

Therapeutic Proteins

Methods and Protocols

Edited by

C. Mark Smales

*Protein Science Group, Department of Biosciences,
University of Kent, Canterbury, Kent, UK*

and

David C. James

*School of Engineering, University of Queensland, St. Lucia,
Queensland, Australia*


HUMANA PRESS  TOTOWA, NEW JERSEY

© 2005 Humana Press Inc.
999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

www.humanapress.com

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise without written permission from the Publisher. *Methods in Molecular Biology*TM is a trademark of The Humana Press Inc.

All papers, comments, opinions, conclusions, or recommendations are those of the author(s), and do not necessarily reflect the views of the publisher.

This publication is printed on acid-free paper. 
ANSI Z39.48-1984 (American Standards Institute)

Permanence of Paper for Printed Library Materials.

Cover illustration: Figure 2 from Chapter 21, "Solid-State Protein Formulation: *Methodologies, Stability, and Excipient Effects*," by Yuh-Fun Maa and Scott P. Sellers.

Cover design by Patricia F. Cleary.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: orders@humanapr.com; or visit our Website: www.humanapress.com

Photocopy Authorization Policy:

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Humana Press Inc., provided that the base fee of US \$30.00 per copy is paid directly to the Copyright Clearance Center at 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license from the CCC, a separate system of payment has been arranged and is acceptable to Humana Press Inc. The fee code for users of the Transactional Reporting Service is: [1-58829-390-4/05 \$30.00].

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging in Publication Data

Therapeutic proteins : methods and protocols / edited by C. Mark Smales and David C. James.

p. ; cm. — (Methods in molecular biology ; 308)

Includes bibliographical references and index.

ISBN 1-58829-390-4 (alk. paper) eISBN 1-59259-922-2

1. Protein drugs—Laboratory manuals. I. Smales, C. Mark. II. James, David C. (David Cameron), 1963- . III. Series: Methods in molecular biology (Clifton, N.J.) ; v. 308.

[DNLM: 1. Proteins—therapeutic use. 2. Proteins—chemical synthesis. 3. Proteins— isolation & purification. QU 55 T3976 2005]

RS431.P75T475 2005

615'.3—dc22 2004021242

Top-Down Characterization of Protein Pharmaceuticals by Liquid Chromatography/Mass Spectrometry

*Application to Recombinant Factor IX Comparability—
A Case Study*

**Jason C. Rouse, Joseph E. McClellan, Himakshi K. Patel,
Michael A. Jankowski, and Thomas J. Porter**

1. Introduction

Recombinant protein pharmaceuticals have revolutionized the treatment of a variety of medical ailments, including cancer, autoimmune diseases, and hemostatic disorders. Proteins manufactured with eukaryotic expression systems may be complex and heterogeneous because of posttranslational modifications (PTMs) and differential proteolytic processing. At one time, detailed characterization and definition of the protein structure were difficult, and the manufacturing process defined the product. If process changes were made, clinical trials were required to demonstrate product equivalence prior to regulatory agency acceptance of the product manufactured by the modified process. To adopt new manufacturing processes in a timely manner, the biopharmaceutical industry and regulatory agencies have worked together over the last few years to develop new guidance documents based on knowledge gained from industry experience in the manufacture and clinical testing of protein pharmaceuticals (1,2). Manufacturers of protein pharmaceuticals consistently strive to deliver the highest quality product in a cost-efficient manner. This can be accomplished through optimization of the production process and incorporation of new technologies to enhance product purity and yield. Process improvements may include a change of the host cell line, enhancement of the cell culture medium or cell culture management, or modifications to the purification process. In some cases, an additional manufacturing site is brought online to augment production capacity.

Steady advancements in analytical techniques and methodology now allow a protein pharmaceutical to be considered a “defined biologic” (formerly known as “well-characterized protein”). This designation provides the opportunity to demonstrate the

From: *Methods in Molecular Biology*, vol. 308: *Therapeutic Proteins: Methods and Protocols*
Edited by: C. M. Smales and D. C. James © Humana Press Inc., Totowa, NJ

“comparability” of the product manufactured by a new process to that produced with the original manufacturing process. Establishing structural and functional comparability by a protocol that includes predefined acceptance criteria can minimize the need to conduct extensive preclinical or clinical testing. Successful demonstration of comparability can therefore improve product development timelines or accelerate the implementation of changes to a commercial manufacturing process.

Mass spectrometry (MS) has emerged as an integral component of comparability programs because it is a rapid, highly informative analytical technique capable of detecting structural changes in peptides and proteins with high selectivity and sensitivity. MS is especially powerful in protein structural analysis when interfaced with reversed-phase high-performance liquid chromatography (RP-HPLC). The structure of recombinant proteins is traditionally assessed through RP-HPLC peptide mapping, which typically involves disulfide bond reduction, cysteine alkylation, and enzymatic digestion, followed by RP-HPLC profiling of the component peptides. The established combination of peptide mapping with MS, which links peptide RP-HPLC elution time with molecular weight and structure, provides verification of the intended amino acid sequence, definition of the N and C termini, and site-specific detection and characterization of the PTMs or degradation products (**Table 1**).

Top-down characterization (**3**) is a newer MS approach for protein identification and characterization that involves accurate mass determination of intact proteins, protein subunits, or large proteolytic fragments with the option of gas-phase ion fragmentation of the intact species. This approach is complementary to peptide mapping with MS (**Note 1**) and, when gas-phase ion fragmentation is employed, the structural endpoints are very similar to in-solution proteolysis. However, in contrast to peptide mapping, the top-down approach provides 100% amino acid sequence coverage, because an accurate mass determination for a particular protein isoform (or related gas-phase fragment ion) is a direct link to its amino acid composition, N and C termini, and PTMs (**Table 1**). Additionally, the top-down strategy has the potential to reduce the number of upstream sample handling and separation steps. If top-down analysis is performed in the liquid chromatography (LC)/MS mode, the simple HPLC profiles, as compared to that of a peptide map, allow protein recovery determination and efficient data analysis. Because intact proteins are more resistant than peptides to the collisional activation processes that occur during MS analysis, labile PTMs (e.g., sulfate and carboxylate on tyrosine-*O*-sulfate and γ -carboxyglutamate, respectively) are largely retained. This phenomenon results in a representative distribution of the protein isoform heterogeneity in the mass spectrum.

Thorough protein characterization by the top-down strategy ideally requires a mass analyzer with high-resolving power and accurate mass-measurement capabilities for unambiguous mass determinations of whole proteins or related gas-phase fragment ions (**6**). With high-resolution tandem mass spectrometry (MS/MS) and specialized fragmentation techniques, initial top-down characterization strategies focused on the generation of extensive, systematic gas-phase fragmentation from intact proteins. This provided almost complete amino acid sequence in several studies for protein identification and site-specific PTM characterization (**Note 2**). Alternatively, the top-down

Table 1
Characterization Methods for Recombinant Glycoproteins

Analyte	Method	Information obtained	Considerations
Component peptides	Reduction, alkylation, proteolytic digestion, C18 RP-HPLC/MS (mass) and nanoESI-QTOF-MS/MS (sequence)	Verification of amino acid sequence Definition of N and C termini Site-specific detection and identification of PTM and degradation products	≤100% Sequence coverage Complicated HPLC profiles with proteolysis artifacts Peptides with PTM may be lost, have low recoveries, or altered distributions
Component <i>N</i> - and <i>O</i> -glycans	Release: PNGase F (<i>N</i> -linked) or hydrazinolysis (<i>N</i> - and <i>O</i> -linked) or β-elimination (<i>O</i> -linked) Profiling: underivatized glycans by HPAEC-PED or derivatization at reducing end: normal phase HPLC with fluorescence detection Mass analysis: MALDI-TOF MS	Elucidation of glycan structures and isomers	Allows for direct analysis of glycans Requires extensive sample manipulation
Intact protein isoforms	C4 or phenyl RP-HPLC/ESI-QTOF MS and nanoESI-QTOF MS	Accurate mass Sequence composition N- and C-terminal processing PTM composition Isoform distribution Extent of heterogeneity New species Correlation to peptide map, <i>N</i> -glycan profile, and other essays	100% Sequence coverage Simple HPLC profiles Retention of labile PTM (i.e., sulfation) Efficient data analysis Determination of protein recovery Reduced sample handling

HPAEC-PED, high-performance anion-exchange chromatography-pulsed electrochemical detection (PED).

Table 2
Mass Analyzer Comparison^a

MS instrument	Ionization mode and mass analyzer	Resolving power ($m/\Delta m$)	Mass accuracy External Calibration (Daltons [%])
Bruker Reflex	MALDI-TOF	900 (Insulin) ^b	±0.6 (0.01)
		600 (Trypsinogen) ^c	±3.0 (0.01)
Waters Micromass Platform II	ESI-Quadrupole	450 (Insulin)	±0.6 (0.01)
		450 (Trypsinogen)	±2.4 (0.01)
Waters Micromass Q-Tof-2	ESI-QTOF (V-Optics)	10,500 (Insulin)	±0.1 (0.002)
		2300 (Trypsinogen)	±0.5 (0.002)
Waters Micromass Q-Tof API-US	ESI-QTOF (W-Optics)	21,000 (Insulin)	±0.06 (0.001)
		2450 (Trypsinogen)	±0.2 (0.001)

^aUsed at Wyeth BioPharma.

^bInsulin = 5729.601 Daltons (bovine).

^cTrypsinogen = 23,981.0 Daltons (bovine).

approach has successfully been implemented on stock and customized mass analyzers with modest resolving power using MS, MS/MS, and the standard collision-induced dissociation (CID) technique, which induces cleavage at the weaker backbone amide bonds (**Note 3**). Many advances in MS during the last 5 yr have centered on improving the performance of conventional mass analyzers (**Table 2**). The latest hybrid quadrupole time-of-flight (QTOF) mass analyzers (**Fig. 1**) feature resolving powers above 10,000 ($m/\Delta m$) with sensitivity in the low-femtomole range, an extended mass range, and stability for accurate mass measurements (**24,25**). These capabilities make this instrument effective for top-down characterization of intact proteins, especially for the exact mass determinations of intact protein isoforms with high sensitivity (**Note 6**).

Fig. 1. (*opposite page*) Schematic of the Waters Micromass Q-Tof-2 mass spectrometer. This instrument was used for all RP-HPLC/MS analyses of activated rFIX samples. FIX activation products in the eluent from the HPLC (flowing 0.100–0.133 mL/min) are desolvated and ionized via the Z-Spray ESI source. The gas-phase analyte ions are accelerated into the QTOF mass analyzer by the potential difference between the ion source block and the extractor cone. All ions are then focused by the hexapole ion bridge. Ions ranging from m/z 800 to 4000 are selected with the quadrupole mass filter (**Note 4**). These ions then enter the collision cell at the appropriate axial translational energy of 5 or 10 eV. Through cooling collisions with argon gas at approx 6×10^{-5} mbar, the ion beam becomes more monoenergetic and collimated prior to orthogonal acceleration into the TOF analyzer. The orthogonal acceleration ion optics (**Note 5**) successively accelerate segments of the quadrupole ion beam into the TOF analyzer for mass measurement every 88 μ s (for this work), which is the necessary time for detection of all ions from m/z 800 to 4000. In the reflectron TOF mass analyzer, the m/z measurement is based on the time it takes an accelerated ion to traverse the total distance from the pusher region through the reflectron to the MCP detector.

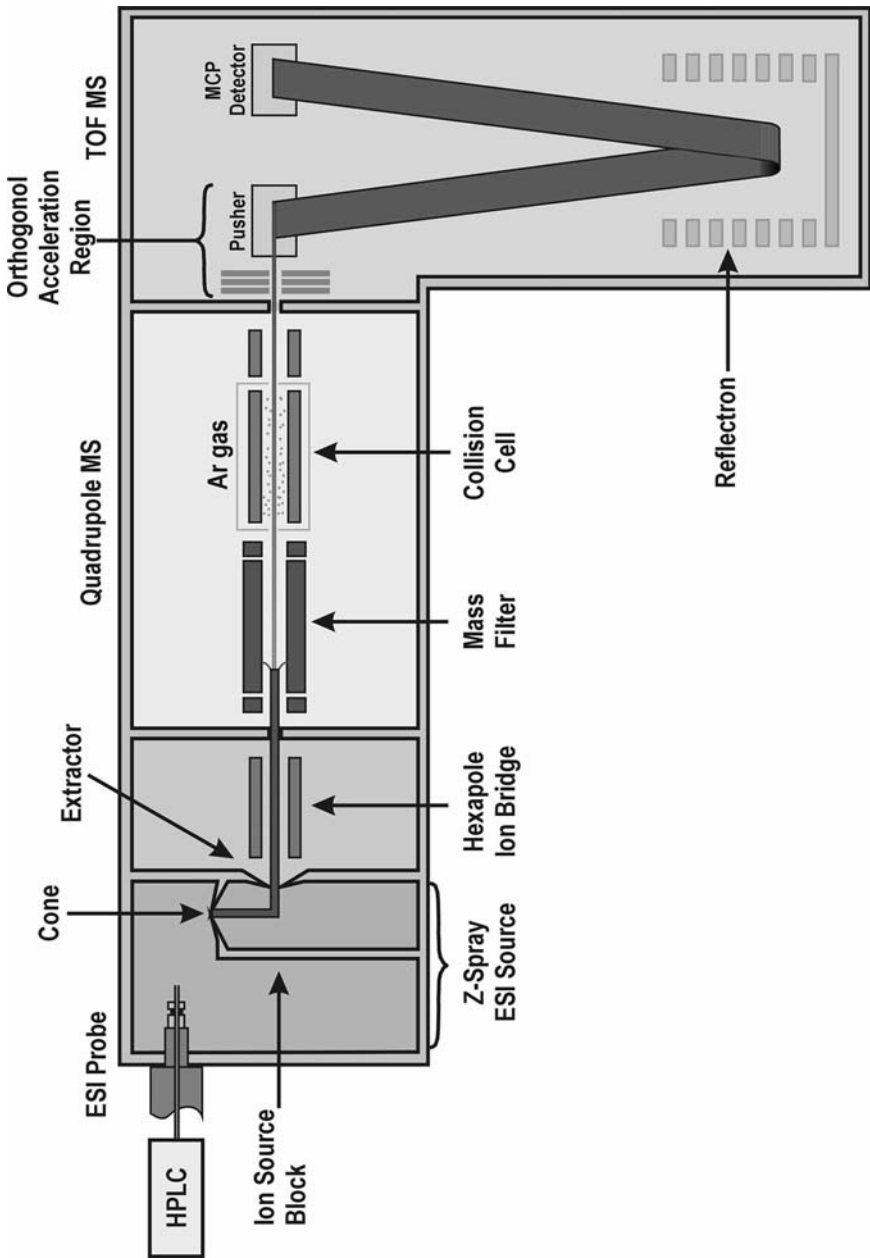


Fig. 1.

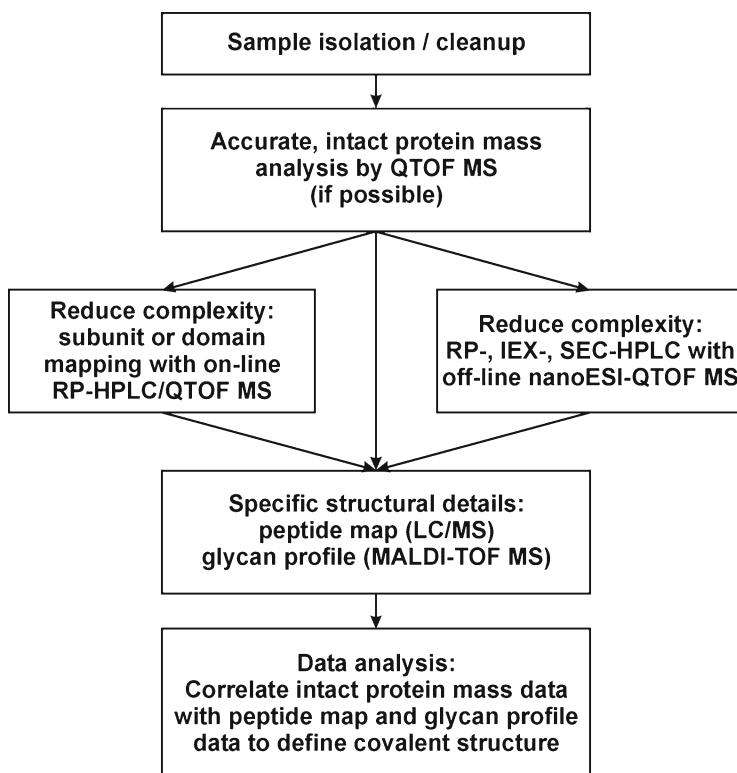


Fig. 2. Our top-down characterization strategy for recombinant proteins.

In our experience, the QTOF mass analyzer, when combined with the powerful maximum entropy (MaxEnt-1) mass deconvolution software (28,29), has provided the unique capability to baseline resolve the mass signals of multiple, variably processed, glycosylated isoforms of recombinant glycoproteins with mass accuracies better than 0.005% (Note 7). The overall isoform profiles of proteins, ranging in mass up to 150 kDa, were found to correlate with structural information from other assays. Thus, our top-down characterization strategy (Fig. 2), which involves the QTOF mass analyzer and MaxEnt-1 software, is built upon accurate mass determinations and protein isoform mass profiles for full-length glycoproteins (if possible), protein subunits, or large proteolytic fragments as an efficient method to survey the global structural heterogeneity. However, instead of fragmentation of intact protein isoforms in the QTOF mass analyzer, our strategy depends on the previously established structural framework, as determined by the MS-characterized peptide map and *N*- and *O*-glycan profiles, to provide the specific structural details that complement the intact mass information (Table 1).

Our version of top-down protein characterization has become an important factor in our structural studies and comparability assessments of product candidates and com-

		Domains
1	Y N S G K L E E F V Q G N L E R E C M E E K C S F	Gla
26	E E A R E V F E N T E R T T E F W K Q Y <u>V</u> D G D Q	
51	C E S N P C L N G G S C K D D I N S Y E C W C P F	EGF's
76	G F E G K N C E L D V T C N I K N G R C E Q F C K	
101	N S A D N K V V C S C T E G Y R L A E N Q K S C E	
126	P A V P F P C G R V S V S Q T S K L T R <u>A E A</u> V F	
151	P D V D Y V N S T E A E T I L D N I T Q S T Q S F	AP
176	N D F T R <u>V</u> V G G E D A K P G Q F P W Q V V L N G	
201	K V D A F C G G S I V N E K W I V T A A H C V E T	SP
226	G V K I T V V A G E H N I E E T E H T E Q K R N V	
251	I R I I P H H N Y N A A I N K Y N H D I A L L E L	
276	D E P L V L N S Y V T P I C I A D K E Y T N I F L	
301	K F G S G Y V S G W G R V F H K G R S A L V L Q Y	
326	L R V P L V D R A T C L R S T K F T I Y N N M F C	
351	A G F H E G G R D S C Q G D S G G P H V T E V E G	
376	T S F L T G I I S W G E E C A M K G K Y G I Y T K	
401	V S R Y V N W I K E K T K L T	

Fig. 3. Sequence and domain structure of FIX. Sites of PTMs are highlighted in shaded boxes. The 12 γ -carboxyglutamates are indicated in the Gla domain. The epidermal growth factor (EGF) domains contain two unique *O*-glycans at Ser⁵³ and Ser⁶¹ and one β -hydroxy-aspartate at Asp⁶⁴. The activation peptide (AP) contains tyrosine-*O*-sulfate at Tyr¹⁵⁵, two N-linked glycans at Asn¹⁵⁷ and Asn¹⁶⁷, and O-linked glycans at Thr¹⁵⁹, Thr¹⁶⁹, Thr¹⁷², and Thr¹⁷⁹ (Notes 12 and 13). Phosphorylation at Ser¹⁵⁸ (open box) is found in plasma-derived FIX but not in rFIX. rFIX is the Ala¹⁴⁸ (underlined) isotype. SP = serine protease domain.

mercial products during evaluation of new reference materials, process scale-up, process improvements, and implementation of new manufacturing facilities. The precise mass determination and isoform profiles acquired by RP-HPLC/electrospray ionization (ESI) QTOF MS, in conjunction with comprehensive product release assay data, have supported the structural comparability and consistency in the samples under evaluation (Note 8). The side-by-side analysis of protein samples with LC/MS minimizes sample handling and provides data sets that are optimal for comparative purposes. We have used this top-down characterization approach in comparability studies with protein pharmaceuticals, such as recombinant human bone morphogenetic protein-2 (30) and the clotting factor, B-domain deleted factor VIII (BDDrFVIII) (31,32; Note 9). Additionally, this technique has been extremely useful in the characterization of recombinant monoclonal antibodies (mAbs) during product development (30,33; Note 10).

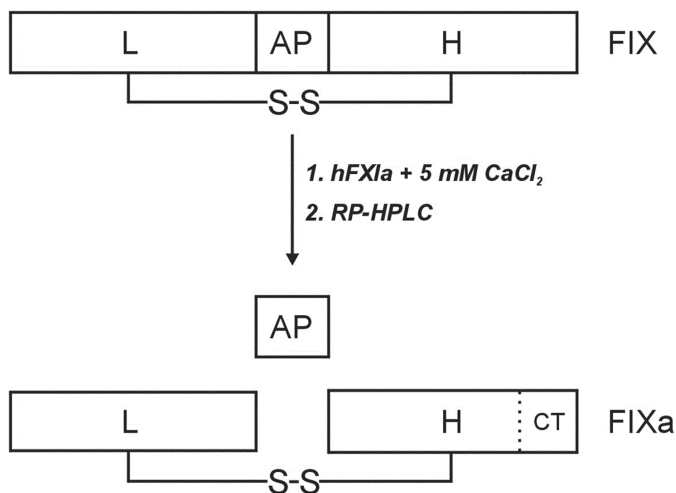


Fig. 4. Activation of FIX with FXIa. L, light chain; AP, activation peptide; H, heavy chain; CT, C-terminal peptide. FXIa cleaves FIX after Arg¹⁴⁵ and Arg¹⁸⁰ to generate FIXa—the enzymatically active product of activation. Under the conditions employed in this work, low-level cleavages were observed in the CT region of the HC, generating CT peptides (Notes 14 and 15).

This chapter presents the application of our top-down characterization approach for assessing structural comparability of a well-characterized glycoprotein, recombinant human factor IX (rFIX; **34**), produced at two manufacturing sites (Note 11). rFIX is a multidomain protein that contains numerous PTMs (Fig. 3). Specific proteolytic cleavage of rFIX at two peptide bonds was employed to liberate the activation peptide (AP; Fig. 4). Chromatographic separation of the rFIX fragments permitted acquisition of mass spectra of reduced complexity when compared to the mass spectrum of intact rFIX. This procedure improved the signal-to-noise ratio and resolution of the rFIX isoforms, which made accurate mass assignments possible for the full range of species present. The experimental methods employed for the top-down characterization of rFIX, including data evaluation and MS acceptance criteria for the determination of structural comparability, are presented here.

2. Materials

2.1. Equipment

1. HPLC System: Alliance 2695 equipped with a 2487 UV/Vis detector (Waters Corporation, Milford, MA).
2. Chromatography column: Symmetry C4 RP-HPLC column (3.9 × 150 mm, 5 μm, part no. 186000286, Waters Corporation).
3. HPLC flow splitter: Low pressure micro-splitter valve (P-451, Upchurch Scientific, Oak Harbor, WA).

4. Syringe pump: Pump 22 (Harvard Apparatus, South Natick, MA).
5. Mass Spectrometer: Q-Tof-2 (Micromass MS Technologies, Waters Corporation, Milford, MA), a hybrid QTOF mass analyzer equipped with a Z-spray ion source for ESI (Fig. 1).

2.2. Software

1. MassLynx 3.5 for NT (Micromass MS Technologies, Waters Corporation) is the Q-Tof-2 instrument control, data acquisition, calibration, and mass spectrum processing software.
2. Probabilistic maximum entropy analysis (MaxEnt-1 module, Micromass MS Technologies, Waters Corporation) is used to deconvolute the multiple-charged mass data (typical for an ESI mass spectrum) into a zero-charge mass spectrum. This software enables straightforward elucidation of molecular weight values for the different protein isoforms without prior knowledge of the protein species present in the sample.
3. Protein Analysis Worksheet (PAWS), version 2000.06.08 for Windows 95/98/NT/2000 (Genomic Solutions, Ann Arbor, MI), is used for the determination of theoretical mass values (monoisotopic and average mass) for intact proteins and subunits based on the predicted amino acid sequences and PTMs.

2.3. Reagents

1. rFIX test materials: multiple batches of rFIX manufactured at sites 1 and 2 were chosen for analysis. Only batches meeting the quality-control release test specifications were eligible for top-down MS analysis and other characterization tests. Representative data obtained for one rFIX batch from each manufacturing site are presented in this chapter.
2. Enzyme: human factor XIa (hFXIa; HCXIA-0160, Haematologic Technologies, Inc.).
3. HPLC-grade solvents: Water (Purelab Plus UV, US Filter, type PL5112 02), trifluoroacetic acid (TFA, 1-mL vials, cat. no. 28904, Pierce), and acetonitrile (HPLC grade, cat. no. AX0145-1, EMD Chemicals, Inc.).
4. Chromatography mobile phases: 0.1% (v/v) TFA in water (mobile-phase A) and 0.1% (v/v) TFA in 95% acetonitrile, 5% water (mobile-phase B).
5. Q-Tof-2 mass axis calibrant: sodium iodide (NaI, Sigma) at 2 mg/mL in 50:50 2-propanol:water (v/v) is prepared daily.
6. Insulin solution: bovine insulin (Sigma) at 2 pmol/ μ L in 50:50 acetonitrile:water (v/v) with 2% (v/v) formic acid is prepared daily.
7. Trypsinogen solution: bovine trypsinogen (Sigma) at 2 pmol/ μ L in 50:50 acetonitrile:water with 2% (v/v) formic acid is prepared immediately prior to each analysis.
8. Formulation buffer: 10 mM L-histidine, 260 mM glycine, 1% (w/v) sucrose, and 0.005% (w/v) polysorbate 80, pH 6.8.
9. NaCl and CaCl₂.

3. Methods

3.1. Activation of rFIX (Fig. 4)

1. Activation of rFIX with hFXIa is performed in a formulation buffer supplemented with 0.1 M NaCl and 5 mM CaCl₂ at an enzyme-to-substrate ratio of 1:100 (w/w) at 37°C for 1 h and 15 min (35,36; Note 16). The final concentration of rFIX is 2.2 mg/mL.
2. After incubation, snap-freeze the activation reaction in liquid N₂, and store at -80°C until LC/MS analysis.

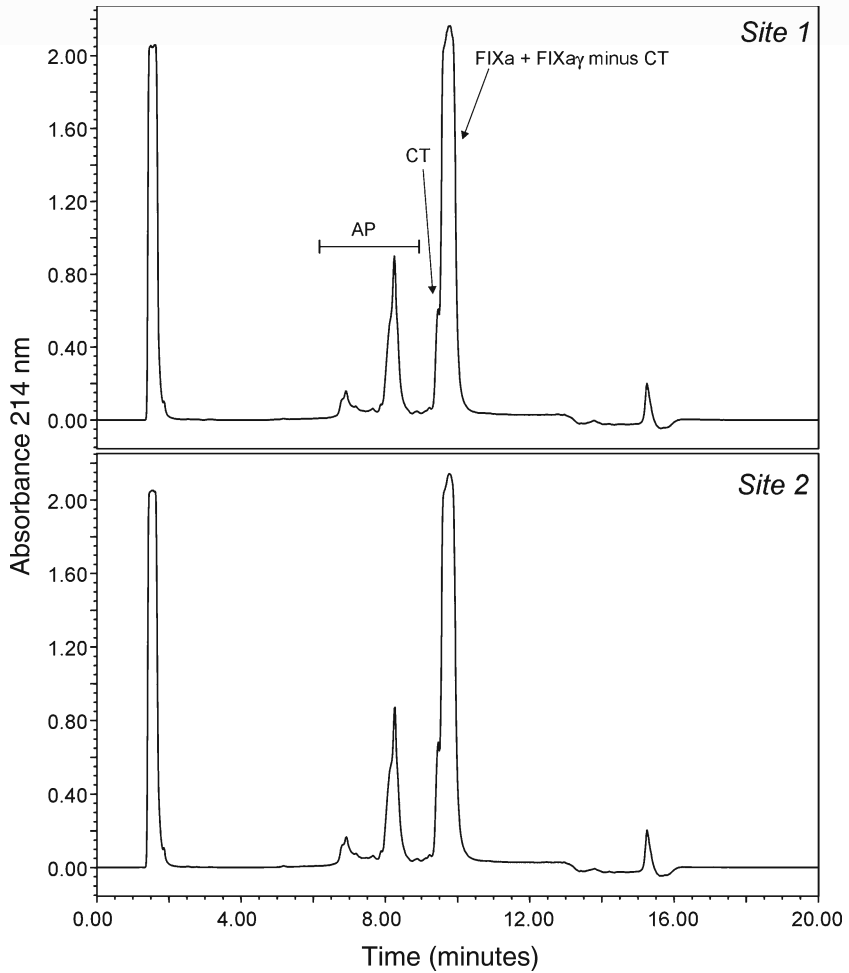


Fig. 5. RP-HPLC separation of the rFIX activation products (Notes 14 and 17).

Table 3
RP-HPLC Gradient

Time (min)	%A	%B
Initial	90	10
1	90	10
10	40	60
10.1	0	100
12	0	100
12.1	90	10
25	90	10

3.2. Chromatography of AP and FIXa (Fig. 5)

1. The activation reaction is thawed and injected (approx 250 μg , 114 μL) onto a Symmetry C4 RP-HPLC column and eluted at a flow rate of 1.0 mL/min with the gradient shown in **Table 3** (see **Fig. 5** and **Note 15**).
2. The effluent (1 mL/min) from the HPLC system is split postcolumn so that 0.100–0.133 mL/min is directed into the ESI probe of the Q-ToF-2 through 0.005 in. inner diameter (ID) peek tubing. The remaining effluent is directed to waste.

3.3. Mass Analysis of AP and rFIXa

The set up of a multiuser Q-ToF-2 for precise protein mass determinations by LC/MS is a 2-d process. Day 1 involves cleaning and configuring the instrument, setting the acquisition parameters, and assessment of instrument performance. At the end of d 1, the instrument temperature is allowed to equilibrate overnight with all voltages on and all gases flowing. Activities on d 2 involve mass calibration of the instrument, activation of temperature compensation mode, and optimization of ESI parameters for protein LC/MS. At the end of d 2, the ion source block and desolvation gas temperatures are increased for LC/MS, and the instrument temperature is allowed to equilibrate overnight. On d 3, the instrument is ready for LC/MS data acquisition.

1. Q-ToF-2 instrument preparations (d 1). All experiments utilize the standard ESI probe and the Z-spray ion source. Flush the metal capillary in the ESI probe with 50:50 acetonitrile:water (v/v), then rinse the exterior of the ESI probe tip with 100% methanol. Close the isolation valve, and remove the external sample cone, cone gas nozzle, and grounding baffle. To ensure full sensitivity, clean these parts in accordance with the manufacturer's instructions, which includes a scrubbing step with 10% (v/v) formic acid in water and sequential ultrasonication steps in 10% (v/v) formic acid in water, 50:50 methanol:acetone (v/v), and 100% 2-propanol. After this cleaning procedure, dry the external sample cone, cone gas nozzle, and grounding baffle with high-purity nitrogen gas, and reinstall all components. Reboot the host and embedded computers with the manufacturer's shutdown and start-up procedures, and reinitialize the MassLynx software before the acquisition of multiple back-to-back LC/MS runs. All ion source, quadrupole, and TOF parameters should match the predetermined historical values. Set the capillary voltage on the ESI probe to 3000 V. Maintain argon gas at a pressure of approx 6×10^{-5} mbar in the collision cell region. Adjust the nitrogen flow rates for the desolvation gas (drying gas) to 400 L/h and cone gas (curtain gas) to 50 L/h. Set the ion source block temperature to 80°C and desolvation gas temperature to 120°C. Set the in/out position of the ESI probe to the furthest distance from the sample cone to minimize ion source contamination. Infuse insulin, trypsinogen, and NaI solutions with a syringe pump through 0.005 in ID PEEK tubing at 5 $\mu\text{L}/\text{min}$ for performance testing and calibration. Use the MassLynx "sample list" interface to acquire all MS data. Optimize the left/right position of the ESI probe to between 1 and 4 mm from the sample cone for the best signal. Adjust the valve that controls the nitrogen gas for ESI nebulization (at the probe tip) for the best signal. Optimal signal is found between 1/4 and 1/2 turn from the closed position when solutions are flowing at 5 $\mu\text{L}/\text{min}$.
2. Q-ToF-2 acquisition parameters (d 1). Activate the manual pusher, which determines the duty cycle of the TOF in sampling the quadrupole ion beam. Set the pusher unit to pulse every 88 μs to enable spectral acquisition out to m/z 4000. Set the MS profile parameters

- to transport all ions between m/z 800 and 4000 through the quadrupole mass filter and hexapole collision cell with optimal sensitivity (**Note 18**). This is achieved with the following settings: dwell time at m/z 1000 is 20%; ramp time from m/z 1000–2000 is 60%; and dwell time at m/z 2000 is 20%, where the indicated percentage is the percent of total scan time. (The dwell and ramp values are indicated as “mass,” not m/z , in MassLynx.) Utilize a total scan time of 5 s for LC/MS analyses of intact glycoproteins; accumulate signal for 4.9 s, and set the interscan delay to 0.1 s for data upload (**Note 19**).
3. Q-ToF-2 performance assessments (d 1). Assess mass analyzer resolving power with the insulin solution using a cone voltage (CV) of 37 V and a collision energy (CE) of 10 eV. Evaluate the resolving powers for the multiple-charged species at m/z 956.4, 1147.5, and 1434.4. Determine the resolving power with the peak-width definition, $m/\Delta m$, where m is the m/z value of a prominent isotope peak, and Δm is the peak width (as m/z) at half height (FWHM). The resolving power calculated for each insulin species should exceed 10,000 and match historical values. Optimize the left/right ESI probe position so that the counts for the most abundant insulin ion (m/z 1147.5) are between 400 and 450 counts per 5 s scan. It is important to keep the counts for the most intense ion below 450 counts per 5 s scan to avoid signal saturation of the time-to-digital converter (TDC) and prevent an apparent loss in resolving power. Ascertain the instrument sensitivity for proteins with the trypsinogen solution using a CV of 50 V and CE of 10 eV. Optimize the left/right ESI probe position so that the counts for the multiple-charged trypsinogen species at m/z 2181 are as intense as possible, and evaluate the number of counts per scan (assuming 5 s total scan time). Accept counts above 250 (minimum); typical values range between 350 and 500 after ESI probe position and nebulizer gas optimization. Ensure that the trypsinogen signals and total ion current (TIC) trace are stable for each successive scan. After performing all cleaning, set-up, and assessment procedures, and prior to mass axis calibration, allow the instrument temperature to equilibrate overnight with all voltages on and all gases flowing.
 4. Q-ToF-2 external calibration and temperature compensation set-up (d 2). With the instrument temperature at equilibrium, ensure that the dynamic calibration (DXC) temperature compensation circuitry is not activated. The DXC temperature compensation adjusts (in real time) mass axis calibration drift during LC/MS caused by changes in instrument and room temperatures. Activate analog channel 4, which connects the TDC to the temperature sensor on the TOF mass analyzer. The temperature sensor monitors the thermal expansion and contraction of the flight tube. Infuse the trypsinogen solution, and acquire data for 5–10 min before calibration to warm up the acquisition electronics. After moving the ESI probe far to the right of the sample cone to prevent contamination with NaI, infuse the NaI calibration solution. Acquire NaI mass data with a CV of 37 V, a CE of 10 eV, and the same instrument and scan settings as described in **Subheadings 3.3.1.** and **3.3.2.** for 3 min. Keep the most intense NaI-related peaks (e.g., m/z 1971.6) below 450 counts (assuming 5 s total scan time) to avoid signal saturation of the TDC acquisition unit but above 400 counts to obtain optimal peak shapes and sensitivity for the species above m/z 3000. Combine (sum together) all NaI mass spectra acquired over the 3 min, smooth the combined spectra (three-channel window, one smooth, and Savitzky–Golay algorithm), and center the data (centroid method at the 80% level by area). Use MassLynx to derive the calibration equation with a fourth order polynomial curve fit. Activate the DXC temperature compensation circuitry immediately after finalizing the calibration parameters in the software. The DXC constant is determined by the manufacturer’s installation engineer, and it typically ranges from 70 to 100 ppm/°C. Repeat the analysis of the trypsinogen

gen solution to reoptimize the ESI probe position (move closer to the sample cone) and nebulization gas flow. Combine the trypsinogen mass spectra acquired over 1 min, smooth (14-channel window, 1 smooth, and Savitzky-Golay algorithm), and center the data (centroid method at the 80% level by area). Check the mass accuracy of the m/z 2181 peak; the measured value should be within 0.002% of the theoretical value of 2181.089. Raise the ion source block temperature to 115°C and desolvation gas temperature to 275°C. These temperatures are optimized for ESI of HPLC effluent flow rates between 0.100 and 0.150 mL/min. Allow the instrument temperature to equilibrate overnight before LC/MS analysis with all voltages on and all gases flowing.

5. LC/MS mass spectral acquisition (d 3). Prior to LC/MS, the instrument temperature must be at equilibrium. Connect the analog output signal from the 2487 UV/Vis detector to analog input 1 on the Q-ToF-2 to provide UV data complementary to the TIC trace. Connect the HPLC outlet tubing to the flow splitter, and set the HPLC to flow at 1 mL/min at initial solvent conditions (**Table 3**). All tubing is 0.005 in ID PEEK. Adjust the split flow rate from the flow splitter to the ESI probe to be between 0.100 and 0.150 mL/min; the remaining 0.85–0.9 mL/min should pass to waste or, if desired, to a fraction collector. Do not connect the flow splitter outlet tubing to the ESI probe at this time. Once the flow splitter is adjusted, maintain the flow from the HPLC at 1 mL/min; otherwise, the flow splitter will have to be readjusted. If necessary, readjust the valve controlling the nitrogen gas for ESI nebulization. The optimal signal is found between one-half and one turn from the closed position for flow rates 10 μ L/min or more for the trypsinogen solution. Upon starting a LC/MS acquisition, infuse the trypsinogen solution at 10 μ L/min from the syringe pump into the ESI probe for 2 min before connecting the flow splitter outlet tubing to the ESI probe. The effluent from the flow splitter should not be connected to the mass spectrometer until after the sample flow-through has passed through the outlet tubing. After this occurs, connect the flow splitter outlet tubing to the ESI probe, and acquire LC/MS data. After all chromatographic peaks have eluted, disconnect the flow splitter outlet tubing from the ESI probe, reconnect the syringe pump outlet tubing, and infuse the trypsinogen solution again (10 μ L/min for 2 min) prior to stopping the LC/MS data acquisition. Acquisition of trypsinogen at the beginning and end of each LC/MS data file provides instant evaluation of run-to-run instrument sensitivity and determination of mass accuracy variation and drift for all experiments (**Note 20**). LC/MS experiments can be performed for several days without much deterioration in mass accuracy and sensitivity. Day-to-day performance is tracked with freshly prepared trypsinogen solutions.
6. LC/MS spectral processing. Each chromatographic peak contains multiple mass spectra. Combine all mass spectra across a chromatographic peak using MassLynx. Ultimately, this mass spectrum is submitted to the MaxEnt-1 module for mass deconvolution. If there is visible chemical background under the mass spectral signals, execute background subtraction in MassLynx in advance of mass deconvolution (**Note 21**). The specific background subtraction parameters for activated rFIX species are shown in **Table 4**. No mass spectral smoothing is performed because it is incompatible with MaxEnt-1.
7. MaxEnt-1 mass deconvolution. Conversion of the multiple-charged data to zero-charge data provided for the straightforward identification of the C-terminal (CT), FIXa, and AP isoforms in each chromatographic peak. Use the “uniform Gaussian” damage model in MaxEnt-1. Select the input m/z range in MassLynx; choose all charge states of the protein(s) by default. For intact proteins, mass accuracy can be increased by limiting the input m/z range to the 10 to 12 most abundant charge states that have well-defined peak shapes. Measure the average peak width at half the height for an abundant, fully resolved,

Table 4
Background Subtraction and MaxEnt-1 Mass Deconvolution Parameters
for the Mass Spectra of the Activated rFIX Species

Procedure/rFIX species	AP ^a	FIXa ^b	CT ^c
Background subtraction	Not required	Required	Not required
Polynomial order		50	
Below curve %		1	
MaxEnt-1 mass deconvolution			
Input <i>m/z</i> range	1700–3000	1039–3814	1000–2200
Output mass range	9000–14,000	43,000–46,000	9000–12,000
Resolution (Dalton/channel)	1	1	1
Peak width at half height (Dalton)	1.45	0.80	1.50
Minimum intensity ratios (left%/right%)	33/33	66/66	33/33

^aSee Fig. 7.

^bSee Fig. 8.

^cSee Table 7.

multiple-charged ion that is centrally located in the selected input *m/z* range. Use the default 33%/33% as the “left/right minimum intensity ratio.” This limits the consideration of adjacent peaks in the same series based on their relative heights; 33% implies that the adjacent peak to the left or right of any particular peak in the multiple-charged envelope is at least 33% as intense. Produce a “survey” zero-charge mass spectrum with MaxEnt-1 at a resolution of 10.0 Daltons per channel and a wide output mass range between 5000 and 160,000 Daltons to establish the presence of all protein components in the chromatographic peak. Stop MaxEnt-1 after several iterations. If necessary, the harmonic artifacts, which are peaks at half and twice the mass of each protein component, may be reduced by raising the left/right minimum intensity ratio to approx 66%/66%. Repeat the MaxEnt-1 survey mass deconvolution until optimized. Use 33%/33% for polypeptides of lower mass, particularly if they are glycosylated, because the abundances of the charged ions are not likely to have the ideal protein-charge envelope with gradual intensity changes. Produce the final zero-charge mass spectrum with MaxEnt-1 for each specific protein component found in the survey at a resolution of 1.0 Daltons per channel with an appropriate narrower output mass range. Set the output mass range to include approx 3000 to 5000 Daltons on either side of the approximate survey-derived mass for each large protein and to include approx 300 to 500 Daltons on either side of the approximate survey-derived mass for smaller protein subunits. The specific MaxEnt-1 parameters used for activated FIX species are shown in **Table 4**. Unless otherwise noted, use the same parameters determined from the survey deconvolution. Allow MaxEnt-1 to iterate to convergence. Center the resulting peaks with the centroid method at the 80% level to obtain masses accurate to the tenths place. Ensure that the MaxEnt-1 deconvolution is a precise representation of the data; the isoform peaks and their ratios in the zero-charge mass spectrum should match or closely resemble an average of those in the multiple-charged data series. Higher mass accuracy and more exact zero-charge profiles are obtained by the judicious use or absence (if possible) of background subtraction (**Note 21**). It is important to use the identical input *m/z* range and MaxEnt-1 deconvolution

parameters (i.e., output mass range, resolution, peak width, minimum intensity ratio, and 80% centroid) for each protein species when processing LC/MS data for comparability.

8. LC/MS analysis for rFIX comparability. Optimize the CV and CE parameters for LC/MS of the AP and FIXa chromatographic species to ensure full retention of labile PTMs and, in severe cases, to prevent peptide-bond fragmentation (**Note 22**). Test a range of CV values at a CE of 5 and 10 eV with repeat injections of activated rFIX. Choose the optimal CV and CE parameter combinations for each FIXa species based on the balance between signal strength and minimal analysis artifacts (e.g., the loss of N-linked glycan antennal arms or sialic acid residues). For AP, a reduced CV of 20 V and CE of 5 eV were required to minimize N-linked glycan trimming caused by energetic collisions (*see Fig. 6 and Note 23*). For FIXa, a CV of 37 V and a CE of 10 eV was required to achieve strong signal strengths (**Fig. 6**). To assess comparability of AP, all batches of activated rFIX are injected for LC/MS analysis with a CV of 20 V and CE of 5 eV. For FIXa and CT, all batches of activated rFIX are reinjected for LC/MS analysis with a CV of 37 V and CE of 10 eV. Perform all LC/MS experiments side-by-side in 1 d, if possible, to minimize retention time and mass drift.

3.4. Evaluation of Mass Data for Comparability

1. Overall comparability assessment. Our comparability programs include product release testing, animal pharmacokinetic studies, forced protein degradation, biophysical analyses, nonroutine HPLC assays, and top-down characterization with MS. A final review of comparability includes the analysis of results from all of these tests.
2. Comparability acceptance criteria for top-down characterization. Because MS analysis is not routinely performed as part of product release testing, an extensive batch history on which to base quantitative and statistical acceptance criteria was not available. However, criteria were set according to the known instrument specifications and qualitative isoform content. Typically, our first approach is to take a bird's-eye view of the mass envelopes generated for each protein sample, then compare sections of the overall mass profiles. If there are new species detected in the new material, these species and their potential impact are evaluated further in the context of the results from other testing. Frequently, a change in the proportion of existing isoforms is acceptable if it is found to be inconsequential in other tests, such as bioactivity or pharmacokinetics. Large differences in the levels of existing species would require further investigation into the manufacturing process parameters. Specific acceptance criteria for this study were:
 - The RP-HPLC chromatograms of rFIX activation products generated from site 2 rFIX must be comparable to those generated from site 1 rFIX with no new species detected.
 - No new species should be detected in the mass spectra of site 2 material when compared with those of site 1 material.
 - Masses of the isoforms found in rFIX from manufacturing site 2 must agree with the masses of the analogous isoforms found in rFIX from site 1 within a specified mass (mass = x Daltons, where $x = 1$ for AP; 2 for FIXa, and 1 for the CT peptides).
 - The measured masses must agree with the theoretical masses according to instrument specifications.
 - Analogous mass spectra of activation products from sites 1 and 2 must contain similar peak distributions.
3. Isoform selection. There is a potential for "data overload" because a large number of isoform signals are generated from top-down MS analysis of complex glycoproteins. To assess comparability, we selected important and representative isoforms across the

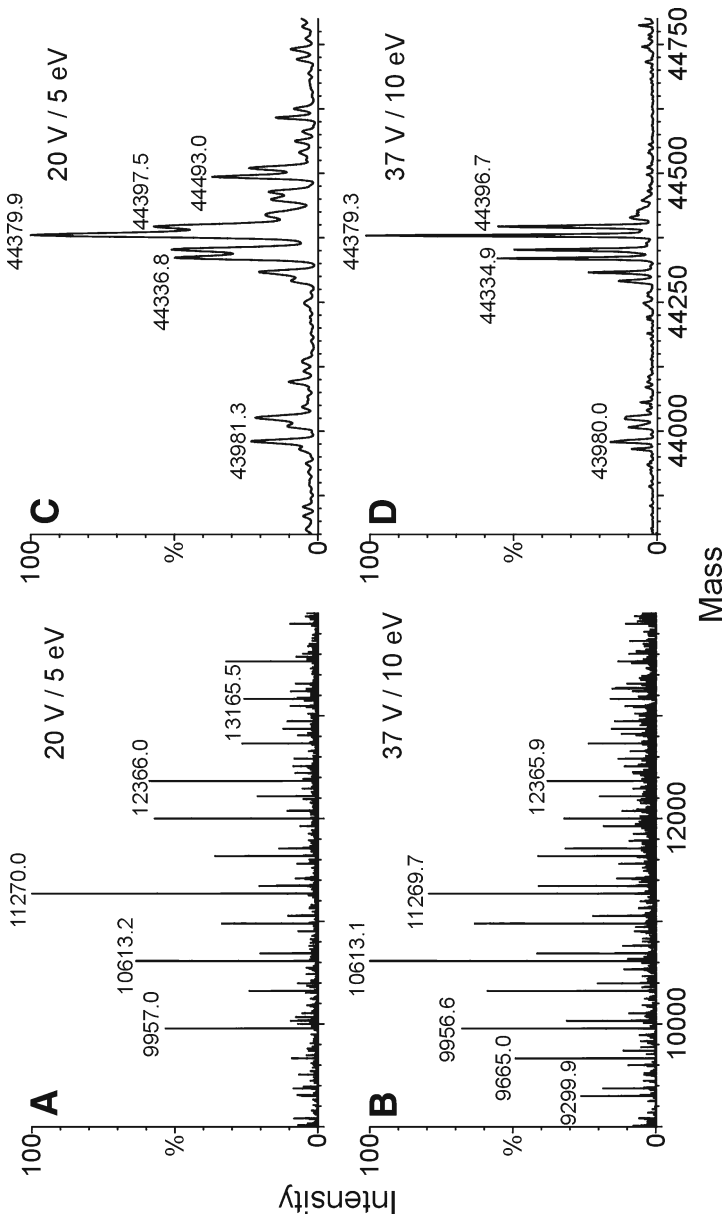


Fig. 6. Optimization and effect of the CV and CE mass spectrometric parameters in the RP-HPLC/ESI-QTOF MS analysis of AP and FIXa (Notes 22 and 23). Shown are the resultant zero-charge mass spectra from (A) AP region acquired with a CV of 20 V and a CE of 5 eV; (B) AP region acquired with a CV of 37 V and a CE of 10 eV; (C) FIXa region acquired with a CV of 20 V and a CE of 5 eV; and (D) FIXa region acquired with a CV of 37 V and a CE of 10 eV. See Tables 5 and 6 for information relating to the identity of the peaks in the mass spectra. For the AP and FIXa regions, the zero-charge mass spectra were deconvoluted with the same MaxEnt-1 parameters as reported in Figs. 7 and 8, respectively.

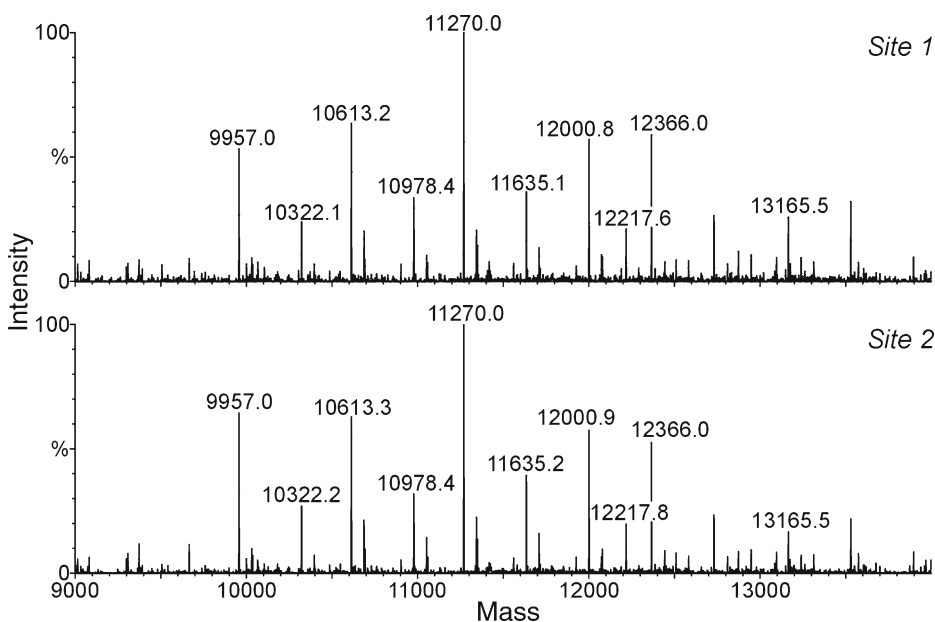


Fig. 7. Analysis of AP from activated rFIX by RP-HPLC/ESI-QTOF MS. Zero-charge mass spectra of the AP region (**Fig. 5**) that compares rFIX material from sites 1 and 2. See **Table 5** for information relating to the identity of the peaks in the mass spectra. The AP region was summed and then deconvoluted to obtain a zero-charge mass spectrum.

mass range. Certainly, specific comparisons of selected mass signals in the absence of an overall evaluation of the mass spectra and in the absence of a screen for previously uncharacterized new species can give misleading results. Identical mass signals can be pulled out of spectra that are otherwise very different, and this should be avoided.

4. rFIX comparability results. The RP-HPLC chromatographic profiles of hFIXa-activated rFIX batches manufactured at site 2 were comparable to those of hFIXa-activated rFIX batches manufactured at site 1 (**Fig. 5**). All minor and major peaks were present in the chromatograms, and no new peaks were detected in chromatograms of activated rFIX from site 2.

Zero-charge mass spectra from the AP region of each respective RP-HPLC chromatogram (**Fig. 5**) are presented in **Fig. 7**. The peak distribution in the mass spectrum of the AP from site 2 rFIX batches was similar to that of the AP from site 1 rFIX, and no new peaks were observed. The major peaks in each mass spectrum represent the microheterogeneity found in the *N*-glycans at the two fully occupied *N*-glycosylation sites. The theoretical and observed masses for selected isoforms are presented in **Table 5**. The observed differences in AP glycoform masses between the site 2 and site 1 batches for analogous species were 1 Dalton or less. Additionally, the measured masses for each AP glycoform when compared to their respective theoretical masses showed mass

Table 5
Masses of AP Glycoforms Observed via RP-HPLC/ESI-QTOF MS

AP identity ^a	Theoretical mass (Dalton) ^b	Site 1 (Dalton) ^c	Site 2 (Dalton)
3A/3SA/1F; 3A/3SA/1F	9956.7	9957.0	9957.0
3A/3SA/1F; 3A/3SA/1F + 1 R	10322.0	10322.1	10322.2
3A/3SA/1F; 4A/4SA/1F	10613.3	10613.2	10613.3
3A/3SA/1F; 4A/4SA/1F + 1 R	10978.6	10978.4	10978.4
4A/4SA/1F; 4A/4SA/1F	11269.9	11270.0	11270.0
4A/4SA/1F; 4A/4SA/1F + 1 R	11635.2	11635.1	11635.2
4A/4SA/1F; 4A/4SA/1F + 2 R	12000.6	12000.8	12000.9
4A/4SA/1F; 4A/4SA/1F + 1 O ^d	12217.7	12217.6	12217.8
4A/4SA/1F; 4A/4SA/1F + 3 R	12365.9	12366.0	12366.0
4A/4SA/1F; 4A/4SA/1F + 2 O ^d	13165.6	13165.5	13165.5

^aAbbreviations: A, antennae; SA, sialic acid (*N*-acetyl neuraminic acid); F, fucose; R, poly-*N*-acetyl lactosamine (GalGlcNAc); O, O-linked glycan.

^bTheoretical masses were calculated for the selected AP glycoforms using PAWS and are reported as average masses.

^cObserved masses were calculated using MaxEnt-1 as described in the text.

^d*O*-glycosylation is found at Thr¹⁵⁹, Thr¹⁶⁹, and Thr¹⁷² with partial occupancy and is composed of Gal-GalNAc (i.e., core 1) with one and two sialic acids (*N*-acetyl neuraminic acid). For the AP region denoted in **Fig. 5**, the minor peak (averages 20% from integration of AP region) represents AP with both *N*- and *O*-glycosylation, whereas the major peak represents AP with *N*-glycosylation only.

differences of 0.005% or less, consistent with the specifications of the Q-Tof-2 mass spectrometer.

Zero-charge mass spectra from the FIXa region of each RP-HPLC chromatogram (**Fig. 5**) are presented in **Fig. 8**. The peak distribution in the mass spectra of the FIXa from site 2 rFIX batches were similar to those of the FIXa from site 1 rFIX, and no new peaks were observed. The major peaks in each mass spectrum represent FIXa with 11 and 12 4-carboxyglutamic acid (Gla) residues and partial β -hydroxylation at Asp⁶⁴. The theoretical and observed masses for selected isoforms are presented in **Table 6**. The observed mass differences between site 1 and 2 batches for analogous FIXa isoforms were 2 Daltons or less. Additionally, the measured masses for each FIXa isoform in comparison to their respective theoretical masses showed mass differences of 0.005% or less, similar to the specifications of the Q-Tof-2 mass spectrometer.

The peak distribution in the mass spectra of the CT peptides derived from site 2 rFIX were like those of the CT peptides derived from site 1 rFIX (data not shown); no new peaks were observed. The theoretical and observed masses of the CT peptides from representative rFIX batches from sites 1 and 2 are presented in **Table 7**. The mass difference between the masses of CT peptides and their respective theoretical masses was less than 1 Dalton. The measured masses for each CT peptide vs their respective theoretical masses showed mass differences of 0.005% or less, which is consistent with the measurements of the Q-Tof-2 mass spectrometer.

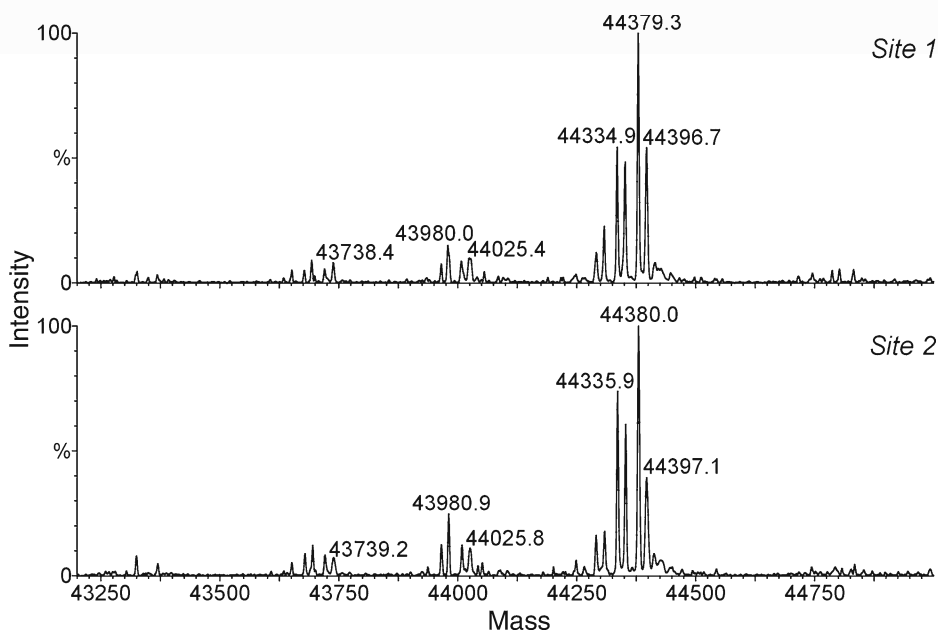


Fig. 8. Analysis of FIXa by RP-HPLC/ESI-QTOF MS. Zero-charge mass spectral profiles of the FIXa peak (**Fig. 5**) comparing rFIX material from sites 1 and 2. See **Table 6** for information relating to the identity of the mass spectral peaks. The FIXa region was summed, background subtracted, and then deconvoluted to obtain a zero-charge mass spectrum.

Table 6
Masses of FIXa Isoforms Observed via RP-HPLC/ESI-QTOF MS

FIXa identity ^a	Theoretical mass (Dalton) ^b	Site 1 (Dalton) ^c	Site 2 (Dalton)
L ¹⁻¹⁴⁵ , H ¹⁸¹⁻⁴¹⁵ , 12 Gla, β-OH ^{d,e,g}	43739.5	43738.4	43739.2
L ¹⁻¹⁴² , H ¹⁸¹⁻⁴¹⁵ , 11 Gla, β-OH ^{d,f,g}	43981.7	43980.0	43980.9
L ¹⁻¹⁴² , H ¹⁸¹⁻⁴¹⁵ , 12 Gla, β-OH ^{d,f,g}	44025.7	44025.4	44025.8
L ¹⁻¹⁴⁵ , H ¹⁸¹⁻⁴¹⁵ , 11 Gla ^{d,f}	44336.1	44334.9	44335.9
L ¹⁻¹⁴⁵ , H ¹⁸¹⁻⁴¹⁵ , 12 Gla ^{d,f}	44380.1	44379.3	44380.0
L ¹⁻¹⁴⁵ , H ¹⁸¹⁻⁴¹⁵ , 12 Gla, β-OH ^{d,f,g}	44396.1	44396.7	44397.1

^aAbbreviations: L, light chain; H, heavy chain; Gla, 4-carboxyglutamic acid; β-OH, β-hydroxylation.

^bTheoretical masses were calculated for the selected isoforms with disulfide bonds included using PAWS and are reported as average masses.

^cObserved masses were calculated using MaxEnt-1 as described in the text.

^dThe O-linked glycan at Ser⁵³ is Glc(Xyl)₂.

^eThe O-linked glycan at Ser⁶¹ is Fuc.

^fThe O-linked glycan at Ser⁶¹ is Fuc-GlcNAc-Gal-NeuAc.

^gβ-hydroxylation is present at Asp⁶⁴.

Table 7
Masses of CT Peptides Observed via RP-HPLC/ESI-QTOF MS

CT peptide identity	Theoretical mass (Dalton) ^a	Site 1 (Dalton) ^b	Site 2 (Dalton)
Val ³²⁸ -Thr ⁴¹⁵	9770.1	9769.9	9770.4
Val ³²⁸ -Thr ^{415c}	9788.1	9787.5	9787.7
Ser ³¹⁹ -Thr ⁴¹⁵	10814.4	10814.4	10814.7

^aTheoretical masses were calculated for the CT peptides with disulfide bonds included using PAWS and are reported as average masses.

^bObserved masses were calculated using MaxEnt-1 as described in the text.

^cThe peptide bond, Arg³³⁸-Ser³³⁹ for this CT peptide is presumed to be hydrolyzed based on the observed 18-Dalton mass increase. However, a disulfide bond between Cys³³⁶ and Cys³⁵⁰ keeps this CT peptide intact as Val³²⁸-Thr⁴¹⁵.

Based on top-down characterization data, the primary structure and PTMs of rFIX manufactured at sites 1 and 2 were judged comparable. The mass spectra of analogous rFIX fragments were similar to each other with no new species detected, and all masses met the acceptance criteria. The LC/MS data agree with the results of forced degradation testing, biophysical analyses, bioactivity assays, and the nonclinical pharmacokinetic study. These results indicate that there is no significant impact on rFIX structure and function when this protein pharmaceutical is manufactured at the new manufacturing facility.

4. Notes

1. Recently, top-down characterization and peptide mapping with MS have been employed in parallel experiments to comprehensively characterize complex mixtures of proteins (4,5). Peptide mapping and sequencing with MS was shown to be more effective than the top-down approach in providing identifications for the mixture of unknown proteins owing to the increased dynamic range of peptide mapping-based techniques. However, the sequence information from the top-down techniques helped to confirm the assignments of the major protein species detected by peptide mapping. Accurate mass determinations from the top-down approach revealed the PTMs in the mass profile and allowed full-length species to be distinguished from N or C-terminally truncated isoforms. Both reports emphasize that the precise mass measurements from the top-down approach provided additional attributes that strengthened the protein identifications and helped define the intact protein structures more thoroughly.
2. Detailed structural information for intact proteins was initially accomplished with Fourier transform-ion cyclotron resonance (FT-ICR) MS/MS (7). To date, extensive contiguous amino acid sequence and PTM information for several intact proteins up to 45 kDa (8-14) have been demonstrated using FT-ICR with activated ion—electron capture dissociation (AI-ECD)—an exclusive technique for FT-ICR instruments (15). The FT-ICR mass analyzer features very high-resolving powers that exceed 10⁵ and mass accuracies less than 0.001% for peptides and less than 0.005% for proteins (7). The AI-ECD technique induces cleavages at NH- α CHR bonds along the protein backbone in an indiscriminate manner and breaks the intramolecular noncovalent bonds between two fragment

- ions to maximize the number of cleaved sites (**15,16**). For proteins larger than 70 kDa, limited proteolysis was performed to generate smaller polypeptide fragments, ranging in size from 5 to 48 kDa, which were more amenable to gas-phase sequencing with FT-ICR MS/MS (**17**).
3. There are several reports specifying the use of a modified quadrupole ion-trap mass analyzer that is capable of systematically converting the unresolved multiple-charged protein fragment ions to all single-charged ions by gas-phase ion/ion proton transfer reactions (**18**). This provides straightforward spectral interpretation in the absence of high resolution, resulting from the uniformity of the charge states (**19**). Abundant sequence information (**20,21**) and site-specific PTM characterization (**22,23**) for whole proteins up to 26 kDa have been demonstrated with this customized system.
 4. In the analysis of large polypeptides, only the ions between m/z 800 and 4000 are allowed to pass through the first-quadrupole mass filter, whereas ions less than m/z 800 are destabilized and filtered out using the MS profile settings.
 5. QTOF mass analyzers have increased sensitivity when compared with other instrument types because of their inherently high-duty cycles. Recent improvements in the orthogonal acceleration ion optics for Q-ToF-2 mass analyzers combined with improved ion beam spatial and kinetic energy distributions have increased resolving powers to above 10,000 ($m/\Delta m$) for ESI-generated ions without compromising sensitivity. In the “pusher” unit of a Q-ToF-2, a voltage gradient is applied across the ion beam through separate “pusher” and “puller” pulses in an attempt to minimize turnaround time and improve resolution.
 6. Fragmentation of intact proteins on a QTOF with CID has resulted in limited stretches of amino acid sequence information, which is not adequate for thorough characterization. However, these “sequence tags” were extremely valuable for the identification of unknown proteins and their attributes when combined with the accurate mass measurement (**26,27**).
 7. Depending on the glycoprotein, these mass signals will coalesce fully or partially when analyzed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS.
 8. Although top-down characterization provides the means for rapid analysis of new species and changes in the isoform profile, a comparability assessment is supported only after the results from all routine testing and nonroutine studies are obtained and evaluated together.
 9. BDDrFVIII, although considerably less complex than full-length rFVIII, still poses an analytical challenge because it has a molecular weight of 170 kDa (80-kDa light chain and 90-kDa heavy chain) and modifications, including tyrosine sulfation, *N*- and *O*-linked glycans, and *N*- and *C*-terminal processing (**31**). The intact light chain was analyzed by online RP-HPLC/ESI-QTOF MS, but the intact heavy chain was too complex and exceeded the capabilities of the QTOF. Proteolytic cleavage with thrombin into 6, 43, 50, and 73 kDa domains followed by online RP-HPLC/ESI-QTOF MS was required for analysis of the entire molecule and permitted a comparability evaluation of materials produced by two manufacturing processes (**32**).
 10. For certain mAbs, minor *N*-terminal processing and low-level modifications were difficult to identify with the initial peptide-map characterization because of small or coeluting peaks in a complex profile. However, these modifications were readily apparent in the mass profile from the complementary top-down approach. mAbs, before and after disulfide bond reduction, were analyzed off-line by nanoelectrospray ionization interfaced to the QTOF mass analyzer. This allowed valid mass determination for the intact mAb tetramer and the light- and heavy-chain subunits, respectively. Modifications detected on the intact mAb tetramer were easily localized on the light or heavy chain. The intact, as

- well as the reduced and alkylated, mAbs were also analyzed online by RP-HPLC/ESI-QTOF MS to obtain complementary mass information from chromatographically resolved species. The accurate mass data verified structural composition, whereas the isoform profiles revealed the *N*-glycosylation patterns in the Fc domain of the heavy chain, C- and N-terminal heterogeneity of both the light and heavy chain, and other PTMs (30,33).
11. rFIX manufactured at the two manufacturing facilities was also compared by product release testing, forced degradation studies, biophysical analyses, bioactivity testing, and nonclinical pharmacokinetics.
 12. Site-specific assignments for the glycans on the AP were previously determined by subdigestion of the RP-HPLC-fractionated, de-*N*-glycosylated AP with endoproteinase Glu-C, followed by N-terminal sequencing and MALDI-TOF MS postsource decay sequencing.
 13. The low-abundance AP species containing the *O*-glycan at Thr¹⁷⁹ was not observed after FIXa activation. This is presumably a result of steric hindrance of the Arg¹⁸⁰-Val¹⁸¹ cleavage site. This form of the AP was found after Lys-C cleavage of rFIX.
 14. Activation of rFIX with hFIXa predominantly generates AP (Ala¹⁴⁶-Arg¹⁸⁰) and FIXa by cleavage of the Arg¹⁴⁵-Ala¹⁴⁶ and Arg¹⁸⁰-Val¹⁸¹ bonds (Fig. 4). FIXa consists of L chain (Tyr¹-Arg¹⁴⁵) linked to H chain (Val¹⁸¹-Thr⁴¹⁵) by one disulfide bond. The heterogeneous AP elutes across two peaks, and the FIXa elutes in the major peak (Fig. 5).
 15. Along with the major cleavages, there are low-level cleavages observed in the HC at the Arg³¹⁸-Ser³¹⁹, Arg³²⁷-Val³²⁸, and Arg³³⁸-Ser³³⁹ peptide bonds, generating a species referred to as FIX γ . FIX γ consists of the light chain linked to HC γ by one disulfide bond. HC γ is Val¹⁸¹-Arg³¹⁸, Val¹⁸¹-Arg³²⁷, or Val¹⁸¹-Arg³³⁸. The CT peptides remain bound noncovalently in nondenaturing conditions. Under the denaturing conditions of RP-HPLC, the FIX γ /CT peptide complex is dissociated. The CT peptides, Ser³¹⁹-Thr⁴¹⁵, Val³²⁸-Thr⁴¹⁵, and Ser³³⁹-Thr⁴¹⁵, elute as a shoulder on the main peak. The N-terminal portion of FIX γ (FIX γ minus the CT peptides) coelutes with FIXa.
 16. To control for any potential experimental variables that might impact the results, all activation reactions should be carried out with the same lot of hFIXa side-by-side in the same experiment.
 17. The early eluting AP peak contains N- and O-linked glycans; the later eluting AP peak contains N-linked glycans with low levels of O-linked glycans. In this work, the mass spectra were summed across the entire AP region of the chromatogram. Representative chromatograms of the activation products of one batch from each site are shown in Fig. 5.
 18. The MS profile provides control of the RF voltages on the multipole devices for transporting a specific *m/z* range of ions to the TOF analyzer. The *m/z* range between 800 and 4000 is the typical region where the multiple-charged ions for intact proteins are detected on the Q-Tof-2. There are sharp (0.75) and gradual (2) cut-off values associated with the low- and high-mass values, respectively, which are important when setting up the MS profile. When the low-mass value is set to 1000, all ions above *m/z* 750 are transmitted, and all ions below *m/z* 750 are filtered out. For the high-mass value, 2000 is required for optimal sensitivity to *m/z* 4000. By not transmitting all of the ions below *m/z* 750, the sensitivity is further maximized for the higher *m/z* protein ions.
 19. More specifically, one mass spectrum is uploaded to the computer hard drive every 4.9 s. This mass spectrum represents the signal accumulation in the TDC from the TOF across *m/z* 800 to 4000 for all orthogonally accelerated ion packets during this period. Longer total scan times of 5 s minimize the repetitive 0.1-s interscan delay, which helps maxi-

mize the signal. Faster total scan times of 2 s (1.9-s signal accumulation and 0.1-s interscan delay) are employed if peptides or both peptides and proteins are anticipated in an LC/MS experiment, because the chromatographic peak widths for peptides are typically more narrow than those of proteins. The identical total scan time that is utilized in LC/MS is also used for instrument calibration. Instrument testing with trypsinogen and insulin is always performed with total scan times of 5 s.

20. Embedded trypsinogen masses in an LC/MS data file also allow the use of the MassLynx “lock mass” feature for correcting mass accuracy drift. Do not use NaI at high-ion source block temperatures, because it will quickly contaminate the RF hexapole and reduce sensitivity.
21. The chemical background is usually minimal in LC/MS, and visible chemical noise may represent either incomplete protein desolvation, small-molecule and metal-ion adducts of the protein, overlapping charge species, or unresolved protein isoforms. Use a “polynomial order” of 0 or 1 with a “below curve %” (i.e., the percentage of data points below the zero line) of 35 or 40 to remove level and constant background from the mass spectrum. Use higher polynomial order values, such as 10, 20, 50, 75, and 99 on a mass spectrum with more complex patterns of chemical noise. Choose the “make graph of fitted polynomial” option to view the straight line or curve fit overlaid with the mass spectral data. The straight line or curve should attempt to mirror the mass spectral data and noise pattern. It should be positioned below the mass spectral data trace but not exceed the root-mean-square of the noise.
22. CV is analyte-dependent, and it represents the ion acceleration potential between the ion source block and extractor cone. The region between the ion source block and the extractor cone in the Q-ToF-2 is turbulent, as the two apertures bridge pressures between approx 760 and approx 10^{-5} to 10^{-4} Torr, respectively. High-CV potentials are usually required for sensitive protein analysis, but these can induce in-source artifacts for proteins with labile PTMs and Asp-Pro peptide bonds. CE represents the axial translational energies of ions that enter the collision cell. The collision cell region contains argon gas at approx 6×10^{-5} Torr (see Fig. 1). For high resolution and accurate mass analysis, CEs of 5 to 10 eV are utilized to collisionally cool ions against argon. At a CE of 10 eV, fragmentation artifacts for sialylated glycopeptides and proteins with labile PTMs can occur from increased collisional activation events.
23. For both CV and CE, careful selection of both parameters is required to ensure mass spectrometric artifacts from energetic collisions are kept to a minimum. The AP glycoform profile in Fig. 6, panel B is skewed as a result of the increased energies of collision caused by the higher voltages employed. New peaks representing the loss of sialic acid residues, entire antennal arms, and O-linked glycans were observed. A CV of 20 V and a CE of 5 eV (Fig. 6, panel A) was used in acquiring AP spectra, because the sensitivity was maintained and the in-source and collision cell fragmentation events were minimized. The glycoform profile at a CV of 20 V and a CE of 5 eV suited well with the N-glycan profile data obtained with high-performance anion-exchange chromatography-PED. Unfortunately, a CV of 20 V and CE of 5 eV was not ideal for the higher molecular weight FIXa species (Fig. 6, panel C); a CV of 37 V and CE of 10 eV is required for proper focusing and optimal detection as shown by the increased resolving power for FIXa in Fig. 6, panel D. Because the ion energetics were different for AP and FIXa, each batch of activated rFIX was analyzed twice by RP-HPLC/ESI-QTOF MS using the parameter sets, 20 V CV and 5 eV CE and 37 V CV and 10 eV CE, for the AP and FIXa, respectively.

Acknowledgments

The authors sincerely thank Drs. Marta Czupryn, Mark Hardy, Lisa Marzilli, and Hubert Scoble at Wyeth BioPharma for their support and critical review of the manuscript. The authors are grateful for the technical support that our colleagues at Waters and Micromass MS Technologies have provided over the years.

References

1. Food and Drug Administration. (2003) Guidance for industry, comparability protocols-protein drug products and biological products-chemistry, manufacturing, and controls information, draft guidance. Available at www.fda.gov/cber/guidelines.htm.
2. ICH. (2003) ICH Q5E, Comparability of biotechnological/biological products, note for guidance on biotechnological/biological products subject to changes in their manufacturing process (CPMP/ICH/5721/03). Available at www.hc-sc.gc.ca/hpfb-dgpsa/tpd-dpt/q5e_step2_notice_e.html.
3. Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., and McLafferty, F. W. (1999) Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J. Am. Chem. Soc.* **121**, 806–812.
4. VerBerkmoes, N. C., Bundy, J. L., Hauser, L., Asano, K. G., Razumovskaya, J., Larimer, F., et al. (2002) Integrating “top-down” and “bottom-up” mass spectrometric approaches for proteomic analysis of *Shewanella oneidensis*. *J. Proteome Res.* **1**, 239–252.
5. Nemeth-Cawley, J. F., Tangarone, B. S., and Rouse, J. C. (2003) “Top Down” characterization is a complementary technique to peptide sequencing for identifying protein species in complex mixtures. *J. Proteome Res.* **2**, 495–505.
6. Balogh, M. P. (2004) Debating resolution and mass accuracy. *LCGC North America* **22**, 118–130.
7. McLafferty, F. W. (2001) Tandem mass spectrometric analysis of complex biological mixtures. *Int. J. Mass Spectrom.* **212**, 81–87.
8. Ge, Y., Lawhorn, B. G., El Nagggar, M., Strauss, E., Park, J. H., Begley, T. P., and McLafferty, F. W. (2002) Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* **124**, 672–678.
9. Sze, S. K., Ge, Y., Oh, H., and McLafferty, F. W. (2002) Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. *Proc. Natl. Acad. Sci. USA* **99**, 1774–1779.
10. Sze, S. K., Ge, Y., Oh, H., and McLafferty, F. W. (2003) Plasma electron capture dissociation for the characterization of large proteins by top down mass spectrometry. *Anal. Chem.* **75**, 1599–1603.
11. Shi, S. D., Hemling, M. E., Carr, S. A., Horn, D. M., Lindh, I., and McLafferty, F. W. (2001) Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. *Anal. Chem.* **73**, 19–22.
12. Ge, Y., El Nagggar, M., Sze, S. K., Oh, H. B., Begley, T. P., McLafferty, F. W., et al. (2003) Top down characterization of secreted proteins from *Mycobacterium tuberculosis* by electron capture dissociation mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **14**, 253–261.
13. Ge, Y., Lawhorn, B. G., El Nagggar, M., Sze, S. K., Begley, T. P., and McLafferty, F. W. (2003) Detection of four oxidation sites in viral prolyl-4-hydroxylase by top-down mass spectrometry. *Protein Sci.* **12**, 2320–2326.
14. Pesavento, J. J., Kim, Y.-B., Taylor, G. K., and Kelleher, N. L. (2004) Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. *J. Am. Chem. Soc.* **126**, 3386–3387.

15. Horn, D. M., Ge, Y., and McLafferty, F. W. (2000) Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal. Chem.* **72**, 4778–4784.
16. Zubarev, R. A. (2003) Reactions of polypeptide ions with electrons in the gas phase. *Mass Spectrom. Rev.* **22**, 57–77.
17. Forbes, A. J., Mazur, M. T., Patel, H. M., Walsh, C. T., and Kelleher, N. L. (2001) Toward efficient analysis of >70 kDa proteins with 100% sequence coverage. *Proteomics* **1**, 927–933.
18. Reid, G. E. and McLuckey, S. A. (2002) “Top down” protein characterization via tandem mass spectrometry. *J. Mass Spectrom.* **37**, 663–675.
19. Stephenson, J. L., McLuckey, S. A., Reid, G. E., Wells, J. M., and Bundy, J. L. (2002) Ion/ion chemistry as a top-down approach for protein analysis. *Curr. Opin. Biotechnol.* **13**, 57–64.
20. Hogan, J. M. and McLuckey, S. A. (2003) Charge state dependent collision-induced dissociation of native and reduced porcine elastase. *J. Mass Spectrom.* **38**, 245–256.
21. Amunugama, R., Hogan, J. M., Newton, K. A., and McLuckey, S. A. (2004) Whole protein dissociation in a quadrupole ion trap: identification of an a priori unknown modified protein. *Anal. Chem.* **76**, 720–727.
22. Reid, G. E., Stephenson, J. L., Jr., and McLuckey, S. A. (2002) Tandem mass spectrometry of ribonuclease A and B: *N*-linked glycosylation site analysis of whole protein ions. *Anal. Chem.* **74**, 577–583.
23. Hogan, J. M., Pitteri, S. J., and McLuckey, S. A. (2003) Phosphorylation site identification via ion trap tandem mass spectrometry of whole protein and peptide ions: bovine alpha-crystallin A chain. *Anal. Chem.* **75**, 6509–6516.
24. Blackburn, R. K. and Moseley, M. A., III. (1999) Quadrupole time-of-flight mass spectrometry: a powerful new tool for protein identification and characterization. *Am. Pharm. Rev.* **2**, 49–59.
25. Morris, H. R., Paxton, T., Dell, A., Langhorne, J., Berg, M., Bordoli, R. S., et al. (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom.* **10**, 889–896.
26. Nemeth-Cawley, J. F. and Rouse, J. C. (2002) Identification and sequencing analysis of intact proteins via collision-induced dissociation and quadrupole time-of-flight mass spectrometry. *J. Mass Spectrom.* **37**, 270–282.
27. Thevis, M., Ogorzalek Loo, R. R., and Loo, J. A. (2003) Mass spectrometric characterization of transferrins and their fragments derived by reduction of disulfide bonds. *J. Am. Soc. Mass Spectrom.* **14**, 635–647.
28. Ferrige, A. G., Seddon, M. J., Green, B. N., Jarvis, S. A., and Skilling, J. (1992) Disentangling electrospray spectra with maximum entropy. *Rapid Comm. Mass Spectrom.* **6**, 707–711.
29. Green, B. N., Hutton, T., and Vinogradov, S. N. Analysis of complex protein and glycoprotein mixtures by electrospray ionization mass spectrometry with maximum entropy processing, in *Protein and Peptide Analysis by Mass Spectrometry* (Chapman, J. R., ed.), Humana, Totowa, NJ, 1996, pp. 279–294.
30. Rouse, J. C., Marzilli, L. A., Johnson, K. A., McClellan, J. E., and Czupryn, M. J. “Top down” glycoprotein characterization by high resolution mass spectrometry and its application to biopharmaceutical development. *Well Characterized Biopharmaceuticals 2004: 8th Symposium on the Interface of Regulatory and Analytical Sciences for Biotechnology Health Products*, Washington DC, 2004.

31. Sandberg, H., Almstedt, A., Brandt, J., Castro, V. M., Gray, E., Holmquist, L., et al. (2001) Structural and functional characterization of B-domain deleted recombinant factor VIII. *Semin. Hematol.* **38(2 Suppl. 4)**, 4–12.
32. Czupryn, M. Current analytical methodologies: application to comparability assessment of recombinant clotting factors. PDA/IABS Conference, State of the art analytical methods for the characterisation of biological therapeutic products and assessment of comparability, Bethesda, MD, 2003.
33. Mehndiratta, P., Grunder, B. C., Marzilli, L. A., McClellan, J. E., Tangarone, B. S., Porter, T. J., and Rouse, J. C. LC/MS characterization of recombinant monoclonal antibodies at the pre-development stage. Sixth International Symposium on Mass Spectrometry in the Health and Life Sciences: Molecular and Cellular Proteomics, San Francisco, CA, 2003.
34. Bond, M., Jankowski, M., Patel, H., Karnik, S., Strang, A., Xu, B., et al. (1998) Biochemical characterization of recombinant factor IX. *Semin. Hematol.* **35(2 Suppl. 2)**, 11–17.
35. DiScipio, R. G., Kurachi, K., and Davie, E. W. (1978) Activation of human factor IX (Christmas factor). *J Clin. Invest.* **61**, 1528–1538.
36. Fujikawa, K., Legaz, M. E., Kato, H., and Davie, E. W. (1974) The mechanism of activation of bovine factor IX (Christmas factor) by bovine factor XIa (activated plasma thromboplastin antecedent). *Biochemistry* **13**, 4508–4516.