

ANALYTICAL CHARACTERIZATION of BIOTHERAPEUTICS

EDITED BY
JENNIE R. LILL
WENDY SANDOVAL



WILEY

Analytical Characterization of Biotherapeutics

Analytical Characterization of Biotherapeutics

Edited by Jennie R. Lill and Wendy Sandoval

*Department of Microchemistry, Proteomics and Lipidomics,
Genentech Inc.
South San Francisco, CA, USA*

WILEY

This edition first published 2017
© 2017 John Wiley & Sons, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Jennie R Lill, Wendy Sandoval to be identified as the authors of the editorial material in this work has been asserted in accordance with law.

Registered Office

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

Editorial Office

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Lill, Jennie R., editor. | Sandoval, Wendy, editor.

Title: Analytical characterization of biotherapeutics / edited by Jennie R. Lill, Wendy Sandoval.

Description: Hoboken, NJ : Wiley, 2017. | Includes bibliographical references and index. |

Identifiers: LCCN 2017013986 (print) | LCCN 2017022128 (ebook) | ISBN 9781119384427 (pdf) |

ISBN 9781119384403 (epub) | ISBN 9781119053101 (hardback)

Subjects: LCSH: Proteins—Therapeutic use. | Proteins—Analysis. | BISAC: SCIENCE / Chemistry / Analytic.

Classification: LCC RM666.P87 (ebook) | LCC RM666.P87 A53 2017 (print) | DDC 615.7—dc23

LC record available at <https://lccn.loc.gov/2017013986>

Cover image: (Background) © Zffoto/Gettyimages;(Illustration) Courtesy of Allison Bruce

Cover design by Wiley

Set in 10/12pt Warnock by SPi Global, Pondicherry, India

Printed in the United States of America

“To Joe and Charlie, thank you for your wonderful smiles and boundless energy”

Jennie R. Lill

“To my daughters, Nicolina and Olivia, who are my daily inspiration”

Wendy Sandoval

Contents

List of Contributors xv

1	Introduction to Biotherapeutics	1
	<i>Jennie R. Lill</i>	
1.1	Introduction	1
1.2	Types of Biotherapeutics and Manufacturing Systems	2
1.3	Types of Analyses Performed	5
1.4	Future perspectives	6
	Acknowledgments	11
	References	11
 2	 Mass Spectrometric Characterization of Recombinant Proteins	 15
	<i>Corey E. Bakalarski, Wendy Sandoval, and Jennie R. Lill</i>	
2.1	Introduction	16
2.1.1	Ionization	16
2.1.1.1	Matrix Assisted Laser Desorption Ionization	17
2.1.1.2	Electrospray Ionization	19
2.1.2	Mass Analyzers for Intact Molecular Weight Measurement of Biotherapeutics	20
2.1.2.1	Time of Flight and Quadrupole Time of Flight Mass Spectrometers	20
2.1.2.2	High-Resolution Intact Mass Measurement and Native MS	21
2.1.2.3	Ion Mobility Spectrometry	22
2.1.3	Software for the Analysis of Intact Molecular Weight Measurements	24
2.1.4	Separation Devices for the Characterization of Biotherapeutics	25
2.1.4.1	High-performance Liquid Chromatography	25
2.1.4.2	Capillary Electrophoresis	26
2.1.4.3	Microfluidic Chromatographic Devices	28
2.2	Peptide Mass Fingerprinting	29

2.3	Tandem Mass Spectrometric Characterization of Biomolecules	30
2.3.1	Bottom-Up MS	33
2.3.2	Proteinformatic Analysis of Bottom-Up Proteomic Data Sets	34
2.3.3	Top-Down MS	36
2.4	Conclusions and Perspectives	37
	References	37
3	Characterizing the Termini of Recombinant Proteins	43
	<i>Nestor Solis and Christopher M. Overall</i>	
3.1	Introduction	44
3.2	Gel Electrophoresis and Edman Sequencing	46
3.3	Mass Spectrometric Approaches for Characterizing True Starts of Proteins	49
3.3.1	Top-Down Approaches	49
3.3.2	Current Caveats in Mass Spectrometric Identification of Protein Termini	54
3.3.3	Bottom-up Approaches for Identification of N- and C-Terminal Peptides	55
3.3.4	Amino Terminal Orientated Mass Spectrometry	56
3.3.5	Determining the True Start of Proteins from ATOMS LC-MS/MS Data	61
3.4	Conclusions	64
	References	66
4	Assessing Activity and Conformation of Recombinant Proteins	73
	<i>Diego Ellerman, Till Maurer, and Justin M. Scheer</i>	
4.1	Introduction	74
4.2	Circular Dichroism	75
4.2.1	Applications of CD	77
4.2.1.1	Thermal Stability Analysis	77
4.2.1.2	Characterization of the Effect of PEGylation	77
4.2.1.3	Formulation and Stability Studies	77
4.2.1.4	Analysis of Biosimilars	78
4.2.2	Technical Improvements	78
4.3	DSC and Isothermal Titration Calorimetry	79
4.3.1	Use of DSC and ITC in Therapeutics Discovery	80
4.3.2	Protein Conjugation	82
4.3.3	Formulation and Stability	82
4.3.4	Analysis of Biosimilars	83
4.4	Hydrogen–Deuterium Exchange–Mass Spectrometry	85

4.4.1	Applications of HDX	86
4.4.1.1	Ligand-induced Conformational Changes and Mapping Interaction Sites	86
4.4.1.2	Applications in Protein Engineering	86
4.4.1.3	Comparability and Biosimilar Studies	88
4.4.1.4	Formulation and Aggregation Analysis	89
4.4.2	Technical Improvements and Challenges	89
4.5	Nuclear Magnetic Resonance	90
4.5.1	Applications of NMR	92
4.5.1.1	Flexible Proteins	92
4.5.1.2	Mapping Protein–Protein Interactions	93
4.5.1.3	Epitope Mapping	94
4.5.1.4	Protein Dynamics	94
4.5.1.5	Protein Conjugates and Complexes	94
4.5.1.6	Posttranslational Modifications	95
4.5.1.7	Biosimilars	95
4.6	Concluding Remarks	96
	References	98
5	Structural Characterization of Recombinant Proteins and Antibodies	111
	<i>Paola Di Lello and Patrick Lupardus</i>	
5.1	Introduction	112
5.2	Antigens, Epitopes, and Paratopes	113
5.2.1	Rationale for Structural Characterization of Epitopes	113
5.3	Choice of Analytical Method for Epitope Mapping	117
5.3.1	EM for Epitope Analysis	117
5.3.2	Epitope and Paratope Mapping by NMR	118
5.3.2.1	Epitope/Paratope Mapping by Chemical Shift Perturbations	119
5.3.2.2	Final Considerations	122
5.3.3	Epitope Mapping by X-ray Crystallography	122
5.4	Recombinant Antigen Generation	123
5.4.1	<i>E. coli</i> Expression of Antigens	124
5.4.2	Insect Cell Expression of Antigens	125
5.4.3	Mammalian Expression of Antigens	126
5.5	N-linked Glycosylation	127
5.5.1	<i>E. coli</i> Expression to Remove Glycosylation as a Factor	128
5.5.2	Manipulating N-linked Glycans on Antigens	128
5.6	Antibody Generation for Crystallography	129
5.7	Crystallization of Antibody/Antigen Complexes	130
5.8	Conclusion	131
	References	131

6	Antibody <i>de novo</i> Sequencing	139
	<i>Natalie Castellana and Adrian Guthals</i>	
6.1	Introduction	139
6.2	Technical Details on Antibody <i>de novo</i> Sequencing	141
6.2.1	Achieving Complete Protein Coverage	141
6.2.2	Achieving High Sequencing Accuracy	142
6.2.3	Handling Protein Modifications	143
6.2.4	Handling Sample Purity	143
6.3	Bioinformatics Workflow	146
6.3.1	Spectral Preprocessing	146
6.3.2	Spectral Alignment-based Approach	146
6.3.3	Sequence Homology-based Approaches	147
6.3.4	Semi-automated and Manual <i>de novo</i> Sequencing	149
6.4	Sequence Validation	149
6.4.1	Mass Spectrometry-based Statistics	149
6.4.2	Intact Mass Comparison	150
6.4.3	Synthetic Peptides	150
6.5	Conclusions	150
	References	151
7	Characterization of Antibody–Drug Conjugates	155
	<i>Yichin Liu</i>	
7.1	Introduction	156
7.2	Characterization of DAR Utilizing MS	157
7.2.1	The Stability of Conjugation Chemistry and the Cleavable Linker of ADC	157
7.2.2	Historical Usage of Hydrophobic Interaction Chromatography in ADC Characterization	158
7.2.3	Intact MS Detection under Denaturing Condition	159
7.2.4	Intact MS Characterization under Native Conditions	159
7.2.5	Middle-down and Bottom-up MS Approach in Mapping Drug Conjugates	161
7.3	Structural Characterization of ADC	162
7.3.1	Ion-Mobility Mass Spectrometry	162
7.3.2	Hydrogen–Deuterium Exchange Mass Spectrometry	163
7.4	Characterization of ADC Catabolism by MS	163
7.5	Conclusions	164
	References	165
8	Characterization of Bispecific or Other Hybrid Molecules	169
	<i>T. Noelle Lombana and Christoph Spiess</i>	
8.1	Introduction	170
8.1.1	Bispecific Antibody Applications	170

8.2	Overview of the Various Bispecific Formats	172
8.2.1	Purification from Mixtures	175
8.2.2	Bispecific Antibodies and Alternative Scaffolds with Tethered Domains	176
8.2.3	Bispecific Molecules with Engineered Mutations	177
8.2.4	Native Bispecific IgG with Dual Binding Behavior	178
8.2.5	Bispecific Antibody Conjugates	179
8.3	Alternatives to Bispecific Antibodies: Antibody Mixtures	179
8.4	Characterization of the Bispecific Molecule	180
8.4.1	Characterization by Bioanalytical Methods	180
8.4.2	Characterization by Mass Spectrometry Methods	183
8.4.2.1	General Considerations	183
8.4.2.2	Purity Analysis of the Final Bispecific Antibody	183
8.4.2.3	Antibody Mixtures	184
8.4.2.4	Increasing Resolution	185
8.4.3	Characterization of Bispecific Antibodies by Binding Assays	185
8.4.4	Developability Assessment of the Bispecific Antibody	186
8.4.4.1	Expression	186
8.4.4.2	Physicochemical Properties	187
8.4.4.3	Chemical Modifications	187
8.4.4.4	Characterization of <i>In Vivo</i> Properties	188
8.5	Conclusions	189
	References	190
9	Bio-Repository	199
	<i>Anne Baldwin, Kurt Schroeder, Lovejit Singh, and Karen Billeci</i>	
9.1	Introduction	199
9.2	Large Molecule Repository Management	202
9.2.1	Informatics	202
9.2.2	Automation	206
9.2.2.1	Automated Refrigerated or Freezer Stores	206
9.2.2.2	Lab Automation	207
9.3	Challenges and Future Perspectives for Working with Diverse Biological Reagent Types	208
	References	209
10	Characterization of Residual Host Cell Protein Impurities in Biotherapeutics	211
	<i>Denise Krawitz, Jason C. Rouse, Justin B. Sperry, Wendy Sandoval, and Martin Vanderlaan</i>	
10.1	Introduction	212
10.2	HCP Measurement and Reporting	212
10.2.1	Antibodies to HCPs	213

10.2.2	Guidance on HCP Limits and Testing	215
10.3	Methods to Characterize Host Cell Impurities	217
10.3.1	HCP-ELISA	217
10.3.2	SDS-PAGE and Western Blots	217
10.3.3	MS Methods for HCP Analysis	219
10.3.3.1	Gel Electrophoresis and MALDI or nanoLC-MS/MS	220
10.3.3.2	Two Dimensional LC-MS/MS	221
10.3.3.3	Targeted MS Analysis	223
10.3.3.4	Ultrahigh-Resolution 1D LC-MS/MS	224
10.3.3.5	Top-down Proteomics	227
10.4	Use of HCP-ELISA and Orthogonal 1D LC-MS/MS in Practice	228
10.4.1	Pros and Cons of MS for Orthogonal HCP Analysis	231
10.4.2	Considerations and MS Evolution	232
10.5	Risk of HCPs Present in Products	232
10.6	Conclusions	233
	References	234
11	Analytical Tools for Biologics Molecular Assessment	239
	<i>Wilson Phung, Wendy Sandoval, Robert F. Kelley, and Jennie R. Lill</i>	
11.1	Introduction to Molecular Assessment	240
11.2	Molecular Assessment	243
11.3	Biotherapeutic Stability	244
11.3.1	Deamidation and Isomerization of Asparagine	246
11.3.2	Oxidation	246
11.4	Physical Degradation	248
11.5	Yield and Structural Stability	249
11.6	Posttranslational Modifications	250
11.7	Analytical Techniques	251
11.8	Summary	252
	References	254
12	Glycan Characterization: Determining the Structure, Distribution, and Localization of Glycoprotein Glycans	257
	<i>John B. Briggs</i>	
12.1	Introduction	258
12.2	Glycan Labeling	264
12.3	Compositional Analysis	266
12.3.1	Neutral Sugar Analysis	267
12.3.2	Sialic Acid Analysis	269
12.4	Glycan Release	272
12.4.1	Release of N-linked Glycans	272
12.4.2	Release of O-linked Glycans	274

12.5	Determining Sites of Glycosylation	276
12.5.1	MS-Based Screening for Glycopeptides	278
12.5.2	Identification of Glycosylation Sites by Analysis of Native Glycopeptides	279
12.5.3	Identification of N-linked Glycosylation Sites by Enzymatic Labeling of Glycosylation Sites	281
12.5.4	Identification of O-linked Glycosylation Sites by Chemical Labeling of Glycosylation Sites	283
12.5.5	Identification of Glycosylation Sites by Edman Degradation	285
12.6	Determining N-linked Glycan Distribution	286
12.6.1	Assessing Glycan Distribution by MS	287
12.6.1.1	Assessing Glycan Distribution by Mass Spectrometric Analysis of Glycoproteins	287
12.6.1.2	Assessing Glycan Distribution by Mass Spectrometric Analysis of Glycopeptides	294
12.6.1.3	Determining Glycan Distribution by Mass Spectrometric Analysis of Native Glycans	294
12.6.1.4	Determining Glycan Distribution by Mass Spectrometric Analysis of Derivatized Glycans	298
12.6.2	Assessing Glycan Distribution by Chromatography and CE	300
12.6.2.1	Analysis of N-linked Glycans by CE	300
12.6.2.2	Analysis of N-linked Glycans by HILIC	303
12.6.2.3	Determining Glycan Distribution by HPAEC	305
12.7	Comparison of Methods Used in Determining Glycan Distribution	307
12.8	Assessing N-linked Glycan Structure	309
12.8.1	Characterization of Glycan Structure Using Standards and Enzymatic Studies	309
12.8.2	Characterization of Glycan Linkage by Methylation Analysis	310
12.8.3	Characterization of Glycan Structure by MS ²	312
12.8.4	Characterization of Glycan Structure by NMR	317
	References	320

Index	333
--------------	------------

List of Contributors

Corey E. Bakalarski

Departments of Microchemistry,
Proteomics & Lipidomics and
Bioinformatics & Computational
Biology
Genentech Inc.
South San Francisco, CA, USA

Anne Baldwin

Departments of Biomolecular
Engineering and Proteomics &
Biological Resources
Genentech Inc.
South San Francisco, CA, USA

Karen Billeci

Department of Proteomics &
Biological Resources
Genentech Inc.
South San Francisco, CA, USA

John B. Briggs

Department of Protein Analytical
Chemistry
Genentech Inc.
South San Francisco, CA, USA

Natalie Castellana

Digital Proteomics LLC
La Jolla
and
Mapp Biopharmaceutical, Inc.
San Diego, CA, USA

Paola Di Lello

Department of Structural Biology
Genentech Inc.
South San Francisco, CA, USA

Diego Ellerman

Department of Protein Chemistry
and Structural Biology
Genentech Inc.
South San Francisco, CA, USA

Adrian Guthals

Digital Proteomics LLC
La Jolla
and
Mapp Biopharmaceutical, Inc.
San Diego, CA, USA

Robert F. Kelley

Drug Delivery Department
Genentech Inc.
South San Francisco, CA, USA

Denise Krawitz

Department of Analytical
Operations
Genentech Inc.
South San Francisco, CA, USA

Jennie R. Lill

Department of Microchemistry,
Proteomics & Lipidomics
Genentech Inc.
South San Francisco, CA, USA

Yichin Liu

Department of Biochemical and
Cellular Pharmacology
Genentech Inc.
South San Francisco, CA, USA

T. Noelle Lombana

Department of Antibody
Engineering
Genentech Research and Early
Development
South San Francisco, CA, USA

Patrick Lupardus

Department of Structural Biology
Genentech Inc.
South San Francisco, CA, USA

Till Maurer

Department of Protein Chemistry
and Structural Biology
Genentech Inc.
South San Francisco, CA, USA

Christopher M. Overall

Center for Blood Research
University of British Columbia
Vancouver, British Columbia,
Canada

Wilson Phung

Department of Microchemistry,
Proteomics & Lipidomics
Genentech Inc.
South San Francisco, CA, USA

Jason C. Rouse

Department of Analytical
Research and Development
Pfizer
Cambridge, MA, USA

Wendy Sandoval

Department of Microchemistry,
Proteomics & Lipidomics
Genentech Inc.
South San Francisco, CA, USA

Justin M. Scheer

Antibody Engineering
Department
Boehringer Ingelheim, Ridgefield,
Connecticut; Department
of Protein Chemistry and
Structural Biology
Genentech Inc.
South San Francisco, CA, USA

Kurt Schroeder

Departments of Biomolecular
Engineering and Proteomics &
Biological Resources
Genentech Inc.
South San Francisco, CA, USA

Lovejit Singh

Departments of Biomolecular
Engineering and Proteomics &
Biological Resources
Genentech Inc.
South San Francisco, CA, USA

Nestor Solis

Center for Blood Research
University of British Columbia
Vancouver, British Columbia,
Canada

Justin B. Sperry

Department of Analytical Research
and Development
Pfizer
Cambridge, MA, USA

Christoph Spiess

Department of Antibody Engineering
Genentech Research and Early
Development
South San Francisco, CA, USA

Martin Vanderlaan

Department of Analytical
Operations
Genentech Inc.
South San Francisco, CA, USA

1

Introduction to Biotherapeutics

Jennie R. Lill

Department of Microchemistry, Proteomics & Lipidomics, Genentech Inc., South San Francisco, CA, USA

Abbreviations

ADAs	antidrug antibodies
ADC	antibody–drug conjugate
ADCC	antibody-dependent cell-mediated cytotoxicity
CDR	complementary-determining region
Fab	antigen binding fragment
Fc	cystallizable fragment
NMR	nuclear magnetic resonance
PEG	polyethyleneglycol
PTM	posttranslational modification

1.1 Introduction

Biotherapeutics, also known as biologics, include protein-based and nucleic acid-based drugs that are commonly derived by recombinant expression in living organisms although a few are made by chemical synthesis. This book focuses on the characterization of protein-based biotherapeutics, exploring the various analytical technologies that have enabled in-depth molecular characterization while discussing current triumphs and limitations.

The first human protein therapeutic derived from recombinant DNA technology was human insulin (Humulin®) created at Genentech, developed by Eli Lilly, and approved by the US Food and Drug Administration (FDA) in 1982. Since that time, major advancements in both recombinant DNA technology

and recombinant protein production have contributed to the development of several hundred biotherapeutics [1] including relatively simple molecules such as interferons, insulin, and the human growth hormone to more complexly engineered moieties including ADCs such as trastuzumab emtansine [2] and brentuximab vedotin [3].

Unlike conventional small molecule (chemical) drugs such as aspirin, antibiotics, and various chemo-therapeutics, the manufacturing process for biotherapeutics is typically far more cumbersome as they are larger compounds with more complex structures and their production can be extremely sensitive to changes in fermentation and environmental conditions. In addition, biotherapeutics are often less stable than many small molecules and can be prone to aggregation [4] or deamidation, oxidation, and other modifications [5]. Since the manufacturing of biotherapeutics is often dependent upon the host cells of living organisms, complex process development is required to ensure reproducible fermentations, isolation, and characterization [6].

1.2 Types of Biotherapeutics and Manufacturing Systems

There are several different types of marketed biotherapeutics including antibody-based drugs, anticoagulants, blood factors, bone morphogenetic proteins, engineered protein scaffolds, enzymes, Fc (cystallizable fragment) fusion proteins, growth factors, hormones, interferons, interleukins, and thrombolytics (Figure 1.1).

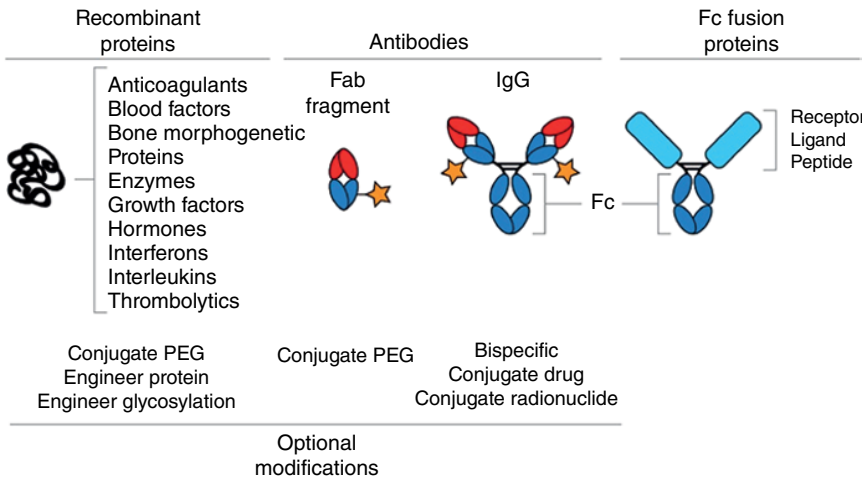


Figure 1.1 Various categories of the main types of biotherapeutics currently marketed. Source: Carter [7]. Reproduced with permission of Elsevier.

Antibody-based drugs represent the largest and most rapidly expanding class of biotherapeutics [1]. Figure 1.2 shows the diverse mechanisms by which the antibody structure can be modified to increase its biotherapeutic potential.

Humanized and other chimeric versions of these antibodies now dominate the market [11] and in the past 5 years have accounted for nearly 30% of all approvals. Various antibody isotypes are now being explored to provide a wealth of functional diversity that is present through the various IgG subclasses that can be exploited to improve clinical safety and performance by increasing stability, reducing adverse events, modulating effector functions, and by the engagement of two antigens by a single antibody [8]. Several variants that have been Fc engineered for reduced effector function have entered the clinic, for example, Eculizumab, a novel engineered IgG isotype, IgG2m4, with reduced Fc functionality. IgG2m4 is engineered based on the IgG2 isotype with four key amino acid residue changes derived from IgG4 (H268Q, V309L, A330S, and P331S). This antibody was demonstrated to have an overall reduction in complement and Fc gamma receptor binding in *in vitro* binding analyses while maintaining the normal *in vivo* serum half-life in rhesus [12].

Biosimilars (biologically identical antibodies, for example) and so-called biobetters (moieties with improved properties such as pharmacodynamic (PD) and pharmacokinetic (PK) readouts, higher potency, longer half-lives, and less immunogenicity, for example) are also starting to emerge, which presents new challenges in terms of testing for the presence of liabilities such as degradative properties, changes in immunogenicity through addition of novel contaminant proteins from new manufacturing processes, and so on. New formats such as glucagon-like peptide 1GLP fused proteins, for example, Eperzan (albiglutide) [13], and PEGylated proteins such as Pegriid (e.g., peginterferon beta-1a) [14] offer improved PK or PD properties but also increased analytical challenges due to their larger masses and increased heterogeneity.

Typically, expression of non-mAb biotherapeutics has been performed in *Escherichia coli* or a noneukaryotic system. This has many advantages for biotherapeutics that are not reliant on PTMs for their optimal activity. Over the years, however, there has been a gradual increase in the prevalence of mammalian expression systems. Of the mammalian expression systems, the Chinese hamster ovary (CHO) cell-based model (reviewed by Krawitz and Sandoval in Ref. [11]) remains the most employed expression system with a smaller percentage of therapeutics manufactured in other mammalian cell lines such as the murine myeloma line, NSO, and baby hamster kidney cells [15, 16]. Nonmammalian eukaryotic expression systems such as yeast [17] are also utilized, each again presenting their own challenges with regard to the correct PTM of the protein, occasionally adding to adverse properties [18, 19].

More recently transgenic animal production systems (e.g., expression of recombinant products in the animals' milk [19, 20], rabbits, and goats) have been explored as a means of biopharmaceutical production although to date

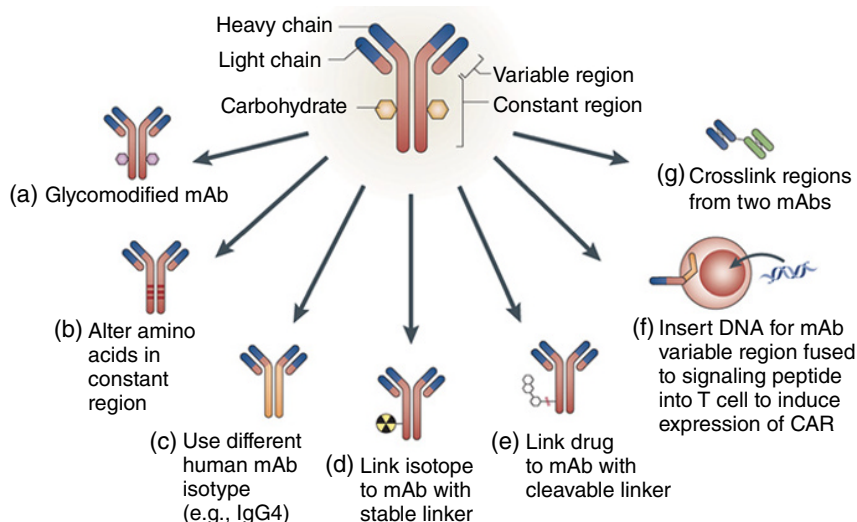


Figure 1.2 Monoclonal antibody (mAb) structure can be modified on the basis of the desired mechanism of action. Immunoglobulin G1 (IgG1) is the most effective naturally occurring human IgG isotype at mediating antibody-dependent cell-mediated cytotoxicity (ADCC). Glycomodified afucosylated mAbs (part a) (such as Obinutuzumab) demonstrate enhanced binding to IgG Fc receptors (FcγRs) and enhanced ADCC. In addition antibody-dependent cellular phagocytosis, a process mediated by macrophages, can also occur [8]. Afucosylated mAbs are produced using cell lines that lack the enzymes responsible for fucosylation. Modifying the amino acid sequence of mAb Fc (part b), as was done to produce ocaratuzumab [9], can also result in enhanced binding to FcγRs and enhanced ADCC. For mechanisms of action in which ADCC is not desirable, IgG4 may be a more appropriate isotype, as IgG4 mAbs do not mediate ADCC to the same degree as IgG1 (part c) although this isotype can still engage macrophage effector function *via* nanomolar affinity binding to FcγRI. Nivolumab, an IgG4 mAb that blocks programmed cell death protein 1 (PD1) on T cells, is one such example. Producing radioimmunoconjugates involves linking the radioisotope to the mAb. A stable linker is most desirable (part d) to limit the leakage of the free radioactive isotope. Conversely, optimal antibody–drug conjugates (ADCs) use a cleavable linker (part e). To avoid nonspecific toxicity, it is desirable for drugs used in ADCs to be cytotoxic once inside the target cell but nontoxic when bound to the mAb in the circulation. Linkers that are pH-sensitive or enzymatically cleaved are now a standard component of ADCs. Chimeric antigen receptor (CAR) T cells get their specificity from mAb variable regions but are a form of gene, not protein, therapy. They are produced by inserting DNA coding for the mAb variable region fused to DNA coding for signaling peptides into T cells (part f). Some bispecific antibodies lack a functional constant region so that they do not nonspecifically crosslink activating receptors and activate T cells (part g). The lack of a constant region on such constructs results in a short half-life, thus requiring continuous infusion to achieve the desired exposure. Source: Weiner [10]. Reproduced with permission of Nature Publishing Group.

there are many challenges associated with this type of biotherapeutic production with few benefits. Throughout this book the challenges of characterizing both the biotherapeutic moiety itself and the contaminant proteins such as CHO-derived proteins are discussed.

1.3 Types of Analyses Performed

Throughout this book a variety of analytical procedures are described. Many of them have been implemented for characterizing biotherapeutic molecules for as long as these moieties have existed. Others have evolved as the need arises. One such example of developing such sets of tools to answer a newly arisen problem is for the *de novo* sequencing of antibodies [21, 22]. Occasionally antibodies are discovered that are of great interest for preclinical testing, for which the cDNA or any genetic information is not available. In these scenarios, researchers have to sequence antibodies at the protein level, one amino acid at a time, and then reverse engineer the antibodies to the nucleotide level. In Chapter 6 Castellana and Guthals provide technical details and review the innovative approaches employed to quickly gain sequence information through a *de novo* approach.

As well as sequence information at the amino acid level, PTM profiling is also an important element in characterizing biotherapeutics [23]. There are a plethora of cotranslational modifications and PTMs that play key roles in the folding of proteins, in their secretion, and in their ultimate stability and effector functionality *in vivo*.

Glycosylation is important both for antibody secretion by B-cells and for *in vivo* antibody effector function. Glyco-engineering is a rapidly growing field, whereby glycosylation sites and composites are engineered to produce antibodies with specific glycoforms which may have an effect on therapeutic efficacy. Obinutuzumab (Gazyva®) [24], for example, is a humanized therapeutic monoclonal antibody that binds to an epitope on the B cell antigen, CD20. This antibody is engineered in a platform that allows control of the proteins' glycosylation, in this case the platform enforces the overexpression of two glycosylation enzymes MGAT3 and the golgi mannosidase 2. This results in the generation of antibodies with bisected nonfucosylated sugars, thereby increasing the antibodies' ability to activate natural killer cells. This means that Obinutuzumab can induce cell death through a dual mechanism of action, both by the antibody directly binding to B cells and by antibody-mediated cytotoxicity by recruiting the immune system to attack B cells. Some types of glycosylation are sometimes not beneficial. For example, Cetuximab, a chimeric mouse-human IgG1 monoclonal antibody against the epidermal growth factor receptor (EGFR) approved for use in colorectal cancer and squamous-cell carcinoma of the head and neck has a high prevalence of hypersensitivity reactions which has been attributed to cross reactivity to a V domain

glycosylation site. In some patients severe adverse events have been observed including anaphylaxis, which was found to be due to the generation of patient-specific antibodies to the galactose- α -1,3-galactose modification [25].

Protein terminal modifications have the effect of modifying a protein's function, half-life, or cellular localization. Pyroglutamate formation, for example, is a highly prevalent modification whereby glutamine and glutamate at the N termini of recombinant monoclonal antibodies can cyclize spontaneously to pyroglutamate (pE) *in vitro* [26]. Proteolytic processing is also an irreversible modification that affects the vast majority of proteins, often with great functional consequences. Intracellular proteolytic processing has distinct effects on the functionality of proteins and can either abrogate or antagonize function, modify half-life, or also determine cellular localization. During protein synthesis, manufacturing, purification, and storage proteolysis events can occur, thereby changing a protein's functionality or stability [27]. Either through direct mass spectrometric analysis as reviewed in Chapter 2, through a variety of historic analytical techniques such as gel electrophoresis or Edman degradation, or by employing a variety of new biochemical-based methodologies for determining the termini of recombinant proteins as reviewed in Chapter 3, the determination of proteolytic processing remains a key analytical need for the characterization of biotherapeutic moieties.

Beyond linear sequence determination, structural analyses are also instrumental in the overall characterization of biotherapeutics. The biomolecular architecture is a vital component in dictating the specificity and overall efficacy of therapeutic proteins. The higher order structure (HOS) of a protein includes the secondary, tertiary, and quaternary structures of a protein that are required for its function. There is a diverse range of biophysical methods including circular dichroism, isothermal calorimetry, which are available for the characterization of a protein HOS, each of them with associated benefits and limitations. Related to conformational analysis is structural analysis as it pertains to epitope and paratope mapping. Again, several well-established techniques such as nuclear magnetic resonance (NMR) [28] and X-ray crystallography as well as some newer techniques such as mass spectrometric-based structural tools [29] including hydrogen deuterium exchange are described in Chapters 4 and 5.

1.4 Future perspectives

Nature has provided us with various types of protein scaffolds to explore as frameworks for building new types of biotherapeutics and there is a growing field of using these scaffolds as alternatives to antibodies [30]. Each of these types of engineered molecular structures offers new advantages in terms of stability and specificity. One example of this is the cystine knot mini proteins/peptides (knottins); these are peptide-based alternative molecules to

monoclonal antibodies which are raised/designed against tumor-associated receptors and other antigens of interest. Knottins contains a disulfide-bonded core that exhibits a high level of resistance to proteolysis and increased thermal stability. Knottins emerged as an attractive molecular candidate for drug development as they fill the niche between small molecule drug design and protein biologics. Knottins have the potential to bind clinical targets with both high selectivity and affinity [31]. There are several naturally occurring knottins that have been approved as biotherapeutics for the treatment of pain [32] and irritable bowel syndrome and for tumor imaging purposes [33].

Elucidating disulfide bonding patterns of any biomolecules, but in particular a structure which relies on disulfide bonding patterns for their folding, stability, and activity, is an important part of molecular characterization. A variety of techniques can be employed for doing this from simple intact molecular weight measurement to more complex top-down proteomic protocols [34] and these methodologies continue to mature as more of these types of molecules emerge on the market.

Another growth area for biotherapeutics is increasing the molecules' *in vivo* half-life. For therapeutics that involves frequent or uncomfortable delivery, for example, injectable ocular therapeutics, or to make drugs that have poor PD properties more tolerable, there is a strong drive to create molecules that have increased *in vivo* stability (and potentially decreased immunogenicity). Some common mechanisms of molecular half-life extension include the generation of Fc fusion proteins, the formulation of biomolecules into various different nanoparticle systems, or the addition of stabilizing peptides. For a comprehensive list of these efforts please refer to Table 1.1.

The addition of albumin to stabilize the half-life of proteins has also been explored. Albumin is the most abundant plasma protein in humans and mice and is highly soluble, extremely stable, and has a circulatory half-life of ~20 days in man [36]. By fusing albumin to therapeutic proteins, these molecules become less susceptible to renal filtration and circulatory clearance and one of the main reasons albumin has a long half-life is due to its ability to bind to FcRn. Association, conjugation, or fusion of therapeutic drugs to albumin has been shown to correlate with superior PK.

In addition to albumin addition, PEGylation, the process of adding polyethyleneglycol (PEG) chains to a molecule by incubating a reactive derivative of PEG with the biomolecule of interest, also appears to work universally for improving therapeutic protein *in vivo* half-life [37]. PEGylation also increases the molecule's hydrodynamic size, thereby prolonging its time in circulation by reducing renal clearance. In addition, PEGylation can also make hydrophobic drugs with poor PK properties more water-soluble. There are a number of PEGylated molecules on the market and many companies are exploring this concept on a wider range of biotherapeutics [38, 39]. Characterizing these modified proteins, such as albumin-conjugated or PEGylated species, brings along increased analytical challenges, some of which are covered by Bakalarski *et al.* in Chapter 2 and by Ellerman *et al.* in Chapter 4.

Table 1.1 The circulatory half-life of a therapeutic protein can be extended by several strategies depending upon the endogenous clearance mechanism of the drug.

Protease degradation	Novel/alternative delivery strategies avoiding need to extend half-life
N-terminal acetylation or C-terminal amidation	Controlled-release depot (subcutaneous, intramuscular, intravenous); e.g. Bydureon—poly(lactic-co-glycolic acid) (PLGA) microspheres (Alkermes)
Nonnatural amino acids at labile sites	
Cyclization using disulfide bonds	
Microspheres or nanoparticles	
Increase size/hydrodynamic volume to prevent clearance by kidneys	Increase size/hydrodynamic volume to prevent clearance by kidneys and attach to protein with a long half-life
Di- or multimers	Attach to Fc (the natural antibody constant region)
Attach PEG—conjugation	Genetic fusion; e.g. Enbrel (Amgen), Mimetibody™ (Centocor), SynFusion (Biogen Idec/Syntonix)
Advanced PEGylation—modification of peptides (and prodrugs) by attaching with specific polymer chains	Covalent attachment (site-specific) (CovX/Pfizer)
Site-specific PEGylation (polytherics)—enables more selective PEGylation, reducing likelihood of protein deactivation upon conjugation and reducing immunogenicity	Attach to human serum albumin
Glyco-PEGylation (Neose)	Genetic fusion, albufuse® (Novozymes Biopharma/GSK/Teva/CSL)
Releasable PEGylation (Enzon)	Conjugation (Novozymes Biopharma, ConjuChem, Cardiovox)
Protein “rPEG”—genetic fusion	Albumin binding peptides/proteins/affinity tags (Genentech/Roche, Ablynx, Philochem, Affibody, Adnexus/BMS)
Poly-glycine	Albumin binding single domain antibodies fused to bioactive peptides (GSK/Domantis, Ablynx, Haptogen/Wyeth/Pfizer)
PASylation (XL-protein)	Albumin binding fatty acids (Novo Nordisk)
XTEN (Amunix)	
Other	
Hyaluronic acid (Novozymes)	

Table 1.1 (Continued)

Increase size/hydrodynamic volume to prevent clearance by kidneys	Increase size/hydrodynamic volume to prevent clearance by kidneys and attach to protein with a long half-life
Hydroxyethyl Starch (HESylation® — Fresenius Kabi)	
Polysialic acid (PolyXen® —Xenetic Biosciences)	
Elastin-like polypeptide (ELP) technology (Phase Bio Pharmaceuticals Inc.)	

Source: Sleep [35]. Reproduced with permission of Elsevier.
 These include the reduction in the endogenous degradation of the drug; slow release/depot formulations; increasing the hydrodynamic volume of the drug by attachment of a large bulky polymer or extension of the therapeutic protein by addition of a linear but unstructured protein; or the addition of a large structured protein which additionally can take advantage of the FcRn-mediated recycling.

PEG polymers themselves are often highly heterogeneous and the increase in molecular weight that occurs when biomolecules are PEGylated means that more specialized mass spectrometric techniques or other analytical methods need to be employed [40]. With an increase in these types of complex molecules being developed for improved drug stability, the analytical challenges posed with their characterization also increase in complexity.

In addition to analyzing the intact biomolecules both *in vitro* and *in vivo* for assessing molecular stability, PK and PD properties, additional types of analyses to assess traits such as immunogenicity are starting to emerge. All biotherapeutics, including monoclonal antibodies and their derivatives, are immunogenic to varying degrees in various patients, with chimeric antibodies representing more of an immunogenic risk than humanized antibodies. These mouse human chimeric antibodies can induce some patients to develop antidrug antibody (ADA) responses [41]. Certain factors are known to influence biotherapeutic immunogenicity including structural homology with respect to human amino acid sequences and various PTMs. Although frameworks can be designed to minimize the potential for immunogenicity, the complementarity-determining regions (CDRs) of antibodies and other variable domains of biotherapeutics can be highly sequence variable making it difficult to predict how immunogenic a reagent maybe [42]. Several companies have adopted strategies to assess immunogenicity retrospectively for therapeutics that have shown an ADA response [43]. Prospectively, a T cell activation assay, whereby CD4+ T cells are monitored for activation by antigen-presenting cells (APCs) loaded

with biotherapeutics, can be employed along side a major histocompatibility complex associated peptide proteomics (MAPPs) assay. The MAPPs assay involves the *in silico* prediction of DR4 or other MHC class II peptide presentation using one of several programs such as SYPETHI, NETMHC, and so on. These algorithms predict potential T cell epitopes derived from the therapeutic protein. Additional analyses whereby MHC class II complexes are immune-precipitated from cells, peptides isolated and analyzed by mass spectrometry are also sometimes performed to complement the results from *in silico* predictions. These types of analyses are starting to be incorporated into early molecular assessment workflows to minimize the potential risk of adverse immunogenicity of new molecules prior to ADAs being reported.

In addition to these types of analyses, additional complexities arise due to chemically modified/conjugated biotherapeutics. Antibody–drug conjugates and antibody–antibiotic conjugates are gaining momentum in the clinic and these hybrid molecules that are composed of an antibody, a cleavable linker, and a chemotherapeutic or potent antibiotic molecule raise their own set of analytical challenges. In Chapter 7, Liu describes the importance of measuring drug-to-antibody ratios (DARs) [44] and the types of chromatographies and mass spectrometric techniques employed for characterizing these complex molecules. Liu describes the applications of increasingly diversified mass spectrometric techniques employed for characterizing the integrity of ADCs from the perspective of production to *in vitro* and *in vivo* testing as well as in terms of elucidating the mechanisms of ADC biotransformations.

So, now we have a variety of different biotherapeutic material in the research environment, manufacturing, and the clinic. How do we organize such materials? Prior to official lot validation that occurs once a biomolecule hits our pipeline, how do we ensure that the antigens we are using for immunization are from the same batch of cells, have been purified in the same manner and already characterized for correct protein sequence, folding, and structure? Biorepositories [45, 46] are growing in popularity and necessity in biotechnology companies. Official curation, stable storage, and easy retrieval of various lots of cDNAs, antigens, antibodies, ADCs and all the associated modifications are now becoming commonplace to allow researchers more consistency in the quality and reproducibility of their assay and other results.

As analytical technologies increase in sensitivity and complexity, low-level heterogeneity of therapeutic biomolecules that were previously unknown becomes revealed. Many of these modifications will be irrelevant in terms of the protein's stability or function; however, when reporting such heterogeneity to the FDA and other regulatory bodies one has to take care to not provide information that might be misconstrued. So how sensitive should the analysis be? At what percentile of the overall nonmodified protein should a mutation or modification be before it is reported or deemed important for functionality?

Some of the techniques involved in the molecular assessment of therapeutic biomolecules are described by Phung *et al.* in Chapter 11. These are all questions that analytical chemists and biochemists face as we embark on the design of new biotherapeutic moieties and see the emergence of biosimilars.

Acknowledgments

Thank you to Wendy Sandoval and Paul J. Carter for informative discussions and editing.

References

- 1 Walsh G (2014) Biopharmaceutical benchmarks 2014. *Nature biotechnology* 32(10):992–1000.
- 2 Niculescu-Duvaz I (2010) Trastuzumab emtansine, an antibody-drug conjugate for the treatment of HER2+ metastatic breast cancer. *Current opinion in molecular therapeutics* 12(3):350–360.
- 3 Borate U, *et al.* (2016) Treatment of CD30-positive systemic mastocytosis with brentuximab vedotin. *Leukemia research* 44:25–31.
- 4 Ratanji KD, Derrick JP, Dearman RJ, & Kimber I (2014) Immunogenicity of therapeutic proteins: influence of aggregation. *Journal of immunotoxicology* 11(2):99–109.
- 5 Jostock T & Knopf HP (2012) Mammalian stable expression of biotherapeutics. *Methods in molecular biology* 899:227–238.
- 6 Li F, Vijayasankaran N, Shen AY, Kiss R, & Amanullah A (2010) Cell culture processes for monoclonal antibody production. *mAbs* 2(5):466–479.
- 7 Carter PJ (2011) Introduction to current and future protein therapeutics: a protein engineering perspective. *Experimental cell research* 317(9):1261–1269.
- 8 Brezski RJ & Georgiou G (2016) Immunoglobulin isotype knowledge and application to Fc engineering. *Current opinion in immunology* 40:62–69.
- 9 Bowles JA, *et al.* (2006) Anti-CD20 monoclonal antibody with enhanced affinity for CD16 activates NK cells at lower concentrations and more effectively than rituximab. *Blood* 108:2648–2654.
- 10 Weiner GJ (2015) Building better monoclonal antibody-based therapeutics. *Nature reviews. Cancer* 15(6):361–370.
- 11 Henricks LM, Schellens JH, Huitema AD, & Beijnen JH (2015) The use of combinations of monoclonal antibodies in clinical oncology. *Cancer treatment reviews* 41(10):859–867.
- 12 An Z, *et al.* (2009) IgG2m4, an engineered antibody isotype with reduced Fc function. *mAbs* 1(6):572–579.

- 13 Trujillo JM & Nuffer W (2014) GLP-1 receptor agonists for type 2 diabetes mellitus: recent developments and emerging agents. *Pharmacotherapy* 34(11):1174–1186.
- 14 Khan UT, Tanasescu R, & Constantinescu CS (2015) PEGylated IFN β -1a in the treatment of multiple sclerosis. *Expert opinion on biological therapy* 15(7):1077–1084.
- 15 Eldering JA, Felten C, Veilleux CA, & Potts BJ (2004) Development of a PCR method for mycoplasma testing of Chinese hamster ovary cell cultures used in the manufacture of recombinant therapeutic proteins. *Biologicals: journal of the International Association of Biological Standardization* 32(4):183–193.
- 16 Thomsen C, *et al.* (1993) A pharmacological characterization of the mGluR1 alpha subtype of the metabotropic glutamate receptor expressed in a cloned baby hamster kidney cell line. *Brain research* 619(1–2):22–28.
- 17 Lee YJ & Jeong KJ (2015) Challenges to production of antibodies in bacteria and yeast. *Journal of bioscience and bioengineering* 120(5):483–490.
- 18 Frenzel A, Hust M, & Schirrmann T (2013) Expression of recombinant antibodies. *Frontiers in immunology* 4:217.
- 19 Spiesberger K, *et al.* (2015) Large-Scale Purification of r28M: A Bispecific scFv Antibody Targeting Human Melanoma Produced in Transgenic Cattle. *PLoS one* 10(10):e0140471.
- 20 Redwan el RM (2009) Animal-derived pharmaceutical proteins. *Journal of immunoassay & immunochemistry* 30(3):262–290.
- 21 Castellana NE, *et al.* (2011) Resurrection of a clinical antibody: template proteogenomic de novo proteomic sequencing and reverse engineering of an anti-lymphotoxin-alpha antibody. *Proteomics* 11(3):395–405.
- 22 Pham V, *et al.* (2006) De novo proteomic sequencing of a monoclonal antibody raised against OX40 ligand. *Analytical biochemistry* 352(1):77–86.
- 23 Heffner KM, *et al.* (2014) Exploiting the proteomics revolution in biotechnology: from disease and antibody targets to optimizing bioprocess development. *Current opinion in biotechnology* 30:80–86.
- 24 Ratner M (2014) Genentech's glyco-engineered antibody to succeed Rituxan. *Nature biotechnology* 32(1):6–7.
- 25 Chung CH, *et al.* (2008) Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *The New England journal of medicine* 358(11):1109–1117.
- 26 Liu H, *et al.* (2014) In vitro and in vivo modifications of recombinant and human IgG antibodies. *mAbs* 6(5):1145–1154.
- 27 Ryan BJ & Henahan GT (2013) Overview of approaches to preventing and avoiding proteolysis during expression and purification of proteins. *Current protocols in protein science* Chapter 5:Unit5 25.
- 28 Stamos J, *et al.* (2004) Convergent recognition of the IgE binding site on the high-affinity IgE receptor. *Structure* 12(7):1289–1301.

- 29 Beck A, Sanglier-Cianferani S, & Van Dorsselaer A (2012) Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Analytical chemistry* 84(11):4637–4646.
- 30 Skrlec K, Strukelj B, & Berlec A (2015) Non-immunoglobulin scaffolds: a focus on their targets. *Trends in biotechnology* 33(7):408–418.
- 31 Moore SJ, *et al.* (2013) Engineered knottin peptide enables noninvasive optical imaging of intracranial medulloblastoma. *Proceedings of the national academy of sciences of the United States of America* 110(36):14598–14603.
- 32 Murray JK, *et al.* (2015) Engineering potent and selective analogues of GpTx-1, a tarantula venom peptide antagonist of the Na(V)1.7 sodium channel. *Journal of medicinal chemistry* 58(5):2299–2314.
- 33 Jiang L, *et al.* (2014) A radiofluorinated divalent cystine knot peptide for tumor PET imaging. *Molecular pharmaceuticals* 11(11):3885–3892.
- 34 Reinwarth M, Avrutina O, Fabritz S, & Kolmar H (2014) Fragmentation follows structure: top-down mass spectrometry elucidates the topology of engineered cystine-knot miniproteins. *PLoS one* 9(10):e108626.
- 35 Sleep D, Cameron J, & Evans LR (2013) Albumin as a versatile platform for drug half-life extension. *Biochimica et Biophysica acta* 1830(12):5526–5534.
- 36 Sleep D (2015) Albumin and its application in drug delivery. *Expert opinion on drug delivery* 12(5):793–812.
- 37 Schmidt PG, Campbell KM, Hinds KD, & Cook GP (2007) PEGylated bioactive molecules in biodegradable polymer microparticles. *Expert opinion on biological therapy* 7(9):1427–1436.
- 38 Moore DJ, Adi Y, Connock MJ, & Bayliss S (2009) Clinical effectiveness and cost-effectiveness of pegvisomant for the treatment of acromegaly: a systematic review and economic evaluation. *BMC endocrine disorders* 9:20.
- 39 Goel N & Stephens S (2010) Certolizumab pegol. *mAbs* 2(2):137–147.
- 40 Gong C, *et al.* (2014) Development and validation of an LC-MS/MS assay for the quantitation of a PEGylated anti-CD28 domain antibody in human serum: overcoming interference from antidrug antibodies and soluble target. *Bioanalysis* 6(18):2371–2383.
- 41 Tatarewicz SM, *et al.* (2014) Strategic characterization of anti-drug antibody responses for the assessment of clinical relevance and impact. *Bioanalysis* 6(11):1509–1523.
- 42 van Schie KA, Wolbink GJ, & Rispens T (2015) Cross-reactive and pre-existing antibodies to therapeutic antibodies—Effects on treatment and immunogenicity. *mAbs* 7(4):662–671.
- 43 Karle A, Spindeldreher S, & Kolbinger F (2016) Secukinumab, a novel anti-IL-17A antibody, shows low immunogenicity potential in human in vitro assays comparable to other marketed biotherapeutics with low clinical immunogenicity. *mAbs* 8(3):536–550.

- 44 Bryant P, *et al.* (2015) In vitro and in vivo evaluation of cysteine rebridged trastuzumab-MMAE antibody drug conjugates with defined drug-to-antibody ratios. *Molecular pharmaceuticals* 12(6):1872–1879.
- 45 Baker M (2012) Biorepositories: Building better biobanks. *Nature* 486(7401):141–146.
- 46 Siwek M (2015) An overview of biorepositories-past, present, and future. *Military medicine* 180(10 Suppl):57–66.

2

Mass Spectrometric Characterization of Recombinant Proteins

Corey E. Bakalarski^{1,2}, Wendy Sandoval¹, and Jennie R. Lill¹

¹ Department of Microchemistry, Proteomics & Lipidomics, Genentech Inc., South San Francisco, CA, USA

² Department of Bioinformatics & Computational Biology, Genentech Inc., South San Francisco, CA, USA

Abbreviations

CCS	collisional cross section
CHCA	alpha-cyano-4-hydroxycinnamic acid
DC	direct current
DHB	2,5-dihydroxybenzoic acid
EMR	extended mass range
ESI	electrospray ionization
FAIMS	field asymmetric ion mobility spectrometry
HPLC	high-performance liquid chromatography
IEC	ion exchange chromatography
IMS	ion mobility spectrometry
ISD	in-source decay
MALDI	matrix assisted laser desorption ionization
MS	mass spectrometry
MS/MS	tandem mass spectrometry
<i>m/z</i>	mass-to-charge ratio
PSM	peptide-spectral match
QTOF	quadrupole time of flight
RF	radio frequency
SEC	size exclusion chromatography

2.1 Introduction

A mass spectrometer in its most simplistic terms is an instrument capable of determining the mass of individual atoms or molecules. In reality one measures the mass (m)-to-charge (z) ratio (m/z) of an ion and since atoms of each element have different masses, the accumulative molecular composition of a molecule (e.g., a protein) can be determined by its m/z .

Antibodies and other recombinant proteins are one of the leading types of therapeutic moieties for the treatment of human disease and often have advantages over traditional small molecule-based drugs in that they are highly specific and typically exhibit fewer off-target side effects. All therapeutic related biomolecules, whether they are antigens to generate immunization campaigns or the actual therapeutic protein itself, may exhibit heterogeneity due to various modifications, either synthetically induced or naturally occurring during the manufacturing process [1].

N-terminal heterogeneity, generated from either truncated or elongated termini from exposure to aberrant proteolysis events during the secretion process, is a common observation on recombinant proteins [2, 3]. Co- or posttranslational modifications (PTMs) on proteins due to expression in and purification from eukaryotic expression systems can occur and often their presence or absence can be indicative of suboptimal manufacturing processes and potentially affect the molecule's activity or stability. Kinases in their active conformation, for example, are often auto-phosphorylated [4]; for IgG antibodies, glycosylation status is important for dictating correct Fc effector function [5]. Due to these inherent heterogeneities extensive analytical characterization is imperative to ensure reproducible, reliable, and safe production of these therapeutic proteins. The fastest, easiest, most accurate, and arguably most sophisticated analytical method to measure each of these events is through intact molecular weight mass measurements using mass spectrometry (MS).

2.1.1 Ionization

The first step in the mass spectrometric analysis of proteins is ionization. After purification and often dialysis into an appropriate MS-compatible buffer (e.g., a composition of an organic solvent such as methanol or acetonitrile in combination with water and an ion pairing reagent), the samples are introduced into the mass spectrometer [6]. To facilitate entry into the mass analyzer, samples must be ionized using one of several techniques; the most common types of ionization for large biomolecules are (i) matrix assisted laser desorption ionization (MALDI) [7], where a sample is spotted onto a MALDI plate for subsequent analysis, and (ii) electrospray ionization (ESI) [8, 9] whereby the sample is injected *via* either direct infusion or elution from an orthogonal analytical

device such as a high-performance liquid chromatograph (HPLC) into an ESI source. Ionization is a critical step in the mass spectrometric analysis of any molecule, as it is this process that converts the molecule(s) of interest from a liquid or solid phase into a gaseous state, which is then amenable to introduction into and detection by the mass analyzer. The methods described herein are considered “soft” ionization techniques, as the actual process of ionization is amenable to keeping the biomolecule of interest intact (and thus retaining molecular weight information) during its conversion to a gaseous ion and entry into the mass spectrometer.

2.1.1.1 Matrix Assisted Laser Desorption Ionization

MALDI is a soft ionization technique typically coupled to MS that was first described by Hillenkamp and Karas [7] and a schematic of a typical MALDI ionization process is depicted in Figure 2.1. During MALDI analysis, the molecule of interest is cocrystallized with an organic matrix which is chosen based on the nature and size of the analyte. For example, intact mass measurement on a MALDI instrument is most often performed using 3,5-dimethoxy-4-hydroxycinnamic acid (a.k.a sinapinic acid) to minimize fragmentation, whereas peptides are commonly analyzed in the presence of alpha-cyano-4-hydroxycinnamic acid (CHCA) or 2,5-dihydroxybenzoic acid (DHB) matrices. The matrix absorbs ultraviolet light generated from a laser (most typically a nitrogen laser light at a wavelength

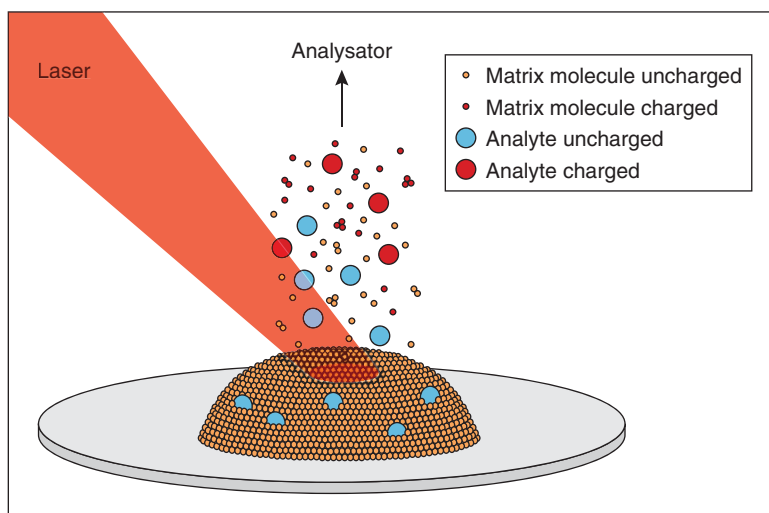


Figure 2.1 MALDI ionization. The uncharged analyte is mixed with an uncharged matrix molecule. After receiving energy *via* the laser pulse the now charged matrix donates a proton to the analyte *via* desorption, therefore allowing the analyte to become charged. The charged analyte is guided through the MALDI mass spectrometer for m/z analysis.