Comparability analysis of anti-CD20 commercial (rituximab) and RNAi-mediated fucosylated antibodies by two LC-MS approaches

Chen Li,1 Anthony Rossomando,2† Shiaw-Lin Wu1,* and Barry L. Karger1,*

1Barnett Institute and Department of Chemistry and Chemical Biology; Northeastern University; Boston, MA USA; 2Alnylam Pharmaceuticals; Cambridge, MA USA

†Current affiliation: Synageva BioPharma Corp.; Lexington, MA USA

Keywords: biosimilar, LC-MS, sequence mutation, free cysteine, disulfide scrambling, structure characterization

Introduction

Rituximab, a therapeutic monoclonal antibody (mAb) targeting CD20 in B cells, is used to treat B cell non-Hodgkin lymphoma and rheumatoid arthritis. The product’s brand names are Rituxan® (in the United States) and MabThera® (in Europe). One important function for the antibody is to induce antibody-dependent cell-mediated cytotoxicity (ADCC), in which the Fc domain, including the glycans, binds specifically to Fc receptors in human effector cells, such as macrophages and natural killer cells, to induce ADCC. Because the glycan structure, particularly the core fucose, is important to mediate ADCC, the reduction of the core fucose (i.e., by RNAi) should enhance the effect. Thus, RNAi-mediated fucosyltransferase (FUT8) and GDP-man-4,6-dehydratase (GMDS) was used to produce anti-CD20 mAb for this purpose. Although the aim was for a biobetter product, the overall structure, except for the level of the core fucose, was intended to be as similar as possible to the reference product to maintain the drug integrity.

In this study, we first used state of the art mass spectrometric methods to characterize the structure of the newly developed RNAi-mediated anti-CD20 mAb and then compared it to the structure of the commercial rituximab molecule. As expected, reduction of core fucosylation was observed for the RNAi-mediated molecule. On the other hand, the primary structure, disulfide linkages, and common modifications such as pyro-Glu formation at the N-terminus, K clipping at the C-terminus, oxidation at Met, and deamidation at Asn, were found to be similar between the two products. The liquid chromatography–mass spectrometry (LC-MS) used for full sequence analysis, however, identified two amino acid residues mutated on the RNAi-mediated molecule. An alternative LC-MS method, using dimethyl labeled with 2CH₂ for rituximab, and isotopically-dimethyl labeled with 2CD₂ for the RNAi-mediated molecule confirmed the amino acid changes on the RNAi-mediated molecule. Moreover, both approaches were in agreement that one variant was fully mutated and the other partially mutated. Small amounts of free cysteines in both molecules were also observed.

*Correspondence to: Shiaw-Lin Wu and Barry L. Karger; Email: si.wu@neu.edu and b.karger@neu.edu
Submitted: 03/09/13; Revised: 04/01/13; Accepted: 04/22/13
http://dx.doi.org/10.4161/mabs.24814
Disulfide scrambling, which could be caused by the free cysteines, was detected in both mAbs. At pH 6.8 or pH 8, which are typical enzymatic digestion conditions, a small amount of disulfide scrambling was observed (a trace amount at pH 6.8 and significantly more at pH 8), but no scrambling was seen at pH 2. The pH used for sample preparation is shown to be critical to measure correctly the free cysteines and disulfide linkages.

Results

To establish identity, the sequence of the newly developed RNAi-mediated molecule was compared with the amino acid sequence of rituximab found in US Patent 5736137. Additionally, disulfide linkages, glycosylation structure, and amino acid modifications in the two mAbs were characterized and compared as described in the following sections.

Peptide mapping. Enzymatic peptide mapping was used for the primary sequence identification. A typical tryptic peptide map of the RNAi-mediated mAb is illustrated in Figure 1, with the identifications of all peptides summarized in the Supplemental Material (Table S1A for the heavy chain and Table S1B for the light chain). As listed in Table S1, several small peptides were identified through miscreavage or by digestion using different enzymes, such as Lys-C or pepsin. Importantly, many peptides with overlapping amino acids were repeatedly identified in the different enzymatic maps. Thus, complete sequence coverage (100%) was successfully achieved by the combined analysis of these enzymatic peptide fragments.

In the comparison of the two molecules, an amino acid mutation of S258Y was found from the tryptic peptide T19H of the RNAi-mediated molecule (Fig. 2). The reason for the mutation was likely due to an error in the PCR procedure (clonal error) as direct DNA sequencing confirmed an X to Y nucleotide substitution at position XXX (data not shown), which changed the serine codon (XXX) to a tyrosine (YYY). Also in the RNAi-mediated molecule, a second mutation F174L, which is partially mutated (~40%), was found for the T13H tryptic peptide (Fig. 3). This mutation could be attributable to extensive passaging (>21 passages) of the CHO cell line producing the RNAi-mediated molecule. To confirm the exact site of the mutation, a non-specific cleavage form of the T13H peptide, which has a shorter length, was chosen for collision induced dissociation-tandem mass spectrometry (CID-MS2) to obtain sufficient backbone cleavage for the site assignment (Fig. 3B). The 100% sequence coverage is critical to confirm the sequence or identify a variation, as in this case. Once the position of variation is identified, the site can be pinpointed for correction.

Because two amino acid mutations were found, additional potential mutations were explored. For this study, we used dimethyl labeling of the digest peptides, with dimethyl (2CH₃) for rituximab, and isotopically-labeled dimethyl (2CD₂) for the RNAi-mediated molecule to examine more carefully the entire amino acid sequence. With this approach, the difference of the mass for each tryptic peptide between the two products should be constant (2CD₂ − 2CH₃ = 4 Da per primary amine) if there is no amino acid mutation. For the mutation of S258Y, the delta mass would be 80 Da (Fig. 4A), and with the corresponding MS2 spectrum indicating the mutation site (Fig. 4B). For the mutation of F174L, the delta mass would be −22 Da (Fig. 5A), and the MS2 spectrum provides site confirmation (Fig. 5B). Importantly, we did not observe any other mutation sites following the complete assignment of the peptides.

Stable isotope labeling by amino acids in cell culture (SILAC) can be used to detect sequence variations between material produced from different manufacturing processes, e.g., light SILAC labeling in a reference material in comparison to another manufacturing lot with heavy SILAC labeling. For the sequence variation in biosimilar or biobetter products, however, it is not practical to use the metabolic approach (i.e., SILAC) to reproduce the cell culture conditions used to make the reference product because these conditions may be known by only the company that markets the product. Thus, we purchased the reference product (rituximab) and used the above external chemical approach (i.e., dimethyl labeling). It should be noted that dimethyl labeling is typically used for relative quantitation, but, in this case, we adopted the labeling method for mutation analysis. The labeling approach provides an additional measure of assurance to the results of peptide mapping sequence analysis. Moreover, the delta mass shift can provide a reasonable indication of the potential sequence variation, which can be subsequently determined by MS2 analysis.

Disulfide linkage analysis. Enzymatic digestion without reduction followed by LC-MS was used for disulfide analysis. All the expected disulfides were identified by accurate precursor mass measurement and CID or electron transfer dissociation (ETD) fragmentation of the precursor ion. The assignment approach is illustrated in Figure 6, with panel A the MS spectrum for the precursor ion, panel B for the CID-MS2 spectrum, and panel C for the ETD-MS2 spectrum. Using this approach, all the disulfide linkages were successfully identified (Figs. S1–S7 and summarized in Table 1). Both anti-CD20 mAbs were found to contain identical disulfide linkages.

In the analysis, unpaired cysteines were also identified (Figs. S8–S15). These unpaired cysteines were found as free cysteines at pH 6.8 with trypsin digestion and at pH 2 with pepsin digestion (Table 2), but they were not observed at pH 8. It is likely that the free cysteines became linked disulfides (scrambled)
Conducted to examine the correlation of scrambling as a function of pH (pH 8, 6.8, and 2). For optimum digestion efficiency, trypsin was used at pH 8 and pH 6.8, and pepsin at pH 2. Scrambled disulfides were found in trace amounts for digestion at pH 6.8, and significant amounts at pH 8, but no scrambling was observed at pH 2 (Table 3). Rituximab appears to have a higher amount in the alkaline pH (Table 3). For accurate results, it is important to measure the free cysteine levels at low pH, and, especially, the digestion pH needs to be optimized to minimize disulfide scrambling.17

To prove that the scrambled disulfides were caused by the sample preparation and were not in the original samples, a study was conducted to examine the correlation of scrambling as a function of pH (pH 8, 6.8, and 2). For optimum digestion efficiency, trypsin was used at pH 8 and pH 6.8, and pepsin at pH 2. Scrambled disulfides were found in trace amounts for digestion at pH 6.8, and significant amounts at pH 8, but no scrambling was observed at pH 2 (Table 3). Rituximab appears to have a higher amount

| Site | Sequence | Disulfide linkages | Enzyme | Rituximab (RT/min) | RNAI-mediated (RT/min) | [MH]+ (Da) |
|------|----------|-------------------|--------|-------------------|----------------------|------------|
| T2H  | 20-23    | MSCK (C22)        | LysC   | 40.41             | 40.69                | 3085.3257  |
|      | 75-98    | SSSTAYMQLSLTSDEASAUYCCAR (C96) | Lys plus trypsin | 57.40             | 57.69                | 7917.9266  |
| T12H | 138-151  | DYPPEPVTVSWNSQLGTSVHTPAVLQSSGLYS | Lys plus trypsin | 3.29              | 33.44                | 2329.1050  |
| T17H | 223-226  | SCCK (C224)       | LysC   | 1.72              | 3.21                | 1261.4936  |
| T178L| 207-213  | SFNQOE (C213)     | LysC   | 56.83             | 57.12                | 5005.4952  |
| T18H | 227-250  | THTCPPAPELLGGPSVFLPPPK (C230C233) | LysC plus trypsin | 39.70             | 40.00                | 3845.8309  |
| T20H | 260-267  | TPEVTCCVVVDVSHDEPKV (C265) | LysC plus trypsin | 39.70             | 40.00                | 3845.8309  |
| T26H | 325-326  | CK (C325)         | LysC   | 3.29              | 33.44                | 2329.1050  |
| T34H | 365-374  | NQVSL1CLV (C371)  | LysC   | 39.70             | 40.00                | 3845.8309  |
| T39H | 421-443  | WQQGVSFCVMSHEALHNHYTKQ (C429) | LysC plus trypsin | 39.70             | 40.00                | 3845.8309  |
| T2L  | 19-23    | VTMICR (C23)      | LysC   | 38.94             | 39.39                | 3528.5505  |
| T5L  | 77-102   | VEAEDAAYCQQWTSNPFQGKT (C87) | LysC plus trypsin | 38.94             | 39.39                | 3528.5505  |
| T9L  | 126-141  | SGTVGCVLNNFYPPR (C133) | LysC plus trypsin | 45.40             | 45.80                | 3556.7563  |
| T16L | 190-206  | VYACEVTHIQGLSSPVTK (C193) | LysC plus trypsin | 45.40             | 45.80                | 3556.7563  |

Table 2. Free cysteine determined at different digestion pHs for rituximab and RNAi-mediated mAb

| pH 8.0 | pH 6.8 | pH 2.0 |
|--------|--------|--------|
| Rituximab (%) | RNAI-mediated (%) | Rituximab (%) | RNAI-mediated (%) | Rituximab (%) | RNAI-mediated (%) |
| T12H (C148) | N.D. | N.D. | 5.4 ± 0.9 | 0.4 ± 0.0 | 1.0 | 0.2 |
| T13H (C204) | N.D. | N.D. | 1.3 ± 0.3 | N.D. | 0.2 | 0.1 |
| T20H (C265) | N.D. | N.D. | 3.0 ± 0.3 | 0.2 ± 0.0 | 3.4 | 1.2 |
| T26H (C325) | — | — | — | — | 3.1 | 0.8 |
| T34H (C371) | N.D. | N.D. | 10.7 ± 1.8 | 0.2 ± 0.1 | 1.6 | N.D. |
| T39H (C429) | N.D. | N.D. | 4.7 ± 0.3 | 0.1 ± 0.0 | 1.0 | N.D. |
| T2L (C23) | N.D. | N.D. | 0.4 ± 0.2 | 0.0 ± 0.0 | N.D. | N.D. |
| T5L (C87) | N.D. | N.D. | 0.2 ± 0.0 | 0.0 ± 0.0 | N.D. | N.D. |
| T9L (C133) | N.D. | N.D. | 1.7 ± 0.2 | 0.1 ± 0.0 | N.D. | N.D. |
| T16L (C193) | N.D. | N.D. | 1.1 ± 0.1 | 0.1 ± 0.0 | 1.0 | N.D. |
| T17L (C213) | — | — | — | — | 0.2 | 0.1 |

The percent of free cysteine at pH 8.0 and 6.8 was obtained by the measured amount of the free cysteine form divided by the total cysteines on the specific tryptic peptide, assuming each cysteine is alkylated. The percent of free cysteine at pH 2 was obtained by the ratio of the observed free cysteines divided by the total cysteines after reduction with TCEP. “N.D.” represents not detectable (the values is too low to be observed), and “–” means not measurable due to the tryptic peptide length (too short to retain in the LC chromatogram).
The amount of free cysteines in rituximab (most likely associated with human IgG2, with some evidence in human IgGs in serum) possess low levels of free sulfhydryl as well.  Thus, it is important to study the disulfide linkages carefully with a sensitive technique, as described here. Nevertheless, scrambling was not observed in either product at pH 2. It can be concluded that if scrambled disulfides are observed when examining a biopharmaceutical, a good test is to explore if scrambling decreases with digestion pH. Any “scrambling” observed at pH 2 likely represents the true structure of the molecule and not an artifact of sample preparation. In this case, a careful study should be conducted to obtain additional data points with different digestion pH to extrapolate the extent of scrambling at the intercept (i.e., extrapolate to pH 0) to reflect the true scrambling.

Table 3. Percent scrambled disulfides at different digestion pHs for rituximab and RNA-mediated mAb

| Glycans | pH 8.0 | pH 6.8 | pH 2.0 |
|---------|--------|--------|--------|
| G0F     | 34.2 ± 1.3 | 14.6 ± 0.8 | N.D.   |
| G0F-NGlc| 7.6 ± 0.6  | 2.3 ± 0.6  | N.D.   |
| G1F     | 35.6 ± 2.7 | 4.2 ± 0.6  | N.D.   |
| G1F-NGlc| 4.7 ± 0.6  | 0.4 ± 0.1  | N.D.   |
| G2F     | 9.6 ± 0.3  | 0.7 ± 0.1  | N.D.   |
| G0      | 2.8 ± 0.1  | 46.8 ± 1.6 | N.D.   |
| G0-NGlc | 1.2 ± 0.1  | 14.6 ± 2.1 | N.D.   |
| G1      | 0.9 ± 0.1  | 9.1 ± 0.6  | N.D.   |
| G1-NGlc | 0.1 ± 0.0  | 1.1 ± 0.1  | N.D.   |
| Man5    | 3.1 ± 0.1  | 5.0 ± 0.4  | N.D.   |
| Man4    | 0.1 ± 0.0  | 0.2 ± 0.1  | N.D.   |
| Man3    | 0.1 ± 0.0  | 0.9 ± 0.0  | N.D.   |

The percent glycosylation was obtained by the measured amount of the specific glycopeptide divided by the total of all forms of the glycopeptide including the non-glycosylated form. The SD was determined from 6 measurements.

Table 4. Comparison of glycan distribution for rituximab and RNA-mediated mAb

| Glycans        | Rituximab ± SD (%) | RNAi-mediated ± SD (%) |
|----------------|--------------------|------------------------|
| G0F            | 34.2 ± 1.3         | 14.6 ± 0.8             |
| G0F-NGlc       | 7.6 ± 0.6          | 2.3 ± 0.6              |
| G1F            | 35.6 ± 2.7         | 4.2 ± 0.6              |
| G1F-NGlc       | 4.7 ± 0.6          | 0.4 ± 0.1              |
| G2F            | 9.6 ± 0.3          | 0.7 ± 0.1              |
| G0             | 2.8 ± 0.1          | 46.8 ± 1.6             |
| G0-NGlc        | 1.2 ± 0.1          | 14.6 ± 2.1             |
| G1             | 0.9 ± 0.1          | 9.1 ± 0.6              |
| G1-NGlc        | 0.1 ± 0.0          | 1.1 ± 0.1              |
| Man5           | 3.1 ± 0.1          | 5.0 ± 0.4              |
| Man4           | 0.1 ± 0.0          | 0.2 ± 0.1              |
| Man3           | 0.1 ± 0.0          | 0.9 ± 0.0              |

The percent glycosylation was obtained by the measured amount of the specific glycopeptide divided by the total of all forms of the glycopeptides including the non-glycosylated form. The SD was determined from 6 measurements.

The Ellman reaction, coupling a chromophore (dithionitrobenzoic acid or DTNB) to a thiol group, has been used to detect free cysteine in a protein. This method, however, is insensitive and often requires hundreds of milligrams of mAb if the free cysteine is at low levels (i.e., ≤1%). Instead, we used the LC-MS approach for free cysteine analysis because it uses minute amounts of mAb (2 μg) and is quite sensitive for detecting free cysteine at low levels. The pH of the formulation buffer is between 5.5 and 6 for both rituximab and the RNAi-mediated molecules. It should be noted that most therapeutic antibodies contain low levels of free cysteines. It has also been reported that human IgGs in serum possess low levels of free sulfhydryl as well (most likely associated with human IgG2, with some evidence in IgG1). Thus, a proper formulation buffer to control the pH is needed to avoid scrambling while still maintaining high solubility. The pH of the formulation buffer is between 5.5 and 6 for both rituximab and the RNAi-mediated molecules.
acid residue. Thus, loss of NH₃ (minus 17 Da) was examined in the N-terminal peptides, and both the non-modified and pyro-Glu forms were found (Figs. S16–S17). The C-terminus of the heavy chain could also be cleaved by carboxypeptidase present during mAb production in the CHO cell. Thus, the K clipping at the C-terminus of the heavy chain was examined (Fig. S18). Both mAbs have the same pyro-Glu formation at the N-terminus, and near complete K clipping at the C-terminus of the heavy chain. Other typical modifications, including Asn deamidation and Met oxidation, were also found. No significant differences were observed for these modifications, except Met286 and pyro-Glu at the light chain (see t-test results in Table S2). The minor difference at Met286 could be due to this amino acid residue being very close to the mutation site (S258Y), and the difference of pyro-Glu on the light chain could be due to the difference in storage time and conditions for the two products.

**Discussion**

The full structure of the two anti-CD20 mAbs was compared using LC-MS analysis of the enzymatic peptide mixtures.
As shown, the full amino acid sequence, including N-terminal, C-terminal, disulfide linkages, and glycosylation site and structure, were successfully identified. The methodology can detect amino acid modifications, mutations, free cysteines, and scrambled disulfides. In such a comparative study, the identification of the entire amino acid sequence is critical. A multi-enzyme digestion strategy is necessary, along with complete LC-MS analysis of all the enzymatically-produced peptides. As shown in this work, dimethyl labeling of H and D isotopically labeled derivatizing agents adds a valuable dimension for verification.

We believe that sensitive LC-MS analysis should detect the structural variants as described. However, it is important not only to detect variants, but also to have a strategy to rule out other possibilities to provide a high degree of assurance that the results are correct. The two orthogonal techniques used in this work independently confirm the findings, but also can determine the degree of similarity of the two products. We believe that such an extensive comparability study by mass spectrometric techniques should be performed early in the research and development process to obtain an accurate sequence and the correct up-stream and down-stream process conditions to generate a highly similar antibody. As suggested in the US Food and Drug Administrations’ biosimilar draft guideline of February 2012, a highly similar antibody could reduce the extent of clinical studies needed for approval, and thus decrease the cost of biosimilar drugs, which could have a significant effect on the cost of healthcare.

For a comparability analysis, it should be emphasized that even 80 to 90% sequence coverage would not be satisfactory for comparison of two antibodies (e.g., reference vs. biosimilar). Although it may be time-consuming and labor-intensive, all peptides must be identified and, where appropriate, quantitated. In the future, top down mass spectrometry\textsuperscript{29-31} may be able to achieve 100% sequence coverage on the intact protein, without enzymatic digestion. With improving sensitivity and enhanced resolution of mass spectrometric instrumentation, the top down approach could ultimately have a great impact in the characterization of biopharmaceuticals.\textsuperscript{32} However, detecting many modifications at once is presently too challenging, as is the identification of all the modifications with associated sites using current software. The combination of both peptide mapping and top down NS may alleviate constraints in both approaches.

![Figure 3](image-url) (A) Precursor mass of normal (top left) and F174L/I mutated tryptic peptide (top right). (B) CID-MS2 of the precursor ion from Figure 3A, normal (top) and F174L/I mutated tryptic peptide (bottom).
The middle down, also known as the extended range proteomic analysis approach,\textsuperscript{33-35} which takes advantage of large peptide fragments for high sequence coverage and with less sophisticated modifications than the entire protein, is at present a useful alternative.

Materials and Methods

Samples. RNAi-mediated mAb was manufactured at Alnylam Pharmaceuticals and provided as 1 mg/mL × 0.8 mL as previously described.\textsuperscript{12} Rituximab was purchased by Alnylam from Imperial College, London and provided as 10 mg/mL × 0.5 mL. The samples were aliquoted as 10 μL per vial for rituximab (100 μg) and 100 μL per vial for RNAi-mediated mAb (100 μg) and stored at −80°C before analysis.

Reagents. Trypsin (sequencing grade, V5111) was purchased from Promega, endoproteinase Lys-C (MS grade, 125–05061) from Wako Chemicals USA, pepsin (Porcine Stomach Mucosa, 0219536701) from MP Biomedicals, and endoproteinase Glu-C (sequencing grade, 11420399001) from Roche Diagnostics. Formaldehyde (CH\textsubscript{2}O) (37% v/v, 252549), labeled formaldehyde (CD\textsubscript{2}O) (20% v/v, 98% D, 492620), sodium cyanoborohydride (20% v/v, 98% D, 98264), triethylammonium bicarbonate (TEAB, T7408), ammonium hydroxide solution (17837), guanidine hydrochloride (Gn-HCl, G3272), dithiothreitol (DTT, D5545), iodoacetamide (IAM, I6125), ammonium bicarbonate (09830) and 1.0 N hydrochloric acid (HCl, 318949) solution were obtained from Sigma-Aldrich. Tris buffer (17–1321–01) was from GE Healthcare, and tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl, 20490) and formic acid (FA, optima LC/MS, A11750) from Fisher Scientific. LC-MS grade water (JT9831–3) and acetonitrile (ACN, EM-AX0145–1) were purchased from VWR. Amicon centrifugal filters (10 kDa molecular weight cut-off, UFC501096) were obtained from EMD Millipore.

Enzymatic digestion. An aliquot of 10 μL of rituximab and 100 μL of RNAi-mediated mAb solution (100 μg) was denatured with 6 M guanidine hydrochloride containing 100 mM ammonium bicarbonate (pH 8), reduced with 5 mM DTT for 30 min at 37°C, and then alkylated with 20 mM IAM in the dark for 45 min at room temperature. The reduced and alkylated protein was buffer exchanged with 100 mM ammonium bicarbonate (pH 8) or 50 mM Tris (pH 6.8) using a 10 kDa molecular weight cut-off Amicon centrifugal filters.

Figure 4. (A) Precursor mass of the peptide with CH\textsubscript{2} dimethyl labeling (top left) and the precursor mass of the peptide with CD\textsubscript{2} dimethyl labeling (top right). The mass difference is not a constant 4 Da but 80 Da, accounting for the difference between S and Y (76 Da), as indicated in the figure. (B) CID-MS\textsuperscript{2} of the precursor ion from Figure 4A, CH\textsubscript{2} dimethyl labeling (top) and CD\textsubscript{2} dimethyl labeling (bottom).
weight cutoff filter to a concentration of 1 mg/mL (100 μL). For tryptic digestion, trypsin (1:50, w/w) was added to the protein solution at room temperature. After 8 h, the enzyme was added a second time (1:50, w/w) and the digestion continued at room temperature for 12 h. For Lys-C digestion, the endoprotease Lys-C (1:50, w/w) was added to the protein solution for 4 h at 37°C. For Lys-C plus trypsin digestion, the protein solution was added with endoprotease Lys-C (1:50 w/w) for 4 h at 37°C, and then trypsin (1:50 w/w) for 20 h at room temperature. For pepsin digestion, the protein solution was dissolved in 10 mM HCl (pH 2). Pepsin (1:10, w/w) was added to the protein solution and incubated at 37°C for 30 min. The reaction was quenched by adjusting the pH to 5 with 100 mM ammonium bicarbonate. For digestion without reduction (for disulfide assignment), the same digestion protocol as above was applied but without the reduction and alkylation steps. For quantitation of free cysteines at pH 2, the protein solution was digested in pepsin, and then half was reduced by 10 mM TCEP at pH 5.0 for an hour at room temperature. In all cases (except pepsin digestion), digestion was terminated by addition of 1% formic acid. An aliquot of 2 μg of the enzyme digest was analyzed per LC-MS run.

**Dimethyl labeling.** After digestion, the digests (20 μL, nearly completely dried) were reconstituted in 20 μL 0.1 mM TEAB, and 1 μL of 37% (v/v) CH₂O was added for the rituximab sample, while 1.9 μL of 20% (v/v) CD₂O was added for the RNAi-Fuc-anti-CD20 mAb sample. Each sample was then added 5.8 μL 1 M NaBH₃CN, and the solution was incubated in a fume hood for 1 h at room temperature. Finally, the reaction was quenched by addition of 4 μL 10% ammonia solution, and then 8 μL of formic acid on an ice top to prevent frothing or heating of the sample. The samples of rituximab and RNAi-Fuc-anti-CD20 mAb were mixed equally for subsequent LC-MS analysis.

**LC-MS.** An Ultimate 3000 nano-LC pump (Dionex) and a self-packed C18 column (Magic C18, 200 Å pore and 5 μm particle size, 75 μm i.d. × 15 cm) (Michrom Bioresources) was coupled online to an LTQ-Orbitrap-ETD XL mass spectrometer.
MS spectrum from m/z 300 to 2000, subsequent CID-MS2 and ETD-MS2 steps were performed on the same precursor ion with a ± 2.5 m/z isolation width. Any incomplete assignment in the CID-MS2 and ETD-MS2 spectra was repeated by targeting the desired ions, e.g., the same precursor but with a different charge state, to gain additional linkage information. This targeted approach was repeated (e.g., targeting multiple charges of a precursor ion or the same disulfide-linked peptide but with different enzymatic cleavages or miscleavages) until the linkage information was complete. In addition, a targeted CID-MS3 after ETD for ions of interest was performed as necessary.

Peptide assignment. The spectra generated in the CID-MS2 step were searched against spectra of theoretical fragmentations (b and y ions) of rituximab sequence with a mass tolerance of ≤ 5 ppm for the precursor ions and with enzyme specificity, using a Sequence probability score (> 95% confidence) as the filter. For peptides with miscleavages or a mass tolerance > 5 ppm (but less than 20 ppm) of the precursor ion, confirmation required

Figure 6. Precursor mass of 1029. 1143 (3+) (A), CID-MS2 of the precursor (B), and ETD-MS2 of the precursor (C) for the disulfide-linked peptide (Cys22-Cys96). For ETD-MS2, a different charge of the precursor ion, 772.59 (4+), was used for fragmentation.

(Thermo Fisher Scientific) through a nanospray ion source (New Objective). Mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in acetonitrile) were used for the gradient consisting of (1) 20 min at 2% B for sample loading and 5 min for desalting at 0.3 μL/min (20 min desalting for pep-sin or dimethylated digests); (2) linear from 2 to 5% B for 2 min; (3) linear from 5 to 35% B for 60 min; (4) linear from 35 to 90% B for 3 min; and finally (5) isocratic at 90% B for 5 min. The flow rate of the column was maintained at 0.2 μL/min, and the mass spectrometer started to record data after 5 min of the gradient. The LTQ-Orbitrap-ETD XL mass spectrometer was operated initially in the data-dependent mode as follows: survey full-scan MS spectra (m/z 300–2000) were acquired in the Orbitrap with a mass resolution of 30,000 at m/z 400 (with an ion target value of 5 × 105 ions), followed by nine sequential MS2 scans using the LTQ. For disulfide mapping, the MS was switched automatically between MS (scan 1 in the Orbitrap), CID-MS2 (scan 2 in the LTQ), and ETD-MS2 (scan 3 in the LTQ). Briefly, after a survey MS spectrum from m/z 300 to 2000, subsequent CID-MS2 and ETD-MS2 steps were performed on the same precursor ion with a ± 2.5 m/z isolation width. Any incomplete assignment in the CID-MS2 and ETD-MS2 spectra was repeated by targeting the desired ions, e.g., the same precursor but with a different charge state, to gain additional linkage information. This targeted approach was repeated (e.g., targeting multiple charges of a precursor ion or the same disulfide-linked peptide but with different enzymatic cleavages or miscleavages) until the linkage information was complete. In addition, a targeted CID-MS3 after ETD for ions of interest was performed as necessary.
manual inspection to match all highly abundant product ions. The sequence coverage generated by each enzymatic map was combined as a total coverage.

Disulide assignment. The anticipated disulide-linked tryptic or multi-enzyme digested peptide masses with different charges were first calculated and then matched to the observed masses in the LC-MS chromatogram. The matched masses (with < 5 ppm mass accuracy) were further confirmed by the corresponding CID-MS2 and ETD-MS2 fragmentation, as well as CID-MS3 fragmentation, as needed.

Glycan structure identification. Theoretical masses of glycan structures such as G0F, G1F and G2F were added to the tryptic peptide backbone (EEQYNSTYR). The anticipated glycopeptide masses with different charges were thus obtained to match the observed masses in the LC-MS chromatogram. The matched masses (with ≤ 5 ppm mass accuracy) were further confirmed by the corresponding CID-MS2 fragmentation.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
This work was supported in part by NIH grant GM 15847. The authors thank Alynam Pharmaceuticals for providing the RNAi-mediated anti-CD20 antibody and rituximab for this study. Contribution Number 1036 from the Barnett Institute.
33. Wu SL, Kim J, Hancock WS, Karger B. Extended Range Proteomic Analysis (ERPA): a new and sensitive LC-MS platform for high sequence coverage of complex proteins with extensive post-translational modifications-comprehensive analysis of beta-casein and epidermal growth factor receptor (EGFR). J Proteome Res 2005; 4:1155-70; PMID:16083266; http://dx.doi.org/10.1021/pr050113n

34. Wu SL, Kim J, Bandle RW, Liotta L, Petricoin E, Karger BL. Dynamic profiling of the post-translational modifications and interaction partners of epidermal growth factor receptor signaling after stimulation by epidermal growth factor using Extended Range Proteomic Analysis (ERPA). Mol Cell Proteomics 2006; 5:1610-27; PMID:16799092; http://dx.doi.org/10.1074/mcp.M600105-MCP200

35. Pipes GD, Campbell P, Bondarenko PV, Kerwin BA, Treuheit MJ, Gadgil HS. Middle-down fragmentation for the identification and quantitation of site-specific methionine oxidation in an IgG1 molecule. J Pharm Sci 2010; 99:4469-76; PMID:20845446; http://dx.doi.org/10.1002/jps.22158