Unveiling a Glycation Hot Spot in a Recombinant Humanized Monoclonal Antibody

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2008

Impact Factor: 5.64 · DOI: 10.1021/ac701810q · Source: PubMed

9 AUTHORS, INCLUDING:

Yi Yang
Genentech
7 PUBLICATIONS **98** CITATIONS
SEE PROFILE

Viswanatham Katta
Medivation, Inc.
65 PUBLICATIONS **3,646** CITATIONS
SEE PROFILE

Anal. Chem. **2008,** *80,* 2379−2390

Unveiling a Glycation Hot Spot in a Recombinant Humanized Monoclonal Antibody

Boyan Zhang,*,† Yi Yang,† Inn Yuk,‡ Roger Pai,§ Patrick McKay,§ Charles Eigenbrot,|| Mark Dennis,⊥ Viswanatham Katta,† and Kathleen Champion Francissen†

Protein Analytical Chemistry, Cell Culture Development, Purification Development, Protein Chemistry, and Antibody Engineering, Genentech, Inc., South San Francisco, California 94080

Biotechnological companies and regulatory agencies are pursuing the complete characterization of protein therapeutics in every detail as a means to mitigate risks of product quality related safety issues. During the characterization of a recombinant humanized monoclonal antibody (referred to as rhuMAb), electrospray mass spectrometric analysis suggested that the light chain was highly glycated. The glycated and unglycated materials, separated using boronate affinity chromatography, were fully characterized using tryptic peptide mapping and tandem mass spectrometry. Using an automatic SEQUEST search of the single protein database for this antibody and extensive manual investigations of the mass spectra of the matched peptides, multiple tentative glycation sites in the light and heavy chains were observed in the highly glycated (>53%) samples. A predominant glycation site was identified and confirmed to be lysine 49 on the light chain, by performing extensive sequence analysis on an isolated glycated peptide utilizing Edman degradation analysis and MALDI-TOF/TOF mass spectrometry. Sequence alignments of rhuMAb with 12 other recombinant monoclonal antibodies and computer modeling of the Fab part of rhuMAb suggest that the unusually high level of glycation of lysine residue 49, which is located adjacent to the second complementarity-determining region (CDR2) in the light chain, is due to a spatial proximity effect in catalyzing the Amadori rearrangement by aspartic acid residue 31 in the CDR1 on the light chain.

Protein glycation is a nonenzymatic glycosylation that usually takes place at the ϵ-amino groups of lysine residues and α-amino groups of amino-terminal residues on proteins by reaction with reducing sugars such as glucose and lactose. A primary amine of the protein reacts with the aldehyde group of glucose to form a labile aldimine (Schiff base) derivative of the protein, which then slowly isomerizes to the more stable ketoamine adduct via the Amadori rearrangement.[1−3] The aldimine and ketoamine derivatives of proteins are referred to as early-stage glycation adducts. Prolonged incubation under aerobic conditions or physiological circulation of the initially glycated proteins may lead to the formation of a variety of advanced glycation end products (AGEs) due to oxidative ketoamine dehydration, fragmentation, or direct reaction of α-oxoaldehydes with amino groups from other proteins to produce more complex and irreversibly covalently cross-linked structures.[4−6] Several AGE structures, such as N^{ϵ}-(carboxymethyl)-lysine, pyrraline, and pentosidine have been previously identified in association with a number of proteins.[6,7]

In reviewing the site specificity studies published to date, a mixed picture of the glycation process is obtained. In some cases, a high degree of specificity has been observed. For example, ribonuclease A exhibited a preference for glycation at Lys41 and Lys7 when treated with glucose in phosphate buffers.[8−10] A preferential glycation site on hemoglobin is the α-amino group of the terminal β-chain valine residue.[9,11] The site specificity of glycation of those two proteins was explained by the catalytic effect of bound phosphate ion in the active sites.[9] The sites of in vivo glycation of human and horse liver alcohol dehydrogenase were identified to be on multiple sites with preference for Lys231.[12] The glycation susceptibility of Lys231 was explained by its unique location close to an imidazole group of His348, and the latter was postulated to be an acid−base catalyst for the Amadori rearrangement of the initially formed Schiff base.[12] Glycation of human serum albumin (HSA) in vivo occurs at multiple sites with preference for Lys525, which accounts for 33% of the overall

* To whom correspondence should be addressed. E-mail: bzhang@gene.com.

† Protein Analytical Chemistry.

‡ Cell Culture Development.

§ Purification Development.

|| Protein Chemistry.

⊥ Antibody Engineering.

(1) Thorpe, S. R.; Baynes, J. W. In *The Glycoconjugates*; Horowitz, M. I.,, Pigman W., Eds.; Academic Press: New York, 1982; Vol. 3, pp. 113−132.

(2) Quan, C. P.; Wu, S.; Dasovich, N.; Hsu, C.; Patapoff, T.; Canova-Davis, E. *Anal. Chem.* **1999**, *71*, 4445−4454.

(3) Watkins, N. G.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1985**, *260*, 10629−10636.

(4) Booth, A. A.; Khalifah, R. G.; Todd, P.; Hudson, B. G. *J. Biol. Chem.* **1997**, *272*, 5430−5437.

(5) Al-Abed, Y.; Kapurniotu, A.; Bucala, R. *Methods Enzymol.* **1999**, *309*, 152−172.

(6) Ahmed, N.; and Thornalley, P. J. *Biochem. Soc. Trans.* **2003**, *31*, 1417−1422.

(7) Zhang, Y.; Cocklin, R. R.; Bidasee, K. R.; Wang M. *J. Biomol. Tech.* **2003**, *14*, 224−230.

(8) Watkins, N. G.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1985**, *260*, 10629−10636.

(9) Watkins, N. G.; Neglia-Fisher, C. I.; Dyer, D. G.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1987**, *263*, 7207−7212.

(10) Brock, J. W.; Hinton, D. J.; Cotham, W. E.; Metz, T. O.; Thorpe. S. R.; Baynes, J. W.; Ames, J. M. *J. Proteome Res.* **2003**, *2*, 506−513.

(11) Shapiro, R.; McManus, M. J.; Zalut, C.; Bunn, H. F. *J. Biol. Chem.* **1980**, *255*, 3120−3127.

(12) Shilton, B. H.; Walton, D. J. *J. Biol. Chem.* **1991**, *266*, 5587−5592.

10.1021/ac701810q CCC: $40.75
Published on Web 02/29/2008

glycation level.[13] The major site of glycation of bovine serum albumin, when it was incubated in vitro with fructose or glucose at physiological temperature and pH in phosphate buffer, is Lys524 (corresponding to Lys525 in HSA).[14] In contrast, Zoellner et al. found that in vitro incubation of HSA with glucose in PBS containing 1 mM EDTA affected a great majority of lysine and arginine residues and glycation of albumin was relatively nonspecific.[15] In addition, Tagami et al. presented data that showed low specificity in reactions of glucose with lysozyme.[16] It is clear that the incubation conditions and length of duration play important roles in the observed differences. Baynes et al. reviewed protein glycation studies published prior to 1988 and suggested a number of factors that are known or are likely to affect the site specificity of the glycation of proteins.[17] These factors include the general structural effects from vicinal side-chain functional groups, allosteric and active binding sites for phosphate, bicarbonate and phosphorylated intermediates, etc., as well as some long-range interactions such as ligand-induced allosteric effects, salt bridges, hydrogen bonding, and the pK_a of the amino groups.[17] These general suggestions are helpful in understanding the process of protein glycation, but they are far from guidelines for predicting the site specificity of protein glycation. It is still uncertain whether specific structural motifs make portions of a protein more susceptible to glycation.

There is an increasing number of recombinant monoclonal antibodies (MAbs) in therapeutic applications.[18] Currently, all of the U.S. Food and Drug Administration (FDA) approved and marketed MAbs are produced from mammalian cell culture. The manufacture of monoclonal antibody therapeutics involves cell culture and purification processes, as well as formulation and packaging. During the cell culture process, a monoclonal antibody is often expressed in Chinese hamster ovary (CHO) cells and secreted into the culture medium. The cell culture process is often performed for 10−14 days before the culture fluid is harvested for purification. As the cell culture medium contains glucose, it is possible that the antibody is glycated to a certain extent. In a number of cases at Genentech Inc., the purified MAbs from cell culture processes exhibited low levels (~5%) of glycation. Due to its low level and wide distribution at multiple sites across the entire antibody, glycation may not be detected during the analytical characterization of an antibody. To date, only a few conference presentations[19,20] on the detection of antibody glycation have been reported.

A full length recombinant humanized MAb, referred to in this work as rhuMAb, was recently determined to be highly glycated (40~60%, as quantified by boronate affinity chromatography), when it was expressed by transient or early stable CHO cell lines. Our preliminary data revealed that the high glycation level was attributed to the predominant modification of a lysine residue on the light chain. This unexpected finding increased our curiosity, and also for the purpose of risk mitigation, efforts were undertaken to identify the glycation pattern and optimize the culture process to reduce the glycation level. The present work reports the discovery, subsequent characterization of the glycated form of rhuMAb, and the identification of a specific lysine residue, which incurs an unexpectedly high level of modification by glucose during the CHO cell culture process. A convenient MALDI TOF/TOF technique is used to collect the MS/MS data to sequence and pinpoint the glycation site of an isolated glycated peptide. In vitro glycation kinetic studies were designed to simulate the cell culture process and to determine the critical parameters affecting the glycation level of antibody product. The concentration of glucose in the culture media is found to be critical. Controlling the glucose concentration at a different level at variuos times, considering the consumption of culture media nutrients and product titer, is suggested in order to minimize and control the glycation level of rhuMAb.

EXPERIMENTAL SECTION

Materials and Reagents. rhuMAb lot I (22 mg/mL, PBS buffer, pH 7.4) and lot II (27 mg/mL, PBS buffer, pH 7.4) bulk materials were produced using stable recombinant CHO cells cultured using in-house proprietary serum-free media with or without controlling the glucose level, respectively. Twelve recombinant monoclonal antibodies, which share complementarity-determining regions (CDRs) highly similar to that of rhuMAb, were produced by small-scale transient CHO cell cultures in-house.

Trypsin, sequencing grade (TPCK treated), was purchased from Promega (Madison, WI), and Peptide:N-glycosidase F (PNGase F) from New England BioLabs Inc (Ipswich, MA). *N*-(2-Hydroxyethyl)piperazine-*N*′-(3-propanesulfonic acid) (EPPS; 99.0%) and dithiothreitol (DTT; ultrapure) were obtained from USB Corp. (Cleveland, OH). D-Sorbitol (minimum 98%), α-D-glucose (ACS reagent), iodoacetic acid (IAA; ~99%), and α-cyano-4-hydroxycinnamic acid (CHCA) were procured from Sigma-Aldrich (St. Louis MO). HPLC grade acetonitrile (ACN) and methanol were purchased from Fisher Scientific (Fair Lawn, NJ). Water used in all experiments was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA).

In Vitro Glycation of rhuMAb. The rhuMAb lot II in PBS buffer, pH 7.4, was further glycated in vitro. A series of mixed solutions, consisting of various concentrations of α-D-glucose (0, 20, 50, 100, 150, and 1200 mM), 2 mg/mL rhuMAb lot II, and 0.05% NaN_3 in 100 mM $NaHCO_3$ buffer, pH 7.4, were prepared to a final volume of 0.5 mL in 0.6-mL Eppenderf tubes. These mixtures were incubated in a water bath at 37 °C for 24 h. The reaction mixtures were submitted for immediate buffer exchange using NAP-5 columns (GE Healthcare) into 0.1 M NaOAc buffer, pH 5.0. The samples were then stored at 2−8 °C until analysis.

(13) Iberg, N.; Fluckiger, R. *J. Biol. Chem.* **1986**, *261*, 13542−13543.
(14) Hilton, D. J. S.; Ames, J. M. *Amino Acids* **2006**, *30*, 425−433.
(15) Zoellner, H.; Hou, J. Y.; Hochgrebe, T.; Poljak, A.; Duncan, M. W.; Golding, J.; Henderson, T.; Lynch, G. *Biochem. Biophys. Res. Commun.* **2001**, *284*, 83−89.
(16) Tagami., U.; Akashi, S.; Mizukoshi, T.; Suzuki, E.; Hirayama, K. *J. Mass Spectrom.* **2000**, *35*, 131−138.
(17) Baynes, J. W.; Watkins, N. G.; Fisher, C. I.; Hull, C. J.; Patrick, J. S.; Ahmed, M. U.; Dunn, J. A.; Thorpe, S. R. In *The Maillard Reaction in Aging, Diabetes, and Nutrition*; Baynes, J. W., Monnier, V. M., Eds.; Alan R. Liss: New Yeak, 1989; pp 43−67.
(18) Reichert, J. M.; Rosensweig, C. J.; Faden, L. B.; and Dewitz, M. C. *Nat. Biotechnol.* **2005**, *23*, 1073−1078.
(19) Quan, C.; Alcala, E.; Ma, S. 10th Symposium on the Interface of Regulatory and Analytical Sciences for Biotechnology Health Products, San Francisco, CA, January 24−26, 2006; WCBP 2006, Abstr. p 86.
(20) Zhang, B.; Yang, Y.; Champion-Francissen, K. 10th Symposium on the Interface of Regulatory and Analytical Sciences for Biotechnology Health Products, San Francisco, CA, January 24−26, 2006; WCBP 2006, Abstr. p 88.

Enzymatic Digestion of rhuMAb. Several enzymatic treatments were performed on rhuMAb samples before they were submitted for HPLC, mass spectrometric analyses, or both, to determine the amino acid sequence and structural integrity of rhuMAb.

PNGase F Treatment. The rhuMAb samples were buffer exchanged into 50 mM Tris buffer, pH 7.5, using a NAP-5 column. The protein concentration was adjusted to 2.5 mg/mL. PNGase F was added in an enzyme-to-substrate ratio of 1:600 (w/w). The digestion was performed at 37 °C overnight (~15 h).

Tryptic Digestion. Prior to tryptic digestion, rhuMAb was reduced with DTT and then alkylated with IAA. Typically, 250 μL of antibody sample (2 mg/mL) was mixed with 20 μL of 1 M DTT in 730 μL of 6 M guanidine, 50 mM Tris, pH 8.0. The mixture was incubated at 37 °C for 1 h. It was then cooled to room temperature, and 50 μL of 1 M IAA in 1 M NaOH was added. The alkylation reaction was incubated at room temperature in the dark for 15−20 min. The residual IAA was quenched by the addition of 10 μL of 1 M DTT. The reduced and carboxymethylated rhuMAb was then buffer exchanged into the digestion buffer, which contained 25 mM Tris, 1 mM $CaCl_2$, pH 8.3, using a PD-10 column (Sephadex G-25 medium, GE Healthcare). Trypsin was added at an enzyme-to-substrate ratio of 1:50 (w/w). The solution was mixed briefly and incubated in a 37 °C water bath for 5 h. The digestion was terminated by adding 0.3% (v/v) trifluoroacetic acid (TFA) to the solution. The digest was then stored at −70 °C until analysis.

HPLC. Solvent delivery, analyte separation, and fraction collection were performed with an Agilent (Santa Clara, CA) 1100 series HPLC equipped with a thermostatic auto-sampler and fraction collector.

Boronate Affinity Chromatography. The UV diode-array detector was set at 280 nm with an 8-nm bandwidth. A TSKgel Boronate 5PW column 7.5 × 75 mm (Tosoh Bioscience, Montgomeryville, PA) was employed, and its temperature was controlled at 40 °C. The mobile phases used were as follows: A, 50 mM EPPS, 10 mM Tris, 200 mM NaCl, 0.05% NaN_3 at pH 8.7; B, 500 mM sorbitol in mobile phase A. After sample injection, the boronate column was eluted isocratically with mobile phase A for 20 min followed by a gradient to 100% B in 5 min. The glycation level was determined by calculating the percentage of integrated peak area of retained peak versus total peak area. For fraction collection, a time-based mode was chosen; both the flow-through peak (unglycated) and retained peak (glycated) fractions were collected from several repetitions and pooled, respectively, for tryptic peptide mapping studies.

Tryptic Peptide Mapping. Tryptic digests of rhuMAb samples were analyzed by RP-HPLC using a Jupitor C18 2.0 × 250 mm column (Phenomenex, Torrance, CA). The column temperature was controlled at 45 °C, and the column effluent was monitored at 214 nm. The flow rate was controlled at 250 μL/min; the mobile phases used were as follows: A, 0.1% TFA in water; B, 0.09% TFA in acetonitrile. The sample injection volume was 30 μL (~30 μg of tryptic peptides). The optimized gradient is as follows (minute/%B): 0/0, 3/0, 23/10, 143/40, 145/95, 148/45, and 150/0. The total analysis time of a single injection is 185 min. Fractions were collected for selected peptides from many repetitions and pooled for further N-terminal sequence and MALDI-TOF−MS/MS analyses.

Edman Degradation Analysis. The isolated peptides were vacuum-dried, reconstituted into 0.1% triethanolamine, and adjusted to pH 7.2 with acetic acid. Aliquots containing 0.5−1.0 nmol of peptides were applied to a preconditioned Polybrene-coated glass fiber disk and subjected to eight cycles of Edman degradation. Molar values for phenylthiohydantoin-amino acids were generated by peak area comparisons to an external standard mixture.

Mass Spectrometry. *LC/ESI-MS.* The molecular masses of the intact, PNGase F-treated, and reduced rhuMAb samples were determined by using a LC/ESI-MS setup coupling an Agilent 1090 HPLC system with a PE Sciex API 3000 electrospray ionization triple-quadrupole mass spectrometer (Applied Biosystems, Foster City, CA). The samples were desalted by RP-HPLC using a capillary column (Poros R1, 0.33 × 200 mm) equilibrated at a flow rate of 200 μL/min with 0.2% formic acid (solvent A) at a column temperature of 40 °C. Samples were eluted using a linear gradient from 25% solvent B (0.2% formic acid in acetonitrile) to 70% solvent B in 16 min. The effluent from the RP-HPLC column was directed into the mass spectrometer operating in the positive ion mode. For analysis after reduction, samples were incubated in 25 mM DTT for 15 min at 37 °C prior to LC/MS analysis.

LC/MS/MS Characterization of Tryptic Peptide Maps. The masses and sequences of the tryptic peptides were determined by using an online-coupled LTQ linear ion trap mass spectrometer (ThermoElectron, San Jose, CA). An Agilent 1100 Series HPLC system was coupled to the mass spectrometer. The effluent from the HPLC was directly infused into the LTQ electrospray ionization source. Electrospray ionization in positive ion mode was achieved by using a needle spray voltage of 4.5 kV and capillary voltage of 44 V. In the LC/MS/MS experiments, nine scan events including a full scan in the m/z range of 300−2000 were followed by four cycles of zoom scans and MS/MS scans on the four most intense ions.

MS/MS spectra interpretation and peptide assignment were accomplished by an automatic database search with a SEQUEST algorithm using Bioworks Browser version 3.2 software (ThermoElectron, San Jose, CA) and manual investigation of each matched product ion spectrum. A *fasta* single protein database, including both light chain and heavy chain sequences of rhuMAb, was created and used as the searching target. For identifying glycation products, glycation-related modifications were defined as variable ones (+162 Da for aldimine and ketoamine) and only allowed for lysine (K) residues, N-terminal aspartic acid (D) residue of light chain, and glutamic acid (E) residue of heavy chain. A special set of filtering parameters, which include probability score of >10 (probability score = $-10 \log[P(\text{pep})]$, where $P(\text{pep})$ is the peptide probability value determined by the search algorithm) and satisfied correlation factor (Xc) values (Xc ≥ 0.5 for singly charged, Xc ≥1.0 for doubly charged, Xc ≥1.5 for triply charged peptides ions), were selected for considering a significant match for a glycation-modified peptide match. These lower than normal scoring values were arbitrarily chosen because of the poor quality of the collision-induced dissociation MS^2 spectra of

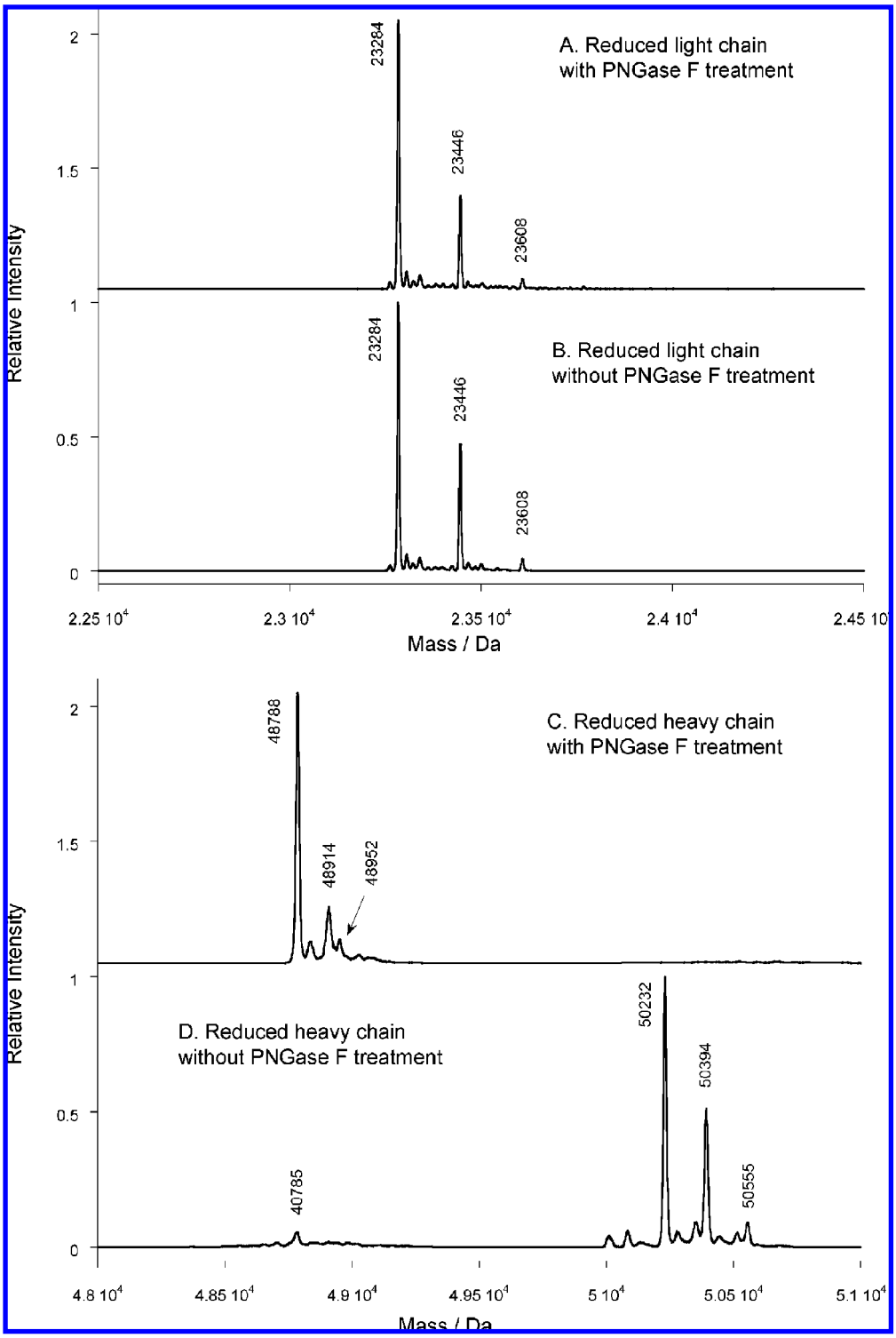

Figure 1. Reconstructed mass spectra for (A) light chain released by disulfide bond reduction of PNGase F digested rhuMAb lot I; (B) light chain released by disulfide bond reduction of rhuMAb lot I; (C) heavy chain released by disulfide bond reduction of PNGase F digested rhuMAb lot I; and (D) heavy chain released by disulfide bond reduction of rhuMAb lot I. The LC/ESI-MS data was acquired on a PE Sciex API 3000 electrospray ionization triple-quadrupole mass spectrometer (Applied Biosystems).

glycated peptides. In general, MS^2 spectra of glycated peptides are dominated by a few fragment ions generated by the loss of water from the precursor ion.

Subsequently, manual investigation of zoom-scan mass spectra and MS/MS spectra of the matched peptide ions was performed to eliminate false positive identifications. For example, matched peptides with two modifications per peptide or with a modification at the C-terminal lysine were considered false positives and were discarded. Zoom-scan MS profiles were examined to confirm the charge state and monoisotopic mass of matched peptides. To estimate the glycation level for each glycation site, the zoom-scan isotopic peaks of a glycated peptide were manually integrated using *Xcalibur* QualBrower. The relative percentage of glycation was subsequently estimated by dividing the peak area of the

glycated peptide ion by the sum of peak areas of glycated and non-glycated peptides.[10]

MALDI-TOF MS/MS. Isolated tryptic peptides were spotted individually on a SCOUT 384 multiprobe plate with a CHCA matrix and analyzed in the positive mode using a Bruker Ultraflex-I MALDI-TOF−TOF mass spectrometer (Bremen/Leipzig, Germany). Selected precursor ions were subjected to collision-induced dissociation with argon as the collision gas. LIFT mode was used to analyze all fragment ions in single sweep, and data were processed with FlexAnalysis software.

Homology Modeling of the Variable Region of rhuMAb. A homology model of the variable region of rhuMAb was produced using the Homology module of Insight II (Accelrys, San Diego, CA) and the X-ray coordinates of the anti-Her2 antibody Herceptin (PDB entry 1FVC)[21] and the anti-peptide antibody PC282 (PDB entry 1KCV).[22] PC282 was used for a segment in the V_H domain (CDR2 and framework 3) in which rhuMAb lacks two "insertion" residues found in Herceptin (52a and 82a). Otherwise, the Herceptin CDRs are the same length as those in rhuMAb and the Herceptin main chain conformations were used.

Data Presentation. All of the chromatograms and mass spectra data were converted to Excel files and reprocessed using KaleidaGraph 3.6 software (Synergy Software, Reading, PA) for presentation.

RESULTS AND DISCUSSION

Discovery of rhuMAb Glycation. Similar to many marketed antibody therapeutics,[18] this antibody (rhuMAb) is composed of two IgG1 heavy chains and two κ light chains that are linked by four pairs of interchain disulfide bonds. When expressed in CHO cells, the Y-type structure of IgG1 antibodies is well conserved. However, structural heterogeneity generated from posttranslational modifications is commonly observed.[23] Comprehensive characterization of rhuMAb revealed the existence of N-glycosylation at Asn297 of the heavy chain, incomplete heavy chain C-terminal lysine processing, and glycation of lysine residues.

During the molecular mass determination of a DTT-reduced sample of rhuMAb lot I, as shown in the reconstructed mass spectra for light chain (Figure 1B), a major peak at 23,284 Da was observed. The difference (1.9 Da) between the observed and predicted average mass of 23 285.9 Da for reduced light chain of rhuMAb is within the instrument detection error of 100 ppm. Incomplete reduction of intrachain disulfide bonds within the light chain can also be ascribed as a source causing the difference. Two other peaks at 23 446 and 23 608 Da, which are 162 and 324 Da greater than the major peak, respectively, were detected. PNGase F treatment did not eliminate these two higher molecular weight species (Figure 1A), indicating that the modification in these two species was not due to enzymatic N-glycosylation. Most commonly, O-linked enzymatic glycosylation is initiated by addition of a *N*-acetyl-α-D-galactosylamine residue (GalNAc, +203 Da) to the hydroxyl oxygen of serine or threonine side chain.[24] Since enzymatic O-glycosylation is not commonly observed for a full-length CHO-derived MAb, it was reasonable to ascribe the source of +162 or +324 Da modifications to nonenzymatic glycation.

Analysis of the heavy chain for possible glycation was complicated by the existing heterogeneity in the N-linked glycosylation at Asn297, with a number of glycoforms differing by multiples of 162 Da already present. MALDI-TOF/MS analysis of N-linked oligosaccharides released by PNGase F digestion of rhuMAb confirmed that they are consistent with those reported for human plasma-derived antibodies[25] and other recombinant antibodies.[26,27] Three observed masses (50 232, 50 394, and 50 555 Da) for the reduced heavy chain (Figure 1D) are consistent, within the detection error, with a heavy chain composed of the three major oligosaccharide structures G_0, G_1, and G_2 with respective predicted average masses of 50 235.4, 50 397.4, and 50 559.4 Da. As shown in Figure 1C, only one major peak was detected upon PNGase F treatment, consistent with the removal of all N-linked glycans on the heavy chain. The observed mass of 48 788 Da has the most intense signal and matches the predicted average mass (48 790.1 Da) within the determination error for reduced heavy chain lacking N-linked glycans. The other observed mass of 48 914 Da is consistent with a partially processed C-terminal lysine variant.[23] The presence of a small peak with a mass of 48 952 Da, 162 Da higher than the predicted average mass of reduced heavy chain, indicates a possible glycation of the heavy chain of rhuMAb.

As shown in Figure 2B, boronate affinity chromatography was able to separate the glycated (retained) from the unglycated one (flow-through) of rhuMAb. The overall glycation level (area percentage of retained peak) of rhuMAb lot I was determined to be 58%.

Glycation of rhuMAb under In Vitro Conditions. The molecular mass determination using ESI-MS and boronate affinity chromatographic analyses of rhuMAb lot I (produced by a stable CHO cell line) revealed its high level of glycation, predominantly on the light chain. To better understand the unusually high susceptibility of the light chain to glycation, we sought to understand the glycation mechanism in order to optimize the culture process to minimize glycation. During preliminary testing, a series of in vitro glycation experiments mimicking cell culture conditions used during rhuMAb production were conducted to identify critical parameters related to glycation. A rhuMAb lot that contained ~41% glycated species was mixed with the cell culture medium containing 8 g/L (49.4 mM) glucose and all other necessary nutrients for cell growth. The buffer was sodium bicarbonate, pH 7.2. The mixture was aliquoted into four tubes. One aliquot was frozen and used as a time zero control, and the other three were incubated at 30, 33, and 37 °C for 5 and 10 days, respectively. Using boronate affinity chromatography, the glycation levels of these samples were found to be in a narrow range from 39 to 42%. It appears that an inhibition of the glycation reaction was occurring. We postulated that the rich content of free amino acids and peptides in the culture medium inhibited the antibody glycation reaction by competing with rhuMAb for glucose in forming the labile aldimine intermediates. To test if free amino acids slow or inhibit glycation of rhuMAb, five reaction

(21) Eigenbrot, C.; Randal, M.; Presta, L.; Carter, P.; Kossiakoff, A. A. *J. Mol. Biol.* **1993**, *229*, 969−995.
(22) Nair, D. T.; Singh, K.; Siddiqui, Z.; Nayak, B. P.; Rao, K. V.; Salunke, D. M. *J. Immunol.* **2002**, *168*, 2371−2382.
(23) Harris, R. J. *Dev. Biol. (Basel)* **2005**, *122*, 117−127.
(24) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opendenakker, G. *Crit. Rev. Biochem. Mol. Biol.*, **1998**, *33*, 151−208.
(25) Jefferis, R.; Lund, J.; Mizutani, H.; Nakagawa, H.; Kawazoe, Y.; Arata, Y.; Takahashi, N. *Biochem. J.,* **1990**, *268*, 529−537.
(26) Lewis, D. A.; Guzzetta, A. W.; Hancock, W. S.; Costello, M. *Anal. Chem.*, **1994**, *66*, 585−595.
(27) Ma, S.; Nashebeh, W. *Anal. Chem.*, **1999**, *71*, 5185−5192.

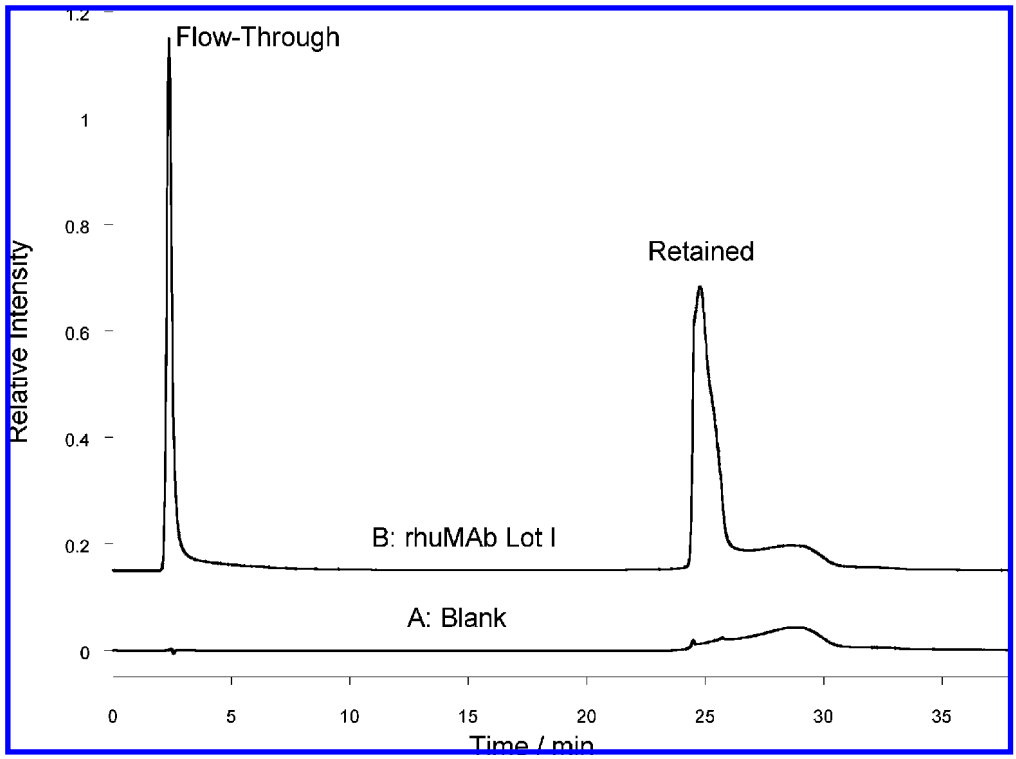

Figure 2. Boronate affinity chromatography of (A) buffer blank and (B) rhuMAb lot I.

mixtures containing 2 mg/mL rhuMAb lot II, 50 mM D-glucose, and various concentrations of L-lysine were prepared in 100 mM sodium bicarbonate buffer at pH 7.4. The rhuMAb lot II starting material contained a low level of glycation (13%). When it was incubated at 37 °C with 50 mM glucose for 24 h in the absence of L-lysine, the glycation level (detected by boronate affinity chromatography) increased to 59%. When 50 mM L-lysine was present in the reaction mixture, the percentage of glycation dropped to 53%. Increasing the lysine concentration from 50 to 500 mM leads to a decrease in the percentage of glycation from 53 to 41%, indicating that free lysine inhibits glycation of the antibody in a concentration-dependent manner. This inhibitory role of lysine and its nonlinear relationship with the glycation level of rhuMAb is graphically described in Figure 3A.

As expected, the concentration of glucose has a direct impact on the glycation level of rhuMAb. This was revealed by the in vitro glycation results. With 20 mM D-glucose in the reaction mixture, ~32% of rhuMAb lot I was glycated after incubation at 37 °C for 24 h. As shown in Figure 3B, when the concentration of D-glucose was increased from 20 to 150 mM, the glycation level increased from 32 to 90%. A linear regression analysis of the data points resulted in a linear relationship between the natural log of the percentage of unglycated rhuMAb (y) and the glucose concentration (x) that can be described by the following equation: $y = -0.0137x + 4.4994$ with fitting coefficient $R^2 = 0.9887$. This supports the conclusion that the reaction is first order with respect to both unglycated protein and glucose. An extreme condition (with as much as 1200 mM D-glucose in the reaction mixture) was tested to verify the impact of the equilibrium concentration of the aldimine intermediate on glycation sites. The glycation of rhuMAb, monitored by boronate affinity chromatography, was driven to completeness (100%). These data clearly demonstrate that the concentration of glucose is critical in rhuMAb glycation. These results suggest that the level of glycation can be controlled to some degree during cell culture processes by modifying glucose feeding strategies.

During the early stage of a cell culture process, the total concentration of free amino acids and primary amine-containing compounds in the culture medium is high and can inhibit the glycation of rhuMAb. Furthermore, there is little accumulation of rhuMAb product early in a cell culture process. Therefore, one would expect that high concentrations of glucose could be fed early during a cell culture process without resulting in high levels of glycation in the final product. In the late stage of a cell culture process when the product accumulates and the concentration of secreted rhuMAb is much higher, then controlling the concentration of glucose at a low level may help minimize glycation of the product.

Identification of Glycation Sites on rhuMAb. The results of in vitro glycation experiments clearly demonstrated that rhuMAb was prone to modification by glucose. Mass spectrometric analyses of the rhuMAb lots produced by cell culture of early stable CHO cell lines revealed that glycation was predominantly located on the light chains. However, there are a total of 14 lysine residues on a light chain. Glycation of a lysine residue modifies the side chain with a sugar unit that prevents tryptic cleavage. When tryptic peptide maps of unglycated (flow-through) and glycated (retained) fractions of rhuMAb lot I were compared, their entire profiles matched well with the exception of two new peaks on the map of the glycated fraction. Figure 4 shows a partial region (30−90 min) of the maps with the two new peaks labeled as N1 and N2. Another peak at retention time of 40.7 min, labeled with an asterisk (*), shows a decrease in signal intensity in the glycated species (Figure 4C). LC/MS/MS characterization of both maps led to an assignment of the identities of all of the peptide peaks except N2. The asterisk (*) labeled peak was found to be

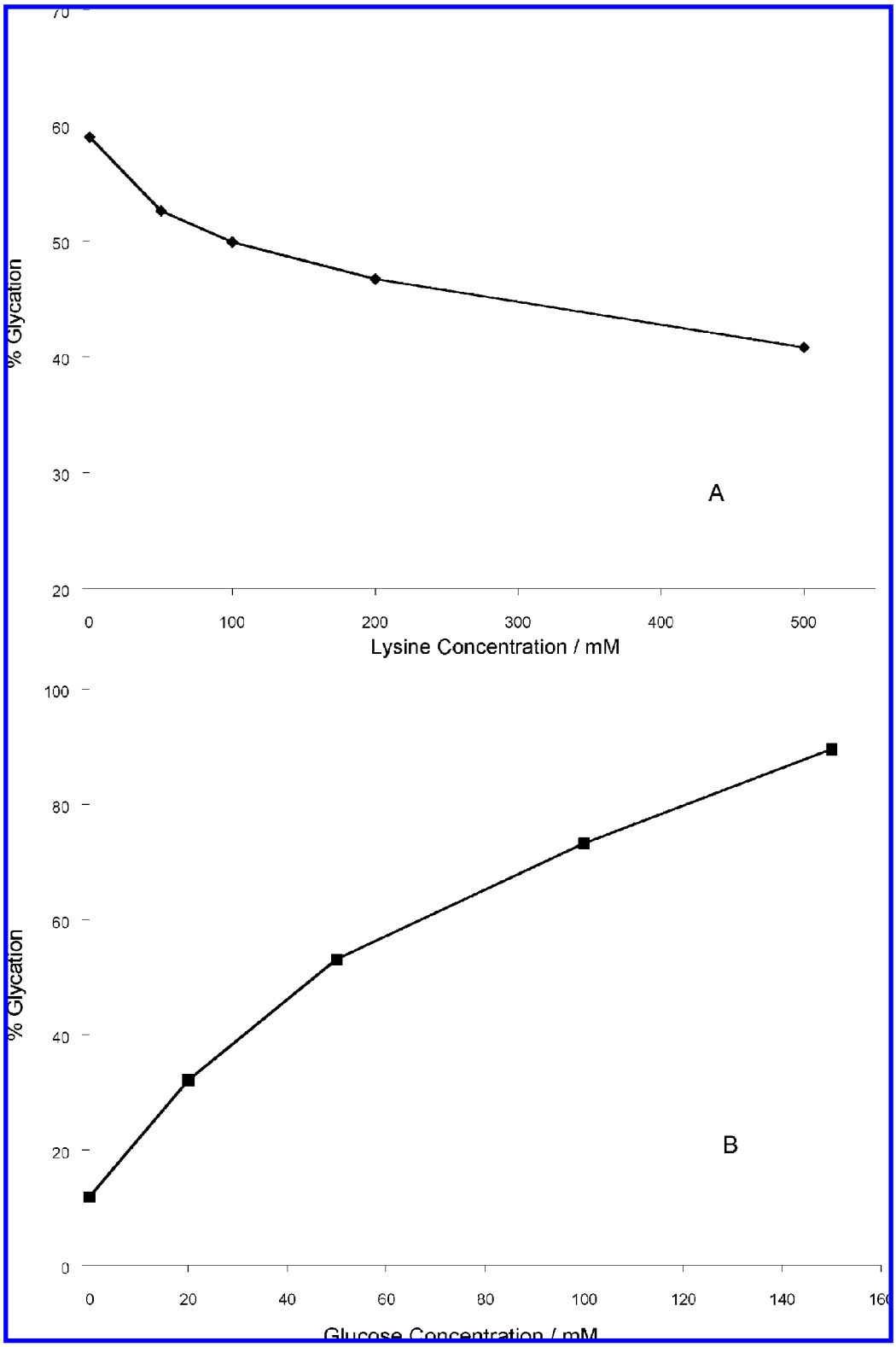

Figure 3. Extent of glycation of rhuMAb lot II determined by boronate affinity chromatography after in vitro incubation with (A) 50 mM glucose in the presence of various concentrations of lysine and (B) various concentrations of glucose.

YASQISGVPSR, a tryptic peptide from the rhuMAb light chain, referred to as T6-LC. The N1 peak was found to be a nonspecifically cleaved peptide with the sequence of ASQISGVPSR, a T6-LC peptide without the N-terminal tyrosine residue, indicating that the decrease of signal intensity of T6-LC was related to the generation of the extra peak N1. The sequence coverage of the tryptic peptide map was 97.7% for intact antibody. Only six small (di-, tri-, or tetraamino acid) peptides were not found, probably due to their highly hydrophilic character, and subsequent elution in the void volume flow-through. The N2 peak was determined to have a monoisotopic mass value of 1880 Da, which is 162 Da higher than the mass of the partial digestion peptide with

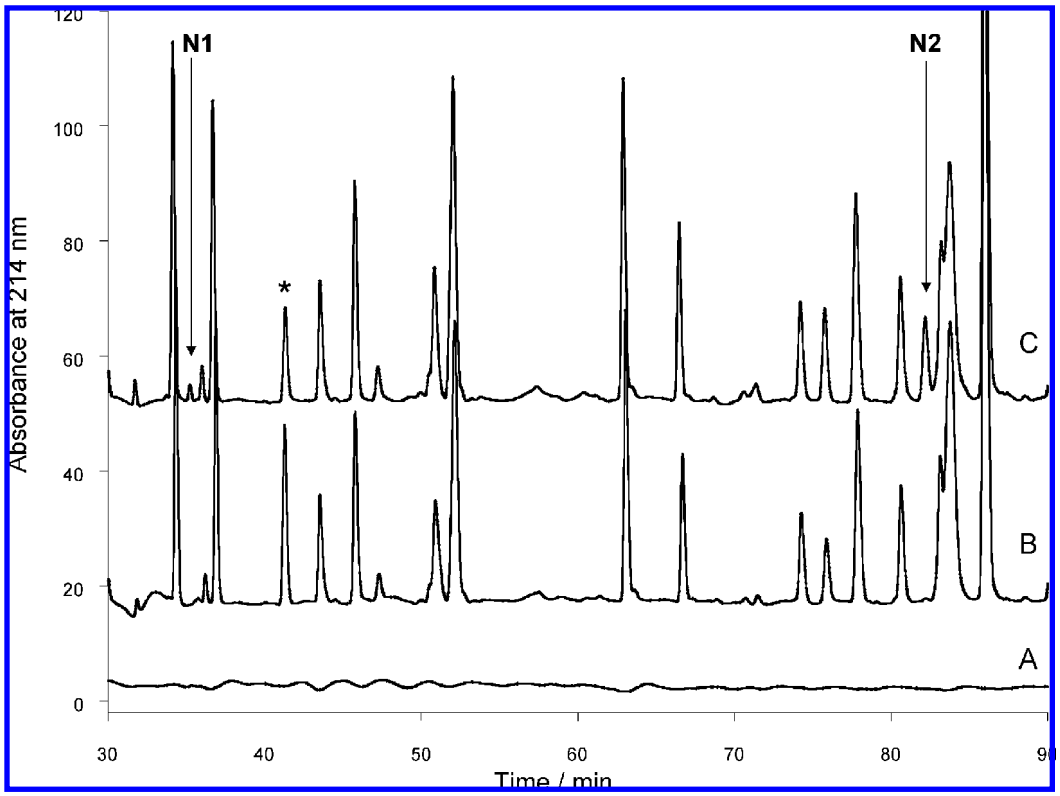

Figure 4. Partial region of the tryptic peptide maps (30−90 min) of (A) buffer blank of PBS, (B) flow-through fraction of rhuMAb lot I on boronate column, and (C) retained fraction of rhuMAb lot I on boronate column.

sequence of LLIKYASQSISGVPSR spanning residues 46−61 of the rhuMAb light chain. Though the mass increase suggested glycation, the MS^2 spectrum did not provide sufficient information. The spectrum is dominated by a peak suggesting loss of four water molecules and sequence-specific b- and y-ions are weak. To confirm the assignment and to identify the glycation site, the N2 peptide was isolated, collected from multiple HPLC peptide mapping analyses, and subjected to further MALDI-TOF MS^2 characterization and Edman degradation analysis.

As shown in Figure 5A, the MALDI-TOF MS^2 spectrum of the N2 peptide displays a series of well-distributed fragment ions. The ion signal at mass-to-charge ratio (m/z) of 1882.1 is a singly charged average mass (MH^+) of the peptide with one glucose modification. The ion with m/z of 1718.9 differs from the MH^+ ion by 163 Da. The strongest ion signal in the spectrum has a m/z value of 1761.0, which was postulated to be due to loss of $[(CHOH)_3CH_2OH]$ from the glycated form. The matched y-ions include y_1 and y_3−y_{12}, confirmed the sequence of the N2 peptide from the fifth residue to the C-terminus (YASQSISGVPSR, residues 50−61 of the rhuMAb light chain). Three matched b-ions, including b_2, b_3, and b_4, were observed. The b_3 and b_4 ions differ by 289.2 Da, which approximately equals the sum of a lysine residue (128.1 Da) and a glucose moiety (162 Da), indicating that the glycation is localized on the lysine residue in this peptide. Therefore, this mass spectrum confirmed that the peptide sequence and glycation site is the following: LLIK(Glc)YASQSISGVPSR. Edman degradation analysis of the N2 peptide was performed for eight cycles and revealed the N-terminal eight amino acid residues as LLIXYASQ. As shown in Figure 5B, there is no signal at 17.7 min, where the fourth residue lysine is expected to elute. A new peak labeled X eluted at 12.6 min on the reversed-phase column. This X residue is postulated to be the glycated lysine residue. Both MALDI-TOF MS^2 fragmentation analysis and Edman degradation analysis provide solid evidence that the N2 peptide is glycated on the K49 residue of the light chain (LC-K49) of rhuMAb.

The comparison of the tryptic peptide UV maps of glycated and unglycated fractions only revealed this one significant change. Using this approach, we were unable to identify any glycated peptides from the heavy chain that were inferred to be slightly glycated by ESI-MS molecular mass determination (Figure 1C). Hence, a proteomic approach was taken to analyze the LC/MS/MS data from the peptide maps. The raw data files acquired with the LTQ mass spectrometer were subjected to extensive data mining including searching of the rhuMAb single protein database using SEQUEST and a manual investigation of all spectra of matched peptides. Summarized in Table 1 are glycation sites identified for rhuMAb lot I and in vitro glycated lot II samples using this proteomic data mining approach. The estimated percent glycation is listed in the parentheses following each glycation site.

It can be seen from the data listed in Table 1 that when the final percentage of glycation of rhuMAb determined by boronate affinity chromatography is low, only the K49 residue on the light chain was found to be glycated. When the final percentage of glycation is higher than 53%, multiple lysine residues from both heavy and light chains were found to be glycated, though at very low percentages. The preference order of glycation cannot accurately be determined for those lysine residues other than LC-K49 based on the estimated glycation extent (the numbers in the parentheses).

The site specificity of protein glycation could be dictated by factors that facilitate either the formation of aldimines, the

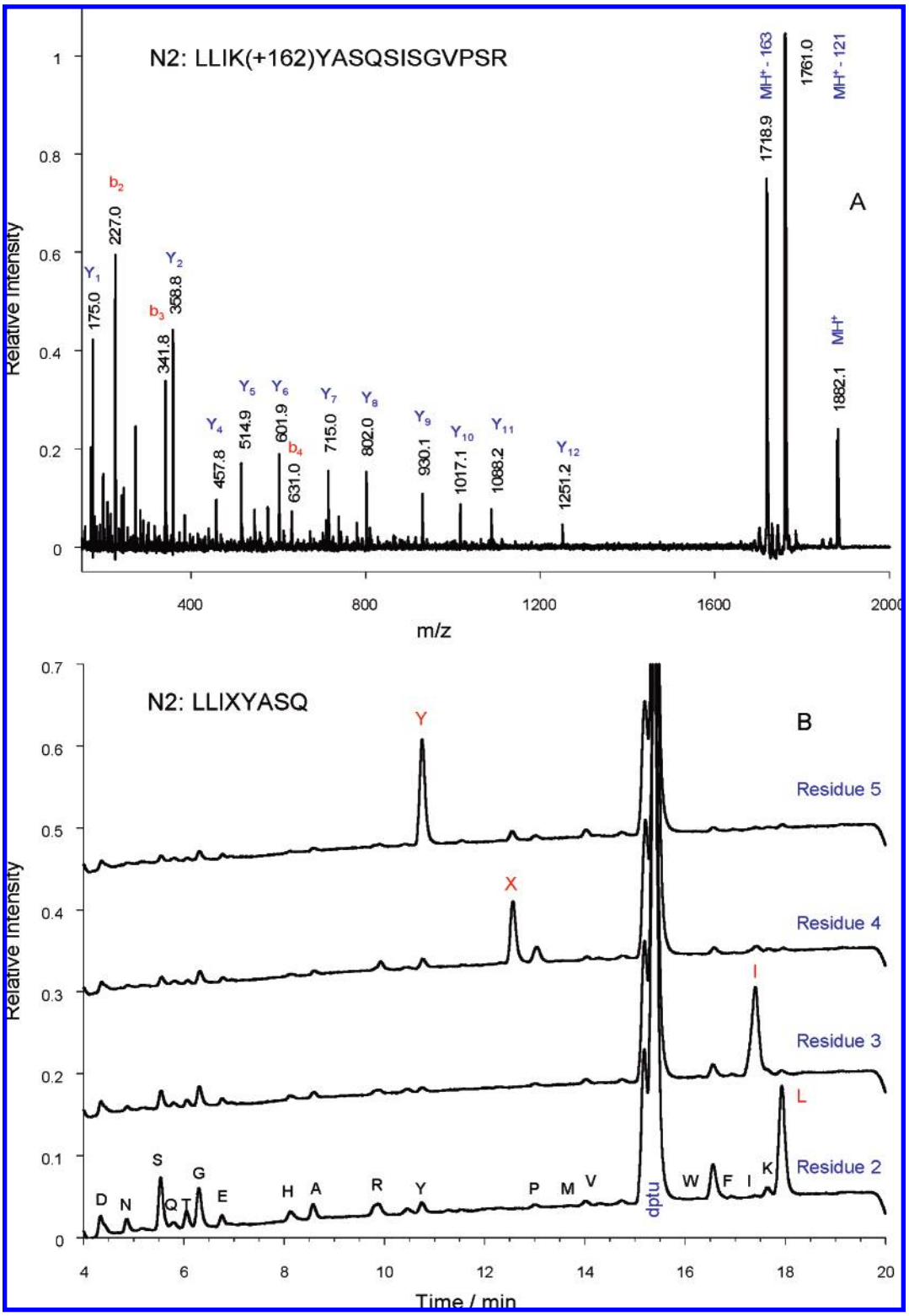

Figure 5. Sequencing analysis of the N2 peptide by (A) MALDI-TOF/TOF MS^2 and (B) eight cycles of Edman degradation analysis.

rearrangement of aldimine to ketoamine, or both.[28,29] The aldimine formation involves the unprotonated form of amines. In spite of the fact that α-amino groups of N-terminal residues in proteins have pK_a values of ~7.8, which is generally much lower than that of the side chain ϵ-amino groups of lysine residues,[29,30] only 3.3% of N-terminal aspartic acid residue (D) of light chain and 1.5% of glutamic acid residue (E) of heavy chain are glycated when rhuMAb lot II was treated with a high concentration of glucose (1.2 M). Under such an extreme condition, as displayed in Table 1, 90.1% of LC-K49 was found to be glycated, while the 10 other

(28) Feeney, R. E.; Blackenhorn, G.; and Dixon, H. B. F. *Adv. Prot. Chem.* **1975**, *29*, 135−203.

(29) Acharya, A. S.; Roy, R. P.; Dorai, B. *J. Prot. Chem.* **1991**, *10*, 345−358.

(30) Zhang, M.; Vogel, H. J. *J. Biol. Chem.* **1993**, *268*, 22420−22428.

Table 1. Identification of Glycation Sites on rhuMAb Using Tryptic Peptide Mapping and LC/MS/MS Characterization

	rhuMAb			
	lot I	lot II		
treatment[a]	none	none	50 mM glucose	1.2 M glucose
overall %glycation[b]	58%	13%	53%	100%
glycation sites (% glycation)[c]	**LC-K49 (57.8%)** HC-K288 (0.6%) LC-K42 (0.4%) HC-K133 (0.3%) LC-K207 (0.2%)	**LC-K49 (6.1%)**	**LC-K49 (45.5%)** HC-K246 (0.5%) LC-K42 (0.3%) HC-K288 (0.2%) LC-K190 (0.2%) LC-K207 (0.2%)	**LC-K49 (90.1%)** HC-K246 (14.1%) LC-K190 (12.1%) HC-K75 (8.8%) LC-K45 (5.8%) LC-K207 (5.3%) LC-K183 (4.8%) HC-K133 (3.5%) LC-D1 (3.3%) HC-K288 (3.2%) LC-K149 (3.1%) HC-K64 (2.8%) HC-E1 (1.5%)

[a] Detailed information of in vitro treatment is provided in the Experimental Section. [b] The overall percentage of glycation was obtained from boronate affinity chromatographic analysis. [c] The numbers provided in parentheses are approximate percentages of glycation at the corresponding amino acid residues. These numbers were calculated based on integrated peak area of extracted ion chromatogram of the ketoamine-modified peptide divided by total peak area of all related peptide ions, which include ketoamine and unmodified tryptic peptides.

Table 2. Glycation Levels Determined by Boronate Affinity Chromatography for rhuMAb and Its Recombinant MAb Mutants Manufactured in Small-Scale Transient CHO Cell Lines

	AA residues[a]		
mutant number	LC31	LC49	% glycation
rhuMAb	D	K	47
1	D	K	28
2	D	K	20
3	D	K	29
4	T	K	24
5	T	K	25
6	S	K	7
7	T	K	4
8	T	K	5
9	D	A	7
10	D	N	8
11	D	R	11
12	T	Y	7

[a] Only two amino acid residues at LC31 and LC49 are displayed.

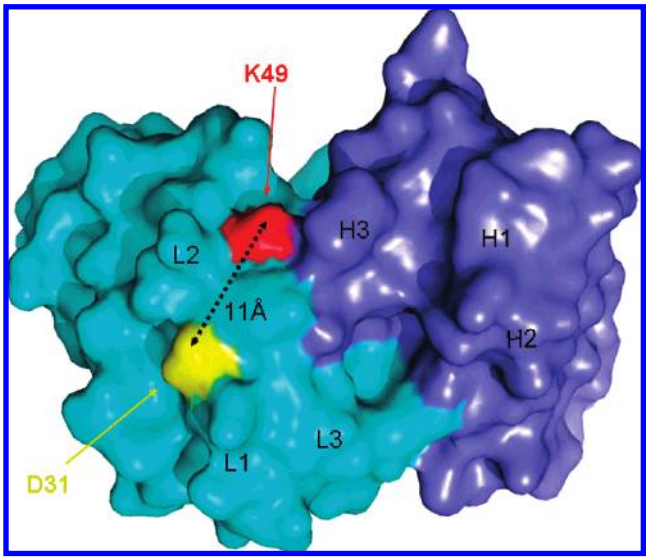

Figure 6. Homology model of the Fab part of rhuMAb. The surface of the structure is displayed using PyMOL Viewer software. The CDR domains on both light and heavy chains are labeled as L1, L2, and L3 and H1, H2, and H3, respectively. LC-K49 (red) and LC-D31 (yellow) are exposed to bulk solvent and located within 11 Å of each other.

lysine residues were found to be glycated to an extent that varied from about 1 to 14%. These data indicate that the pK_a of the amino group is not a major determinant of the site specificity of rhuMAb glycation. Therefore, it is reasonable to deduce that the Amadori rearrangement is the rate-limiting step in the glycation of rhuMAb. This supports the notion Watkins et al. obtained from the studies of ribonuclease A.[8]

Glycation Hot Spot on rhuMAb. As shown above, regardless of the method of glycation, the K49 residue on the light chain of the antibody is the first and predominant glycation site. To better understand why this lysine residue on rhuMAb is highly susceptible to modification by glucose, we examined the amino acid sequences and glycation level of rhuMAb and 12 other human or murine versions of recombinant MAbs that were prepared using the same transient CHO cell line at Genentech, Inc. All of the monoclonal antibody sequences examined have similar CDRs. The presence or absence of aspartate at residue 31 (middle of the first CDR region on the light chain (L1)) and lysine at residue 49 (adjacent to the second CDR on light chain (L2)) and the corresponding overall level of glycation are shown in Table 2. These MAbs can be considered in four groups depending on whether they contain D31 or K49. The first group consists of four Mabs, including rhuMAb, having amino acid residues at LC31 and LC49 as aspartic acid and lysine, respectively. Their overall glycation levels ranged from 20 to 47%. For the five MAbs in the second group, the amino acid residue at LC31 was either threonine (T) or serine (S) and LC49 was lysine; their overall glycation levels ranged from 4 to 25%. For the three Mabs in the third group, the amino acid residue at LC31 was aspartic acid but the amino acid residues at LC49 were changed to residues other than lysine; their overall glycation levels ranged from 7 to 11%. The fourth group

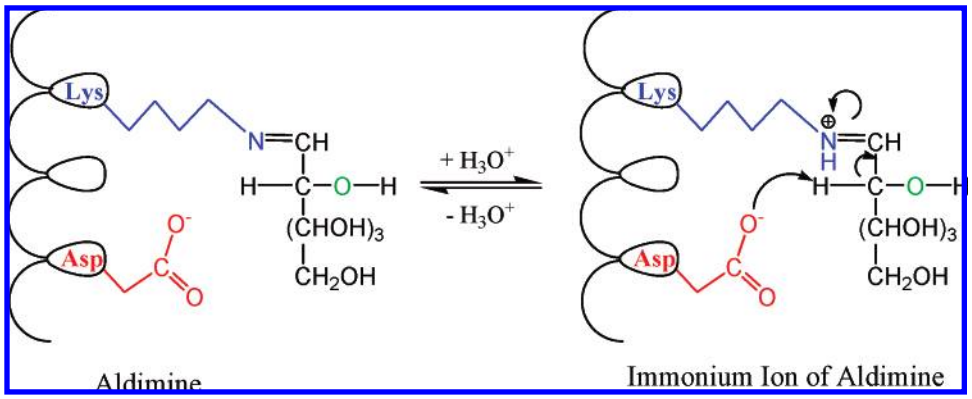

Figure 7. Schematic representation describing the internally catalyzed Amadori rearrangement of aldimine (Schiff base) adduct of LC-K49 by LC-D31 on surface of rhuMAb. Only the aldimine and its imonium ion chemical structures are displayed; the remaining two steps (enolization and tautomerization to generate the ketoamine adduct) are similar to the machanism proposed by Watkins et al.[9]

only has one MAb in which both LC-D31 and LC-K49 were changed to other amino acid residues; the glycation level was determined to be 7%. These MAb research materials were prepared under similar cell culture conditions, but because of the possible variations in cell growth rate and glucose concentration, some variation in the extent of glycation is expected. However, a trend can be seen in the data in Table 2 that suggests the coexistence of LC-D31 and LC-K49 in a MAb leads to high glycation; loss of either residue leads to decreased glycation. This trend suggests that LC-D31 may assist the glycation at LC-K49. A homology model of the variable region of rhuMAb was used to explore the structural relationship between these two residues. In Figure 6, a model of the surface of the variable domains of rhuMAb was generated based on the published crystal structure of a portion of rhuMAb-her2. Both LC-K49 (red) and LC-D31 (yellow) are exposed to bulk solvent and located within 11 Å of each other like two "hills" facing each other over a "valley".

We propose that the LC-D31 residue acts as an internal catalyst facilitating the Amadori rearrangement of the aldimine adduct on LC-K49. The average N−O distance in a hydrogen bond in proteins is ∼3 Å.[31] Because of the side chain flexibility, the distance between the oxygen atom of the side chain carboxylate group of Asp31 residue and the nitrogen atom of the Schiff base in the aldimine intermediate is much less than 11 Å but probably larger than 3 Å. As described in Figure 7, although they are not close enough to form a hydrogen bond, it is reasonable to postulate that the oxygen atom of the carboxylate group of LC-D31 acts as a nucleophile and approaches closely enough to abstract a proton from the α-carbon atom in the protonated aldimine (imonium ion) side chain. Figure 7 only displays the first two steps of the proposed catalytic mechanism of Amadori rearrangement in glycation of rhuMAb. The remaining two steps, namely, the formation of enolamine and tautomerisation to give rise to the relatively stable ketoamine adduct, are suggested to be extremely similar to the catalytic mechanism of glycation of ribonuclease A by a bound-phosphate anion that was proposed by Watkins et al. in 1987.[9] Unlike ribonuclease A, in which there is a unique phosphate-binding microenvironment in its active site,[9] the glycation susceptibility of LC-K49 in rhuMAb is independent of the buffer components tested that included phosphate, bicarbonate, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES) (data not shown). Therefore, the intramolecular catalytic mechanism is a reasonable explanation of the early glycation hot spot at LC-K49. In fact, as early as 1958, Isbell and Frush proposed a mechanism for carboxylic acid promoting Amadori rearrangement.[32] This proposed catalytic mechanism is also consistent with previous observations of glycation preference for lysine residues present in the vicinity of a proton abstracting group.[10,12] A similar catalytic role of aspartic acid residues was also suggested by Venkatraman et al. to explain the experimental glycation results of model peptides.[33] In that study, five synthetic stereochemically defined peptide helices were utilized to demonstrate that aspartic acid was an efficient catalytic residue in the Amadori arrangement.

CONCLUSIONS

The results presented here provide solid evidence that Lys49, adjacent to CDR-2 on the light chain, is the most susceptible glycation site in rhuMAb. The major factor for this site preference is the spatial proximity effect for catalysis by the Asp31 residue in the region of CDR-1 on the light chain. The unusually high glycation level observed in rhuMAb materials produced in transient and stable CHO cell lines can be explained by the unusual susceptibility of LC-K49 to modification by glucose. Glycation at other sites is detected at higher levels of total glycation.

The details of the mechanism of early-stage glycation of rhuMAb during its production have been delineated by the present study. When synthesized inside the CHO cells and secreted into the cell culture fluid, multiple lysine residues in rhuMAb react with free glucose in the media to form labile aldimine derivatives. This first step of glycation is in a concentration-dependent dynamic equilibrium between reactants (rhuMAb and glucose) and aldimine products. Higher concentrations of free glucose will induce higher concentrations of aldimine intermediates and greater numbers of modified lysine residues. Fast conversion of aldimine adducts to ketomine adducts will drive the

(31) Xu, D.; Tsai, C.-J.; Nussinov, R. *Protein Eng.* **1997**, *10*, 999−1012.

(32) Isbell, H. S.; Frush H. L. *J. Org. Chem.* **1958**, *23*, 1309−1319.

(33) Venkatraman, J.; Aggarwal, K.; Balaram, P. *Chem. Biol.* **2001**, *8*, 611−625.

overall reaction forward. Unconverted aldimine adducts are unstable and can return to free antibody when it is harvested for purification, especially since no glucose is present in the elution solutions. The Amadori rearrangement of aldimine to form a ketoamine appears to be the rate-limiting step of the overall glycation reaction. In rhuMAb, because of the catalytic activity provided by LC-D31, the aldimine adduct formed at LC-K49 is converted to the more stable ketoamine form much faster than those on other sites. The so-called site-specific glycation seems to be caused by the catalytic effect of nucleophiles such as aspartic acid. To minimize and control the glycation level of rhuMAb, it is suggested that the glucose concentration in the culture fluid should be controlled at a low level, especially when it is close to the harvest day when the competitions for glucose by amino acid and peptide nutrients become low and product titer becomes high.

ACKNOWLEDGMENT

The authors thank Milly Milton and Sonia Pollitt for the purification of the standard material (rhuMAb lot I) used in the study; Louisette Basa and Baojen Shyong for providing technical support in ESI-MS determination of molecular masses and LC/MS/MS analysis of tryptic peptide maps; Armando Codorba for technical support in N-terminal sequencing of the isolated N2 peptide; Cindy Quan for information on boronate chromatography; Dr. Sherman Fong and Henry Chiu for information about 12 human and murine versions of MAb; and Drs. Eleanor Canova-Davis and Brad Snedecor for critical review and helpful discussion of the manuscript.

Received for review August 27, 2007. Accepted November 5, 2007.

AC701810Q