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

Oksana Tyshchuka, Hans Rainer Völger, Claudia Ferrara, Patrick Bulau, Hans Koll, and Michael Mølhøj

Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Roche Diagnostics GmbH, Penzberg, Germany; Oncology Discovery & Translational Area, Roche Innovation Center Zurich, Schlieren, Switzerland; 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 [(G4S)3] of a fused heavy chain containing in total 2 (G4S)2-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

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. 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. 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. 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. A widely applied protease-resistant GS-linker has the sequence of (G4S)n as proposed by a study of naturally occurring linkers in multi-domain proteins. 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 G4S-units. (G4S)n-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 (VH) and the light chain variable domain (VL) and hereby facilitating the association of the 2 domains. The flexibility of the (G4S)3-linker allows the correct orientation of the VH and VL domains and does not interfere with their folding. Other antibody formats and fusions involving GS-linkers have been reported and are currently in clinical trials.

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. Generally, unexpected heterogeneities are unwanted in biotherapeutic proteins.
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. 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.

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 (G4S)2-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 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 [(G4S)2] 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 X9GGGGSGGGGSR covering a GS-linker [(G4S)2] and the presence of the same peptide with a mass corresponding to a 79.97 Da modification.
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-xylene 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 localized to the C-terminal serine residue (bold).

![Figure 4](image-url)

**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 (X9GGGGSGGGGSR).

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.33 If a peptide contains more hydrophilic than non-phosphorylated peptides. However, the phosphorylation is specific to the glycine-serine linker I

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 (X9GGGGSGGGGpSR) was spiked at 0.5 and 1.0 µM 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 µM and 1.0 µM 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 phospate group was removed by incubating a freeze dried and rebuffered 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.38-41 Following treatment with the phosphatase, the phosphorylated tryptic linker peptide X9GGGGSGGGGpSR 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 (X9GGGGSGGGGpSR), the UPLC-MS/MS analysis of a tryptic digest of mAb1-F did not allow differentiation between
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 GGGGSGGGGSR, and thereby differentiate linker I (GGGGSGGGGSR) and II (GGGGSGGGGSR) (Fig. 2).

One thermolysin peptide (XGGGGSGGGGSR) 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). As ETD is better suited for peptide analysis containing precursor ions with charge states >2, the generated thermolysin peptides were analyzed by HCD.42-43 Orbitrap HCD of the unmodified and modified GS-linker I specific peptide XGGGGSGGGGSR localized the +79.97 Da modification to the C-terminal serine residue based on mass differences between the unmodified ions, y6C (m/z 674.3) to y15C (m/z 1217.5) (Fig. 7B), and the corresponding modified fragment ions, y6C (m/z 754.3) to y15C (m/z 1297.5) (Fig. 7C). No modified y4C (m/z 431.2) and y5C (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 XGGGGSGGGGSR 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,
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 C79 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 C79.97 Da modified GS linker I peptide (XGGGGSGGGGSREX3) at 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 of the thermolysin peptide. Non-sequence losses from the sequence type product ions b11(−98)+(m/z 726.3), b12(−98)+ (m/z 882.5), b13(−98)+ (m/z 1011.4), y7(−98)+ (m/z 713.3), y8(−98)+ (m/z 770.4), y9(−98)+ (m/z 827.4), and y10(−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 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 (XGGGGSGGGGSREX3) 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.
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. 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.

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 C1, C2, C3 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/) produces neural network predictions for serine, threonine and tyrosine phosphorylation sites in eukaryotic proteins. 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 GGSRT 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. 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 however, evaluation by GPS 3.0 (Group-based Prediction System, online available at http://gps.biocuckoo.org) 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.

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 (XGGGGSGGGGSRE3). 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.
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 phosphorylation site upon fusion with the GS-linker.


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 CaCl2, 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 an Acquity BEH300 C18 column, 1x 150 mm, 1.7 µm, 300 A (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 × 10^6, resolution: 6 × 10^5, m/z range: 300-2000, maximum injection time: 100 ms); MS/MS (AGC: 1 × 10^5, 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 5×10^5 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 ETnCD 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: 1x $10^6$, maximum injection time: 200 ms. ETD supplemental activation was enabled. Supplemental activation collision energy was set to 25%. The AGC target was set to $1 \times 10^6$, 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 (Lighthouse data) was used to calculate theoretical masses and theoretically correct peptide was analyzed by LC-MS/MS.

**Spiking of a synthetic phosphopeptide**

Based on a calculated amount of the modified tryptic peptide in mAb1-F, 0.5, and 1.0 μM of a synthetic pSer-containing peptide NH$_2$-X$_9$GGGGSGGGGpSR-COOH was spiked into the tryptic digest. 2.4 μg 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 ~62 μg tryptic digest. The peptides were resuspended in 25 μL 100 mM Tris-HCl, 5 mM 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 $-80°C$.

**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](http://orcid.org/0000-0002-3470-4191)
A novel tetravalent 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

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:11908887; http://dx.doi.org/10.1016/S1044-0395(01)00361-0

Fresé CK, Zhou H, Taus T, Alehaart AF, Mechtler K, Heck AJ, Mohammed S. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (ETcID). J Proteome Res 2013; 12:1520-5; PMID:23347405; http://dx.doi.org/10.1021/pr301130k

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

Morton RK. The action of purified alkaline phosphatases on di- and tri-phosphorodine nucleotides. Biochem J 1955; 61:240-4; PMID:13260204; http://dx.doi.org/10.1042/bj0610240

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

Good DM, Wirtala M, McAllister 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

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

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002; 298:1912-34; PMID:12747124; http://dx.doi.org/10.1126/science.1075762

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:12356807; http://dx.doi.org/10.1016/S0968-0004(02)01759-5

Übersax 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

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

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

Blom N, Sicheritz-Ponten 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/pmc.200300771

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