 Da modified (elution time 37.6 min) tryptic glycine-serine linker peptides (X<sub>9</sub>GGGGSGGGGS) of the fused heavy chain of mAb1-F transiently expressed in human embryonic kidney cells. The modified peptide was better retained on the reverse-phase column. A relative comparison of the integrated EIC chromatograms quantified the modification to 5.5% (including 1.1% with *O*-xylose). MA, manually integrated peak; NL, normalized intensity level.

(Fig. 3). Relative quantitative evaluation of the extracted ion current (EIC) chromatograms of the unmodified tryptic peptide ( $z = 2$  and  $3$ ,  $m/z$  windows 561.61-561.63 and 841.91-841.93) (Fig. 3A) and the modified tryptic peptide ( $z = 2$  and  $3$ ,  $m/z$  window 588.26-588.28 and 881.89-881.91) (Fig. 3B), determined the modification to be present at 5.5% relative to the unmodified peptide (including 1.1% with *O*-xylose at the GSG motif (data not shown)). The EIC chromatograms also revealed the modified peptide eluting at 37.6 min to be more hydrophobic on the reverse-phase column compared with the unmodified peptide eluting at 36.5 min (Fig. 3).

To determine the amino acid position of the +79.97 Da modification, we conducted collision-induced dissociation

(CID) experiments and evaluated the fragmentation data for the modified and unmodified tryptic linker peptides. The CID mass spectra of the triply protonated unmodified and modified peptides are shown in Fig. 4A and 4B, respectively. Mass differences between the following fragments of the modified peptide:  $y_2^+(m/z\ 342.1)$ ,  $y_3^+(m/z\ 399.2)$ ,  $y_4^+(m/z\ 456.1)$ ,  $y_5^+(m/z\ 513.2)$ ,  $y_7^+(m/z\ 657.3)$ ,  $y_8^+(m/z\ 714.3)$ ,  $y_9^+(m/z\ 771.4)$ , and  $y_{11}^+(m/z\ 885.4)$  (Fig. 4B), and the corresponding unmodified ions:  $y_2^+(m/z\ 262.1)$ ,  $y_3^+(m/z\ 319.2)$ ,  $y_4^+(m/z\ 376.3)$ ,  $y_5^+(m/z\ 433.3)$ ,  $y_7^+(m/z\ 577.3)$ ,  $y_8^+(m/z\ 634.4)$ ,  $y_9^+(m/z\ 691.4)$ , and  $y_{11}^+(m/z\ 805.4)$  (Fig. 4A) were identified. In addition, no modified  $y_1^+(m/z\ 175.1)$  fragment ion for both peptides was detected (Fig. 4B). This suggests the 79.97 Da modification was



**Figure 4.** Ion trap MS/MS data obtained by collision induced dissociation of the triple protonated (A) unmodified and (B) +79.97 Da modified tryptic glycine-serine linker peptides of the fused heavy chain of mAb1-F transiently expressed in human embryonic kidney cells, and (C) MS/MS spectrum by electron-transfer/higher-energy collision dissociation of the triple protonated modified peptide. The additional 79.97 Da is localized to the C-terminal serine residue (bold).

localized to the C-terminal serine residue of the tryptic peptide (X<sub>9</sub>GGGGSGGGSR).

### 98 Da neutral loss of the modified tryptic linker peptide

Based on the literature, the location of the +79.97 Da modification (4 decimal places: +79.9663 Da) suggests a phosphorylated (monoisotopic mass: +79.9663 Da) serine rather than a sulfated (monoisotopic mass: +79.9568 Da) serine. The possibility of a sequence variant (i.e., an amino acid misincorporation) of the affected serine residue could be ruled out because the theoretical mass shift of any sequence variant would not match the observed mass shift. The modified peptide was found to be more hydrophobic than the unmodified tryptic peptide. Since phosphorylations add anionic/acidic phosphate groups to the respective amino acid, phosphopeptides are assumed to be more hydrophilic than non-phosphorylated peptides. However, although a phosphorylation lowers the isoelectric point compared to the non-phosphorylated peptide, it does not necessarily lead to an increase in hydrophilicity.<sup>33</sup> If a peptide contains positively charged basic amino acid residues, an increase in nominal hydrophilicity is overcompensated by charge neutralization, thereby reducing the overall hydrophilicity.<sup>33</sup> With the C-terminal arginine, the affected tryptic GS-linker peptide contains only one basic amino acid, which could explain why the phosphorylated GS-linker peptide is better retained on the reverse-phase column (Fig. 3).

As a phosphoester bond is weaker than a peptide bond, the analysis of protonated phosphorylated peptides by ion trap CID-MS/MS often give rise to prominent non-sequence product ions corresponding to the neutral loss of 98 Da from the precursor ions, yielding  $[M+nH-98]^{n+}$  products.<sup>34-36</sup> Consequently, this product whereby the phosphate group is dissociated from the phosphopeptide diagnostically indicates the presence of a phosphorylated peptide ion. The neutral loss involves 2 competing cleavages of the phosphoester bond resulting in product ions with different structures but identical  $m/z$  values. First, there is a direct loss of H<sub>3</sub>PO<sub>4</sub> (98 Da) from the phosphorylated residue, secondly there is a combined loss of HPO<sub>3</sub> (80 Da) and H<sub>2</sub>O (18 Da) from the phosphorylation site and from an additional site within the peptide, respectively.<sup>35</sup> Indeed, the CID-MS/MS spectrum of the modified tryptic GS-linker peptide included an intense precursor ion signal with a -98 Da loss ( $\text{pre}(-98)^+$ ,  $m/z$  1664.8) (Fig. 4B) not present in the CID-MS/MS spectrum of the unmodified peptide (Fig. 4A). Also, -98 Da non-sequence losses from the sequence type product ions  $\gamma 2(-98)^+$  ( $m/z$  244.2),  $\gamma 4(-98)^+$  ( $m/z$  358.1),  $\gamma 5(-98)^+$  ( $m/z$  414.7), and  $\gamma 6(-98)^+$  ( $m/z$  472.2) could be demonstrated (Fig. 4B), which were not present in the CID-spectrum of the unmodified tryptic peptide (Fig. 4A).

Compared to CID, electron-transfer/higher-energy collision dissociation (EThcD) may include both b-, c-, y- and z-ions and has been shown to give more complete fragmentation and more reliable phosphorylation site localization compared to electron-transfer dissociation (ETD) or higher-energy collision dissociation (HCD) alone.<sup>37</sup> When performing EThcD-MS/MS on the triply protonated modified peptide both the N-terminal fragment ion  $c19^+$  ( $m/z$  1605.7) and the following C-terminal fragments

ions proved to contain the modification:  $z'2^+$  ( $m/z$  326.1),  $\gamma 2^+$  ( $m/z$  343.2),  $z'3^+$  ( $m/z$  383.2),  $z'4^+$  ( $m/z$  440.2),  $\gamma 4^+$  ( $m/z$  456.2),  $z'5^+$  ( $m/z$  496.5),  $z'6^+$  ( $m/z$  554.2),  $z'7^+$  ( $m/z$  641.3),  $z'8^+$  ( $m/z$  698.3),  $\gamma 8^+$  ( $m/z$  714.3),  $z'9^+$  ( $m/z$  755.4),  $\gamma 9^+$  ( $m/z$  771.4),  $z'10^+$  ( $m/z$  812.4),  $\gamma 10^+$  ( $m/z$  828.4), and  $z11^+$  ( $m/z$  867.8) (Fig. 4C), confirming the phosphoserine in C-terminal position.

### Confirmation of phosphoserine by synthetic peptide spiking

To confirm the phosphoserine suggested by MS, a synthetic peptide representing the tryptic linker peptide with the phosphoserine in the proposed position (X<sub>9</sub>GGGGSGGGGpSR) was spiked at 0.5 and 1.0  $\mu\text{M}$  concentration into the mAb1-F tryptic digest and the mixtures were analyzed by LC-MS/MS. A CID-MS/MS spectra of the spiked phosphopeptide alone is shown in the supplemental data (Fig. S1). EICs including the unmodified and modified tryptic peptides ( $z = 2$  and  $3$ ,  $m/z$  windows 561.61-561.63 + 588.26-588.28 + 841.91-841.93 + 881.89-881.91) of tryptic digests without and with addition of 0.5  $\mu\text{M}$  and 1.0  $\mu\text{M}$  synthetic phosphopeptide are illustrated in Fig. 5A-C, respectively. The peak area of the EIC peak at 36.8 min. (Fig. 5A) increased in the spiked samples (Fig. 5B-C), demonstrating the co-elution of the +79.97 Da modified tryptic linker peptide of mAb1-F and the spiked synthetic phosphopeptide, confirming the correct assignment of the phosphoserine in the GS-linker of the IgG-based fusion protein.

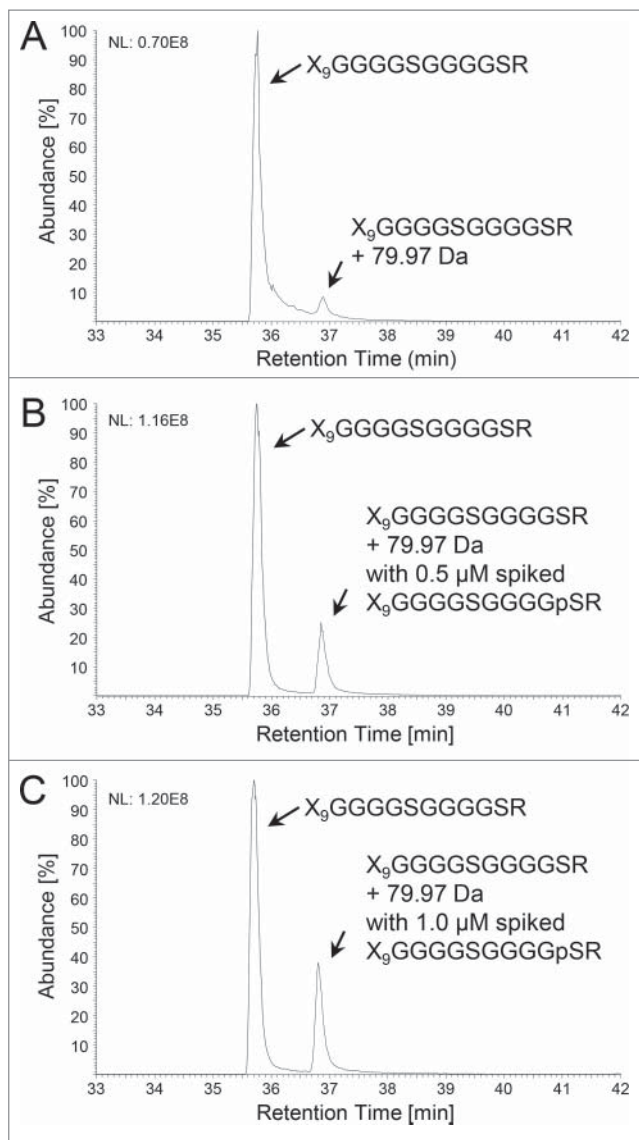
### Enzymatic dephosphorylation of the modified tryptic linker peptide

To test an alternative approach for phosphoserine identification in the tryptic linker peptide, the phosphate group was removed by incubating a freeze dried and rebuffed tryptic digest with alkaline phosphatase prior to LC-MS and CID-MS/MS experiments. Alkaline phosphatase has a broad specificity for phosphate esters of alcohols, amines, pyrophosphate, and phenols, and is routinely used to dephosphorylate proteins, peptides and nucleic acids.<sup>38-41</sup> Following treatment with the phosphatase, the phosphorylated tryptic linker peptide X<sub>9</sub>GGGGSGGGGpSR was no longer detectable by LC-MS (Fig. 6A), verifying the enzymatic removal of the phosphate group (XIC:  $z = 2$  and  $3$ ,  $m/z$  windows 561.61-561.63 + 588.26-588.28 + 841.91-841.93 + 881.89-881.91). A control reaction without alkaline phosphatase verified that the level of the phosphorylated peptide was not influenced by freeze drying the tryptic digest (Fig. 6B). Thus, the dephosphorylation reaction at tryptic peptide level represents an alternative approach for the identification of the phosphoserine in the GS-linker.

### The phosphorylation is specific to the glycine-serine linker I

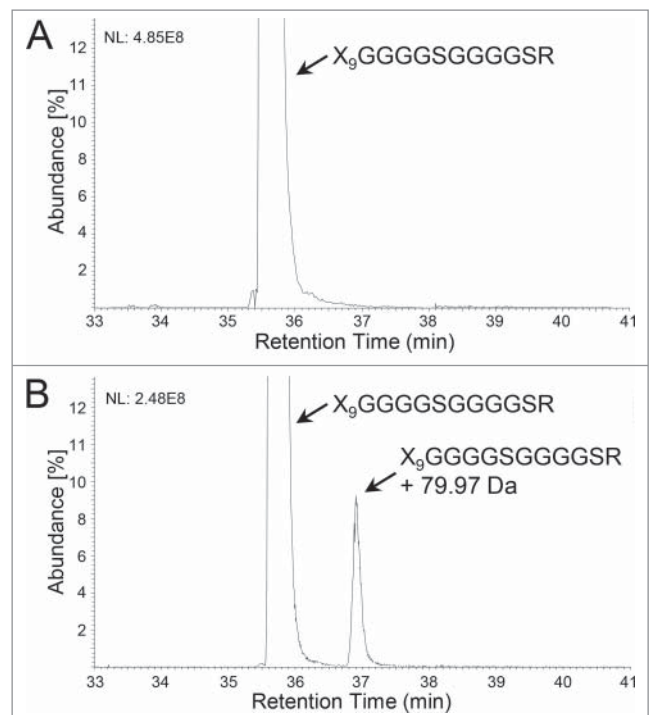
Because the tryptic peptides covering the GS-linkers I and II possess the same amino acid sequence (X<sub>9</sub>GGGGSGGGSR), the UPLC-MS/MS analysis of a tryptic digest of mAb1-F did not allow differentiation between





**Figure 5.** Spiking of a synthetic phosphopeptide to the tryptic digest. Extracted ion current chromatograms ( $z = 2$  and  $3$ ) of the unmodified and the  $+79.97$  Da modified tryptic linker peptide ( $X_9GGGGSGGGSR$ ) of the fused heavy chain of mAb1-F transiently expressed in human embryonic kidney cells (A) without, (B) with  $0.5 \mu M$ , and (C) with  $1.0 \mu M$  spiked synthetic phosphopeptide ( $X_9GGGGSGGGpSR$ ). The spiked samples demonstrated increased peak area of the phosphorylated peptide, supporting the correct identification of the modified tryptic peptide. NL, normalized intensity level.

the 2 linkers (Fig. 2). Also, no tryptic peptides with missed cleavages that would allow a localization of the phosphoserine to one or both of the GS-linkers could be identified in the LC-MS/MS data sets. To definitively elucidate the position of the phosphoserine within the fused heavy chain, further LC-MS/MS analysis was conducted based on a digestion with the endoprotease thermolysin to provide additional fragmentation data C-terminal to the GS-linker sequence GGGGSGGGSR, and thereby differentiate linker I (GGGGSGGGSR) and II (GGGGSGGGSR) (Fig. 2). One thermolysin peptide ( $XGGGGSGGGSR$ ) present in charge state  $z = 2$  and specific for the GS-linker I was present unmodified ( $m/z$  window 665.79–665.81, elution time: 12.5 min) and with a  $+79.97$  Da modification ( $m/z$  window 705.77–705.79, elution time: 13.5 min) (Fig. 7A).



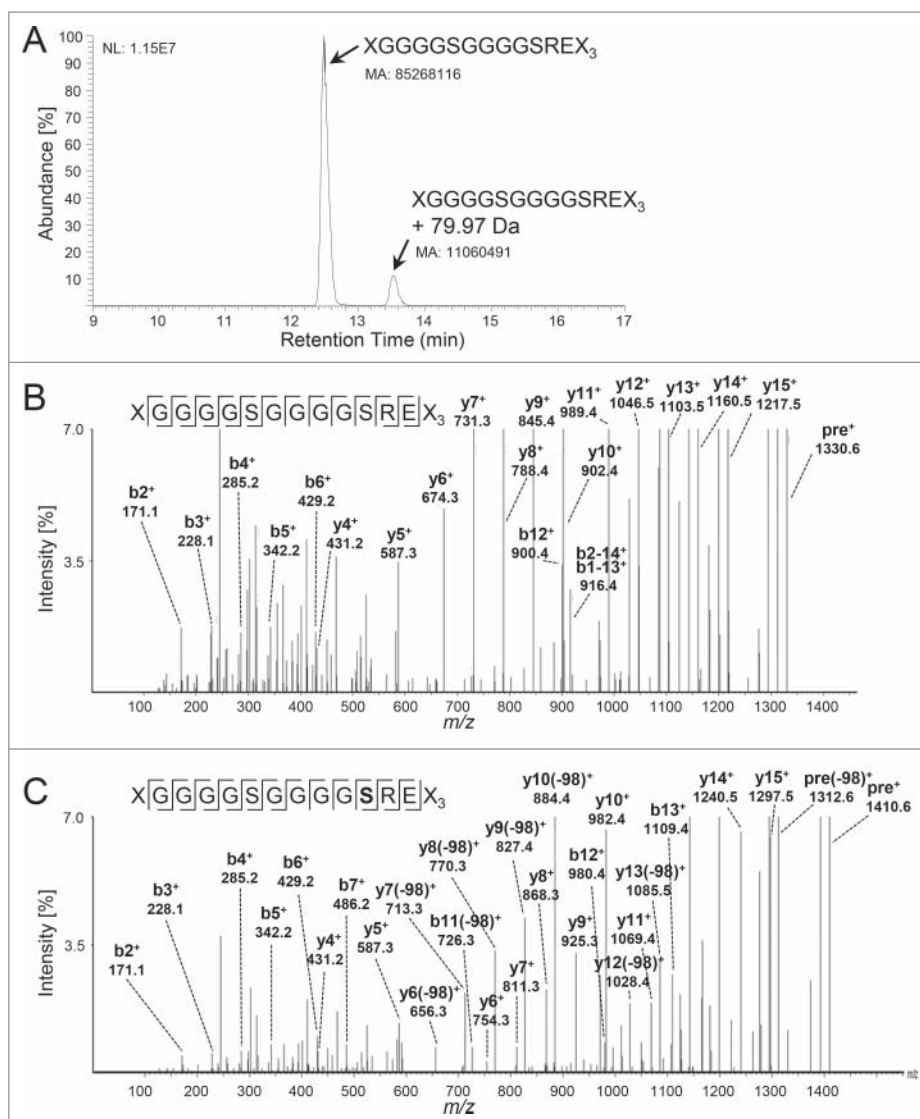
**Figure 6.** Enzymatic dephosphorylation of the modified tryptic linker peptide with alkaline phosphatase. Extracted ion current chromatograms ( $z = 2$  and  $3$ ) of the unmodified and the  $+79.97$  Da modified tryptic linker peptide ( $X_9GGGGSGGGSR$ ) of mAb1-F from human embryonic kidney cells incubated (A) with and (B) without alkaline phosphatase (control reaction). The modified tryptic peptide was no longer detectable following the enzymatic dephosphorylation. NL, normalized intensity level.

As ETD is better suited for peptide analysis containing precursor ions with charge states  $>2$ , the generated thermolysin peptides were analyzed by HCD.<sup>42–43</sup> Orbitrap HCD of the unmodified and modified GS-linker I specific peptide  $XGGGGSGGGSR$  localized the  $+79.97$  Da modification to the C-terminal serine residue based on mass differences between the unmodified ions,  $y_6^+(m/z 674.3)$  to  $y_{15}^+(m/z 1217.5)$  (Fig. 7B), and the corresponding modified fragment ions,  $y_6^+(m/z 754.3)$  to  $y_{15}^+(m/z 1297.5)$  (Fig. 7C). No modified  $y_4^+(m/z 431.2)$  and  $y_5^+(m/z 587.3)$  fragment ions were verified for both peptides.

Integration of the corresponding EIC peaks (Fig. 7A) revealed a phosphorylation of the GS-linker I specific peptide of 11.3% (including 1.4% with *O*-xylose at the GSG motif (data not shown)). Several thermolysin peptides covering the C-terminal of the GS-linker II of the fused heavy chain were also detected, but none were found to be modified with additional  $+79.97$  Da. With this observation, we conclude the phosphorylation to be exclusively present in the GS-linker I. Since the tryptic  $X_9GGGGSGGGSR$  sequence is present in GS-linker I and II of the fused heavy chain, the 11.3% level of phosphorylation for the GS-linker I peptide is in good agreement with 5.5% phosphorylation of the tryptic linker peptide consisting of both GS-linkers.

#### Formation of phosphoserine in CHO versus HEK cells

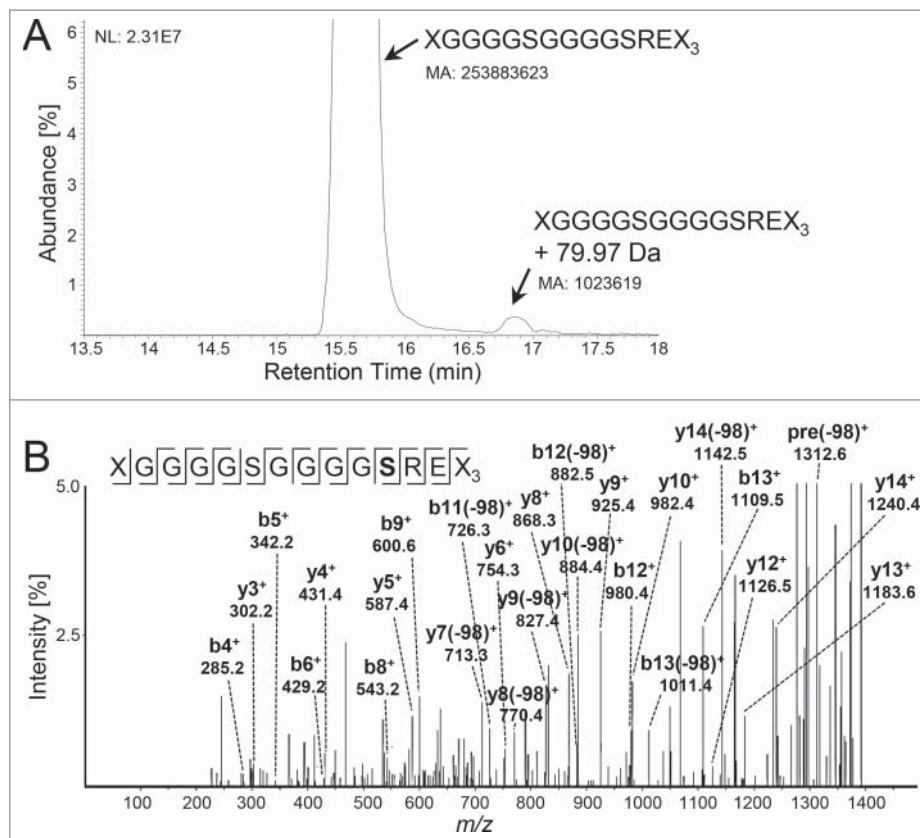
To determine if the extent of phosphoserine in the GS-linker I was affected by the host cell expression system,



**Figure 7.** Specific phosphorylation of the glycine-serine linker I. (A) Extracted ion current chromatogram (EIC) ( $z = 2$ ) of the unmodified (elution time 12.5 min) and the phosphorylated (elution time 13.5 min) thermolysin glycine-serine linker I peptide (XGGGGSGGGGSREX<sub>3</sub>) of mAb1-F transiently expressed in human embryonic kidney cells. A relative comparison of the integrated EIC chromatograms quantified the modification to 11.3% (including 1.4% with *O*-xylose at the GSG motif (data not shown)). Orbitrap HCD-MS/MS data of the double protonated (B) unmodified and (C) modified thermolysin glycine-serine linker I peptide. The +79.97 Da modification is localized to the C-terminal serine residue (bold). MA, manually integrated peak; NL, normalized intensity level.

mAb1-F was stably expressed in CHO cells, and the purified fusion proteins analyzed by UHR-ESI-QTOF-MS. Total mass determination did not demonstrate the presence of an intense signal corresponding to a +79 Da modification in the CHO cell-expressed fusion proteins (data not shown). Thereupon, we analyzed the material by LC-MS/MS peptide mapping. To prevent any potential contamination from other samples containing the phosphorylated GS-linker I, a new reverse phase UPLC column was used to analyze the CHO material by peptide mapping. LC-MS/MS following a thermolysin digest revealed the presence of the +79.97 Da modified GS linker I peptide (XGGGGSGGGGSREX<sub>3</sub>) to a relative level of 0.4% (including 0.3% with *O*-xylose at the GSG motif (data not shown)) (Fig. 8A). The slightly different elution times compared to the equivalent peptides from mAb1-F expressed in HEK cells (Fig. 7A) are explained by use of the new reverse phase column resulting in marginally later elution times. The site of modification, determined by

HCD MS, was localized to the C-terminal serine residue in the linker peptide (Fig. 8B). Modified  $y_6^+$  ( $m/z$  754.3),  $y_8^+$  to  $y_{10}^+$  ( $m/z$  868.3; 925.4; 982.4) and  $y_{12}^+$  to  $y_{14}^+$  ( $m/z$  1126.5; 1183.6; 1240.4) and un-modified  $y_3^+$  to  $y_5^+$  ( $m/z$  302.2; 431.4; 587.4) fragment ions confirm the 79.97 Da modification to be localized to the C-terminal serine residue of the thermolysin peptide. Non-sequence losses from the sequence type product ions  $b_{11}(-98)^+$  ( $m/z$  726.3),  $b_{12}(-98)^+$  ( $m/z$  882.5),  $b_{13}(-98)^+$  ( $m/z$  1011.4),  $y_7(-98)^+$  ( $m/z$  713.3),  $y_8(-98)^+$  ( $m/z$  770.4),  $y_9(-98)^+$  ( $m/z$  827.4), and  $y_{10}(-98)^+$  ( $m/z$  884.4), via the modified amino acid side chain further confirms the modified serine residue (Fig. 8B). As with the enzymatic dephosphorylation of the modified tryptic linker peptide of mAb1-F from HEK cells (Fig. 6), treatment with alkaline phosphatase removed the +79.97 Da modification of the thermolysin GS-linker I peptide of the fusion protein from the stably expressed CHO cells (data not shown).



**Figure 8.** Linker phosphorylation in mAb1-F stably expressed in Chinese hamster ovary cells. (A) Extracted ion current (EIC) chromatograms ( $z = 2$ ) of the unmodified (elution time 15.6 min) and the phosphorylated (elution time 16.85 min) thermolysin glycine-serine linker I peptide (XGGGSGGGGSREX<sub>3</sub>). The phosphorylated peptide was quantified to 0.4%. (B) HCD-MS/MS data of the double protonated modified thermolysin peptide. MA, manually integrated peak; NL, normalized intensity level.

## Discussion

Reversible post-translational phosphorylation, principally on serine, threonine or tyrosine residues, is one of the most widespread and important PTMs for regulating protein activity in eukaryotic cells. Many signal transduction pathways and cellular processes are regulated by signals due to protein phosphorylation.<sup>44-46</sup> More than 500 kinases have been predicted from the human genome, and the phosphorylation of the hydroxy group in serine to produce phosphoserine is catalyzed by various types of kinases.<sup>44</sup>

The analysis by LC-MS/MS following thermolysin digests of mAb1-F produced in CHO and HEK cells demonstrated the phosphorylation to be specific to the GS-linker I. Because the amino acids in the positions  $-1$  to  $-9$  relative to the affected C-terminal serine residues in the GS-linker I and II are identical, we speculated that neighboring residues at positions  $+1$ ,  $+2$ ,  $+3$  etc. determine the susceptibility in this case. Several protein phosphorylation site prediction tools have been developed and are available online. The NetPhos 2.0 server (online available at [www.cbs.dtu.dk/services/NetPhos/](http://www.cbs.dtu.dk/services/NetPhos/)) produces neural network predictions for serine, threonine and tyrosine phosphorylation sites in eukaryotic proteins.<sup>47</sup> Using NetPhos 2.0, we analyzed the likelihood of phosphorylation of the C-terminal serine residues of the GS-linker I (position: GGSRE) and II (position: GGSRT). NetPhos 2.0 had a score of 0.970 for GGSRE, indicating a very likely phosphorylation site, whereas

GSRT had a score of only 0.069, indicating that the confidence for this site being a true phosphorylation site is quite low.

The sequence logo of serine phosphorylation sites emphasizes amino acid residues that are frequently found, and indicate the position  $+2$  and  $+3$  to be dominated by glutamic acid.<sup>47</sup> The  $+2$  position of the phosphorylated serine residue in the GS-linker I is occupied by a glutamic acid and, when substituting the glutamic acid in the  $+2$  position with an alanine, NetPhos 2.0 had a low score of only 0.144, indicating the  $+2$  glutamic acid to be detrimental for the serine residue to be a site prone to phosphorylation. Consequently, the reported post-translational phosphorylation could be attributed to a motif (SXE) predicted to be a site of phosphorylation in eukaryotic proteins and introduced upon fusing the GS-linker with the  $>10$  kDa non-IgG protein.

The specific kinase(s) responsible for the phosphorylation of the GS-linker I of mAb1-F produced in HEK and CHO cells is (are) unknown. The substrate specificities of protein kinases are based on the target amino acid and also on flanking consensus sequences being dominated by acidic, basic or hydrophobic residues. Several online tools are available for the prediction of kinase-specific phosphorylation sites;<sup>48-49</sup> however, evaluation by GPS 3.0 (Group-based Prediction System, online available at <http://gps.biocuckoo.org><sup>50</sup>), and NetPhos 3.1 (online available at <http://www.cbs.dtu.dk/services/NetPhos/>) suggested multiple kinases to be potentially involved in the phosphorylation of the GS-linker I.



The level of phosphorylated GS-linker I was greatly affected by whether mAb1-F was expressed transiently in HEK cells (11.3%) or stably expressed in CHO cells (0.4%). The reasons for the difference in phosphorylation could include differences in kinases expressed in the 2 cells, levels of specific kinase activities, localization of kinases to distinct subcellular compartments, or differences in the availability of co-factors needed by the enzymes. Also, the size of the phosphoproteome in a given cell is dependent upon the temporal and spatial balance of kinase and phosphatase concentrations in the cell. Moreover, the level of linker phosphorylation may differ between potential kinase-specific consensus sequences and types of kinases involved. In this instance, there was no detectable difference in the target bindings or biological activities of the HEK and CHO mAb1-F proteins (data not shown), nor could any differences be determined in the storage stability of the 2 protein materials. This may not be surprising as the phosphorylation is distant from the active sites of the > 10 kDa protein of the fused heavy chain. Nevertheless, compared with the *O*-xylosylation of GS-linkers, which cannot be prevented with the consensus sequence GSG being present, the reported phosphorylation and heterogeneity in therapeutic proteins can be prevented by avoiding a potential phosphorylation site upon fusion with the GS-linker.

## Materials and methods

### Reagents and enzymes

Trypsin, thermolysin, and alkaline phosphatase from bovine intestinal mucosa were ordered from Sigma-Aldrich. The 20-mer synthetic peptide NH<sub>2</sub>-X<sub>9</sub>GGGGSGGGGpSR-COOH (98% HPLC-purity) was synthesized at Biosyntan GmbH. PNGase F was obtained from Roche Diagnostics GmbH, Cus-  
tom Biotech.

### Recombinant expression and purification of the fusion protein

The coding regions for the chains comprising mAb1-F were cloned into expression vectors enabling secretory expression in HEK or CHO cells. Plasmid transfections into HEK293-F cells (Invitrogen) were performed according to the cell supplier's instructions using Maxiprep (Qiagen) preparations of the antibody vectors, Opti-MEM I medium, 293fectin, and an initial cell density of 1–2 × 10<sup>6</sup> viable cells/mL in serum-free FreeStyle 293 expression medium (all reagents from Invitrogen). Fusion protein containing cell culture supernatants were harvested after 7 days of cultivation in shake flasks by centrifugation at 14,000 × *g* for 30 min and filtered through a 0.22 μm sterile filter. The fusion protein was purified directly from the supernatant, or the supernatant was stored at –80°C until chromatographic purification. Stable expression in CHO cells was performed following the selection and 250 L fed batch fermentation of 2 stable transfected cell lines.

### UHR-ESI-QTOF-MS

The fusion protein was deglycosylated using PNGase F and reduced in 100 mM TCEP and desalted by HPLC on a Sephadex

G25 5×250 mm column (Amersham Biosciences) using 40% acetonitrile with 2% formic acid (v/v). The total mass was determined by ESI-QTOF MS on a maXis 4G UHR-QTOF MS system (Bruker Daltonik) equipped with a TriVersa NanoMate source (Advion). Calibration was performed with the ESI-L Low Concentration Tuning Mix (Agilent Technologies). For the deglycosylated and reduced fusion protein, data acquisition was done at 600–2000 *m/z* (ISCID: 0.0 eV). For visualization of the results, an in-house developed software tool was used to transform the *m/z* spectra into deconvoluted mass spectra.

### UPLC-MS/MS peptide mapping

mAb1-F was denatured and reduced in 0.3 M Tris-HCl pH 8, 6 M guanidine-HCl and 20 mM dithiothreitol (DTT) at 37° C for 1 hour, and alkylated by adding 40 mM iodoacetic acid (C13: 99%) (Sigma-Aldrich) and incubating at room temperature in the dark for 15 min. Excess iodoacetic acid was inactivated by adding further 20 mM DTT to the reaction. The alkylated fusion protein was buffer exchanged using NAP5 gel filtration columns and a proteolytic digestion with trypsin performed in 50 mM Tris-HCl, pH 7.5 at 37° C for 16 hours. The reaction was stopped by adding formic acid to 0.4% (v/v). Digestions with thermolysin were performed in 25 mM Tris-HCl, 1 mM CaCl<sub>2</sub>, pH 8.3 at 25° C for 30 minutes and stopped by adding EDTA to 8 mM. The digested samples were stored at –80°C and analyzed by UPLC-MS/MS using a nanoAcquity UPLC (Waters) coupled to a TriVersa NanoMate (Advion) and an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). About 2.4 μg digested fusion protein was injected in 5 μL. Chromatographic separation was performed by reversed-phase on a Acquity BEH300 C18 column, 1x 150 mm, 1.7 μm, 300 Å (Waters) using a flow rate of 60 μL/min. The mobile phase A and B contained 0.1% (v/v) formic acid in UPLC grade water and acetonitrile, respectively. A column temperature of 50°C was used and a gradient of 1% to 40% mobile phase B over 90 min followed by an increase to 99% mobile phase B for 2 min and a re-equilibration step at 1% mobile phase B for 6 min was applied. Two injections of mobile phase A were performed between sample injections using a 50 min gradient to prevent carry-over between samples. The effluent was split post column using the TriVersa Nanomate, and a nanoliter flow portion directed into the mass spectrometer.

High-resolution MS spectra were acquired with the Orbitrap mass analyzer, and parallel detection of CID MS/MS fragment ion spectra in the ion trap with dynamic exclusion enabled (repeat count of 1, exclusion duration of 15s (±10 ppm)). The Orbitrap Fusion was used in the data-dependent mode. Essential MS settings were: full MS (AGC: 2 × 10<sup>5</sup>, resolution: 6 × 10<sup>4</sup>, *m/z* range: 300–2000, maximum injection time: 100 ms); MS/MS (AGC: 1 × 10<sup>4</sup>, maximum injection time: 100 ms, isolation width: 2 Da). Normalized collision energy was set to 35%, activation *p*: 0.25, isolation width: 2 Da. For methods where exclusively HCD MS/MS spectra were acquired, an Orbitrap full MS scan was followed by up to 20 HCD Orbitrap MS/MS spectra on the most abundant ions. The AGC for MS/MS experiments was set to 5×10<sup>4</sup> at a maximum injection time of 500 ms. Normalized collision energy was set to 20%, and



HCD fragmentation ions were detected in the Orbitrap at a resolution setting of  $15 \times 10^3$ . All other settings were as described for the method using exclusively CID fragmentation. A complementary EThcD method based on HCD and ETD as data dependent fragmentation techniques involved full scan MS acquired with the Orbitrap mass analyzer, and parallel detection of ETD and HCD fragment ion spectra in the ion trap and Orbitrap mass analyzer, respectively. A fixed cycle time was set for the full scan with as many as possible data dependent MS/MS scans. Full MS: same setting as for CID and HCD. For HCD, the MS/MS setting were the same as listed above. For ETD, MS/MS the settings were as follows: reaction time was set to 50 ms, ETD reagent target:  $1 \times 10^6$ , maximum injection time: 200 ms. ETD supplemental activation was enabled. Supplemental activation collision energy was set to 25%. The AGC target was set to  $1 \times 10^4$ , the precursor isolation width was 2 Da and the maximum injection time was set to 250 ms.

The analysis of the LC-MS/MS data and the PTM identification was performed using the PEAKS studio 6.0 and 7.5 software (Bioinformatics Solutions Inc.) with the preprocessing option used and PepFinder software (Thermo Fisher Scientific). Manual data interpretation and quantification was performed using XCalibur software (Thermo Fisher Scientific). GPMAW (Lighthouse data) was used to calculate theoretical masses and the XICs were generated with the most intense isotope mass using a mass tolerance of 8 ppm.

### Spiking of a synthetic phosphopeptide

Based on a calculated amount of the modified tryptic peptide in mAb1-F, 0.5, and 1.0  $\mu\text{M}$  of a synthetic pSer-containing peptide  $\text{NH}_2\text{-X}_9\text{GGGSGGGGpSR-COOH}$  was spiked into the tryptic digest. 2.4  $\mu\text{g}$  tryptic digest with or without spiked synthetic peptide was analyzed by LC-MS/MS.

### Enzymatic dephosphorylation

Enzymatic dephosphorylation of the +79.97 Da modified tryptic linker peptide was performed by freeze drying  $\sim 62 \mu\text{g}$  tryptic digest. The peptides were resuspended in 25  $\mu\text{L}$  100 mM Tris-HCl, 5 mM  $\text{MnCl}_2$ , pH 8.0, and incubated with 250 units of alkaline phosphatase at 37°C for 1 h. The digested samples were stored at  $-80^\circ\text{C}$ .

### Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

### Acknowledgments

We are grateful to all Roche colleagues who contributed with insightful discussions and guidance.

### ORCID

Claudia Ferrara  <http://orcid.org/0000-0002-3470-4191>

### References

- Spieß C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Mol Immunol* 2015; 67:95-106; PMID:25637431; <http://dx.doi.org/10.1016/j.molimm.2015.01.003>
- Weidle UH, Kontermann RE, Brinkmann U. Tumor-antigen-binding bispecific antibodies for cancer treatment. *Semin Oncol* 2014; 41:653-60; PMID:25440609; <http://dx.doi.org/10.1053/j.seminoncol.2014.08.004>
- Kontermann RE, Brinkmann U. Bispecific antibodies. *Drug Discov Today* 2015; 20:838-47; PMID:25728220; <http://dx.doi.org/10.1016/j.drudis.2015.02.008>
- Pardridge WM. Biopharmaceutical drug targeting to the brain. *J Drug Target* 2010; 18:157-67; PMID:20064077; <http://dx.doi.org/10.3109/10611860903548354>
- Niewoehner J, Bohrmann B, Collin L, Urich E, Sade H, Maier P, Rueger P, Stracke JO, Lau W, Tissot AC, Loetscher H, Ghosh A, Freskgård PO. Increased brain penetration and potency of a therapeutic antibody using a monovalent molecular shuttle. *Neuron* 2014; 81:49-60; PMID:24411731; <http://dx.doi.org/10.1016/j.neuron.2013.10.061>
- Amet N, Lee HF, Shen WC. Insertion of the designed helical linker led to increased expression of tf-based fusion proteins. *Pharm Res* 2009; 26:523-8; PMID:19002568; <http://dx.doi.org/10.1007/s11095-008-9767-0>
- Bai Y, Shen WC. Improving the oral efficacy of recombinant granulocyte colony-stimulating factor and transferrin fusion protein by spacer optimization. *Pharm Res* 2006; 23:2116-21; PMID:16952003; <http://dx.doi.org/10.1007/s11095-006-9059-5>
- Zhao HL, Yao XQ, Xue C, Wang Y, Xiong XH, Liu ZM. Increasing the homogeneity, stability and activity of human serum albumin and interferon-alpha2b fusion protein by linker engineering. *Protein Expr Purif* 2008; 61:73-7; PMID:18541441; <http://dx.doi.org/10.1016/j.pep.2008.04.013>
- Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 2013; 65:1357-69; PMID:23026637; <http://dx.doi.org/10.1016/j.addr.2012.09.039>
- Chichili VPR, Kumar V, Sivaraman J. Linkers in the structural biology of protein-protein interactions. *Protein Sci* 2013; 22:153-67; PMID:23225024; <http://dx.doi.org/10.1002/pro.2206>
- Klein JS, Jiang S, Galimidi RP, Keeffe JR, Bjorkman PJ. Design and characterization of structured protein linkers with differing flexibilities. *Protein Eng Des Sel* 2014; 27:325-30; PMID:25301959; <http://dx.doi.org/10.1093/protein/gzu043>
- Argos P. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J Mol Biol* 1990; 211:943-58; PMID:2313701; [http://dx.doi.org/10.1016/0022-2836\(90\)90085-Z](http://dx.doi.org/10.1016/0022-2836(90)90085-Z)
- Huston JS, Levinson D, Mudgett-Hunter M, Tai MS, Novotný J, Margolies MN, Ridge RJ, Brucoleri RE, Haber E, Crea R, et al. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc Natl Acad Sci U S A* 1988; 85:5879-83; PMID:3045807; <http://dx.doi.org/10.1073/pnas.85.16.5879>
- Desplancq D, King DJ, Lawson AD, Mountain A. Multimerization behaviour of single chain Fv variants for the tumour-binding antibody B72.3. *Protein Eng* 1994; 7:1027-33; PMID:7809029; <http://dx.doi.org/10.1093/protein/7.8.1027>
- Hagemeyer CE, von Zur Muhlen C, von Elverfeldt D, Peter K. Single-chain antibodies as diagnostic tools and therapeutic agents. *Thromb Haemost* 2009; 101:1012-9; PMID:19492141; <http://dx.doi.org/10.1160/TH08-12-0816>
- Spieß C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Mol Immunol* 2015; 67:95-106; PMID:25637431; <http://dx.doi.org/10.1016/j.molimm.2015.01.003>
- Orcutt KD, Ackerman ME, Cieslewicz M, Quiroz E, Slusarczyk AL, Frangioni JV, Wittrup KD. A modular IgG-scFv bispecific antibody topology. *Protein Eng Des Sel* 2010; 23:221-8; PMID:20019028; <http://dx.doi.org/10.1093/protein/gzp077>
- Wen D, Foley SF, Hronowski XL, Gu S, Meier W. Discovery and investigation of O-xylosylation in engineered proteins containing a

- (GGGG)n linker. *Anal Chem* 2013; 85:4805-12; PMID:23581628; <http://dx.doi.org/10.1021/ac400596g>
19. Castoldi R, Jucknischke U, Pradel LP, Arnold E, Klein C, Scheiblich S, Niederfellner G, Sustmann C. Molecular characterization of novel trispecific ErbB-cMet-IGF1R antibodies and their antigen-binding properties. *Protein Eng Des Sel* 2012; 25:551-9; PMID:22936109; <http://dx.doi.org/10.1093/protein/gz048>
  20. Mayer K, Baumann AL, Grote M, Seeber S, Kettenberger H, Breuer S, Killian T, Schäfer W, Brinkmann U. TriFabs–Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery. *Int J Mol Sci* 2015; 16:27497-507; PMID:26593903; <http://dx.doi.org/10.3390/ijms161126037>
  21. Spahr C, Shi SD, Lu HS. O-glycosylation of glycine-serine linkers in recombinant Fc-fusion proteins: attachment of glycosaminoglycans and other intermediates with phosphorylation at the xylose sugar subunit. *mAbs* 2014; 6:904-14; PMID:24927272; <http://dx.doi.org/10.4161/mabs.28763>
  22. Reusch U, Burkhardt C, Fucek I, Le Gall F, Le Gall M, Hoffmann K, Knackmuss SH, Kiprijanov S, Little M, Zhukovsky EA. A novel tetravalent bispecific TandAb (CD30/CD16A) efficiently recruits NK cells for the lysis of CD30+ tumor cells. *mAbs* 2014; 6:728-39; PMID:24670809; <http://dx.doi.org/10.4161/mabs.28591>
  23. Reusch U, Duell J, Ellwanger K, Herbrecht C, Knackmuss SH, Fucek I, Eser M, McAleese F, Molkenhuth V, Gall FL, et al. A tetravalent bispecific TandAb (CD19/CD3), AFM11, efficiently recruits T cells for the potent lysis of CD19C tumor cells. *mAbs* 2015; 7:584-604; PMID:25875246; <http://dx.doi.org/10.1080/19420862.2015.1029216>
  24. Mack M, Riethmüller G, Kufer P. A small bispecific antibody construct expressed as a functional single-chain molecule with high tumor cell cytotoxicity. *Proc Natl Acad Sci U S A* 1995; 92:7021-5; PMID:7624362; <http://dx.doi.org/10.1073/pnas.92.15.7021>
  25. Turner J, Schneider SM. Blinatumomab: A New Treatment for Adults With Relapsed Acute Lymphocytic Leukemia. *Clin J Oncol Nurs* 2016; 20:165-8; PMID:26991709; <http://dx.doi.org/10.1188/16.CJON.165-168>
  26. Arndt MA, Krauss J, Vu BK, Newton DL, Rybak SM. A dimeric angiogenin immunofusion protein mediates selective toxicity toward CD22+ tumor cells. *J Immunother* 2005; 28:245-51; PMID:15838381; <http://dx.doi.org/10.1097/01.cji.0000161396.96582.10>
  27. Weimer T, Wormsbächer W, Kronthaler U, Lang W, Liebing U, Schulte S. Prolonged in-vivo half-life of factor VIIa by fusion to albumin. *Thromb Haemost* 2008; 99:659-67; PMID:18392323; <http://dx.doi.org/10.1160/TH07-08-0525>
  28. Friedman M, Lindström S, Ekerljung L, Andersson-Svahn H, Carlsson J, Brismar H, Gedda L, Frejd FY, Ståhl S. Engineering and characterization of a bispecific HER2 x EGFR-binding antibody molecule. *Bio-technol Appl Biochem* 2009; 54:121-31; PMID:19492986; <http://dx.doi.org/10.1042/BA20090096>
  29. Habberger M, Bomans K, Diepold K, Hook M, Gassner J, Schlothauer T, Zwick A, Spick C, Kepert JF, Hienz B, et al. Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: Susceptibility vs. functionality of critical quality attributes. *mAbs* 2014; 6:327-39; PMID:24441081; <http://dx.doi.org/10.4161/mabs.27876>
  30. Li Y, Monine M, Huang Y, Swann P, Nestorov I, Lyubarskaya Y. Quantitation and pharmacokinetic modeling of therapeutic antibody quality attributes in human studies. *mAbs*. 2016 Aug-Sep;8(6):1079-87; PMID:27216574; <http://dx.doi.org/10.1080/19420862.2016.1186322>
  31. Spencer D, Novarra S, Zhu L, Mugabe S, Thisted T, Baca M, Depaz R, Barton C. O-xylosylation in a recombinant protein is directed at a common motif on glycine-serine linkers. *J Pharm Sci* 2013; 102:3920-4; PMID:24105735; <http://dx.doi.org/10.1002/jps.23733>
  32. Spahr C, Kim JJ, Deng S, Kodama P, Xia Z, Tang J, Zhang R, Siu S, Nuanmanee N, Estes B, et al. Recombinant human lecithin-cholesterol acyltransferase Fc fusion: analysis of N- and O-linked glycans and identification and elimination of a xylose-based O-linked tetrasaccharide core in the linker region. *Protein Sci* 2013; 22:1739-53; PMID:24115046; <http://dx.doi.org/10.1002/pro.2373>
  33. Steen H, Jebanathirajah JA, Rush J, Morrice N, Kirschner MW. Phosphorylation analysis by mass spectrometry: myths, facts, and the consequences for qualitative and quantitative measurements. *Mol Cell Proteomics* 2006; 5:172-81; PMID:16204703; <http://dx.doi.org/10.1074/mcp.M500135-MCP200>
  34. DeGnore JP, Qin J. Fragmentation of phosphopeptides in an ion trap mass spectrometer. *J Am Soc Mass Spectrom* 1998; 9:1175-88; PMID:9794085; [http://dx.doi.org/10.1016/S1044-0305\(98\)00088-9](http://dx.doi.org/10.1016/S1044-0305(98)00088-9)
  35. Cui L, Yapici I, Borhan B, Reid GE. Quantification of competing H<sub>3</sub>PO<sub>4</sub> versus HPO<sub>3</sub> + H<sub>2</sub>O neutral losses from regioselective <sup>18</sup>O-labeled phosphopeptides. *J Am Soc Mass Spectrom* 2014; 25:141-8; PMID:24249041; <http://dx.doi.org/10.1007/s13361-013-0744-4>
  36. Moyer SC, Cotter RJ, Woods AS. Fragmentation of phosphopeptides by atmospheric pressure MALDI and ESI/Ion trap mass spectrometry. *J Am Soc Mass Spectrom* 2002; 13:274-83; PMID:11908807; [http://dx.doi.org/10.1016/S1044-0305\(01\)00361-0](http://dx.doi.org/10.1016/S1044-0305(01)00361-0)
  37. Frese CK, Zhou H, Taus T, Altelaar AF, Mechtler K, Heck AJ, Mohammed S. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (ET<sub>h</sub>CD). *J Proteome Res* 2013; 12:1520-5; PMID:23347405; <http://dx.doi.org/10.1021/pr301130k>
  38. Fernley HN. Mammalian alkaline phosphatase. In: Boyer PD (ed.), *The Enzymes*, 3rd ed, Academic Press New York London 1971:417-47; [http://dx.doi.org/10.1016/S1874-6047\(08\)60378-9](http://dx.doi.org/10.1016/S1874-6047(08)60378-9)
  39. Morton RK. The substrate specificity and inhibition of alkaline phosphatases of cow's milk and calf intestinal mucosa. *Biochem J* 1955; 61:232-40; PMID:13260203; <http://dx.doi.org/10.1042/bj0610232>
  40. Morton RK. The action of purified alkaline phosphatases on di- and tri-phosphopyridine nucleotides. *Biochem J* 1955; 61:240-4; PMID:13260204; <http://dx.doi.org/10.1042/bj0610240>
  41. McLachlin DT, Chait BT. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr Opin Chem Biol* 2001; 5:591-602; PMID:11578935; [http://dx.doi.org/10.1016/S1367-5931\(00\)00250-7](http://dx.doi.org/10.1016/S1367-5931(00)00250-7)
  42. Good DM, Wirtala M, McAlister GC, Coon JJ. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol Cell Proteomics* 2007; 6:1942-51; PMID:17673454; <http://dx.doi.org/10.1074/mcp.M700073-MCP200>
  43. Molina H, Matthiesen R, Kandasamy K, Pandey A. Comprehensive comparison of collision induced dissociation and electron transfer dissociation. *Anal Chem* 2008; 80:4825-35; PMID:18540640; <http://dx.doi.org/10.1021/ac8007785>
  44. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298:1912-34; PMID:12471243; <http://dx.doi.org/10.1126/science.1075762>
  45. Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 2002; 27:514-20; PMID:12368087; [http://dx.doi.org/10.1016/S0968-0004\(02\)02179-5](http://dx.doi.org/10.1016/S0968-0004(02)02179-5)
  46. Ubersax JA, Ferrell JE Jr. Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol* 2007; 8:530-41; PMID:17585314; <http://dx.doi.org/10.1038/nrm2203>
  47. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 1999; 294:1351-62; PMID:10600390; <http://dx.doi.org/10.1006/jmbi.1999.3310>
  48. Pawson T, Nash P. Assembly of cell regulatory systems through protein interaction domains. *Science* 2003; 300:445-52; PMID:12702867; <http://dx.doi.org/10.1126/science.1083653>
  49. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 2004; 4:1633-49; PMID:15174133; <http://dx.doi.org/10.1002/pmic.200300771>
  50. Xue Y, Ren J, Gao X, Jin C, Wen L, Yao X. GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol Cell Proteomics* 2008; 7:1598-608; PMID:18463090; <http://dx.doi.org/10.1074/mcp.M700574-MCP200>